

Liver 4DMRI: A retrospective image-based sorting method

Chiara Paganelli^{a)}

Dipartimento di Eletttronica, Informazione e Bioingegneria, Politecnico di Milano, Milano 20133, Italy

Paul Summers

Division of Radiology, Istituto Europeo di Oncologia, Milano 20133, Italy

Massimo Bellomi

Division of Radiology, Istituto Europeo di Oncologia, Milano 20133, Italy and Department of Health Sciences, Università di Milano, Milano 20133, Italy

Guido Baroni and Marco Riboldi

Dipartimento di Eletttronica, Informazione e Bioingegneria, Politecnico di Milano, Milano 20133, Italy and Bioengineering Unit, Centro Nazionale di Adroterapia Oncologica, Pavia 27100, Italy

Purpose: Four-dimensional magnetic resonance imaging (4DMRI) is an emerging technique in radiotherapy treatment planning for organ motion quantification. In this paper, the authors present a novel 4DMRI retrospective image-based sorting method, providing reduced motion artifacts than using a standard monodimensional external respiratory surrogate.

Methods: Serial interleaved 2D multislice MRI data were acquired from 24 liver cases (6 volunteers + 18 patients) to test the proposed 4DMRI sorting. Image similarity based on mutual information was applied to automatically identify a stable reference phase and sort the image sequence retrospectively, without the use of additional image or surrogate data to describe breathing motion.

Results: The image-based 4DMRI provided a smoother liver profile than that obtained from standard resorting based on an external surrogate. Reduced motion artifacts were observed in image-based 4DMRI datasets with a fitting error of the liver profile measuring 1.2 ± 0.9 mm (median \pm interquartile range) vs 2.1 ± 1.7 mm of the standard method.

Conclusions: The authors present a novel methodology to derive a patient-specific 4DMRI model to describe organ motion due to breathing, with improved image quality in 4D reconstruction.

Key words: time-resolved imaging, 4DMRI, retrospective sorting, radiotherapy, treatment planning

1. INTRODUCTION

Accurate quantification of breathing-induced anatomical changes is a key factor in high precision extra-cranial radiotherapy^{1,2} and particle therapy.³ When respiratory motion is not adequately accounted for during imaging, significant artifacts may appear, thus jeopardizing accurate quantification. In order to manage organ motion due to breathing, different procedures have been developed.⁴ A simple prospective method to avoid respiratory motion is to interrupt breathing during image acquisition (i.e., breath-hold),^{5,6} with the disadvantage that patients might not be able to hold their breath in a consistent position (for multiple-breath-hold imaging) or for the duration of the image acquisition.⁷ Conversely, in prospective gated acquisition, a snapshot image of a specific respiratory phase is acquired for reconstructing a motion-free volume with the use of an external surrogate.^{8,9} Both these methods acquire an image volume in a single specific respiratory phase, requiring repeated acquisition at multiple reference phases to cover the whole breathing cycle.^{10,11} An alternative approach is time-resolved retrospective 4D imaging, such as four dimensional computed tomography (4DCT),^{12–14} which has become a basic component for respiratory motion assess-

ment.^{4,15} Due to the absence of ionizing radiation, increased soft tissue contrast, and high temporal resolution, 4D magnetic resonance imaging (4DMRI) is an attractive technology for studying organ motion.^{15–18} However, due to the limited frequency at which full 3D volumes can be acquired, rapidly acquired 2D multislice MRI data are resorted and stacked into a 4DMRI image.¹⁵ In conventional approaches, derived from 4DCT, images are sorted retrospectively using an external surrogate.¹² This latter may not correlate well with the actual internal motion^{19,20} and the limited accuracy in breathing state identification can cause artifacts.^{21,22} The use of multiple respiratory related signals²³ or internal surrogates^{24–26} is expected to improve 4D reconstruction.^{22,27} As an alternative, the use of a motion model²⁸ has shown promising results but is intrinsically limited by the computational cost of a model-based 4D reconstruction.

Several techniques have been reported specifically for 4DMRI sorting. An interleaved multislice 2D sequence with 2D “navigator slices” acquired at a fixed location was introduced by von Siebenthal and colleagues,²⁹ where retrospective sorting was based on the interpolation of a motion signal representing the similarity between “navigator slice” samples. Similarly, Wachinger and colleagues³⁰ applied manifold learning

techniques to navigator slices in order to reconstruct a gating signal, which required images acquired at the same anatomical location. Another approach³¹ provides the acquisition of multiple volumes with a navigator signal used to define multiple gating windows. Disadvantages of these methods include the need for sequence modification and a longer acquisition time, as navigator images are acquired purely for sorting purposes. Recently, an anatomy-based 4D sorting technique^{32,33} was proposed relying on the “body area.” In this case, the consistency of the breathing signal across multiple slices is an open issue especially with axial slices, as phase shifts between thorax and abdomen need to be accurately accounted for.^{32,33} It has been shown, conversely, that phase shifts in sagittal slices are less of an issue in 4D reconstruction.³⁴ Tryggestad *et al.*,³⁵ instead proposed a retrospective 4DMRI technique, in which an external respiratory trace acquired simultaneous with imaging data was used to derive a set of “average MRI volumes.” A graph-based retrospective 4DMRI sorting was also proposed:³⁶ a graph was constructed to represent the relationship among the slice in terms of spatial contiguity and the 3D image was derived by searching for an optimal path along the slice location but performing a manual selection of the respiratory phase. Finally, in a recent work,³⁷ the construction of a preoperative 4D image was proposed by interpolating along the time axis, the transformation obtained via nonrigid registration of 3D images acquired in known respiratory phase (i.e., end-exhale, mid-inhale, end-inhale), thus leading to an estimation of the intermediate phases.

The aim of this study is to propose and validate a novel image-based 4DMRI resorting method in order to derive a patient-specific 4D model, describing organ motion due to breathing. For evaluation purposes, we assessed the accuracy of the proposed resorting method by computing the fitting error of the liver profile.

2. MATERIALS AND METHODS

2.A. Testing dataset

Dynamic multislice MRI images were acquired in six healthy volunteers (mean age 24.8 yr) and 24 patients at the

Istituto Europeo di Oncologia (Milano, Italy) with a 1.5 T scanner (Magnetom Avanto, Siemens Medical Systems). A balanced steady-state free precession sequence (TrueFISP) was used during free-breathing to repeatedly acquire oblique sagittal 2D images of the liver in an interleaved-sequential slice sampling order (i.e., slice 2,4,6...1,3,5...2,4,6...), in order to avoid magnetization recovery and interslice cross talk effects.

The MRI parameters were optimized to achieve the trade-off between temporal and spatial resolution:

- repetition time (TR)/echo time (TE): 2.9 ms/1.26 ms
- flip angle: 68°
- bandwidth: 601 Hz/pixel
- scan matrix: 256 × 224 pixels with spacing of 1.28 × 1.28 mm
- slice thickness of 5 mm
- acquisition time: 180 ms/slice (i.e., slice sampling 3.6 s).

A total of 20 slices for 20 frames were acquired (i.e., imaging duration of 72 s).

An external surrogate for liver motion was acquired with the scanner’s respiratory belt (Siemens Physiologic Monitoring Unit). Using the respiratory belt for triggering, we also acquired gated volumes at the expiration and inspiration phases. For consistency, the sequence parameters were the same of the dynamic acquisition.

2.B. Image-based 4DMRI sorting

The proposed method involved a retrospective sorting of MRI images according to the image content. First [Fig. 1(A)], a reference volume (i.e., *reference_mi*) corresponding to the expiration phase (i.e., most stable phase) was created instead of considering a gated acquisition at the expiration phase, as this may be affected by artifacts due to irregular respiration and belt inaccuracies. The reference volume was constructed by (i) selecting the first slice of the imaging frame as initialization, (ii) choosing among all possible slices at the following slice location, the one with highest mutual information (MI) with respect to the previous slice, and (iii) repeating this process for all slice location values.

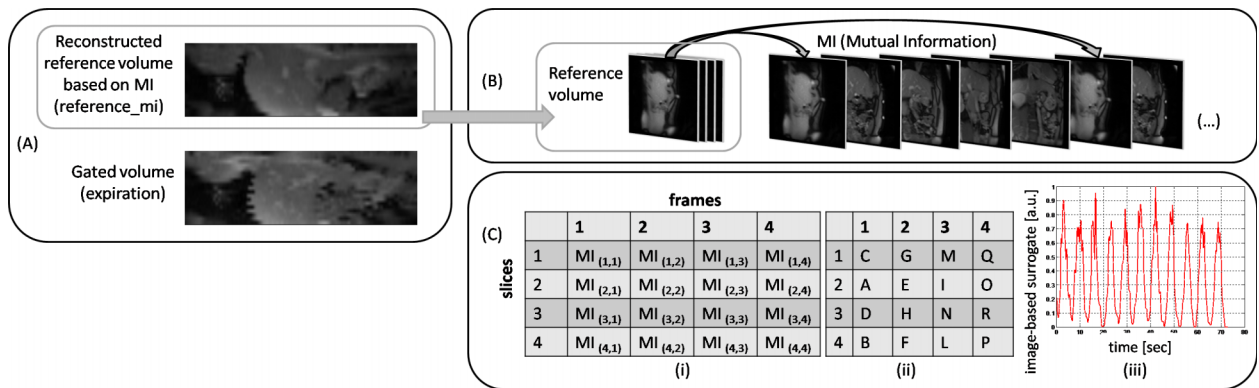


FIG. 1. Workflow of the image-based method. (A) Reference volume constructed by stacking slices with highest MI; (B) computation of the MI between the slice of the reference volume and the corresponding slice of the series; (C) organization of the MI values for each slice of each frame (i) according to the image acquisition time (ii) and derivation of the image-based surrogate (iii) for the subsequent 4D reconstruction in phase/amplitude binning.

Subsequently, we computed the MI between a slice of the reference volume and the corresponding slice [Fig. 1(B)] within the image sequence, obtaining a MI value for each frame. Then, we arranged the MI values of all slices for each frame according to the image acquisition time (i.e., the interleaved slice acquisition 2,4,...1,3...2,4,... led to an interleaved arrangement of the MI values A,B,...C,D,...E,F,...), thus obtaining an image-based respiratory surrogate signal [Fig. 1(C)]. Interpolation in the time domain was applied to replace MI values equal to 1, due to the comparison of the slice in the reference volume and the same slice in the image series.

Finally, a 4D retrospective sorting was performed according to the image-based surrogate (i.e., *mi_sorting*) in both phase and amplitude binnings.³⁸ Among all possible slices falling in a specific bin, we selected the slice with the highest MI value with respect to the adjacent previous slice, to minimize shift between slices.

The resorting method in amplitude and phase binnings was applied also to the external surrogate approach (i.e., *belt_sorting*), as clinically utilized.^{12,13} To reduce the influence of noise, the respiratory trace was prefiltered by applying a tenth order low pass filter with a 2 Hz cutoff frequency.

A mixed approach where the external surrogate was complemented with the image content information was also implemented (i.e., *belt+mi_sorting*). In this case, slices were searched within the bins defined by the respiratory belt and stacked by selecting the one with highest similarity (i.e., highest MI) with respect to the previous one. This led to an external surrogate resorting method driven by the image content (i.e., *belt+mi_sorting*).

2.C. Experiments

4D MRI reconstruction was performed with both external (i.e., belt) and image-based surrogates. We constructed eight bins in phase and amplitude for both surrogates. The same binning approach was applied to the *belt+mi_sorting*.

To evaluate the performance of the retrospective sorting, we quantified:

- The Pearson’s correlation (p -value < 0.05) and the mean absolute phase shift between the external surrogate detected by the belt and the image-based surrogate.
- The root mean square fitting error (RMSE) of the liver profile in the reference volume. Specifically, each slice of the volume was filtered with a vertical derivative and the sum of the pixel values was taken in a region of interest.³⁹ Finally, a second-order fitting was performed among the slices, to evaluate the regularity of the liver profile (i.e., of the retrospective slice sorting). We compared the RMSE of the gated acquisition in expiration and the one of the *reference_mi* volumes. A Wilcoxon rank sum test ($\alpha = 5\%$) was performed to evaluate the significant difference between the two error populations. Furthermore, in the six volunteers, a feature extraction method⁴⁰ was applied between the *reference_mi* volume and the gated acquisition at exhale, and between the *reference_mi* volume and the gated acquisition at inhale, in order to show the consistency between the *reference_mi* volume and the exhale phase. A Wilcoxon rank sum test ($\alpha = 5\%$) was performed to evaluate

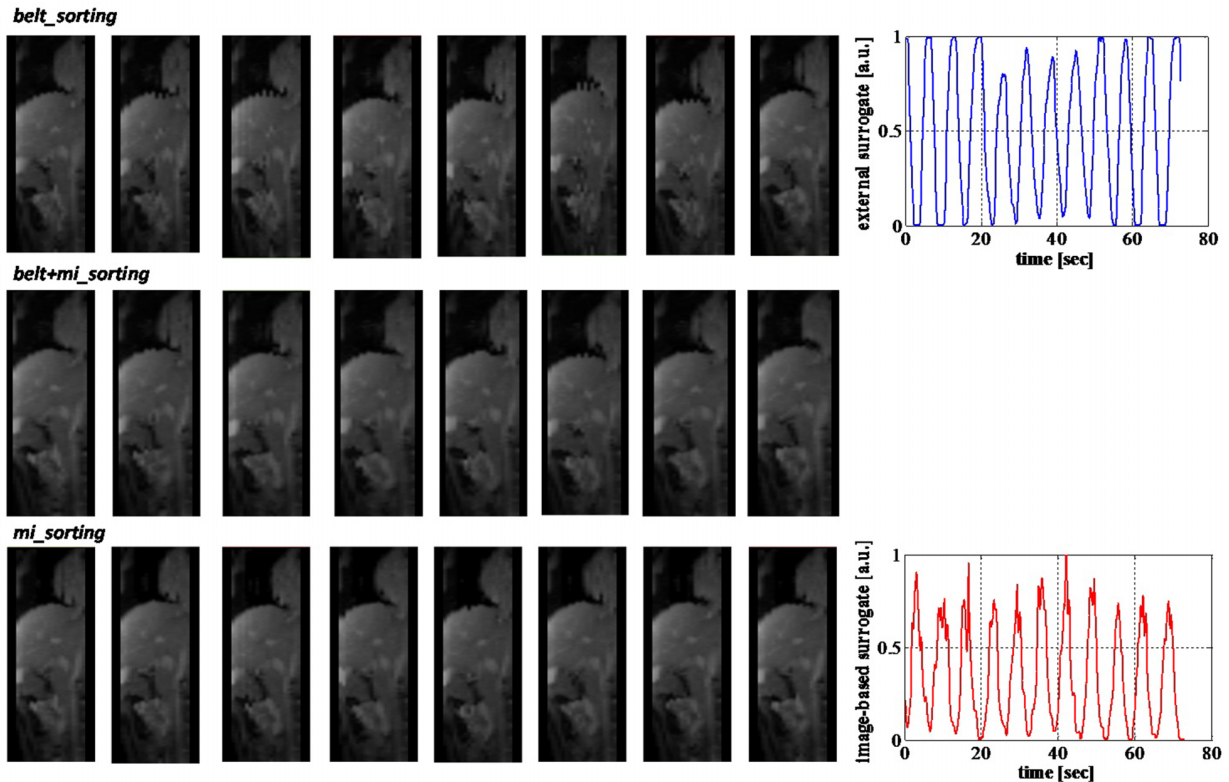


FIG. 2. Coronal 4DMRI with corresponding surrogate signals. First row: *belt_sorting* in amplitude binning; second row: *belt+mi_sorting* in amplitude binning; third row: *mi_sorting* in amplitude binning.

TABLE I. Median fitting error [mm] of the liver profile. For each volunteer, the median \pm interquartile range (maximum) of the RMSE among all bins is reported for *belt_sorting*, *belt+mi_sorting*, and *mi_sorting* for phase/amplitude binning.

	<i>belt_sorting</i>		<i>belt+mi_sorting</i>		<i>mi_sorting</i>	
	Phase binning	Amplitude binning	Phase binning	Amplitude binning	Phase binning	Amplitude binning
V1—8bins	2.1 \pm 3.2 (6.2)	2.9 \pm 2.1 (6.8)	2.0 \pm 2.2 (3.9)	2.8 \pm 2.1 (4.4)	2.4 \pm 1.6 (3.5)	1.8 \pm 2.0 (5.5)
V2—8bins	1.7 \pm 0.6 (2.8)	2.1 \pm 0.2 (2.7)	1.4 \pm 1.0 (2.8)	1.8 \pm 0.4 (3.1)	1.9 \pm 0.4 (2.3)	1.1 \pm 0.2 (1.9)
V3—8bins	2.4 \pm 0.4 (3.0)	1.1 \pm 0.7 (2.3)	2.2 \pm 0.1 (2.7)	1.0 \pm 0.5 (1.8)	1.8 \pm 0.4 (2.1)	1.6 \pm 0.2 (1.9)
V4—8bins	1.9 \pm 0.9 (2.4)	1.8 \pm 0.4 (2.8)	1.5 \pm 0.4 (2.0)	1.4 \pm 0.6 (2.0)	1.3 \pm 0.6 (2.1)	1.3 \pm 0.6 (1.8)
V5—8bins	3.3 \pm 0.8 (5.7)	2.8 \pm 1.1 (4.3)	3.2 \pm 1.8 (4.6)	2.6 \pm 0.8 (3.3)	3.1 \pm 1.3 (4.0)	2.4 \pm 1.2 (3.6)
V6—8bins	1.9 \pm 0.6 (2.9)	2.9 \pm 0.4 (3.0)	1.6 \pm 0.8 (2.7)	1.4 \pm 1.0 (2.7)	1.7 \pm 0.5 (2.8)	1.7 \pm 1.3 (2.8)
Grouped systematic error	2.3 \pm 1.2	2.2 \pm 1.2	2.0 \pm 1.3	1.7 \pm 1.4	1.9 \pm 0.9	1.6 \pm 0.9
V1 irregular	5.0 \pm 3.9 (9.5)	6.4 \pm 1.9 (7.1)	3.6 \pm 2.1 (4.8)	4.5 \pm 1.4 (5.6)	2.3 \pm 1.6 (4.4)	2.12 \pm 2.2 (4.7)

the significant difference between *reference_mi/gated* at exhale and *reference_mi/gated* at inhale.

- The RMSE of the 4DMRI liver profile in all the eight bins. The median \pm interquartile range of the RMSE was calculated for each subject. Similarly to van Herk,⁴¹ we defined as grouped systematic error the median of RMSE among all subjects. In order to discriminate between the group systematic error obtained from *belt_sorting*, *belt+mi_sorting*, and *mi_sorting* for both phase and amplitude binning approaches, we performed a Friedman test at 5% significance level. For volunteers, we also applied a feature extraction method^{40,42} between image pairs, keeping the first phase as reference and evaluating the motion obtained in the reconstructed 4DMRI as distance between features.

3. RESULTS

3.A. Volunteer study

Figure 2 shows a coronal view of a 4DMRI for a representative volunteer as derived with *belt_sorting* and *mi_sorting* amplitude binning, along with the two corresponding surrogates. The correlation values between the respiratory trace and the MI values ranged from -0.6 to -0.8 in all patients, indicating a negative correlation significantly different from 0 (p -value $<$ 0.05). The mean absolute phase shift was quantified as 0.79 ± 0.20 s.

The median value of the RMSE of the reference volume liver profile was quantified as 1.4 ± 1.5 and 1.0 ± 0.3 mm for all subjects in the gated and in the *reference_mi* volume, respectively. The Wilcoxon test confirmed that the two populations were significantly different. Furthermore, the distance of the extracted features was quantified as 1.3 ± 0.4 mm and 7.3 ± 2.2 mm between *reference_mi* and gated exhale volume and *reference_mi* and inhale gated volume, respectively (alpha $<$ 5%).

For most volunteers (Table I), the median error and the corresponding maximum error were lower with *mi_sorting* than with *belt_sorting* and *belt+mi_sorting*. The grouped systematic error decreased moving from *belt_sorting* to *mi_sorting*: in phase/amplitude binning, an error of $2.3 \pm 1.2/2.2 \pm 1.2$ mm for *belt_sorting* was measured against $1.9 \pm 0.9/1.6 \pm 0.9$ mm for *mi_sorting*, whereas *belt+mi_sorting* produced an error of $2.0 \pm 1.3/1.7 \pm 1.4$ mm. Friedman test confirmed a significant difference between *mi_sorting* and *belt_sorting* both in phase and amplitude binnings.

Furthermore, we acquired an irregular breathing pattern from the first volunteer (V1). The median RMSE (Table I-last row) of the 4DMRI liver profile was about 5 mm for *belt_sorting*, whereas under the maximum voxel dimension (i.e., 5 mm) for *belt+mi_sorting* and *mi_sorting*.

Figure 3(A) shows the number of features matches between the respiratory phases and the motion estimated from the extracted features for volunteers. A breathing pattern trend is

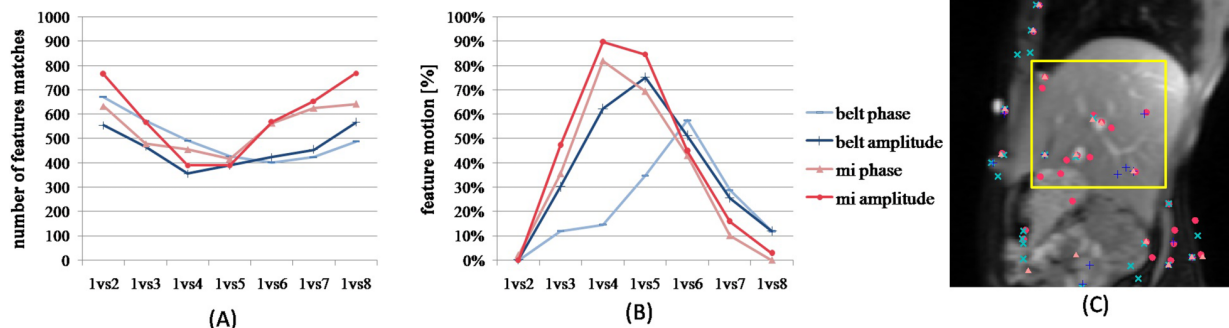


FIG. 3. Internal anatomy. (A) Number of matches and (B) average feature motion [%] between the different bins of the 4DMRI (1 vs 2, 1 vs 3, 1 vs 4, 1 vs 5, 1 vs 6, 1 vs 7, and 1 vs 8) for both *mi_sorting* and *belt_sorting*, for all volunteers. (C) An example of identified matches between exhale-inhale in the liver (yellow square) in one slice of the exhale phase. Specifically, red circle for *mi_sorting* with amplitude binning, pink triangle for *mi_sorting* with phase binning, blue cross for *belt_sorting* with amplitude binning, and light blue asterisk for *belt_sorting* with phase binning.

more visible in the image-based sorting than using the external surrogates, with a higher number of matches for contiguous respiratory phases (e.g., 1 vs 2, 1 vs 3, 1 vs 7, and 1 vs 8) and a more regular respiratory motion for *mi_sorting* rather than *belt_sorting*. In addition, a greater number of matches were found in the liver [Fig. 3(C)] in *mi_sorting* than *belt_sorting*, thus better describing the respiratory motion range [Fig. 3(B)].

3.B. Patient study

For all patients, as for the volunteers, the RMSE in gated acquisition was larger than in the constructed *reference_mi* volume, with median values of 1.9 ± 1.5 mm and 0.9 ± 0.7 mm, respectively.

The grouped systematic errors (Table II) with *mi_sorting* were lower than with the other retrospective sorting methods and were significantly different from *belt_sorting* ones (Fried-man test, $\alpha = 5\%$). For patient *P6*, a six bins 4DMRI was constructed because neither *belt_sorting* nor *belt+mi_sorting* in amplitude binning were possible in eight bins. Conversely, the complete dataset with 8 bins was successfully reconstructed with *mi_sorting*, with a median RMSE of 1.6 ± 0.4 mm. The correlation between the external surrogate and the image-based one was -0.1 for *P6*, as opposed to other patients where values ranged from -0.6 to -0.8 , with a mean absolute phase shift of 0.78 ± 0.60 s overall. 4DMRI images of *P6* with *belt_sorting* and *mi_sorting* in amplitude binning are shown in Fig. 4. A patient with an irregular breathing pattern (*P16*) is also reported in Fig. 5.

4. DISCUSSION

In this work, we propose a retrospective image-based sorting method of dynamic multislice MRI images to create a patient-specific 4DMRI. The proposed method is based on the use of MI as a surrogate for retrospective sorting in phase/amplitude binning.

With respect to other techniques,^{32–35} our approach does not require image preprocessing and provides a surrogate signal across the whole image. Furthermore, no “average 4DMRI”³⁵ and no navigator slice²⁹ are required, thus avoiding to affect the signal to noise ratio in the reconstructed 4D image dataset and reducing the overall acquisition time and the specific absorption rate, respectively. Derivation of the surrogate signal directly from the image series does not require external monitoring devices,³⁵ which suffer from extrinsic (i.e., operator variability) and intrinsic inaccuracies, providing technical and procedural simplifications to the clinical workflow. In addition, it does not need to acquire additional volumes and interpolate between the respiratory phases³⁷ and a previous manual identification of the respiratory phases is not required in contrast to Ref. 36.

For testing purposes, the proposed image-based 4DMRI (i.e., *mi_sorting*) was compared to a standard resorting method based on an external surrogate (i.e., *belt_sorting*) and to a mixed approach where the external surrogate was complemented with the image content information (i.e., *belt+mi_sorting*). It is worth considering that, as reported in Ref. 35, a gain adjustment is typically present in the belt surrogate, in addition to variability in positioning and setup and to an indirect correlation with internal anatomy. The two

TABLE II. Median fitting error [mm] of the liver profile. For each patient, the median \pm interquartile range (maximum) of the RMSE among all phases is reported for *belt_sorting*, *belt+mi_sorting*, and *mi_sorting* for phase/amplitude binning.

	<i>belt_sorting</i>		<i>belt+mi_sorting</i>		<i>mi_sorting</i>	
	Phase binning	Amplitude binning	Phase binning	Amplitude binning	Phase binning	Amplitude binning
<i>P1</i> —8bins	3.4 ± 2.0 (5.8)	1.9 ± 1.1 (4.4)	2.3 ± 2.1 (3.9)	1.5 ± 1.1 (4.3)	1.7 ± 1.1 (3.4)	1.6 ± 0.9 (3.5)
<i>P2</i> —8bins	4.5 ± 0.4 (5.3)	4.6 ± 0.7 (5.7)	4.6 ± 0.4 (4.8)	4.6 ± 0.5 (5.2)	1.0 ± 0.4 (1.6)	1.0 ± 0.1 (1.1)
<i>P3</i> —8bins	1.2 ± 0.2 (1.5)	1.2 ± 0.2 (1.4)	1.1 ± 0.4 (1.7)	1.2 ± 0.6 (1.7)	1.0 ± 0.5 (1.6)	1.1 ± 0.4 (1.4)
<i>P4</i> —8bins	1.7 ± 1.0 (3.0)	1.6 ± 0.7 (2.9)	1.3 ± 0.6 (2.0)	1.4 ± 0.6 (2.4)	1.1 ± 0.4 (1.7)	1.3 ± 0.2 (1.6)
<i>P5</i> —8bins	1.4 ± 0.7 (2.0)	1.9 ± 0.9 (2.4)	1.4 ± 0.6 (2.01)	1.5 ± 0.5 (1.9)	1.0 ± 0.4 (2.1)	1.1 ± 0.3 (2.4)
<i>P6</i> —6bins	1.7 ± 0.9 (2.2)	2.5 ± 0.3 (2.8)	1.0 ± 0.3 (1.4)	1.4 ± 1.1 (2.4)	0.9 ± 0.1 (1.3)	0.8 ± 0.0 (1.1)
<i>P7</i> —8bins	1.6 ± 0.9 (2.6)	1.9 ± 0.8 (4.0)	2.0 ± 0.5 (2.6)	1.8 ± 0.5 (2.5)	1.9 ± 0.4 (2.9)	2.0 ± 0.3 (2.3)
<i>P8</i> —8bins	2.3 ± 1.0 (4.2)	2.3 ± 0.6 (4.4)	2.2 ± 1.2 (4.2)	2.0 ± 0.7 (4.5)	2.3 ± 0.6 (3.7)	2.0 ± 1.3 (3.8)
<i>P9</i> —8bins	1.2 ± 0.9 (2.1)	1.2 ± 0.7 (2.1)	0.8 ± 0.2 (1.2)	0.8 ± 0.3 (1.9)	0.7 ± 0.2 (1.6)	0.7 ± 0.1 (0.9)
<i>P10</i> —8bins	3.6 ± 1.8 (6.4)	3.2 ± 1.8 (3.7)	1.6 ± 1.4 (4.1)	2.4 ± 1.6 (3.5)	1.5 ± 1.2 (3.5)	1.2 ± 0.7 (2.3)
<i>P11</i> —8bins	2.3 ± 0.6 (3.4)	3.1 ± 0.7 (3.3)	2.0 ± 0.9 (3.4)	2.6 ± 0.9 (3.9)	1.7 ± 0.6 (3.4)	1.9 ± 0.7 (3.3)
<i>P12</i> —8bins	1.3 ± 1.6 (3.5)	1.8 ± 0.3 (2.7)	1.4 ± 0.5 (2.2)	1.1 ± 0.5 (1.7)	0.8 ± 0.2 (1.2)	1.0 ± 0.7 (2.0)
<i>P13</i> —8bins	1.3 ± 0.4 (2.1)	1.4 ± 0.1 (1.6)	1.3 ± 0.2 (1.5)	1.3 ± 0.4 (1.8)	1.4 ± 0.3 (1.9)	1.2 ± 0.3 (1.8)
<i>P14</i> —8bins	4.1 ± 1.9 (5.7)	3.8 ± 0.6 (4.9)	3.4 ± 0.2 (4.1)	3.5 ± 0.4 (4.0)	3.3 ± 0.6 (3.9)	3.3 ± 0.5 (3.7)
<i>P15</i> —8bins	1.6 ± 0.5 (2.4)	2.6 ± 0.6 (4.7)	1.3 ± 0.3 (2.1)	1.9 ± 1.1 (3.4)	1.6 ± 1.2 (3.3)	1.2 ± 0.3 (1.6)
<i>P16</i> —8bins	3.2 ± 0.8 (3.8)	3.2 ± 0.9 (3.5)	3.1 ± 0.3 (3.6)	3.2 ± 1.1 (3.4)	2.4 ± 0.8 (4.2)	1.8 ± 0.9 (3.6)
<i>P17</i> —8bins	2.5 ± 0.6 (3.8)	2.5 ± 1.3 (3.7)	1.8 ± 0.6 (3.7)	1.7 ± 0.9 (3.4)	1.6 ± 0.2 (2.9)	1.4 ± 0.5 (2.1)
<i>P18</i> —8bins	3.7 ± 3.6 (6.8)	3.5 ± 3.7 (6.7)	2.8 ± 3.0 (6.3)	2.7 ± 0.9 (6.4)	2.2 ± 2.6 (5.8)	2.4 ± 2.3 (3.7)
Grouped systematic error	2.1 ± 1.8	2.1 ± 1.7	1.6 ± 1.6	1.5 ± 0.9	1.4 ± 1.1	1.2 ± 0.9

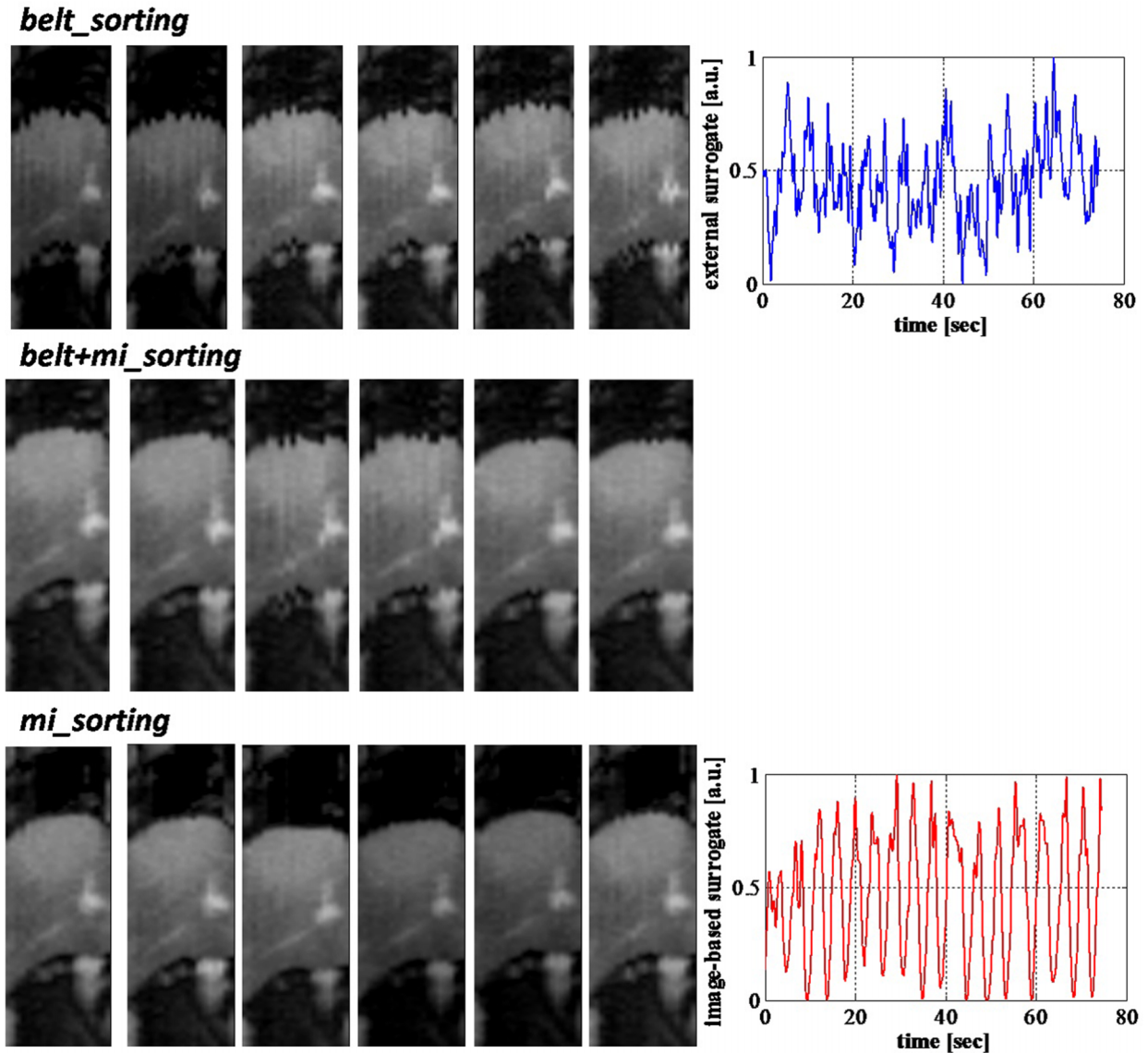


FIG. 4. Coronal six bins 4DMRI with corresponding surrogate signals of patient *P6*. First row: *belt_sorting* in amplitude binning; second row: *belt+mi_sorting* in amplitude binning; third row: *mi_sorting* in amplitude binning.

surrogates showed high negative correlation (i.e., due to the comparison of the image series with the *reference_mi* volume) with p -value < 0.05 , attesting the effectiveness of the image-based signal in breathing motion description. A phase shift was measured between the belt and the image-based surrogates for both volunteers and patients, with average values within 0.8 s. Overall, the image-based surrogate performed better than the respiratory belt, which was not robust, as already reported in Ref. 35. This is particularly evident in the results for the irregular breathing pattern of volunteer *V1*: in this case, the image-based method allowed reducing errors from more than 5 mm (*belt_sorting*) to 2 mm. Noise in the respiratory trace was evident also in patient *P6*, in which *belt_sorting* and *belt+mi_sorting* did not produce a 4DMRI featuring eight bins, whereas the *mi_sorting* method succeeded. Therefore, six bins were reconstructed instead of eight and reduced artifacts (i.e., more regular liver profile) were observed. A smoother liver

profile of the 4DMRI obtained with the *mi_sorting* is also visible in Fig. 5, where an irregular breathing patient is shown.

In both volunteers and patients, the grouped systematic error was systematically lower in *mi_sorting* than in *belt_sorting*, with a statistical significant difference between the two populations. In our work, no phantoms³² were available to quantify 4D sorting artifacts and exhale/inhale gated acquisitions were used to evaluate the correspondence of the *reference_mi* volume to the exhale phase. In addition, the assessment of artifacts by using a similarity measure across neighboring slice²⁹ was not considered due to the construction of the time-resolved imaging with one of this measure (i.e., mutual information). Therefore, we proposed the method developed by Rit *et al.*³⁹ based on the quantification of the variability of diaphragm motion to evaluate quantitatively the regularity of the diaphragm profile in the coronal direction (i.e., to quantify artifacts due to the retrospective reconstruction). The

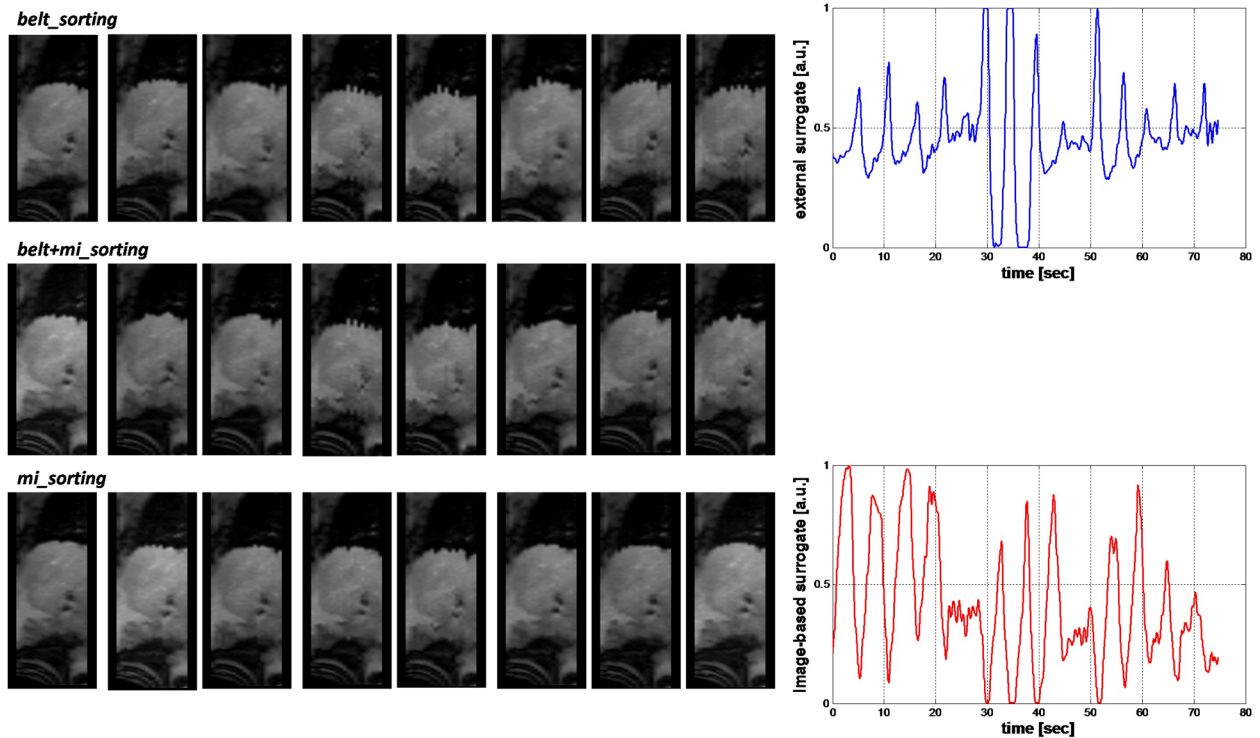


Fig. 5. Coronal eight bins 4DMRI with corresponding surrogate signals of patient *P16* (irregular breathing pattern). First row: *belt_sorting* in amplitude binning; second row: *belt+mi_sorting* in amplitude binning; third row: *mi_sorting* in amplitude binning.

drawback in this method is the evaluation of a region of interest around the diaphragm, losing the whole volume information. On the other hand, the application of the feature extraction method⁴⁰ to volunteers data on the whole volume confirmed the better retrospective reconstruction with *mi_sorting*, able to preserve internal anatomy motion rather than *belt_sorting*. Conversely, the *belt+mi_sorting* represented an intermediate situation between *belt_sorting* and *mi_sorting*, showing that results provided by the respiratory belt could be improved by the image content. Artifacts were nonetheless present in some image-based 4DMRI, due to individual respiratory variations and to the trade-off between imaging duration and number of frames per slice location.

Belt inaccuracies were also identified in the gated acquisition, resulting in a volume with an irregular liver profile, due to irregularity in the breathing pattern (i.e., wrong identification of the expiration phase by the respiratory trigger) and variability in the acquisition (i.e., acquired separately from the multislice data and with a qualitative selection of the gating window). For this purpose, the construction of the *reference_mi* volume by analyzing the whole 4D image series yielded a more stable reference at the expiration phase than the gated volume, also confirmed by the identified features⁴⁰ in the six volunteers database. However, future works should analyze the best starting slice for the construction of the reference volume. Preliminary tests on one volunteer showed similar results by selecting the last slice as initialization, with a slightly better performance by using the first one. Such a result may be due to the acquisition process: oblique sagittal slices were acquired, in which the first slice described the right side of the

liver near the skin, representing a more stable starting point for the construction of a reference volume.

In addition, the intensity change effects in MRI images were also minimized by acquiring multiple interleaved slices (i.e., ensuring full magnetization recovery) and using a similarity measure based on the image information (i.e., MI) instead of absolute pixel values (e.g., sum of squared difference). However, a deep analysis on intensity standardization⁴³ and nonuniformity correction⁴⁴ should be taken in consideration, as well as a comparison with other metrics.²⁹

The average CPU time required for *belt_sorting*, *belt+mi_sorting*, and *mi_sorting* was quantified in 20, 42, and 123 s, respectively (2.2 GHz Intel Core i7 processor). Conversely, the setup of the respiratory belt required ~4–5 min, as multiple adjustments in belt position and tightening are typically needed to achieve satisfactory visual feedback of the breathing trace on the scanner monitor. Additional time (~1–2 min) was also typically required to start the respiratory trace acquisition before scanning.

5. CONCLUSION

In this work, we proposed an image-based 4DMRI retrospective sorting, which was compared to the standard method based on the acquisition of an external surrogate. The better performance of the proposed method would provide a patient-specific 4DMRI able to accurately describe organ motion, improving clinical margins delineation in radiotherapy treatment planning, and providing a model for the latest

developments in MRI-guided technologies.⁴⁵ Future works should be focused on the optimization of parallel imaging sequences to avoid the retrospective reconstruction, thus acquiring full volumes in dynamic mode.

ACKNOWLEDGMENTS

The work was supported by AIRC, Italian Association for Cancer Research. The authors would like to thank Giulia Buizza and Simona Cacciato for their contribution in the liver profile extraction. The authors report no conflicts of interest in conducting the research.

a) Author to whom correspondence should be addressed. Electronic mail: chiara.paganelli@polimi.it; Telephone: +390223999022.

¹L. A. Dawson and M. B. Sharpe, "Image-guided radiotherapy: Rationale, benefits, and limitations," *Lancet Oncol.* **7**, 848–858 (2006).

²D. A. Jaffray, "Image-guided radiotherapy: From current concept to future perspectives," *Nat. Rev. Clin. Oncol.* **9**, 688–699 (2012).

³M. Riboldi, R. Orecchia, and G. Baroni, "Real-time tumour tracking in particle therapy: Technological developments and future perspectives," *Lancet Oncol.* **13**, 383–391 (2012).

⁴P. J. Keall *et al.*, "The management of respiratory motion. Radiation oncology report of AAPM Task Group 76," *Med. Phys.* **3**, 3874–3900 (2006).

⁵J. Hanley *et al.*, "Deep inspiration breath-hold technique for lung tumors: The potential value of target immobilization and reduced lung density in dose escalation," *Int. J. Radiat. Oncol., Biol., Phys.* **45**, 603–611 (1999).

⁶S. S. Lee, J. H. Byun, H. S. Hong, S. H. Park, H. J. Won, Y. M. Shin, and M. G. Lee, "Image quality and focal lesion detection on T2-weighted MR imaging of the liver: Comparison of two high-resolution free-breathing imaging techniques with two breath-hold imaging techniques," *J. Magn. Reson. Imaging* **26**, 323–330 (2007).

⁷J. M. Blackall, "Respiratory motion in image-guided interventions of the liver," Ph.D. thesis, University of London, 2002.

⁸E. C. Ford, G. S. Mageras, E. Yorke, K. E. Rosenzweig, R. Wagman, and

C. C. Ling, "Evaluation of respiratory movement during gated radiotherapy using film and electronic portal imaging," *Int. J. Radiat. Oncol., Biol., Phys.* **52**, 522–531 (2002).

⁹R. George, T. D. Chung, S. S. Vedam, V. Ramakrishnan, R. Mohan, E. Weiss, and P. J. Keall, "Audio-visual biofeedback for respiratory-gated radiotherapy: Impact of audio instruction and audio-visual biofeedback on respiratory-gated radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* **65**, 924–933 (2006).

¹⁰M. von Siebenthal, G. Székely, A. Lomax, and P. Cattin, "Systematic errors in respiratory gating due to intrafraction deformations of the liver," *Med. Phys.* **34**, 3620–3629 (2007).

¹¹S. A. Nehme *et al.*, "Quantification of respiratory motion during 4D-PET/CT acquisition," *Med. Phys.* **31**, 1333–1338 (2004).

¹²S. S. Vedam *et al.*, "Acquiring a four-dimensional computed tomography dataset using an external respiratory signal," *Phys. Med. Biol.* **48**, 45–62 (2003).

¹³E. Rietzel, T. Pan, and G. T. Y. Chen, "Four-dimensional computed tomography: Image formation and clinical protocol," *Med. Phys.* **32**, 874–889 (2005).

¹⁴P. J. Keall, "4-Dimensional computed tomography imaging and treatment planning," *Semin. Radiat. Oncol.* **14**, 81–90 (2004).

¹⁵G. D. Hugo and M. Rosu, "Advances in 4D radiation therapy for managing respiration: Part I—4D imaging," *Med. Phys.* **22**, 258–271 (2012).

¹⁶H. U. Kauczor and C. Plathow, "Imaging tumor motion for radiotherapy planning using MRI," *Cancer Imaging* **6**, 140–144 (2006).

¹⁷J. M. Blackall *et al.*, "MRI-based measurements of respiratory motion variability and assessment of imaging strategies for radiotherapy planning," *Phys. Med. Biol.* **51**, 4147–4169 (2006).

¹⁸J. Biederer, C. Hintze, M. Fabel, and J. Dinkel, "Magnetic resonance imaging and computed tomography of respiratory mechanics," *J. Magn. Reson. Imaging* **32**, 1388–1397 (2010).

¹⁹M. J. Murphy, "Tracking moving organs in real time," *Semin. Radiat. Oncol.* **14**, 91–100 (2004).

²⁰S. S. Vedam *et al.*, "Quantifying the predictability of diaphragm motion during respiration with a noninvasive external marker," *Med. Phys.* **30**, 505–513 (2003).

²¹Y. D. Mutaf, J. A. Antolak, and D. H. Brinkmann, "The impact of temporal inaccuracies on 4DCT image quality," *Med. Phys.* **34**, 1615–1622 (2007).

²²D. Han, J. Bayouth, S. Bhatia, M. Sonka, and X. Wu, "Characterization and identification of spatial artifacts during 4D-CT imaging," *Med. Phys.* **38**, 2074–2087 (2011).

²³C. Gianoli *et al.*, "A multiple points method for 4D CT image sorting," *Med. Phys.* **38**, 656–667 (2011).

²⁴R. Zeng, J. A. Fessler, J. M. Balter, and P. A. Balter, "Iterative sorting for four-dimensional CT images based on internal anatomy motion," *Med. Phys.* **35**, 917–926 (2008).

²⁵R. Li, J. H. Lewis, L. Cervino, and S. B. Jiang, "4D CT sorting based on patient internal anatomy," *Phys. Med. Biol.* **54**, 4821–4833 (2009).

²⁶E. Johnston *et al.*, "Reducing 4D CT artifacts using optimized sorting based on anatomic similarity," *Med. Phys.* **38**, 2424–2429 (2011).

²⁷E. Rietzel and G. T. Y. Chen, "Improving retrospective sorting of 4D computed tomography data," *Med. Phys.* **33**, 377–379 (2006).

²⁸J. Hinkle, M. Szegedi, B. Wang, B. Salter, and S. Joshi, "4D CT image reconstruction with diffeomorphic motion model," *Med. Image Anal.* **16**, 1307–1316 (2012).

²⁹M. von Siebenthal, G. Székely, U. Gamper, P. Boesiger, A. Lomax, and P. Cattin, "4D MR imaging of respiratory organ motion and its variability," *Phys. Med. Biol.* **52**, 1547–1564 (2007).

³⁰C. Wachinger, M. Yigitsoy, E. J. Rijkhorst, and N. Navab, "Manifold learning for image-based breathing gating in ultrasound and MRI," *Med. Image Anal.* **16**, 806–818 (2012).

³¹J. Tokuda *et al.*, "Adaptive 4D MR imaging using navigator-based respiratory signal for MRI-guided therapy," *Magn. Reson. Med.* **59**, 1051–1061 (2008).

³²J. Cai *et al.*, "Four-dimensional magnetic resonance imaging (4D-MRI) using image-based respiratory surrogate: A feasibility study," *Med. Phys.* **38**, 6384–6394 (2011).

³³J. Yang *et al.*, "Four-dimensional magnetic resonance imaging using axial body area as respiratory surrogate: Initial patient results," *Int. J. Radiat. Oncol., Biol., Phys.* **88**, 907–912 (2014).

³⁴Y. Liu *et al.*, "Investigation of sagittal image acquisition for 4D-MRI with body area as respiratory surrogate," *Med. Phys.* **41**, 101902 (13pp.) (2014).

³⁵E. Tryggstad *et al.*, "Respiration-based sorting of dynamic MRI to derive representative 4D-MRI for radiotherapy planning," *Med. Phys.* **40**, 051909 (12pp.) (2013).

³⁶Y. B. Tong, J. K. Udupa, K. C. Ciesielski, J. M. McDonough, A. Mong, and R. M. Campbell, "Graph-based retrospective 4D image construction from free-breathing MRI slice acquisitions," *Proc. SPIE* **9038**, 90380I–1–90380I-7 (2014).

³⁷C. Weon, W. H. Nam, D. Lee, J. Y. Lee, and J. B. Ra, "Position tracking of moving liver lesion based on real-time registration between 2D ultrasound and 3D preoperative images," *Med. Phys.* **42**, 335–345 (2015).

³⁸N. Wink, C. Panknin, and T. D. Solberg, "Phase versus amplitude sorting of 4D-CT data," *J. Appl. Clin. Med. Phys.* **7**, 77–85 (2006).

³⁹S. Rit, M. van Herk, L. Zijp, and J. J. Sonke, "Quantification of the variability of diaphragm motion and implications for treatment margin construction," *Int. J. Radiat. Oncol., Biol., Phys.* **82**, e399–e407 (2012).

⁴⁰C. Paganelli *et al.*, "Scale invariant feature transform in adaptive radiation therapy: A tool for deformable image registration assessment and re-planning indication," *Phys. Med. Biol.* **58**, 287–299 (2013).

⁴¹M. van Herk, "Errors and margins in radiotherapy," *Semin. Radiat. Oncol.* **14**, 52–64 (2004).

⁴²C. Paganelli *et al.*, "Scale invariant feature transform as feature tracking method in 4D imaging: A feasibility study," in *Annual International Conference of the IEEE on Engineering in Medicine and Biology Society (IEEE, San Diego, CA, 2012)*.

⁴³Y. Tong and J. K. Udupa, "Interactive non-uniformity correction and intensity standardization of MR images," *Proc. SPIE* **9415**, 94151N–1–94151N-6 (2015).

⁴⁴L. G. Nyúl and J. K. Udupa, "New variants of a method of MRI scale standardization," *IEEE Trans. Med. Imaging* **19**, 143–150 (2000).

⁴⁵B. G. Fallone *et al.*, "First MR images obtained during megavoltage photon irradiation from a prototype integrated Linac-MR system," *Med. Phys.* **36**, 2084–2088 (2009).