

Ultrasound-driven 4D MRI

Alina Giger^{1,2}, Marc Stadelmann³, Frank Preiswerk⁴,
Christoph Jud^{1,2}, Valeria De Luca^{5‡}, Zarko Celicanin^{1,6},
Oliver Bieri^{1,6}, Rares Salomir⁷ and Philippe C. Cattin^{1,2}

¹ Department of Biomedical Engineering, University of Basel, Allschwil, Switzerland

² Center for medical Image Analysis & Navigation, University of Basel, Allschwil, Switzerland

³ Institute for Surgical Technology & Biomechanics, University of Bern, Bern, Switzerland

⁴ Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

⁵ Computer Vision Laboratory, ETH Zurich, Zurich, Switzerland

⁶ Division of Radiological Physics, Department of Radiology, University of Basel Hospital, Basel, Switzerland

⁷ Faculty of Medicine, Radiology, University of Geneva, Geneva, Switzerland

E-mail: alina.giger@unibas.ch

Abstract. We present an ultrasound-driven 4D magnetic resonance imaging (US-4DMRI) method for respiratory motion imaging in the thorax and abdomen. The proposed US-4DMRI comes along with a high temporal resolution, and allows for organ motion imaging beyond a single respiratory cycle. With the availability of the US surrogate both inside and outside the MR bore, 4D MR images can be reconstructed for 4D treatment planning and online respiratory motion prediction during radiotherapy.

US-4DMRI relies on simultaneously acquired 2D liver US images and abdominal 2D MR multi-slice scans under free respiration. MR volumes are retrospectively composed by grouping the MR slices corresponding to the most similar US images. We present two different US similarity metrics: an intensity-based approach, and a similarity measure relying on predefined fiducials which are being tracked over time. The proposed method is demonstrated on MR liver scans of 8 volunteers acquired over a duration of 5.5 min each at a temporal resolution of 2.6 Hz with synchronous US imaging at 14 Hz to 17 Hz. Visual inspection of the reconstructed MR volumes revealed satisfactory results in terms of continuity in organ boundaries and blood vessels. In quantitative leave-one-out experiments, both US similarity metrics reach the performance level of state-of-the-art navigator-based approaches.

Keywords: 4D imaging, 4D MRI, hybrid ultrasound/MR imaging, radiotherapy, respiratory motion

1. Introduction

Time-resolved volumetric imaging, or 4D imaging, is a key factor in respiration induced organ motion quantification and finds use in several clinical applications, such as high-intensity focused ultrasound (HIFU) ablation (Schwenke, Strehlow, Haase, Jenne, Tanner, Langø, Loeve, Karakitsios, Xiao, Levy et al. 2015) or radiotherapy of thoracic and abdominal tumours (Buerger, Clough, King, Schaeffter & Prieto 2012, Stemkens, Tijssen, de Senneville, Lagendijk & van den Berg 2016). In radiotherapy, accurate knowledge about respiratory motion is crucial for both 4D treatment planning and precise tumour tracking during dose delivery. Therefore, 4D imaging is inevitable when high-precision radiotherapy techniques should be translated from static to mobile treatment targets. For example, in active scanning proton therapy, respiration-induced organ deformations in combination with dynamic dose delivery cause interplay effects that adversely affect treatment quality and may lead to inhomogeneous dose distributions (Phillips, Pedroni, Blattmann, Boehringer, Coray & Scheib 1992, Bert & Durante 2011). In this context, 4D imaging can form the basis for respiratory motion models and eventually enable highly conformal radiation procedures.

At present, 4D computed tomography (4D CT) is the preferred imaging modality for treatment planning in radiotherapy (Hugo & Rosu 2012). The main difficulty, however, lies in the inevitable trade-off between radiation dose delivered to the patient and image quality. As 4D CT methods generally assume a constant respiration period, respiration variabilities often lead to serious image artifacts and volume inconsistencies (Yamamoto, Langner, Loo, Shen & Keall 2008, Johnston, Diehn, Murphy, Loo & Maxim 2011). In order to overcome these issues, recent developments have increasingly focused on 4D magnetic resonance imaging (4D MRI). The absence of radiation dose in MRI allows for longer scan durations and superior soft-tissue contrast enhances tumour/tissue delineation (Neto, Elazzazzi, Altun & Semelka 2008). Moreover, MRI offers the possibility of selecting the imaging plane in the main direction of motion, i.e. in sagittal direction, which further reduces motion artifacts (Liu, Yin, Chang, Czito, Palta, Bashir, Qin & Cai 2014).

Although possible, dynamic 3D MRI is not favourable for respiratory motion imaging since it suffers from low temporal or spatial resolution if large volumes need to be scanned (Blackall, Ahmad, Miquel, McClelland, Landau & Hawkes 2006, Dinkel, Hintze, Tetzlaff, Huber, Herfarth, Debus, Kauczor & Thieke 2009). Consequently, various methods based on retrospective binning of partial image data have been proposed. 2D MR images at varying slice positions are continuously acquired while the patient's respiratory motion is recorded either using external surrogates, such as belt respiration sensors (Tryggstad, Flammang, Han-Oh, Hales, Herman, McNutt, Roland, Shea & Wong 2013), or internal image-based metrics which are derived from MR signals directly. Examples of the latter are body area (Cai, Chang, Wang, Segars & Yin 2011, Liu et al. 2014), body boundaries in combination with low-frequency components in the Fourier domain (Hui, Wen, Stemkens, Tijssen, van den Berg,

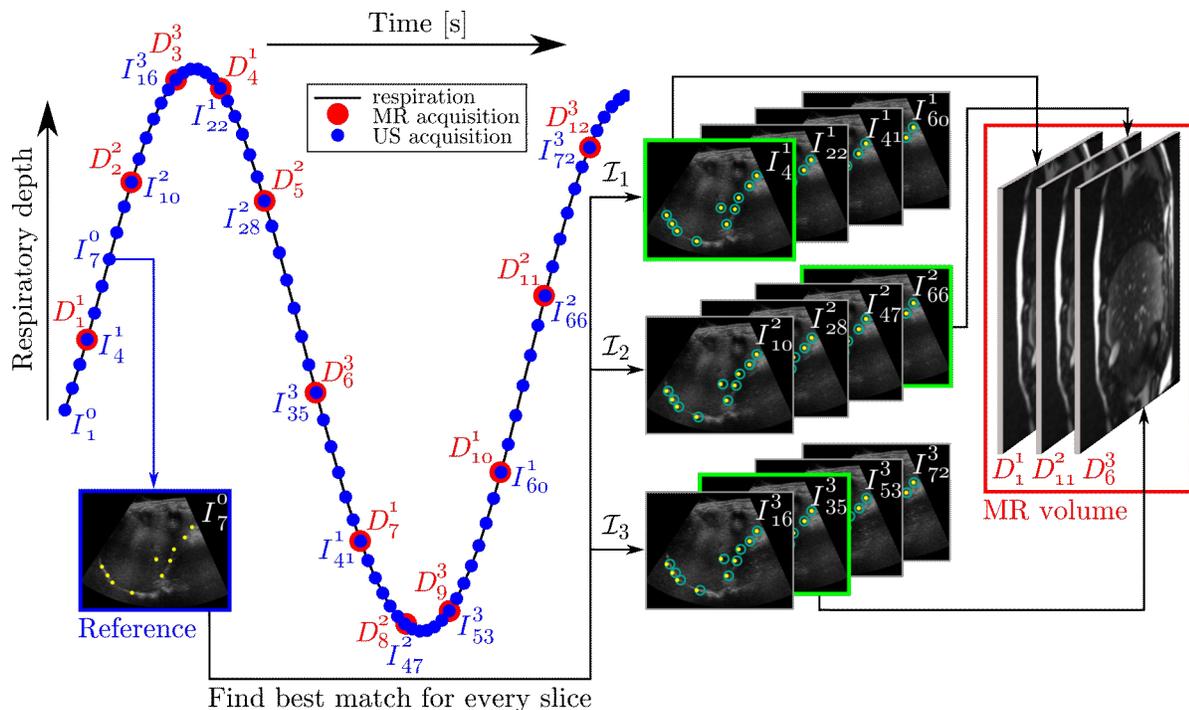


Figure 1: Principle of US-4DMRI: simplified acquisition scheme with $n = 3$ slices per MR volume and US fiducial tracking as surrogate data. For each point of the respiratory cycle (blue dots), one MR volume is reconstructed. To do so, the US images I_{iUS}^s , which were acquired simultaneously with an MR scan D_{iMR}^s , are grouped into subsets \mathcal{I}_s according to the MR slice position s . For a given reference (blue frame), the best matching US image for each slice position (green frames) is identified based on automatically tracked US fiducials (marked as yellow dots and green circles). Finally, the corresponding MR scans are stacked into a 3D image (red frame).

Hwang & Beddar 2016), mutual information (Paganelli, Summers, Bellomi, Baroni & Riboldi 2015), manifold learning (Wachinger, Yigitsoy, Rijkhorst & Navab 2012), and deformation fields of interleaved navigator slices (Von Siebenthal, Gamper, Boesiger, Lomax, Cattin et al. 2007). Based on the surrogate data, the MR scans are attributed to different respiratory states and 3D MR volumes are reconstructed. Similarly, retrospective sorting of 3D-acquired k -space data based on MR navigators or self-navigation techniques has also been proposed (Buerger et al. 2012, Stemkens, Tijssen, de Senneville, Heerkens, van Vulpen, Lagendijk & van den Berg 2015, Rank, Heußer, Buzan, Wetscherek, Freitag, Dinkel & Kachelrieß 2017, Mickevicius & Paulson 2017).

External surrogates rely on the correlation between the external signal and internal organ motion which is not always valid (von Siebenthal, Székely, Lomax & Cattin 2007, Stemkens et al. 2015). Conversely, image-based metrics and self-navigation are powerful tools for MR-guided interventions and radiotherapy (Stemkens et al. 2016). However, when it comes to proton therapy, the connection between pretreatment 4D MRI and motion estimation during dose delivery is not straightforward as these

surrogates are not available outside the MR bore. To overcome these issues, hybrid ultrasound (US) and MRI acquisition offers a promising approach since US imaging provides internal organ motion data while being available both inside and outside the MR bore. Only recently, single-element US signals, MRI and regression-based methods have been combined for generating synthetic MR images (Preiswerk, Toews, Hoge, Chiou, Panych, Wells III & Madore 2015, Preiswerk, Toews, Cheng, Chiou Jr, Mei, Schaefer, Hoge, Schwartz, Panych & Madore 2016).

In this work, we introduce 2D US surrogates of the liver and the diaphragm for retrospectively stacking abdominal MR images, as shown in figure 1. By using hybrid US/MRI imaging, we propose an ultrasound-driven 4D MRI (US-4DMRI) with increased temporal resolution and reduced overall acquisition duration as compared to navigator-based approaches. We investigate the applicability of US surrogates for retrospective 4D MRI by formulating two different US similarity metrics, and compare the performance of US-4DMRI with the navigator-based 4D MRI in (Von Siebenthal, Gamper, Boesiger, Lomax, Cattin et al. 2007).

2. Materials and methods

2.1. Simultaneous US and MR data acquisition

The data used in the present work have been acquired as part of a related study (Preiswerk, De Luca, Arnold, Celicanin, Petrusca, Tanner, Bieri, Salomir & Cattin 2014). For details on the experimental setup and acquisition parameters, the reader is referred to (Petrusca, Cattin, De Luca, Preiswerk, Celicanin, Auboiroux, Viallon, Arnold, Santini, Terraz et al. 2013) and (Preiswerk et al. 2014).

Simultaneous US/MR acquisitions were performed in a 1.5 T MR scanner. We applied a balanced steady-state free precession MR pulse sequence and, for validation purposes, an interleaved acquisition scheme (Von Siebenthal, Gamper, Boesiger, Lomax, Cattin et al. 2007), where data and so-called navigator slices were acquired alternately. 2D multi-slice MR scans were acquired at a frame rate of $f_{\text{MR}} \approx 5.2$ Hz, or $f_{\text{MR}}^d = f_{\text{MR}}/2 \approx 2.6$ Hz for data slices d only, with a spatial resolution of 1.82 mm and a slice thickness of 4.02 mm. The US frame rate was 14 Hz to 17 Hz.

Eight US/MR liver data sets were recorded for a total duration of 5.5 min each. No subject showed signs of hepatic disorders. Sagittal MR slices were positioned to cover the right liver lobe and scanned the region of interest (ROI) in sequential order. The navigator was acquired at a fixed location in the sagittal plane in order to minimize out-of-plane motion of liver vessels during respiration (Rohlfing, Maurer, O'dell & Zhong 2004). The MR-compatible US probe was either placed on the 6th or 7th right intercostal space, respectively, or frontally on the abdomen resulting in sagittal-oblique images. The volume acquisition time t_{vol} , defined as the time needed for a complete scan of the target volume, ranged from 10.1 s to 11.7 s. The signal properties for all subjects k are summarized in table 1.

Table 1: Signal properties of the liver data sets. At each slice position $s \in \{1, \dots, n\}$, a total of $|\mathcal{I}_s|$ MR data scans were acquired; the overall number of data scans is $N^d = n|\mathcal{I}_s|$; the volume acquisition time is $t_{vol} = n/f_{MR}^d$.

| Subject k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|--|------|------|------|------|------|------|------|------|
| <i>General information</i> [-] | | | | | | | | |
| No. of respiratory cycles | 97 | 61 | 71 | 94 | 96 | 52 | 58 | 63 |
| No. of MR slice positions, n | 30 | 30 | 30 | 30 | 30 | 30 | 26 | 28 |
| No. of MR images per slice pos., $ \mathcal{I}_s $ | 28 | 28 | 28 | 28 | 28 | 28 | 30 | 28 |
| No. of MR data images, N^d | 840 | 840 | 840 | 840 | 840 | 840 | 780 | 784 |
| Volume acquisition time, t_{vol} [s] | 11.7 | 11.7 | 11.6 | 11.7 | 11.6 | 11.7 | 10.1 | 10.9 |
| <i>Temporal resolution</i> [Hz] | | | | | | | | |
| f_{MR} (data and navigator scans) | 5.12 | 5.12 | 5.16 | 5.12 | 5.16 | 5.12 | 5.16 | 5.16 |
| f_{MR}^d (data scans only) | 2.56 | 2.56 | 2.58 | 2.56 | 2.58 | 2.56 | 2.58 | 2.58 |
| f_{US} | 16.0 | 16.0 | 17.0 | 17.0 | 17.0 | 14.0 | 15.0 | 15.0 |

2.2. Ultrasound-driven 4D MRI

The main principle of US-4DMRI is shown in figure 1. Let f_{MR}^d and f_{US} represent the MR and US acquisition frame rates. Let i_{MR} and i_{US} define the indices of MR data scans and US images, respectively. Then,

- $D_{i_{MR}}^s$ denotes the MR scan acquired at time $t_{i_{MR}} = i_{MR}/f_{MR}^d$ and slice position s ,
- $I_{i_{US}}^s$ denotes the US image corresponding to $D_{i_{MR}}^s$, acquired at time $t_{i_{US}} = i_{US}/f_{US}$,
- $I_{i_{US}}^o$ denotes the remaining US images at times $t_{i_{US}}$ with no corresponding MR scan.

Note that $t_{i_{MR}}$ and $t_{i_{US}}$ do not necessarily coincide. However, since f_{US} is considerably higher than f_{MR}^d , we assume that the MR slice $D_{i_{MR}}^s$ and its temporally closest US image $I_{i_{US}}^s$ represent the liver in a sufficiently similar state.

For every $t_{i_{US}}$ of the respiratory motion (blue dots in figure 1), a complete MR volume is reconstructed, resulting in a 4D image frame rate of $f_{4D} = f_{US}$. We generate n subsets \mathcal{I}_s each of them comprising all US images associated with the given MR slice position s , that is $\mathcal{I}_s = \{I_{i_{US}}^{s'} \mid s' = s\}$. The US image $I_{i_{US}}$, being either $I_{i_{US}}^s$ or $I_{i_{US}}^o$, serves as reference (blue frame in figure 1) in order to find the best matching image $I_{j_{US}}^s \in \mathcal{I}_s$ for each slice position s (green frames in figure 1) by solving

$$j^* = \arg \min_j S(I_{i_{US}}, I_{j_{US}}^s), \quad (1)$$

with similarity measure S . Then, the corresponding MR scans $D_{j_{MR}}^s$ build the final MR volume for time $t_{i_{US}}$. The optimal US image $I_{j_{US}}^s$ is determined by exhaustive search on the subset \mathcal{I}_s . Alternatively, the N_{avg} best matching scans could be averaged in order to increase the signal-to-noise ratio (Von Siebenthal, Gamper, Boesiger, Lomax, Cattin et al. 2007).

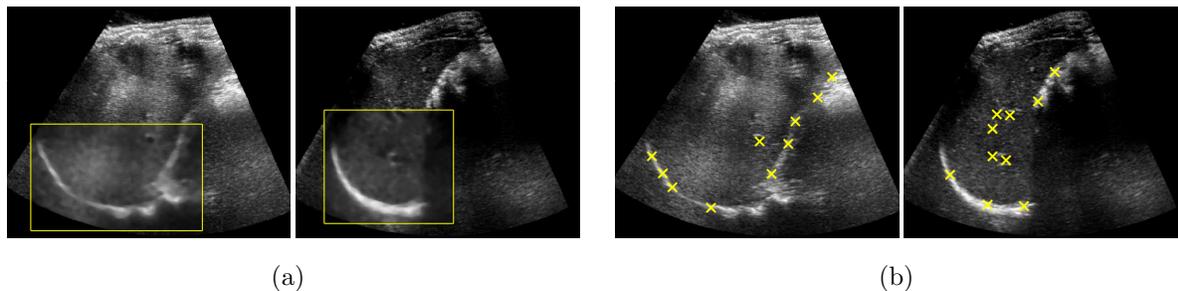


Figure 2: US surrogate data for two different subjects. (a) The intensity-based similarity metric was computed as the MAD of the image area inside the highlighted rectangle. (b) The selected fiducials are marked as yellow crosses. Since the acoustic window of the left image reveals only few vessel structures, most fiducials were placed on the liver boundary. Conversely, in the right image the fiducials were equally distributed on vessel centers and the diaphragm.

2.2.1. Similarity measure Automatically tracked US fiducials have shown to be suitable surrogates for the respiratory motion model in (Preiswerk et al. 2014). In this work, we examine the applicability of US feature tracking in the context of 4D MRI. Below, we additionally introduce a comparison method which is directly computed on image intensities.

Intensity-based Dealing with mono-modal surrogate images, the mean absolute intensity difference (MAD) represents a computationally efficient similarity metric. In order to further reduce computational time and since the respiratory motion closely correlates with the motion of the diaphragm, we selected a ROI around the liver boundary. A median filter with a fixed window size of 10 was then applied to reduce speckles. Let \mathcal{R} represent the ROI and $\tilde{I}_{iUS}(r) \in \{0, \dots, 255\}$ the intensity value of the median filtered US image at pixel location r . Then, the similarity term can be formulated as

$$S_{\text{MAD}}(I_{iUS}, I_{jUS}^s) = \sum_{r \in \mathcal{R}} |\tilde{I}_{iUS}(r) - \tilde{I}_{jUS}^s(r)|. \quad (2)$$

Figure 2a shows sample US images with the cropped and median-filtered ROI overlaid.

Fiducial tracking For each US sequence, we manually selected between 5 and 11 fiducials on a reference image at end-exhalation as shown in figure 2b. These fiducials correspond to dominant anatomical structures, e.g. points on the diaphragm and vessel centers, and their positions were tracked over time. Considering the target application of US-4DMRI in real-time tumour tracking, a computationally fast image registration method was used (De Luca, Tanner & Székely 2012), which enables real-time position prediction by taking advantage of the recurrence in organ position between different respiratory cycles. We assume the liver at times t_{iUS} and t_{jUS} to be in a comparable

state if the fiducials of $I_{i_{\text{US}}}$ and $I_{j_{\text{US}}}^s$ coincide. Moreover, similar velocities foster a grouping of images at comparable characteristic points of the respiratory cycle: the spatial variation between two temporally consecutive US images near end-exhalation or end-inhalation will be much smaller as compared to the spatial difference of a mid-cycle pair of images. In addition, including the direction of the trajectories avoids matching exhalation frames during inhalation, and vice versa.

Let $\mathbf{x}_l(t_{i_{\text{US}}}) \in \mathbb{R}^2$ denote the position of the l -th fiducial at time $t_{i_{\text{US}}}$, with $l \in \{1, \dots, m\}$ and m being the total number of fiducials. Similarly, let $\Delta \mathbf{x}_l(t_{i_{\text{US}}}) = \mathbf{x}_l(t_{i_{\text{US}}}) - \mathbf{x}_l(t_{i_{\text{US}}-1})$ be the spatial variation of the l -th fiducial between two temporally consecutive US images. We define the similarity measure as

$$S_{\text{Fid}}(I_{i_{\text{US}}}, I_{j_{\text{US}}}^s) = \sum_{l=1}^m \|\mathbf{x}_l(t_{i_{\text{US}}}) - \mathbf{x}_l(t_{j_{\text{US}}})\| + \omega \sum_{l=1}^m \|\Delta \mathbf{x}_l(t_{i_{\text{US}}}) - \Delta \mathbf{x}_l(t_{j_{\text{US}}})\|, \quad (3)$$

with Euclidean norm $\|\cdot\|$ and weight ω . The latter was empirically determined using grid search on a training set.

2.3. Evaluation methods

2.3.1. Leave-one-out experiments Given the image pair $I_{i_{\text{US}}}^s$ and $D_{i_{\text{MR}}}^s$, we identified the best matching US image $I_{j_{\text{US}}}^s \in \mathcal{I}_s \setminus \{I_{i_{\text{US}}}^s\}$ such that equation (1) held. The actual MR image $D_{i_{\text{MR}}}^s$ at $t_{i_{\text{US}}}$ served as ground truth and was compared to the optimal solution $D_{j_{\text{MR}}}^s$ acquired at the same slice position s but at a different point of time, $t_{j_{\text{US}}}$.

The data was split into a training set acquired at even slice positions $s_{\text{even}} \in \{2, 4, \dots, n\}$, and a validation set with $s_{\text{odd}} \in \{1, 3, \dots, n-1\}$. Quantitative validation involved leave-one-out experiments for all s_{odd} and for each image in the associated subset, that is $\forall I_{i_{\text{US}}}^{s_{\text{odd}}} \in \mathcal{I}_{s_{\text{odd}}}$. The difference between the selected MR image $D_{j_{\text{MR}}}^s$ and the true left-out image $D_{i_{\text{MR}}}^s$ was quantified as the MAD of the pixel intensities contained in a predefined ROI around the liver. For each subject, the ROI was manually specified as the minimum bounding box which contains the entire liver for all respiratory states. In order to achieve optimal image contrast within the ROI, we applied a low-pass image filter.

Note that the design of the leave-one-out experiments implies a reconstruction rate of f_{MR}^d since the MR ground truth images $D_{i_{\text{MR}}}^s$ cannot be acquired at higher frame rates. The supplementary material to this paper provides movies of 4D MR images at different frame rates ranging from $f_{4\text{D}} = f_{\text{MR}}^d$ to $f_{4\text{D}} = f_{\text{US}}$.

2.3.2. Slice selection Having both US images and MR navigators as surrogates for the given data, we investigated whether the same MR images were selected for either of the two methods. Assume we aim to reconstruct an MR volume at time $t_{i_{\text{US}}}$. Let $u \in \{1, \dots, |\mathcal{I}_s|\}$ denote the index of the optimal MR scan $D_{j_{\text{MR}}}^s$ within the subset

Table 2: Statistical results of leave-one-out experiments for tuning of ω . The best results are highlighted in bold font.

| ω | 0.00 | 0.25 | 0.50 | 0.75 | 1.00 | 1.25 | 1.50 |
|--------------------|------|-------------|-------------|------|------|------|------|
| mean | 2.32 | 2.22 | 2.22 | 2.23 | 2.25 | 2.26 | 2.27 |
| median | 2.14 | 2.08 | 2.08 | 2.10 | 2.12 | 2.13 | 2.13 |
| standard deviation | 0.93 | 0.81 | 0.81 | 0.82 | 0.83 | 0.84 | 0.84 |

\mathcal{I}_s determined by US-4DMRI. Analogously, let $v \in \{1, \dots, |\mathcal{I}_s|\}$ represent the index of the selected MR scan for navigator-based stacking. We computed the joint sample distribution $p(u, v) \in [0, 1]$ for the complete 4D image, not taking into account the cases where a ground-truth MR image $D_{i_{\text{MR}}}^s$ exists. The probabilities were computed for the validation set only.

3. Experiments and results

The reconstruction was based on $|\mathcal{I}_s| = 30$ images per slice position for Subject 7 and $|\mathcal{I}_s| = 28$ images for the remaining subjects. We omitted noise reduction through averaging and set $N_{\text{avg}} = 1$ for all results below. For an in-depth analysis of the fiducial tracking approach, we refer to (Preiswerk et al. 2014) and (De Luca et al. 2012).

3.1. Parameter tuning

The weight ω in (3) was determined based on leave-one-out experiments on the training set. Table 2 summarizes the mean, median and standard deviations of the resulting error distributions. The results indicate that the stacking performance improves when taking the spatial variations $\Delta \mathbf{x}_l$ into account ($\omega > 0$). Table 2 further shows that the proposed similarity measure is robust against the choice of ω . Equal results were obtained for weighting factors $\omega = 0.25$ and $\omega = 0.5$, and the statistical measures increased only slightly for $\omega \geq 0.75$. In the following, $\omega = 0.25$ was selected.

3.2. Slice selection

For the sake of conciseness, we focus on US-4DMRI relying on fiducial tracking for slice selection analysis. Figure 3a shows from left to right the reference US image $I_{i_{\text{US}}}^s$, the best matching US image $I_{j_{\text{US}}}^s$ and the absolute intensity difference between the two for a representative case. The tracked US fiducials are indicated with yellow dots and green circles, respectively. Figure 3b displays the corresponding MR ground truth $D_{i_{\text{MR}}}^s$, the optimal data scan $D_{j_{\text{MR}}}^s$ and the difference between these. Visual comparison between the US ground truth and the optimal solution reveals a good alignment of the tracked fiducials. Only small intensity differences in both US and MR images can be observed.

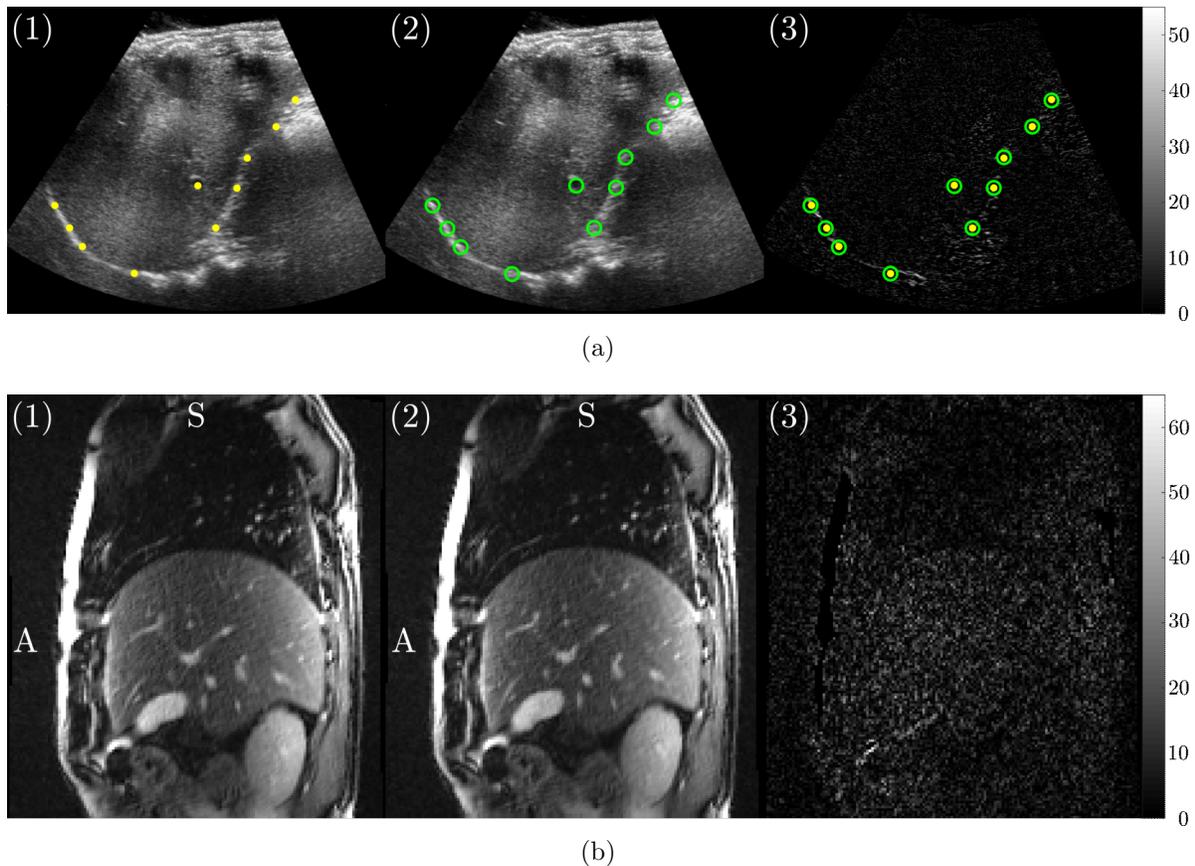


Figure 3: Slice selection based on US fiducial tracking for Subject 1 and slice position $s = 9$. The images represent from left to right, (1) the ground truth image, (2) the selected optimal image and (3) the absolute intensity difference between (1) and (2). The US and MR image intensities are in the range of 0–255; the intensity differences are defined by the color bars in (3). (a) US images with tracked fiducials, and (b) the corresponding MR scans acquired in sagittal direction.

When comparing the 4D MR images obtained by US-4DMRI and the navigator-based approach, respectively, we find that overall 22.9% of the selected MR scans are identical. Figure 4 shows the 2D histogram of $p(u, v)$ for each subject k separately. The probability value in the upper left corner indicates the percentage of identical MR images for the respective subject, i.e. $p_k := p_k(u = v)$. Note that the histograms do not give information about the superiority of either method. Yet, the joint histograms reveal some interesting insights: for Subject 6, pronounced off-diagonal lines can be observed on either side of the diagonal (highlighted with arrows). The periodic occurrence of the off-diagonal lines suggests that Subject 6 shows a regular respiration pattern with distinctive organ states. Conversely, a less pronounced respiration pattern can be observed for Subject 1 which shows a more uniformly distributed joint probability $p_1(u, v)$.

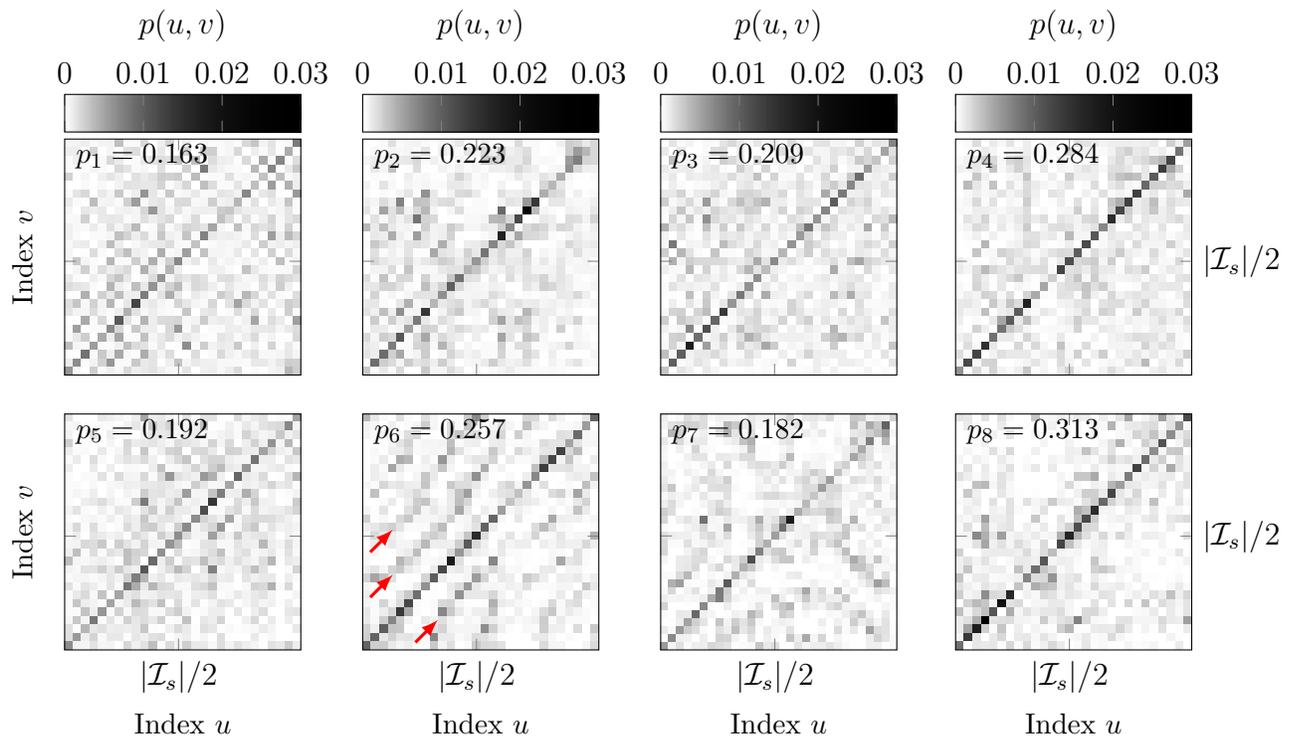


Figure 4: Slice selection illustrated as joint histogram for navigator-based 4D MRI and US-4DMRI, respectively, and for each subject k separately. The probability values p_k denote the percentage of equally selected MR images by both methods, i.e. $p_k = p_k(u = v)$.

3.3. Qualitative validation

Figure 5 shows orthogonal cuts through reconstructed MR volumes of Subject 1 and Subject 6, respectively, at end-exhalation and end-inhalation of a sample respiratory cycle. From left to right, the results for the three stacking methods are presented: (1) navigator-based 4D MRI (Von Siebenthal, Gamper, Boesiger, Lomax, Cattin et al. 2007), (2) US-4DMRI relying on the fiducials' positions and velocities, and (3) intensity-based US-4DMRI. Furthermore, random stacking (4) is shown as upper baseline. If properly stacked, smooth organ boundaries and vessel structures in axial and coronal plane are to be expected.

In both figures, a vertical stripe artifact can be observed as indicated with arrows in the leftmost columns. This artifact is a consequence of the interleaved acquisition scheme and indicates the position of the navigator: as every second MR scan was acquired at the same position, the tissue in that region was excited more frequently and therefore underwent saturation. In the case of US-4DMRI, the stripe artifact will no longer be visible.

Figure 5a and figure 5b reveal additional stacking artifacts at end-inhalation. For US-4DMRI, the dominant vascular structure and the anterior abdominal wall are

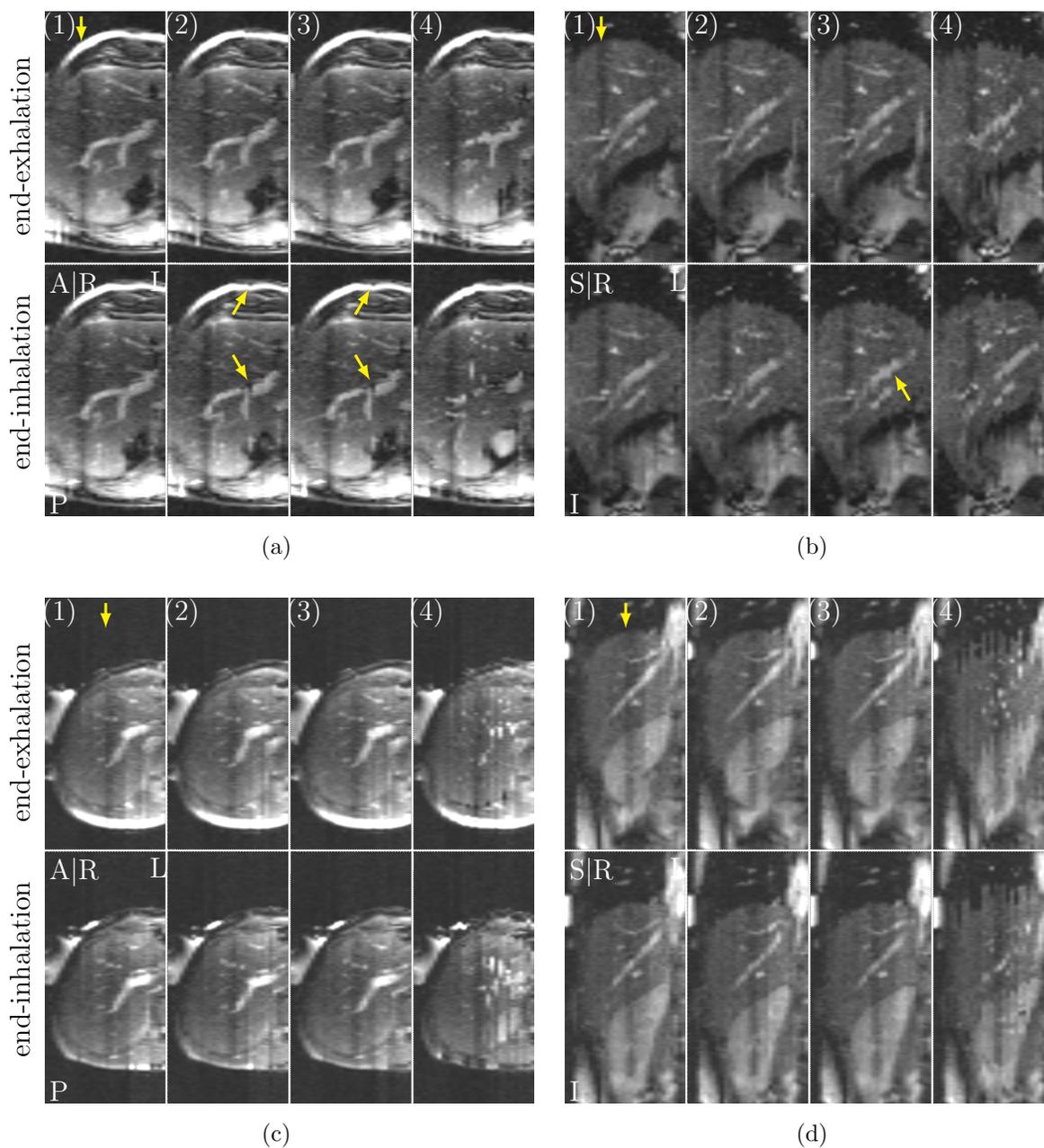


Figure 5: Axial and coronal cuts through stacked MR volumes of (a),(b) Subject 1 and (c),(d) Subject 6 for (1) navigator-based 4D MRI, (2) US-4DMRI relying on fiducial tracking, (3) intensity-based US-4DMRI, and (4) random stacking. The images show the liver at end-exhalation (upper rows) and end-inhalation (lower rows) of a sample respiratory cycle. The arrows in the leftmost column point at the saturation artifacts caused by the interleaved acquisition scheme, while the remaining arrows indicate stacking artifacts.

distorted due to an improperly chosen MR slice. At end-exhalation, no substantial differences were detected between results for the navigator-based approach and US-

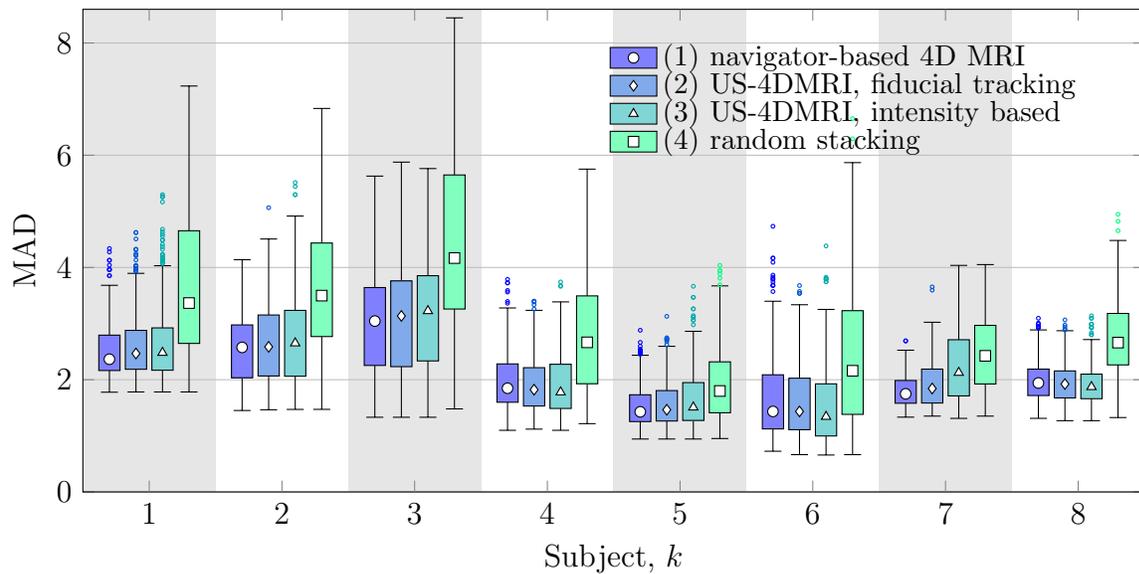


Figure 6: Error distributions of the leave-one-out experiments for each subject k and sorting method (intensity range 0–255). The whiskers comprise all samples within 1.5 times the interquartile range (IQR) above and below the box.

4DMRI, respectively. Note that the qualitative results shown in figure 5a and figure 5b depict the results of Subject 1 for which the navigator-based 4D MRI and US-4DMRI differ most according to the joint histogram in figure 4. Conversely, axial and coronal cuts of Subject 6 in figure 5c and figure 5d show smooth liver boundaries and vascular structures both at end-exhalation and end-inhalation for all stacking methods.

3.4. Quantitative validation

The statistical evaluation presented below is based on paired t -tests; p -values and effect size in terms of Cohen’s d are reported. The significance level was set to 1%.

3.4.1. Results per subject Figure 6 shows the error distributions obtained for stacking methods (1)–(4), as defined in section 3.3, and for each subject k individually. The error distributions contain between $N^d/2 = 390$ and $N^d/2 = 420$ test samples.

In general, methods (1)–(3) performed significantly better than random stacking ($p < 0.001$, $0.35 < d < 1.61$). Comparable results were obtained for stacking methods (1) and (2) where the most significant differences were observed for Subject 1 ($p < 0.001$, $d = 0.22$) and Subject 7 ($p < 0.001$, $d = 0.30$), while for the remaining subjects either no significant difference ($p > 0.03$ $k \in \{6, 8\}$) or a negligible effect size ($d < 0.15$) were found. The same applies when comparing (1) and (3): Cohen’s d was in the range of 0.24 and 0.30 for $k \in \{1, 2, 5\}$ and below 0.17 for the remaining data sets. The only exception is Subject 7, where the navigator-based approach clearly outperformed intensity-based US-4DMRI ($p < 0.001$, $d = 0.90$).

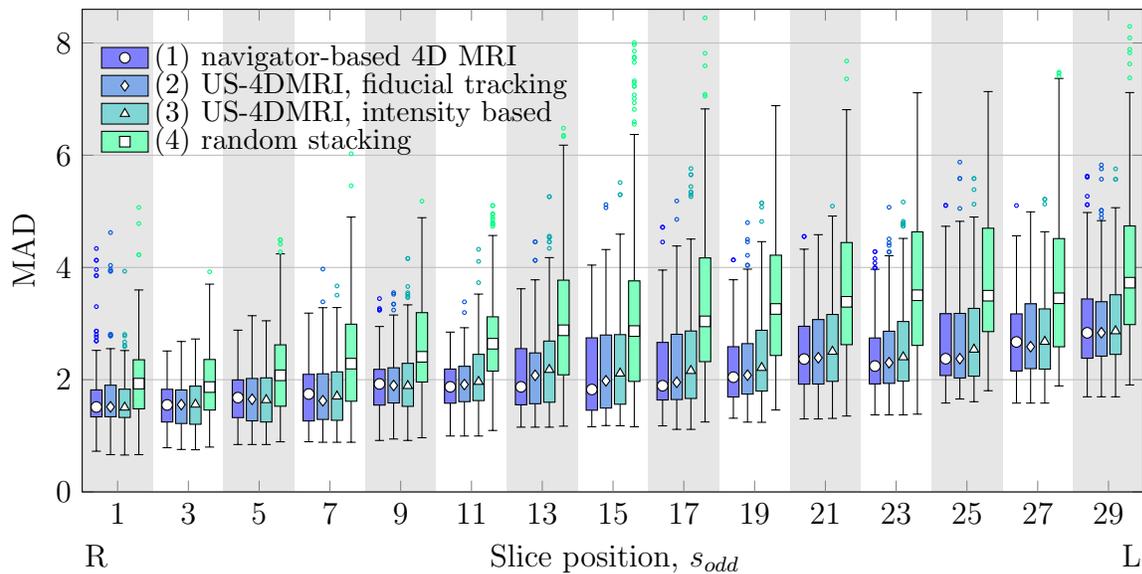


Figure 7: Error distributions of the leave-one-out experiments for each slice position s_{odd} and sorting method (intensity range 0–255). The whiskers have a maximum length of 1.5 IQR.

When comparing (2) and (3), the former outperformed intensity-based US-4DMRI in five out of eight cases. Yet, the difference was either not significant ($p > 0.02$, $k \in \{4, 8\}$) or negligible ($d < 0.20$), with the exception of Subject 7 ($p < 0.001$, $d = 0.65$).

3.4.2. Results per slice position The box plots in figure 7 display the error distributions cumulated over all subjects and for each s_{odd} separately with a total number of 168–226 samples per slice position. Low slice numbers represent the anatomical right (R) side; with increasing s_{odd} the MR slice moves towards the umbilical fissure on the left side (L) of the right liver lobe.

The error for the navigator-based approach and both US-4DMRI was significantly lower than for random stacking ($p < 0.001$, $0.56 < d < 1.23$). In general, the navigator-based approach achieved better results as compared to the US-driven 4D MRI methods. This was observed most clearly for slice positions 11–19 and 23: for these slice positions, (1) outperformed (3) with small yet non-negligible effect size ($0.23 < d < 0.35$). Comparing (1) and (2) either revealed no statistical significant difference ($p > 0.02$, $s_{odd} = 23$) or an effect size below 0.15 ($s_{odd} \in \{11, \dots, 19\}$).

3.5. Computation times

In order to reconstruct one MR volume, navigator-based 4D MRI and intensity-based US-4DMRI required on average 120 ms and 84 ms, respectively. It is important to note that navigator-based 4D MRI additionally requires image registration for all navigators

since the stacking is based on deformation field differences (Von Siebenthal, Gamper, Boesiger, Lomax, Cattin et al. 2007). The computational time for the registration is not included in the value reported above. For US fiducial tracking, a mean computation time of 9 ms per US image was reported (De Luca et al. 2012). Based on these fiducials, volume reconstruction was performed in 3 ms on average, resulting in a mean computation time of 12 ms per volume. All values were computed for a MATLAB (The MathWorks, Inc., Natick, Massachusetts, USA) implementation on a standard CPU.

4. Discussion

Qualitative analysis of sample MR volumes showed only minor distortions and discontinuities in liver boundaries, emphasizing the feasibility of the proposed US-driven stacking methods. When comparing end-inhalation with end-exhalation volumes, more satisfactory results were obtained in the latter case. We explain this observations with the hysteresis in respiratory motion (Seppenwoolde, Shirato, Kitamura, Shimizu, Van Herk, Lebesque & Miyasaka 2002). Longer time intervals at end-exhalation lead to an increased number of MR scans acquired near exhalation and to improved stacking results.

With regard to quantitative results, small effect sizes in terms of Cohen’s d evince that the proposed US-4DMRI is able to compete with the navigator-based 4D MRI independent of the choice of the US similarity measure. A comparison of the two US surrogates suggests that US-4DMRI based on fiducial tracking achieves better results than intensity-based US-4DMRI. Yet, a relevant effect size was measured for one subject only. In terms of computation times, fiducial-based US-4DMRI clearly outperformed the other approaches. Although manual fiducial selection introduces additional complexity to US-4DMRI as compared to the purely intensity-based approach, the shorter computation times might outweigh the extra effort for real-time applications, such as image-guided proton therapy.

Figure 7 shows an increase in the error from the anatomical right to the left side. We explain this observation by the increased presence of vessel structures in the region of the inferior vena cava and the portal vein. Since blood vessels show high image contrast in MRI, the MAD will be higher for slices farther away from the right liver boundary. Cardiac motion might have additional adverse effects on the stacking quality in the region of large blood vessels.

Limitations and future work The number of MR images per slice position available for stacking was $|\mathcal{I}_s| = 30$ or $|\mathcal{I}_s| = 28$ and thus 6 to 8 times smaller than in (Von Siebenthal, Gamper, Boesiger, Lomax, Cattin et al. 2007). The relatively short acquisition duration of 5.5 min per subject implies that long-term effects cannot conclusively be discussed in the present work. However, similar to the navigator-based approach, the use of internal surrogate data potentially enables the proposed US-4DMRI to obtain satisfactory results even in the case of organ drift. Additional experiments with longer scan durations will

be performed in order to evaluate the capability of US-4DMRI to cope with long-term deformations and organ drift.

The number of MR images per slice position $|\mathcal{I}_s|$ further has a direct impact on the achievable temporal resolution of the 4D image: since the MR slices for volume reconstruction are repeatedly selected from constant and finite image sets \mathcal{I}_s , the variation in MR volumes is confined too. The supplementary material to this paper showed that visually no difference between the 4D images at $f_{4D} = 8$ Hz and $f_{4D} = 16$ Hz could be observed for the given amount of data. In future work, we aim for a temporal resolution of up to 10 Hz or approximately 40 MR volumes per respiratory cycle given larger data sets. The clinical benefit of increased temporal resolution on motion models, however, remains to be shown.

A common limitation of all retrospective slice stacking approaches is their failure of reconstructing anatomically plausible MR volumes if the desired organ state was not captured at each slice position. This issue becomes accentuated for extreme respiration depths or in the presence of coughing. As an extension, (Preiswerk et al. 2016) presented a cough-detector based on the US signal itself. For US-4DMRI an upper threshold to the similarity measure could be applied in order to exclude deficient 3D images (Von Siebenthal, Gamper, Boesiger, Lomax, Cattin et al. 2007). Further extensions to the proposed US-4DMRI will include cardiac motion when the US acoustic window captures the inferior vena cava and its pulsation. Uniquely from other respiratory surrogate methods, US imaging provides internal motion signals at a temporal resolution substantially higher than normal resting heart rate in adults (Aladin, Whelton, Al-Mallah, Blaha, Keteyian, Juraschek, Rubin, Brawner & Michos 2014). Colour Doppler imaging can provide additional information on the speed and the direction of blood flow through vessels (Crowe, Manasseh, Chmielewski, Hachulla, Speicher, Greiser, Müller, De Perrot, Vallée & Salomir 2017). Yet taking cardiac motion into account requires even larger data and might lead to increased scan durations.

With the possibility to monitor internal organ motion outside the MR bore, US-4DMRI bears the potential to enhance online tumour tracking during radiotherapy. However, one major simplification of the proposed setup for image-guided proton therapy is the assumption that the US probe remains attached to the patient’s chest for both pretreatment 4D MR imaging and motion estimation during dose delivery. As the US-based similarity measures are sensitive to the imaging plane, a repositioning of the US probe will have adverse effects on the stacking results and motion estimates. Further effort will be needed in order to relax this constraint, for example by investigating the use of skin tattoos to facilitate the repositioning of the US probe. The application of US-4DMRI in radiotherapy demands further examinations on how the US transducer affects the dose delivery.

5. Conclusion

Ultrasound imaging has shown to be a feasible and promising surrogate for retrospective 4D MRI. While most prevalent methods are restricted to reconstruct one single respiratory cycle, US-4DMRI is capable of capturing varying respiration patterns in terms of respiratory depth and frequency far beyond one cycle. Satisfactory MR volumes have been reconstructed showing only minor discontinuities in blood vessels and liver boundaries. Moreover, competitive results were achieved in quantitative leave-one-out experiments when compared to navigator-based 4D MRI. The presented US-4DMRI outperforms navigator-based approaches with respect to reduced acquisition time, increased temporal resolution and availability of the surrogate data outside the MR bore. We envision the application of US-driven 4D MRI in patient-specific motion modelling for dose delivery planning and online tumour tracking in highly conformal, active scanning proton therapies for thoracic and abdominal cancer.

Acknowledgments

This work was supported by the Swiss National Science Foundation, SNSF (CRSII2-127549 and 320030_163330/1).

References

- Aladin, A. I., Whelton, S. P., Al-Mallah, M. H., Blaha, M. J., Keteyian, S. J., Juraschek, S. P., Rubin, J., Brawner, C. A. & Michos, E. D. (2014). Relation of resting heart rate to risk for all-cause mortality by gender after considering exercise capacity (the Henry Ford exercise testing project), *The American journal of cardiology* **114**(11): 1701–1706.
- Bert, C. & Durante, M. (2011). Motion in radiotherapy: particle therapy, *Physics in medicine and biology* **56**(16): R113.
- 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, *Physics in Medicine & Biology* **51**(17): 4147.
- Buerger, C., Clough, R. E., King, A. P., Schaeffter, T. & Prieto, C. (2012). Nonrigid motion modeling of the liver from 3-D undersampled self-gated golden-radial phase encoded MRI, *IEEE transactions on medical imaging* **31**(3): 805–815.
- Cai, J., Chang, Z., Wang, Z., Segars, W. P. & Yin, F.-F. (2011). Four-dimensional magnetic resonance imaging (4D-MRI) using image-based respiratory surrogate: A feasibility study, *Medical physics* **38**(12): 6384–6394.
- Crowe, L. A., Manasseh, G., Chmielewski, A., Hachulla, A.-L., Speicher, D., Greiser, A., Müller, H., De Perrot, T., Vallée, J.-P. & Salomir, R. (2017). Spatially Resolved MR-Compatible Doppler Ultrasound: Proof of Concept for Triggering of Diagnostic Quality Cardiovascular MRI for Function and Flow Quantification at 3T, *IEEE Transactions on Biomedical Engineering* .
- De Luca, V., Tanner, C. & Székely, G. (2012). Speeding-up image registration for repetitive motion scenarios, *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, IEEE, pp. 1355–1358.
- Dinkel, J., Hintze, C., Tetzlaff, R., Huber, P. E., Herfarth, K., Debus, J., Kauczor, H. U. & Thieke, C. (2009). 4D-MRI analysis of lung tumor motion in patients with hemidiaphragmatic paralysis, *Radiotherapy and Oncology* **91**(3): 449–454.

- Hugo, G. D. & Rosu, M. (2012). Advances in 4D radiation therapy for managing respiration: Part I—4D imaging, *Zeitschrift für Medizinische Physik* **22**(4): 258–271.
- Hui, C., Wen, Z., Stemkens, B., Tijssen, R., van den Berg, C., Hwang, K.-P. & Beddar, S. (2016). 4D MR imaging using robust internal respiratory signal, *Physics in medicine and biology* **61**(9): 3472.
- Johnston, E., Diehn, M., Murphy, J. D., Loo, B. W. & Maxim, P. G. (2011). Reducing 4D CT artifacts using optimized sorting based on anatomic similarity, *Medical physics* **38**(5): 2424–2429.
- Liu, Y., Yin, F.-F., Chang, Z., Czito, B. G., Palta, M., Bashir, M. R., Qin, Y. & Cai, J. (2014). Investigation of sagittal image acquisition for 4D-MRI with body area as respiratory surrogate, *Medical physics* **41**(10): 101902.
- Mickevicius, N. J. & Paulson, E. S. (2017). Investigation of undersampling and reconstruction algorithm dependence on respiratory correlated 4D-MRI for online MR-guided radiation therapy, *Physics in medicine and biology* **62**(8): 2910.
- Neto, J. A. G., Elazzazzi, M., Altun, E. & Semelka, R. C. (2008). When should abdominal magnetic resonance imaging be used?, *Clinical Gastroenterology and Hepatology* **6**(6): 610–615.
- Paganelli, C., Summers, P., Bellomi, M., Baroni, G. & Riboldi, M. (2015). Liver 4DMRI: A retrospective image-based sorting method, *Medical physics* **42**(8): 4814–4821.
- Petrusca, L., Cattin, P., De Luca, V., Preiswerk, F., Celicanin, Z., Auboiroux, V., Viallon, M., Arnold, P., Santini, F., Terraz, S. et al. (2013). Hybrid ultrasound/magnetic resonance simultaneous acquisition and image fusion for motion monitoring in the upper abdomen, *Investigative radiology* **48**(5): 333–340.
- Phillips, M. H., Pedroni, E., Blattmann, H., Boehringer, T., Coray, A. & Scheib, S. (1992). Effects of respiratory motion on dose uniformity with a charged particle scanning method, *Physics in medicine and biology* **37**(1): 223.
- Preiswerk, F., De Luca, V., Arnold, P., Celicanin, Z., Petrusca, L., Tanner, C., Bieri, O., Salomir, R. & Cattin, P. C. (2014). Model-guided respiratory organ motion prediction of the liver from 2D ultrasound, *Medical image analysis* **18**(5): 740–751.
- Preiswerk, F., Toews, M., Cheng, C.-C., Chiou Jr, y. G., Mei, C.-S., Schaefer, L. F., Hoge, W. S., Schwartz, B. M., Panych, L. P. & Madore, B. (2016). Hybrid MRI-Ultrasound acquisitions, and scannerless real-time imaging, *Magnetic Resonance in Medicine* .
- Preiswerk, F., Toews, M., Hoge, W. S., Chiou, J.-y. G., Panych, L. P., Wells III, W. M. & Madore, B. (2015). Hybrid Ultrasound and MRI Acquisitions for High-Speed Imaging of Respiratory Organ Motion, *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, pp. 315–322.
- Rank, C. M., Heußer, T., Buzan, M. T., Wetscherek, A., Freitag, M. T., Dinkel, J. & Kachelrieß, M. (2017). 4D respiratory motion-compensated image reconstruction of free-breathing radial MR data with very high undersampling, *Magnetic resonance in medicine* **77**(3): 1170–1183.
- Rohlfing, T., Maurer, C. R., O’dell, W. G. & Zhong, J. (2004). Modeling liver motion and deformation during the respiratory cycle using intensity-based nonrigid registration of gated mr images, *Medical physics* **31**(3): 427–432.
- Schwenke, M., Strehlow, J., Haase, S., Jenne, J., Tanner, C., Langø, T., Loeve, A. J., Karakitsios, I., Xiao, X., Levy, Y. et al. (2015). An integrated model-based software for FUS in moving abdominal organs, *International Journal of Hyperthermia* **31**(3): 240–250.
- Seppenwoolde, Y., Shirato, H., Kitamura, K., Shimizu, S., Van Herk, M., Lebesque, J. V. & Miyasaka, K. (2002). Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy, *International Journal of Radiation Oncology Biology Physics* **53**(4): 822–834.
- Stemkens, B., Tijssen, R. H., de Senneville, B. D., Heerkens, H. D., van Vulpen, M., Lagendijk, J. J. & van den Berg, C. A. (2015). Optimizing 4-dimensional magnetic resonance imaging data sampling for respiratory motion analysis of pancreatic tumors, *International Journal of Radiation Oncology* Biology* Physics* **91**(3): 571–578.

- Stemkens, B., Tijssen, R. H., de Senneville, B. D., Lagendijk, J. J. & van den Berg, C. A. (2016). Image-driven, model-based 3D abdominal motion estimation for MR-guided radiotherapy, *Physics in medicine and biology* **61**(14): 5335.
- Tryggestad, E., Flammang, A., Han-Oh, S., Hales, R., Herman, J., McNutt, T., Roland, T., Shea, S. M. & Wong, J. (2013). Respiration-based sorting of dynamic MRI to derive representative 4D-MRI for radiotherapy planning, *Medical physics* **40**(5): 051909.
- Von Siebenthal, M., Gamper, U., Boesiger, P., Lomax, A., Cattin, P. et al. (2007). 4D MR imaging of respiratory organ motion and its variability, *Physics in medicine and biology* **52**(6): 1547.
- von Siebenthal, M., Székely, G., Lomax, A. J. & Cattin, P. C. (2007). Systematic errors in respiratory gating due to intrafraction deformations of the liver, *Medical physics* **34**(9): 3620–3629.
- Wachinger, C., Yigitsoy, M., Rijkhorst, E.-J. & Navab, N. (2012). Manifold learning for image-based breathing gating in ultrasound and MRI, *Medical image analysis* **16**(4): 806–818.
- Yamamoto, T., Langner, U., Loo, B. W., Shen, J. & Keall, P. J. (2008). Retrospective analysis of artifacts in four-dimensional CT images of 50 abdominal and thoracic radiotherapy patients, *International Journal of Radiation Oncology Biology Physics* **72**(4): 1250–1258.