



Original paper

A clustering approach to 4D MRI retrospective sorting for the investigation of different surrogates



Giorgia Meschini^{a,*}, Chiara Paganelli^a, Chiara Gianoli^b, Paul Summers^c, Massimo Bellomi^{c,d}, Guido Baroni^{a,e}, Marco Riboldi^b

^a Department of Electronics, Information and Bioengineering, Politecnico di Milano, Piazza Leonardo da Vinci, 32, 20133 Milan, Italy

^b Chair of Experimental Physics – Medical Physics, Ludwig-Maximilians-Universität München, Am Coulombwall 1, 85748 Garching bei München, Germany

^c Department of Imaging and Radiological Science, European Institute of Oncology, Via Giuseppe Ripamonti, 435, 20141 Milan, Italy

^d Department of Oncology and Eματο-oncology, University of Milan, Via Festa del Perdono, 7, 20122, Italy

^e Bioengineering Unit, CNAO Foundation, Str. Campeggi, 53, 27100 Pavia, Italy

ARTICLE INFO

Keywords:

4D MRI

Retrospective sorting

Liver

ABSTRACT

Purpose: In retrospective 4-Dimensional Magnetic Resonance Imaging (4D MRI) sorting, respiratory surrogate selection affects the image quality of reconstructed volumes. We propose a method for retrospective 4D MRI sorting based on clustering, which allowed us to compare the performance of single or multiple internal surrogates vs. a conventional external signal.

Methods: A k-medoids clustering algorithm was exploited for sorting 2D MRI into 4D MRI, relying on (A) multiple or (B) single automatically tracked internal landmarks or (C) respiratory belt signal. 4D MRI reconstructions for seven liver cancer patients were compared to those of the state-of-the-art mutual information (MI) approach. Sorting artifacts were measured by the root mean square error (RMSE) between the diaphragm profile and a fitted second order curve. Diaphragm and tumor motions were evaluated.

Results: The median RMSEs ranged 0.97–1.66 mm, 1.24–1.89 mm, 1.43–2.27 mm, 1.74–3.72 mm for the MI, (A), (B) and (C) methods, respectively. Significant differences (Friedman, $\alpha = 5\%$) were found between (C) and all other methods, and between (B) and MI approaches. The discrepancies between (A) and MI approaches ranged 1.1–6.2 mm and 0.7–5.3 mm respectively in diaphragm and tumor motions. Methods (A) and (B) showed similar ranges of motion.

Conclusion: With multiple internal points, our method yielded the description of a higher range of motion and similar image quality with respect to the MI approach. The single point method led to more artifacts, suggesting the superior suitability of multiple internal surrogates for retrospective 4D MRI sorting. Considering internal rather than external information favored superior performance.

1. Introduction

The efficacy of external beam radiotherapy relies on accurate delivery of the therapeutic dose to the target, while sparing surrounding healthy tissues. To achieve this aim, an imaging representation of the irradiated anatomy is a key requirement for conformal planning, and aids in considering possible uncertainties in the geometrical localization of the tumor and organs at risks [1]. For thoraco-abdominal tumors, respiratory-correlated organ motion represents a crucial issue and requires the adoption of specific mitigation techniques to reduce deviations from the planned dose distribution [2,3].

In these cases, treatment planning is based on a time-resolved or

four-dimensional (4D) representation of the patient anatomy, i.e. a set of 3D volumes, each related to a specific phase of an average breathing cycle. Although 4D computed tomography (4D CT) represents the standard imaging technique for motion quantification and planning [2], more recently 4D magnetic resonance imaging (4D MRI) has been explored [4,5]. This latter technique provides better soft tissue contrast and avoids additional dose to the patient.

As far as 4D MRI techniques are concerned, one option for obtaining a 4D MRI is to repeatedly acquire 3D volumes [6,7], but this results in poor spatio-temporal resolution and image quality due to motion-related blurring. Prospective 4D MRI sorting based on 3D volumes has been also proposed [8–10], entailing the acquisition of MR data only in

* Corresponding author at: Politecnico di Milano – Department of Electronics, Information and Bioengineering, Via G. Colombo, 40, 20133 Milano, Italy.

E-mail address: giorgia.meschini@polimi.it (G. Meschini).

URL: <https://www.cartcas.polimi.it> (G. Meschini).

<https://doi.org/10.1016/j.ejmp.2019.02.003>

Received 12 October 2018; Received in revised form 19 January 2019; Accepted 6 February 2019

Available online 13 February 2019

1120-1797/ © 2019 Associazione Italiana di Fisica Medica. Published by Elsevier Ltd. All rights reserved.

correspondence to specific phases of the respiratory cycle. In this case, image quality is strongly related to the efficiency of acquisition triggering, which is based on a respiratory surrogate (e.g. respiratory belt). Respiratory surrogates also play a key role in retrospective 4D MRI sorting, which involves the continuous acquisition of MR data during free breathing and the subsequent sorting into different breathing phases. While external surrogates have been applied for this purpose [11–13], it is recognized that in case of poor correlation between external and internal motion [14], inaccuracy in the identification of the breathing phase may cause severe artifacts. The use of internal surrogates can be therefore considered by exploiting the capability of MRI to provide reliable navigator images [15], image-based surrogates [16–19] or k-space data [20–22]. Navigator images and k-space strategies require longer acquisition time and are not always available in the clinical scanners. The data for image-based approaches on the other hand, can generally be obtained with routine equipment, but so far the sorting has tended to rely on the extraction of motion information from a single image-based surrogate. This may result in inadequate respiratory correlation and spatial information between the considered surrogate and internal anatomy, thus again producing image artifacts. The general consensus is that artifact reduction in reconstructed 4D MRI volumes can be expected if internal surrogates are used [13] or if multi-dimensional information is integrated in the image resorting process [23,19]. This raises the issue of optimizing surrogate selection in terms of location and dimensionality as a key requirement for artifact-free 4D MRI reconstruction in a clinical scenario.

In this study, we propose a novel approach for 4D MRI reconstruction relying on k-medoids clustering, which allowed us to investigate 4D MRI sorting performance as a function of different surrogates. Specifically, we selected surrogates based on location (external vs. internal) and dimensionality (multiple vs. single), relying on an external respiratory bellow and automatic internal landmarks tracking. The implemented technique for 4D MRI sorting is validated (see Supplementary Material) using digital phantom data [24], and applied to liver patient data, on which the proposed approach is compared with respect to a reference state-of-the-art image-based technique.

Supplementary data associated with this article can be found, in the

online version, at <https://doi.org/10.1016/j.ejmp.2019.02.003>.

2. Materials and methods

2.1. Patient dataset

Oblique multi-slice sagittal images of the liver of seven patients were acquired during free-breathing with the TrueFISP sequence using a 1.5 T scanner (Magnetom Avanto, Siemens Healthineers, Erlangen, Germany). Interleaved-sequential slice sampling (i.e. slice 2, 4, 6 ... 1, 3, 5 ... 2, 4, 6) was used in order to avoid interslice cross talk effects [17]. The imaged volume consisted of 20 slices, and for each slice 20 frames were acquired (acquisition time: 180 ms/slice). Pixel size was 1.28 mm × 1.28 mm, with a 5 mm slice thickness. In addition, an external surrogate for respiratory motion was provided by the respiratory belt of the scanner (Siemens Physiologic Monitoring Unit).

2.2. Clustering approach for 4D MRI retrospective sorting

The 4D MRI sorting was performed by means of a k-medoids clustering technique which exploited (i) multiple or (ii) single automatically tracked internal landmarks, and (iii) an external signal.

2.2.1. Investigated surrogates

In the “multiple points” method, for each slice location (Fig. 1A), a set of stable points within the abdomen were automatically selected as surrogates via the scale-invariant feature transform (SIFT, [25]) which consists of two steps: a set of landmarks (i.e. features, points) is first extracted from each image based on local contrast, and correspondences between the features sets are defined by minimizing the distance between so-called feature descriptors, which encode the local image information in terms of intensity and gradient magnitude/direction. The SIFT was applied on the first two frames of the image time-series and regions of interest (ROI) were defined to track points throughout the acquisition. Each ROI was centered in one of the points selected by the SIFT. Points tracking was then accomplished through template matching between the ROIs and the remaining frames.

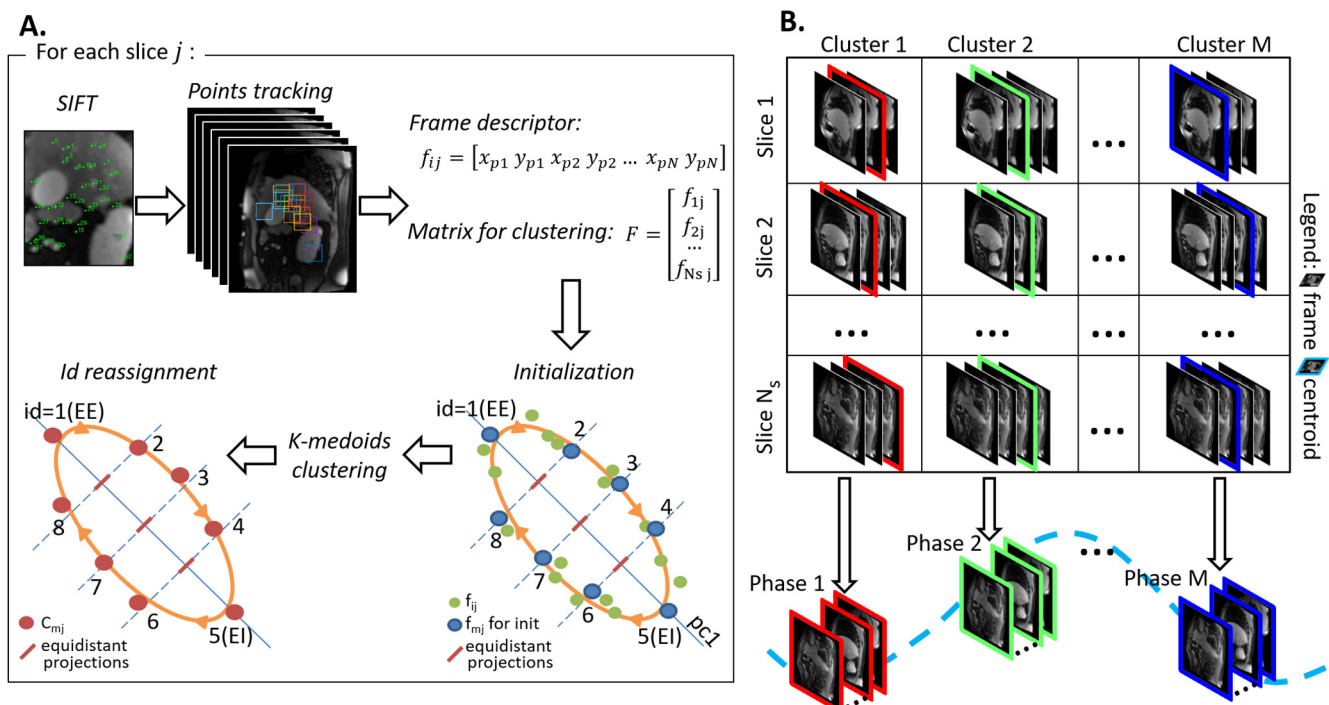


Fig. 1. Multiple-points method. (A) Procedure repeated for each slice location. (B) 4D MRI reconstruction procedure.

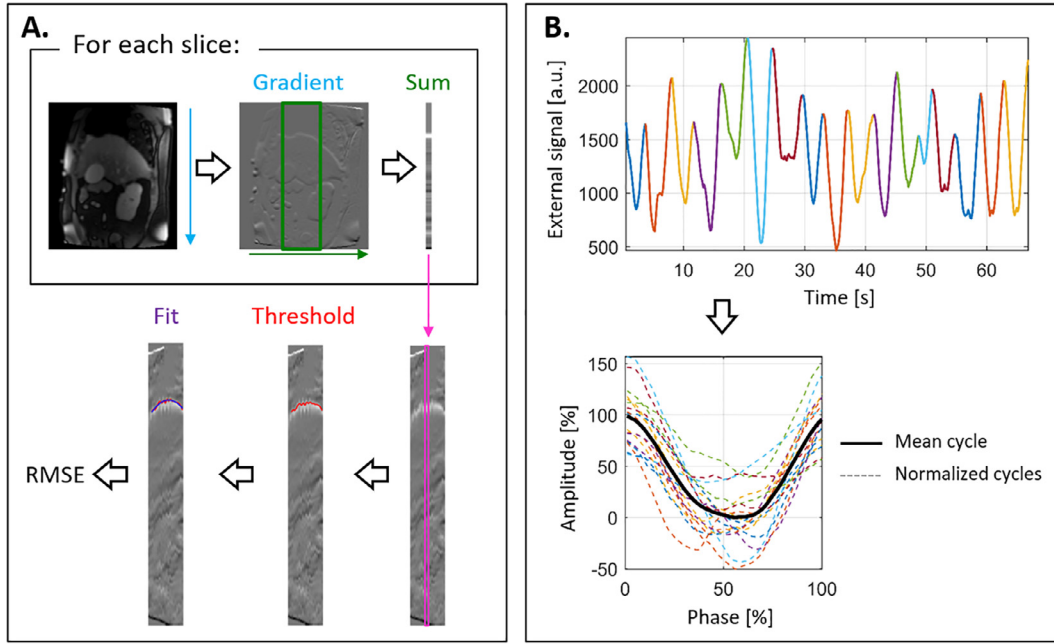


Fig. 2. A. Diaphragm profile extraction. B. Mean cycle estimation.

Template matching is an efficient and widely applied technique in 2D tracking [26–28]. In this study, we set the template (ROI) size to $4 \text{ cm} \times 4 \text{ cm}$ in order to mitigate the effect of recurrent patterns in the image (such as similar vessels) and of out-of-plane motion. Moreover, a manual filtering of template matching outliers was performed: as the ROIs partially overlap, neighboring points are expected to show similar motion, thus points presenting motion uncorrelated with their neighbors' were discarded. Each frame (i) of each slice (j) was represented by the $2N$ -dimensional vector (f_{ij}) with the coordinates (x_p, y_p) belonging to the N points (p) tracked within the slice:

$$f_{ij} = [x_{p1} y_{p1} x_{p2} y_{p2} \dots x_{pN} y_{pN}]$$

Only one point was considered as a surrogate for k-medoids clustering in the “single point” method:

$$f_{ij} = [x_p y_p]$$

Specifically, the most cranial point among the multiple extracted ones was automatically chosen, which is the closest to the diaphragm apex, and thus expected to exhibit the maximum range of motion.

In the “external surrogate” method, the internal motion information was not considered, and the external respiratory signal (s) was associated to each frame of each slice before performing the k-medoids clustering:

$$f_{ij} = s_i$$

2.2.2. K-medoids clustering and 4D MRI reconstruction

An initialization procedure was implemented to achieve consistent starting conditions for clustering (Fig. 1A). The frames with the minimum and maximum projection on the first principal component (pc1) of motion were considered to initialize the end-inhale (EI) and end-exhale (EE) centroids, respectively. The remaining initialization centroids were assigned considering frames with the projections closer to equally spaced segments (red segments) along pc1, exploiting the external surrogate to distinguish between the inspiration and the expiration phases. Fig. 1A shows a schematic representation in case of one tracked point (i.e. $N = 1$): among all f_{ij} frames (green dots), M frames (f_{mj} , with $m = 1, \dots, M$, blue circles) were identified for initialization. For all reconstructions on patient data, the parameter M was set to 8.

For each investigated surrogate, the k-medoids clustering was

applied to the matrix F , containing the f_{ij} vectors as rows. The clustering technique aimed to assign each observation to one of the M clusters. The k-medoids clustering entails an iterative procedure, during which centroids are changed and observations are reassigned to clusters, in order to minimize the sum of Euclidean distances between the observations and the centroid of each cluster [29]. The advantage of k-medoids clustering with respect to k-means clustering is that centroids exactly correspond to observations, i.e. acquired 2D MR images in our case.

Each cluster was associated to a number from 1 to M , called the identifier. At the end of the iterative clustering procedure, reassignment of cluster identifiers was needed to ensure the correspondence between the identifiers and the represented breathing phases among all slices. As for the initialization procedure described above, the final identifiers were assigned considering the projections of clusters centroids (C_{mj} corresponding to red circles in Fig. 1A, $N = 1$) onto the EE to EI axis.

Finally, the centroids frames with corresponding identifiers were stacked in 3D volumes related to each breathing phase of the 4D MRI (Fig. 1B).

2.3. 4D MRI reconstruction evaluation

The proposed approach was first validated (see Supplementary Material) in an in-silico framework [30], that provided a ground truth mean cycle as well as a direct comparison with the methods proposed in the literature.

As ground truth was not available for the evaluation of patients' 4D MRI reconstructions, the results of a mutual information (“MI”) technique were taken as the reference standard for evaluating the performance of the clustering approach in the reconstruction of 4D MRI on the basis of the different surrogates. The MI approach consists of the following steps: (i) defining a starting 3D volume, by stacking together the frames at different slice location with the highest mutual information with respect to the previous slice, starting from the first frame of the first slice location [17], (ii) deriving an image-based surrogate by computing the MI values between each acquired frame of each slice and the corresponding slice in the 3D volume, (iii) ordering the MI values as a function of the image acquisition time and (iv) sorting according to the surrogate values to derive the 4D MRI.

To evaluate the quality of 4D MRI reconstruction of patients' data,

the following indices were considered:

- Diaphragm fitting error. This was evaluated to describe sorting artifacts. For each 3D volume, each sagittal slice was subjected to a vertical derivative filter and the sum of the pixel values was taken in a region of interest, as schematized in Fig. 2A. Then, a second order curve was fit to the diaphragm profile as extracted in the coronal view [17,31]. The diaphragm fitting error was computed as the root mean square error (RMSE) distance between the diaphragm profile and the fitted curve. The median and interquartile range of the RMSE values were calculated for each subject.
- Maximum diaphragm motion (MD) depicted in the 4D reconstruction.
- Tumor Dice coefficient and range of motion (TM). The tumor was manually segmented in the end-exhale and end-inhale 3D volumes of each reconstructed 4D MRI. The Dice coefficient between corresponding phases was evaluated and the 3D range of tumor motion computed for each reconstructed 4D MRI.
- Image difference. A set of transversal slices centered on the tumor forming a 10 cm thick slab was selected and 4D MRI sets were compared, phase by phase, to the reference reconstruction (i.e. MI). Specifically, the root mean squared difference was computed and normalized by the range of intensity values within the reference volume, thus quantifying an image difference (imgDiff) metric:

$$\text{imgDiff}_{ph,m} = \frac{\sqrt{\sum_i^{Nv} (V_{ph,m} - V_{ph,MI})^2 / Nv}}{\max(V_{ph,MI}) - \min(V_{ph,MI})}$$

where Nv is the number of voxel in each volume, the subscript ph refers to the respiratory phase and m to the reconstruction method.

- Respiratory regularity of the external signal. Period variation was defined as its standard deviation, whereas amplitude variation was quantified by considering the mean respiratory pattern, which was computed as the mean between the phase-normalized respiratory cycles (Fig. 2B), and by scoring the root mean square amplitude variation [32].

Non-parametric statistical analysis was performed across reconstructions considering (i) the fitting RMSE, (ii) MD, (iii) TM and (iv) imgDiff. Specifically, the Friedman test [33] was used with a significance level of $\alpha = 5\%$: this test aims at detecting significant differences between the reconstruction methods, as applied to the same patients' group. In the subsequent multiple pairwise comparisons, the Bonferroni correction was used to compensate for multiple tests (that would otherwise increase the likelihood of false positives).

3. Results

Table 1 reports breathing regularity indices for the seven patients. The diaphragm fitting RMSE values of the reconstructed 4D MRI are represented in Fig. 3A and summarized in Table 2. Across subjects, median of the diaphragm fitting RMSEs ranged from 0.97 to 1.66 mm,

Table 1

Breathing period variation: RMSE (mean) [s]. Breathing amplitude variation: RMSE [%].

	Period variation [s]	Amplitude variation [%]
P.1	0.3 (3.5)	12.4
P.2	0.2 (3.0)	16.7
P.3	0.9 (4.7)	41.4
P.4	0.7 (4.0)	24.5
P.5	0.6 (4.3)	22.8
P.6	1.6 (5.6)	38.4
P.7	0.9 (4.7)	16.3

1.24 to 1.89 mm, 1.43 to 2.27 mm and 1.74 to 3.72 mm for the MI, multiple points, single point and external approaches, respectively. No significant correlation (Spearman rank correlation index) was found between the median diaphragm fitting RMSEs and respiratory regularity indices. The median values of the diaphragm fitting RMSEs for the MI and multiple points 4D MRI reconstructions were comparable to the sagittal pixel size (1.28 mm) and remained below 2 mm for all patients. A significant statistical difference in the fitting RMSE was seen between the external surrogate and all the other methods, and between the single point and the MI methods.

As seen in Fig. 3B, the external surrogate approach tended to underestimate both diaphragm and tumor movements, whereas the multiple points and single point approaches yielded similar overestimates of MD and TM with respect to MI reconstruction. The Friedman test highlighted a significant difference in MD values between the MI and the multiple points reconstructions, as well as between the external surrogate and both the multiple and single point reconstructions. For TM, on the other hand, the Friedman test suggested no statistically significant difference between MI and all other methods, whereas the external surrogate was associated with TM significantly different from both the multiple and the single point reconstructions.

Despite the non-significant difference in tumor range of motion, the tumor Dice values (Fig. 3C) imply poor overlap of corresponding structures. The multiple and single point methods both presented their lowest tumor Dice values for the end-inhale phase of P6, who showed the highest TM and the smallest tumor volume (1.2 cm³, Fig. 3C).

The median imgDiff presented a minimum value for the end-exhale phase (Fig. 3D) and, across frames, the imgDiff ranged from 1.8 to 4.2%, 1.9 to 4.2% and 3.1 to 4.8% for the multiple points, the single point and the external surrogate approaches, respectively. The imgDiff values associated with the external surrogate reconstruction resulted significantly different from those corresponding to both the multiple and single internal surrogate methods.

In Fig. 4, coronal views of the 4D MRI reconstructed with the various methods for the two patients (P4 and P6) exhibiting the highest MD values illustrate the residual artifacts, as seen in the diaphragm profile. In this respect, the image quality of the MI, multiple points and single point approaches was similar, whereas the artifacts for the external surrogate reconstruction were more severe. Moreover, the multiple and single point methods were able to describe a wider range of motion with respect to the MI 4D MRI (see diaphragm compared to the pink dashed line, which is aligned to the diaphragm at end-inhale phase of the MI reconstruction).

On an Intel Core i7-6700HQ CPU (2.60 GHz), points tracking for the multiple points method required 2.1 s per tracked point (total duration of 24 min on average) and ~10 min for template matching outliers filtering, leading to a total reconstruction time of ~35 min. The single point method required ~11 min, whereas the external method required less than 3 s. The average CPU time required for MI reconstruction was 123 s [17].

4. Discussion

In this study, we present a novel technique for retrospective 4D MRI sorting based on k-medoids clustering. The interesting feature of the implemented approach is the possibility to exploit different respiratory surrogates, which allowed us to compare the use of single/multiple internal tracked points and of an external respiratory signal.

The proposed technique was first validated in-silico with a digital phantom (see Supplementary Material) and then evaluated on clinical datasets. The validation on the digital phantom highlighted better performance of the internal multiple/single point surrogates with respect to the external surrogate. Internal multiple/single point surrogates produced errors comparable to the ones of current state-of-the-art image-based retrospective sorting [30], confirming the capability of the proposed approach to derive accurate 4D MRI.

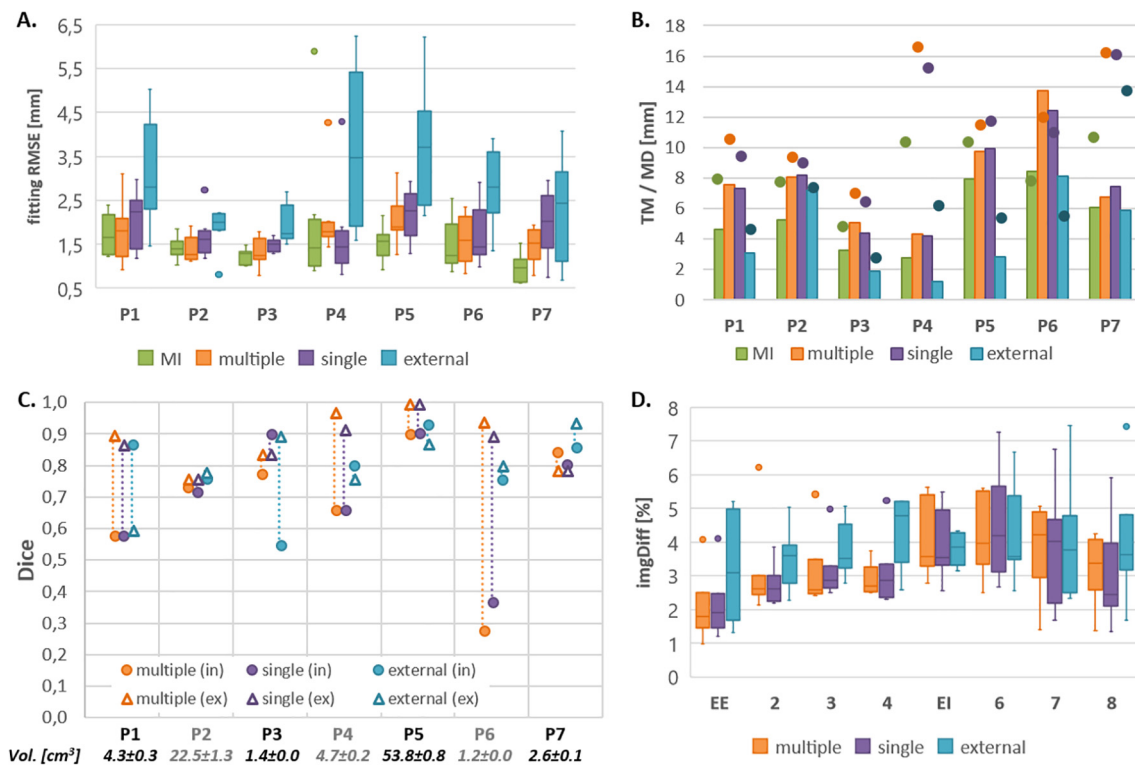


Fig. 3. A. Boxplot representation of the RMSE of diaphragm profile fitting; B. MD (dots) and TM (bars) for each patient. C. Dice values between the tumor segmented on the MI and on the other reconstructions in correspondence to end-inhale (circles) and end-exhale (triangles); on the bottom, tumor volumes as mean $\pm$ standard deviation values among all segmented volumes are listed; D. Boxplot representation of the imgDiff with respect to the MI 4D MRI for each respiratory phase.

Table 2

Mean $\pm$ standard deviation over all patients of: diaphragm fitting error (RMSE), maximum diaphragm motion (MD), tumor range of motion (TM) and Dice coefficient (Dice), image difference (imgDiff).

	RMSE [mm]	MD [mm]	TM [mm]	Dice	imgDiff [%]
MI	1.5 $\pm$ 0.7	8.6 $\pm$ 2.1	5.5 $\pm$ 2.2	–	–
multiple	1.7 $\pm$ 0.6	12.0 $\pm$ 3.5	7.9 $\pm$ 3.1	0.79 $\pm$ 0.19	3.3 $\pm$ 1.2
single	1.8 $\pm$ 0.7	11.4 $\pm$ 3.4	7.7 $\pm$ 2.9	0.79 $\pm$ 0.16	3.3 $\pm$ 1.3
external	2.8 $\pm$ 1.3	6.6 $\pm$ 3.5	4.3 $\pm$ 2.7	0.80 $\pm$ 0.11	3.9 $\pm$ 1.3

Due to the lack of a ground truth 4D MRI for the patient dataset, the mutual information approach [17] was adopted as a reference to compare 4D MRI volumes reconstructed with surrogates which differed in dimensionality (single vs. multiple) and location (internal vs. external). 4D MRI reconstructions with the MI method were based on amplitude sorting, since this has been demonstrated to achieve better results than phase sorting [31]. Other state-of-the-art techniques, such as those based on body area, navigator slice or k-space surrogates were not investigated. During the separate acquisitions required for these techniques [30], different breathing patterns would occur, introducing uncertainty in the comparison of 4D MRI reconstructions.

We note that, for all patients, the described motion resulted higher with the multiple and single point approaches than with the MI method: the difference in diaphragm and tumor motion ranged 1.1–6.2 mm and 0.7–5.3 mm, respectively. This is likely due to the k-medoids algorithm, in which the frames are sorted into clusters (i.e. breathing phases) according to the distance between each frame (described by the surrogate value f_{ij}) and the cluster centroid. As a result, clusters centroids tend to be distant from each other and respiratory phases are optimally differentiated. This allows a better description of the range of motion exhibited during the MR scan with respect to a binning on predefined amplitude values. Also, the derivation of the starting 3D volume for the MI approach by stacking slices with highest MI value could affect

motion description and further investigations on this are required, as defined in Paganelli et al. [17].

The single point approach resulted in ranges of motion comparable to those of the multiple points approach but led to more sorting artifacts relative to the reference MI 4D MRI. To justify the similar described range of motion, we note that the single point was chosen as the one expected to express the maximum displacement (i.e. the closest to the diaphragm) from amongst those used in the multiple points 4D MRI reconstruction. Nevertheless, in the studied patient cases, multiple internal points were seen to be a more robust surrogate for 4D MRI sorting with respect to the single point approach. In fact, no significant difference was found between the MI and the multiple points reconstruction in terms of artifacts in the diaphragm profile, suggesting the capability of the clustering approach based on multiple internal points to provide a retrospective sorted 4D MRI with image quality comparable to the state-of-the-art technique.

The external surrogate approach showed significantly worse performance compared to other methods. Specifically, it was associated to a greater image difference (Fig. 3D) and higher image artifacts (Fig. 4) with respect to the MI reconstruction. Moreover, a higher variability of the fitting RMSE for different patients was observed for the external surrogate with respect to other approaches (Table 2 and Fig. 3A). This is explained by the lack of the external surrogate to correctly describe internal motion, thus leading to severe artifacts in the reconstructed respiratory phases. Consequently, the external signal is not the optimal choice for describing frames f_{ij} in terms of internal anatomical motion. In our analysis, we considered only a single external signal, due to procedural challenges in recording multiple external surrogates during MRI acquisition. A comparative analysis of multiple vs. single external surrogates has been reported previously for 4D CT image acquisition, highlighting the advantage of using redundant information [23]. As expected, general results show that internal information performs better than the external surrogate. Although the lack of multiple external surrogates for 4D MRI reconstruction might represent a caveat of

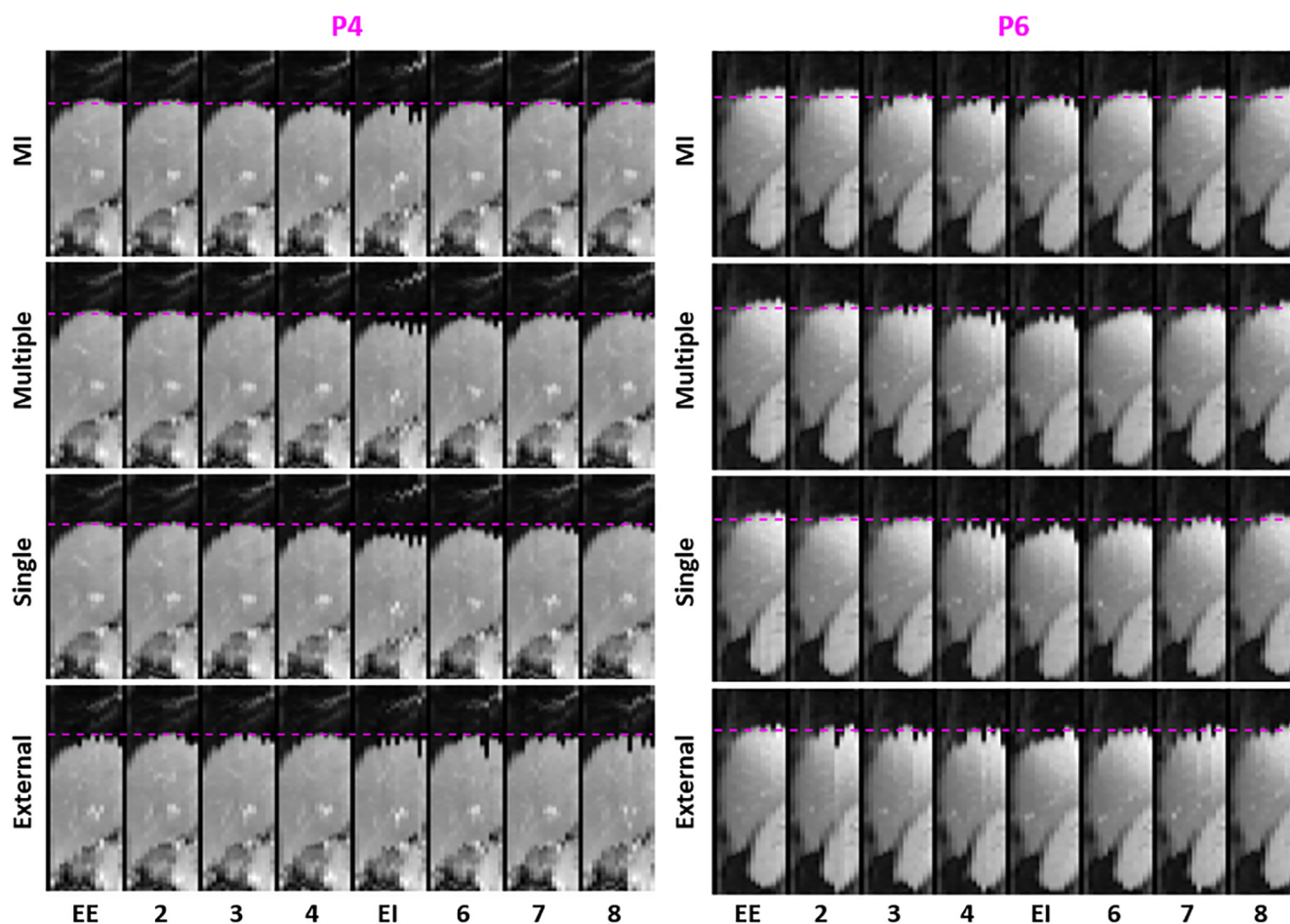


Fig. 4. Coronal views of the reconstructed 4D MRI with the four methods for two sample patients. A pink dashed line in correspondence to diaphragm position in the end-inhale of the MI dataset is reported to highlight diaphragm motion. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

our analysis, we believe that the use of internal respiratory surrogate may largely overcome the performance obtainable with complex external surrogates' configurations.

No significant correlation between the median fitting RMSE and respiratory regularity indices was found for all the methods, meaning that the performances of the k-medoid clustering approach are not dependent on breathing motion regularity, as also found in the phantom study (see Supplementary Material).

Concerning limitations of the present work, we recognize that in the current implementation of the multiple points and single point methods for 4D MRI sorting, we exploited the available external surrogate to distinguish exhale and inhale phases. In future studies, we will seek to replace the external signal with image-based surrogates derived directly from the acquired 2D images for a 4D MRI reconstruction driven entirely by internal anatomy [17]. A further improvement would be to automate the individuation of template matching outliers, for instance by performing k-mean clustering on internal surrogates and considering only those within the largest cluster [19]. This automatization together with the selection of a limited number of points (e.g. close to the tumor) are expected to reduce the duration of 4DMRI reconstruction. Finally, an extension of the proposed method to a larger patient cohort and other surrogates would be considered, as well as the application of the method to other anatomical sites, such as the lower abdomen (e.g. for pancreas patients) and the thorax.

5. Conclusion

The implemented clustering approach for 4D MRI reconstruction combined with the use of multiple internal surrogates can improve range of motion description while also obtaining the same image quality as a state-of-the-art technique [17]. In our cohort of seven patients, a range of motion similar to that obtained with the multiple points approach was achieved by choosing the single internal surrogate that best described breathing motion, but the resulting 4D MRI exhibited more artifacts than the reference MI reconstruction. This suggests a greater accuracy in the motion description derived from multiple internal points compared to a single point approach. The use of the external surrogate led to more sorting artifacts in the reconstructed 4D MRI, thus demonstrating the superior performances of internal surrogates for retrospective 4D MRI sorting.

Acknowledgement

This project was supported by AIRC, Italian Association for Cancer Research.

References

- [1] van Herk M. Errors and margins in radiotherapy. *Semin Radiat Oncol* 2004;14:52–64.
- [2] Keall PJ, Mageras GS, Balter JM, Emery RS, Forster KM, Jiang SB, et al. The management of respiratory motion in radiation oncology report of AAPM Task Group 76. *Med Phys* 2006;33:3874–900.
- [3] Bert C, Durante M. Motion in radiotherapy: particle therapy. *Phys Med Biol*

- 2011;56(16):R113–44. <https://doi.org/10.1088/0031-9155/56/16/R01>.
- [4] Paganelli C, Whelan B, Peroni M, Summers P, Fast M, van de Lindt T, et al. MRI-guidance for motion management in external beam radiotherapy: current status and future challenges. *Phys Med Biol* 2018;63:22TR03.
 - [5] Stemkens B, Paulson ES, Tijssen RHN. Nuts and bolts of 4D-MRI for radiotherapy. *Phys Med Biol* 2018;63:21TR01.
 - [6] Plathow C, Schoebinger M, Fink C, Ley S, Puderbach M, Eichinger M, et al. Evaluation of lung volumetry using dynamic three-dimensional magnetic resonance imaging. *Invest Radiol* 2005;40(3):173–9.
 - [7] Blackall JM, Ahmad S, Miquel ME, McClelland JR, Landau DB, Hawkes DJ. MRI-based measurements of respiratory motion variability and assessment of imaging strategies for radiotherapy planning. *Phys Med Biol* 2006;51:4147–69.
 - [8] Tokuda J, Morikawa S, Haque HA, Tsukamoto T, Matsumiya K, Liao H, et al. Adaptive 4D MR imaging using navigator-based respiratory signal for MRI-guided therapy. *Magn Reson Med* 2008;59:1051–61.
 - [9] Hu Y, Caruthers SD, Low DA, Parikh PJ, Mutic S. Respiratory amplitude guided 4-dimensional magnetic resonance imaging. *Int J Radiat Oncol Biol Phys* 2013;86(1):198–204. <https://doi.org/10.1016/j.ijrobp.2012.12.014>.
 - [10] Du D, Caruthers SD, Glide-Hurst C, Low DA, Li HH, Mutic S, et al. High-quality T2-weighted 4-dimensional magnetic resonance imaging for radiation therapy applications. *Int J Radiat Oncol Biol Phys* 2015;92:430–7. <https://doi.org/10.1016/j.ijrobp.2015.01.035>.
 - [11] Tryggestad E, Flammang A, Han-Oh S, Hales R, Herma J, McNutt T, et al. Respiration-based sorting of dynamic MRI to derive representative 4D-MRI for radiotherapy planning. *Med Phys* 2013;40:051909.
 - [12] Liu Y, Yin FF, Chen NK, Chu ML, Cai J. Four dimensional magnetic resonance imaging with retrospective k-space reordering: a feasibility study. *Med Phys* 2015;42(2):534–41. <https://doi.org/10.1118/1.4905044>.
 - [13] Li G, Wei J, Olek D, Kadbi M, Tyagi N, Zakian K, et al. Direct comparison of respiration-correlated four-dimensional magnetic resonance imaging reconstructed using concurrent internal navigator and external bellows. *Int J Radiat Oncol Biol Phys* 2017;97(3):596–605. <https://doi.org/10.1016/j.ijrobp.2016.11.004>.
 - [14] Vedam SS, Keall PJ, Kini VR, Mostafavi H, Shukla HP, Mohan R. Acquiring a four-dimensional computed tomography dataset using an external respiratory signal. *Phys Med Biol* 2003;48:45–62.
 - [15] von Siebenthal M, Székely G, Gamper U, Boesiger P, Lomax A, Cattin P. 4D MR imaging of respiratory organ motion and its variability. *Phys Med Biol* 2007;52:1547–64.
 - [16] Cai J, Chang Z, Wang Z, Segars WP, Yin F-F. Four-dimensional magnetic resonance imaging (4D-MRI) using image-based respiratory surrogate: a feasibility study. *Med Phys* 2011;38:6384–94.
 - [17] Paganelli C, Summers P, Bellomi M, Baroni G, Riboldi M. Liver 4DMRI: a retrospective image-based sorting method. *Med Phys* 2015;42:4814–21.
 - [18] Uh J, Ayaz Khan M, Hua C. Four-dimensional MRI using an internal respiratory surrogate derived by dimensionality reduction. *Phys Med Biol* 2016;61(21):7812–32.
 - [19] Hui C, Wen Z, Stemkens B, Tijssen RH, van den Berg CA, Hwang KP, et al. 4D MR imaging using robust internal respiratory signal. *Phys Med Bio*. 2016;61(9):3472–87. <https://doi.org/10.1088/0031-9155/61/9/3472>.
 - [20] Buerger C, Clough RE, King AP, Schaeffter T, Prieto C. Nonrigid motion modeling of the liver from 3-D undersampled self-gated golden-radial phase encoded MRI. *IEEE Trans Med Imaging* 2012;31(3):805–15. <https://doi.org/10.1109/TMI.2011.2181997>.
 - [21] Yang W, Fan Z, Tuli R, Deng Z, Pang J, Wachsmann A, et al. Four-dimensional magnetic resonance imaging with 3D radial sampling and self-gating based K-space sorting: early clinical experience on pancreatic cancer patients. *Int J Radiat Oncol Biol Phys* 2015;93(5):1136–43. <https://doi.org/10.1016/j.ijrobp.2015.08.028>.
 - [22] Deng Z, Pang J, Yang W, Yue Y, Sharif B, Tuli R, et al. Four-dimensional MRI using three-dimensional radial sampling with respiratory self-gating to characterize temporal phase-resolved respiratory motion in the abdomen. *Magn Reson Med* 2016;75:1574–85.
 - [23] Gianoli C, Riboldi M, Spadea MF, Travaini LL, Ferrari M, Mei R, et al. A multiple points method for 4D CT image sorting. *Med Phys* 2011;38:656–67.
 - [24] Paganelli C, Summers P, Gianoli C, Bellomi M, Baroni G, Riboldi M. A tool for validating MRI-guided strategies: a digital breathing CT-MRI phantom of the abdominal site. *Med Biol Eng Compu* 2017;55(11):2001–14. <https://doi.org/10.1007/s11517-017-1646-6>.
 - [25] Paganelli C, Peroni M, Riboldi M, Sharp GC, Ciardo D, Alterio D, et al. Scale invariant feature transform in adaptive radiation therapy: a tool for deformable image registration assessment and replanning indication. *Phys Med Biol* 2013;58:287–99.
 - [26] Bjerre T, Crijns S, af Rosenschöld PM, Aznar M, Specht L, Larsen R, et al. Three-dimensional MRI-linac intra-fraction guidance using multiple orthogonal cine-MRI planes. *Phys Med Biol* 2013;58:4943–50.
 - [27] Brix L, Ringgaard S, Sørensen TS, Poulsen P. 3D liver motion tracking using real-time 2D MRI. *Med Phys* 2014;41:042302.
 - [28] Fast MF, Eiben B, Menten MJ, Wetscherek A, Hawkes DJ, McClelland JR, et al. Tumour auto-contouring on 2d cine MRI for locally advanced lung cancer: a comparative study. *Radiother Oncol* 2017;125:485–91.
 - [29] Park H-S, Jun C-H. A simple and fast algorithm for K-medoids clustering. *Expert Syst Appl* 2009;36:3336–41.
 - [30] Paganelli C, Kipritidis J, Lee D, Baroni G, Keall P, Riboldi M. Image-based retrospective 4D MRI in external beam radiotherapy: a comparative study with a digital phantom. *Med Phys* 2018;45(7):3161–72. <https://doi.org/10.1002/mp.12965>.
 - [31] Rit S, van Herk M, Zijp L, Sonke JJ. Quantification of the variability of diaphragm motion and implications for treatment margin construction. *Int J Radiat Oncol Biol Phys* 2012;82:e399–407.
 - [32] Venkat RB, Sawant A, Suh Y, George R, Keall PJ. Development and preliminary evaluation of a prototype audiovisual biofeedback device incorporating a patient-specific guiding waveform. *Phys Med Biol* 2008;53:N197–208.
 - [33] Hollander M, Wolfe DA. Nonparametric statistical methods. Hoboken, NJ: John Wiley & Sons, Inc.; 1999.