



## On-line 3D motion estimation using low resolution MRI

Markus Glitzner, Baudouin Denis de Senneville, Jan Lagendijk, Bas Raaymakers, Sjoerd Crijs

### ► To cite this version:

Markus Glitzner, Baudouin Denis de Senneville, Jan Lagendijk, Bas Raaymakers, Sjoerd Crijs. On-line 3D motion estimation using low resolution MRI. Physics in Medicine and Biology, 2015. hal-01181385

**HAL Id: hal-01181385**

**<https://hal.science/hal-01181385>**

Submitted on 30 Jul 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On-Line 3D Motion Estimation Using Low Resolution MRI

M Glitzner<sup>1</sup>, B Denis de Senneville<sup>1,2</sup>, J J W Lagendijk<sup>1</sup>,  
B W Raaymakers<sup>1</sup>, S P M Crijns<sup>1</sup>

<sup>1</sup> Department of Radiotherapy, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands

<sup>2</sup> Mathematical Institute of Bordeaux, UMR 5251 CNRS/University of Bordeaux, 33405 Talence Cedex, France

E-mail: m.glitzner@umcutrecht.nl

## Abstract.

Image processing such as deformable image registration finds its way into radiotherapy as a means to track non-rigid anatomy. With the advent of magnetic resonance imaging (MRI) guided radiotherapy, intrafraction anatomy snapshots become technically feasible.

magnetic resonance (MR) imaging provides the needed tissue signal for high-fidelity image registration. However, acquisitions, especially in 3D, take a considerable amount of time. Pushing towards real-time adaptive radiotherapy, MR imaging needs to be accelerated without degrading the quality of information.

In this paper, we investigate the impact of image resolution on the quality of motion estimations. Potentially, spatially undersampled images yield comparable motion estimations. At the same time, their acquisition times would reduce greatly due to the sparser sampling. In order to substantiate this hypothesis, an exemplary 4D dataset of the abdomen is downsampled gradually. Subsequently, spatiotemporal deformations are extracted consistently using the same motion estimation for each downsampled dataset. Errors between the original and the respectively downsampled version are then evaluated.

Compared to ground-truth, results show high similarity of deformations estimated from downsampled image data. Using a dataset with  $(2.5mm)^3$  voxel size, deformation fields could be recovered well up to a downsampling factor of 2, i.e.  $(5mm)^3$ . In a therapy guidance scenario MRI, imaging speed would accordingly increase approximately fourfold, with acceptable loss of estimated motion quality.

*Keywords:* deformable image registration, local deformations, resampling, data sufficiency *Phys. Med. Biol.*

## 1. Introduction

Recent developments enabled on-line magnetic resonance imaging (MRI) during radiotherapy (Raaymakers et al. 2009, Lagendijk et al. 2014). In further consequence, new possibilities for elements in the feedback chain of adaptive radiotherapy (ART) arise: in contrast to currently available real-time positioning devices which provide only surrogate displacement data, MRI returns 3D anatomical information of both target volume and organs at risk (OAR).

After acquisition, 3D displacements need to be estimated from the dynamic volumes using image-based motion estimation algorithms (Oliveira & Tavares 2012). The displacements become 3D deformation vector field (DVF) which can be used as feedback variable for e.g. dose calculations or target tracking during intervention (de Senneville et al. 2011). However, MRI is an inherently slow image modality and acquisition speed depends directly on image resolution. Additionally, MRI acquisition speedup is limited by both technical (gradient strength/noise bandwidth) and physiological (specific absorption rate (SAR), peripheral nerve stimulation (PNS)) constraints (?). Similarly, model based acceleration techniques such as parallel imaging (Blaimer et al. 2004) or compressed sensing (Lustig et al. 2007) are limited by the validity of their underlying model and the demand for reconstruction time.

The aspiration of this work is to investigate the impact of image resolution on the quality of motion estimation. Potentially, with lower sufficient resolution for comparable motion estimation, magnetic resonance (MR) acquisition could be accelerated significantly. The gain in acquisition speed would enable applications which demand for high-frequency, real-time MRI, such as 3D target and dose tracking for ART.

## 2. Methods

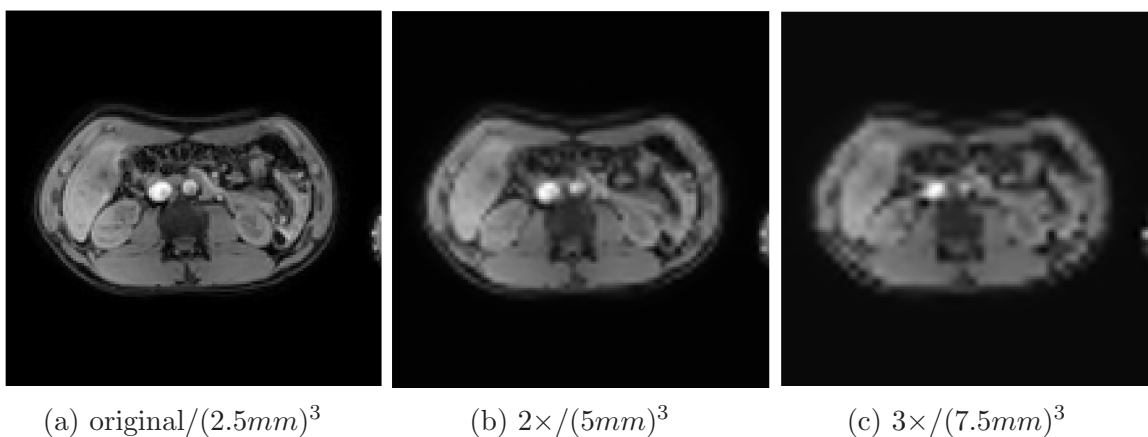


Figure 1: Downsampled images resampled back into the original grid. Images exemplarily show the image information loss with higher downsampling factors

### 2.1. Image data and imaging

As ground truth for all experiments, ~~a 4D-MRI dataset depicting a volunteer's abdomen~~ was datasets depicting volunteers' abdomen were used. The use of volunteer data was chosen for the lack of satisfying synthetic phantoms for deformable image registration.

To avoid image modulation from breathing motion, respiratory gating was employed with a gating window of  $5mm$  using navigator echoes on the interface between lung and left liver-dome. Gating in end-of-expiration was chosen because of an expected low residual breathing motion uncertainty for abdominal sites (Berbeco et al. 2006).

Since isotropic resolution is a prerequisite of the utilized image processing, the acquisition was tailored to acquire data with  $(2.5mm)^3$  voxel size.  $k$ -space was filled in 3D manner in order to avoid malicious slice effects, such as slice overlap and inter-slice motion. Additionally, 3D  $k$ -space sampling yields higher signal to noise ratio (SNR) compared to sliced acquisitions.

The navigator triggered an echo planar imaging (EPI) readout with an EPI-factor of 21,  $T_R/T_E/\alpha = 47ms/9.8ms/25^\circ$ . Binomial pulses  $(1 - 2 - 1)$  (Hore 1983) were used for selective water excitation, concurrently suppressing the fat signal. Furthermore, moderate sensitivity encoding (SENSE) ( $P = 1.5$ ,  $S=1.5$ ) was used to further accelerate imaging.

~~In total, -~~

<u>Volunteer</u>	<u># of 3D dynamics</u>	<u>total scan duration</u>
<u>1</u>	100	<u>17min</u>
<u>2</u>	<u>150</u>	<u>15min</u>
<u>3</u>	<u>80</u>	<u>8min</u>

Table 1: Overview of the acquired amount of data for each volunteer.

The amount of data acquired for each volunteer is stated in table 1. From the individual total scan duration and the amount of 3D dynamics~~were measured over 17min, with an approximate time per dynamic of 3s. The volunteer was,~~ it is possible to infer gating efficiencies for each volunteer: While for the first volunteer, an approximate amount of 6 images could be extracted per minute, scanning the other two yielded 10 images per minute, on average. During the scans the volunteers were breathing freely and ~~was~~ were not commanded any breathing alteration.

### 2.2. Image Processing

To simulate coarser sampling, downsampled data was synthetically generated using MATLAB's (The Mathworks, Natick, MA, USA) tri-cubic reconstruction kernels. The cubic kernel was chosen over linear interpolation to more closely resemble the ideal reconstruction kernel.

The original grid with  $(2.5mm)^3$  voxel size was downsampled with factors  $k_i = [1.5, 2, 2.5, 3, 4]$  to simulate the loss of image content by coarser imaging. After

downsampling, the data was again upsampled by the respective inverse factor in order to be able to compare data on a common (the original) grid (see figure 1).

For the registration, baseline intensity variations throughout the imaging stack had to be removed. ~~Accordingly, all images were equalized into the same dynamic range by normalizing the (magnitude) images to the median of the maxima found in all individual images. Subsequently, intensity values were capped to the estimated median-max.~~ This was necessary because the  $T_1$ -weighted imaging sequence yields the brightest signal for blood. Consequently, since the cardiac cycle has a similar time scale as the dynamic imaging, these signal maxima appeared modulated in vicinal vessels. Accordingly, the median of the maxima found in all individual (magnitude) images was calculated. Subsequently, the intensity values were capped to the estimated median-max multiplied with a cut-off factor  $k_{cutoff} = 0.8$ . Finally, the mean of every individual image dynamic was normalized to the mean of the reference image. This proofed to be more robust to grey level variations, compared to referencing to the dynamic range of the reference image.

### 2.3. Motion estimation

Motion throughout the imaging session was estimated with an implementation (Roujol et al. 2010) of non-rigid registration based on optical flow. As derived by Horn & Schunck (1981), the optical flow's objective function is formulated as

$$\iiint_{\Omega} (I_x v_x + I_y v_y + I_z v_z + I_t)^2 + \alpha^2 (\|\nabla v_x\|_2^2 + \|\nabla v_y\|_2^2 + \|\nabla v_z\|_2^2) dx dy dz \quad (1)$$

The left side of the functional states that every spatio-temporal gray level variation  $\nabla I$  and  $I_t$  is attributed to motion  $\begin{pmatrix} v_x & v_y & v_z \end{pmatrix}^T$ . In addition, the regularization term on the right side is introduced to penalize sudden spatial changes of motion, which was validated for organ deformations (Østergaard Noe et al. 2008). Accordingly, the unknown DVF is then extracted by minimization of equation 1 with respect to  $\begin{pmatrix} v_x & v_y & v_z \end{pmatrix}^T$ .

In this work, the regularization parameter was set to  $\alpha = 0.3$ . This value was chosen phenomenologically, in order to avoid instabilities in the motion estimation, that would lead to non-continuous estimations of the DVFs (Roujol et al. 2011).

Motion was estimated between the reference image (first image dynamic) and every consecutive imaging dynamic~~—~~. This yielded yielding 100 DVFs for the  $N = 100$  dynamics. This process was repeated for every 4D stack of the respectively applied resampling factor with index  $i^{th}$ .

### 2.4. Error assessment

As depicted in figure 2, the quality of the motion estimation for the  $i^{th}$  dataset was assessed by taking the differences between the DVF from resampled data and the DVF

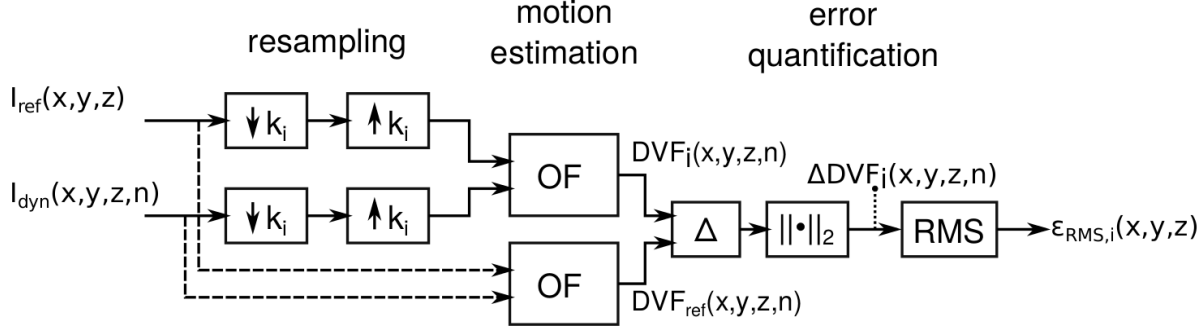


Figure 2: Evaluation workflow for motion estimation and error quantification:  $k_i$  indicates the respective factors for the resampling operators ( $\downarrow, \uparrow$ ), OF means the motion estimation using optical flow. Subsequent  $L_2$ -norm ( $\|\cdot\|_2$ ),  $\Delta$  and  $RMS$  operators work on a per-pixel basis.  $x,y,z$  are the spatial coordinates,  $n$  denotes the temporal index and  $i$  is the index for the individual downsampling factors.

of the original data. However, the multidimensional, 4D vector fields

$$\mathbf{DVF}_i(\mathbf{x}, n) = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} (\mathbf{x}, n) \quad (2)$$

require integral metrics for convenient quantification of the differences. For clarity, the analyses are done solely in terms of absolute, component-wise differences (endpoint error (Baker et al. 2011))

$$\Delta DVF_i(\mathbf{x}, n) = \|\mathbf{DVF}_i(\mathbf{x}, n) - \mathbf{DVF}_{\text{ref}}(\mathbf{x}, n)\|_2, \quad (3)$$

with  $n$  denoting the temporal index of a 3D-dynamic.

Locally, two salient voxels were selected and analyzed over time [for the longest dataset \(volunteer 1 in table 1\)](#). Furthermore, for generating global statistics, we make use of the root mean square (RMS) which is applied on the absolute component-wise differences between two DVFs

$$\epsilon_{\text{RMS},i}(\mathbf{x}) = \sqrt{\frac{1}{N} \sum_{n=1}^N \Delta DVF_i(\mathbf{x}, n)}. \quad (4)$$

## 2.5. Data selection

In order to limit the sample size, comparatively mobile voxels including continuous and spontaneous motion events were elected for the global statistics.

Using linear regression, coefficients  $\beta(\mathbf{x})$  were calculated for every voxel over time as

$$\underset{\alpha(\mathbf{x}), \beta(\mathbf{x})}{\operatorname{argmin}} \{ \alpha(\mathbf{x}) + \beta(\mathbf{x})n = \|\mathbf{DVF}_{\text{ref}}(\mathbf{x}, n)\|_2 \}. \quad (5)$$

Voxels with  $\beta(\mathbf{x}) > 0.02\text{mm}(\text{dynamic})^{-1}$  were considered in the global statistics. In combination with the metric in eq. 5 this threshold extracts the elected dynamic voxels.

In addition to regression coefficients, the RMS of the DVFs could be a feasible quantifier of the target mobility in a voxel over time. However, for the chosen MR acquisition type, it did not prove feasible because of high false-positive rates. These were caused by the previously mentioned pulsation artifacts in the aorta and vena cava, randomly modulating the individual dynamics and thus the extracted DVF.

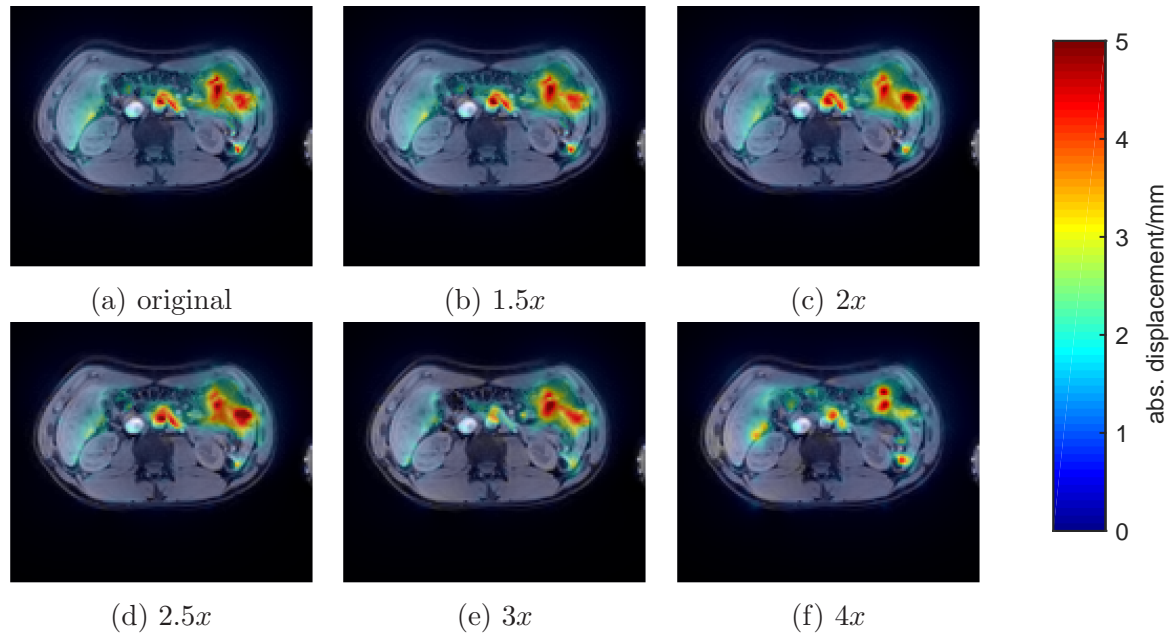


Figure 3: Absolute DVF in a selected slice of volunteer 1 for increasing downsampling (a-f). Qualitative similarity is apparent until downsampling with factor 2.5.

### 3. Results

Figure 3 shows a selected slice of ~~anatomy~~ the anatomy of volunteer 1 in the original image stack. A transparent overlay is applied onto anatomy, representing absolute displacement fields estimated in the last (100<sup>th</sup>) imaging dynamic for either of the 6 (resampled) datasets~~(+)~~.

Over the selected slice, the DVFs extracted from the individually resampled data appear visually congruent. The stretched displacement at the edge of the liver, next to the right kidney, appears consistently throughout all datasets and holds its structure up to a two-fold downsampling before it disintegrates.

Large deformations in the region between stomach and duodenum are prominent in all scenarios and retain shape. However, starting from factor 2.5, a visible blur is introduced to the DVF distribution. At 4x-downsampling, formerly prominent spots significantly decrease in motion magnitude.



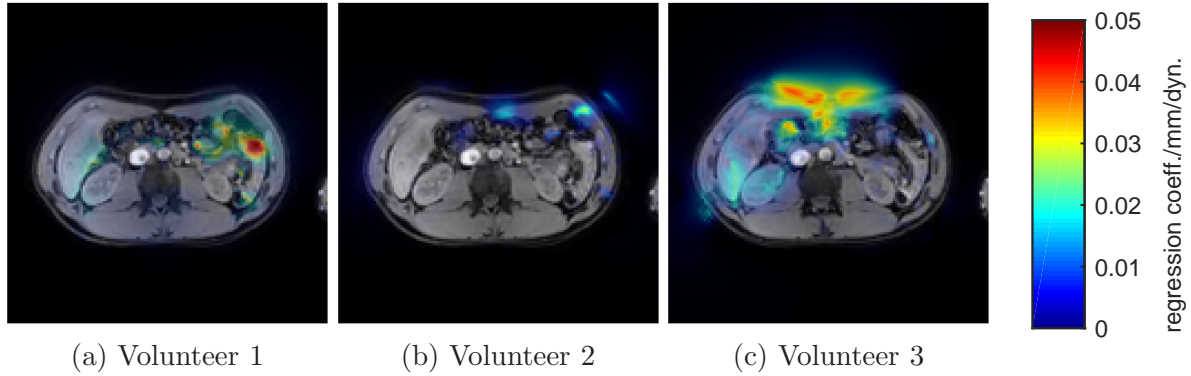


Figure 4: Global error quantification: distribution of the extracted regression coefficients  $\beta(\mathbf{x})$  over the anatomies of 3 volunteers

In figure fig. 4, the intersubject difference of mean motion  $\beta$  is apparent. While volunteers 1 and 3 have significant motion hotspots, the second volunteer's anatomy remains mostly steady throughout the scan duration. This reduces points which are in line with the inclusion criterium for the global RMS error statistics ( $\beta > 0.02\text{mm/dynamic}$ ).

According to what was observed qualitatively for volunteer 1 in figure 3, qunatitative error statistics in figure 5 show the monotonous increase in RMS error with increasing undersampling. This effect occurs repeatedly for the other two subjects. In addition to the average error, higher moments of the error statistics increase equally, which leads to outliers  $\epsilon_{RMS,i}(\mathbf{x}) > 2\text{mm}$  for downsampling with factors greater than 2.

In addition to the global statistics of many points (figure 5), local performance evaluation is shown in figure 6, illustrating the transient impact of downsampling. In the exemplarily selected point at the liver/kidney (figure 6a), a constant drift of about  $2.5\text{mm}/17\text{min}$  is expressed over time. With coarser sampling starting from factor 3, the deformation appears greatly deteriorated with RMS errors proportional to magnitude. However, until factor 2.5, errors appear independent of the magnitude and stay lower than  $1\text{mm}$ .

Slightly different in the intestinal sample (figure 6b), absolute errors start to exceed  $1\text{mm}$  already at factor 2.5. Undersampling with factors 3, as before, results in a greatly deteriorated DVF with errors exceeding  $2\text{mm}$ .

#### 4. Discussion

Although removing image content by spatial downsampling, the used image registration algorithm was able to recover deformation fields to a high degree. In the synthetic experiments the most finely resolved data  $((2.5\text{mm})^3)$  served as ground truth; with this, congruent DVF were able to be extracted up to a downsampling factor of 2  $((5\text{mm})^3$  voxel size), with a ~~local error of less than  $1\text{mm}$  and~~ global RMS deviations staying beneath  $1.5\text{mm}$ . Exemplary samples of local points resulted in errors of less than  $1\text{mm}$ .



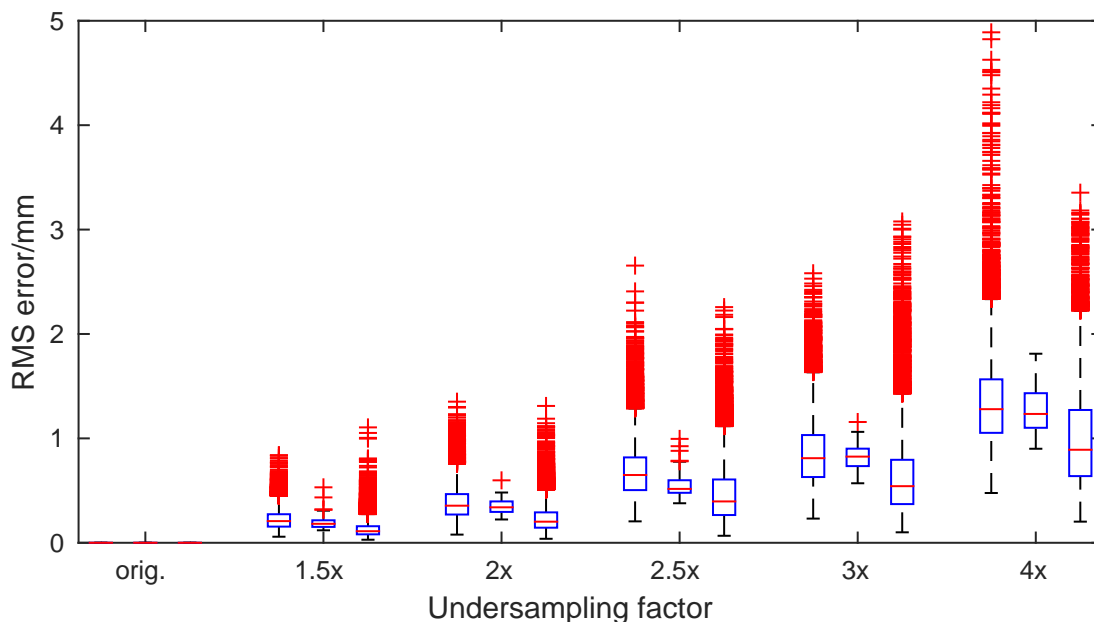


Figure 5: Global RMS error statistics: The boxplots illustrate the evolution of the RMS error within the elected voxels ( $\beta(\mathbf{x}) > 0.02mm(dyn.)^{-1}$ ) for the various resampling factors and volunteers. In each cluster the RMS error statistics from volunteer 1, 2 and 3 is plotted from left to right. For completeness, a boxplot for the original, non-resampled case was added.

This observation contradicts the idea that only images of diagnostic quality are eligible for image registration purposes. Just the opposite, deformable image registration is much more data sufficient. Accordingly, imaging for the generation of displacement data is less demanding on image resolution than e.g. diagnostic imaging.

The fact that estimation of non-rigid motion requires spatial regularization is backing this hypothesis. Adding a spatial smoothness constraint (eq. 1, right side) enables stable motion estimation for regions with small image gradients (Bruhn et al. 2005). Additionally, to suppress the impact of outliers, literature suggests the application of low-pass filters similar to those used in this work for image preconditioning in order to obtain better motion estimations (Barron et al. 1994).

The simulated low-resolution MRI proposed here inherently implements both spatial (velocity) regularization and image low-pass filtering.

In real-time MRI, where image resolution can and must be traded off against imaging time, the proposed method could be used to significantly accelerate the acquisition. Using  $2x$  spatial undersampling, the imaging time of about 3s for a  $(2.5mm)^3$ -acquisition could ideally be sped up fourfold while retaining the quality of extracted DVF.

It is important to note that offline resampling using the cubic reconstruction kernels is solely an approximation of the effect of an actual low resolution MR acquisition, due to accompanying changes in contrast and point spread function (PSF). However, in first approximation, the used reconstruction kernels are methodologically sound.

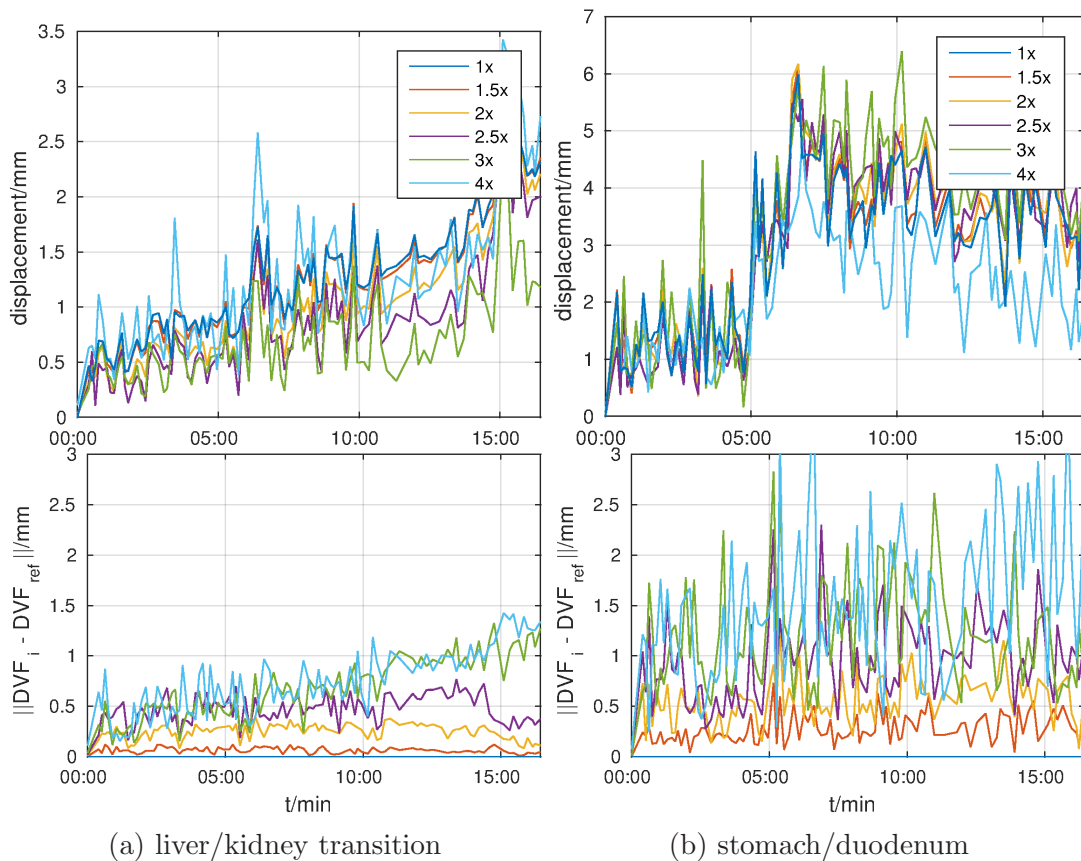


Figure 6: Two manually selected points of interesting DVF evolution in volunteer 1 for multiple downsampling factors

Combined with gated MRI, this technique enables fast 3D anatomy updates of abdominal volumes during the exhalation plateau. Used in an acquisition without respiratory trigger, i.e. free breathing, motion estimates would be less impacted by intra-scan artifacts (blurring) due to the shorter acquisition time.

Due to respiratory gating in the exemplary data, only medium- to long-term drift motion in the abdomen was monitored. Thus, observed deformations within the body are limited and spatially smooth to a first approximation. The image registration is constrained equally, having a spatial smoothness regularization.

Accordingly, similar results are expected with slowly changing pelvic sites. However, results may be different when observing steep gradients in DVF over time, such as when regarding abdominal breathing excursions. This would put stress on the differences between the algorithmic and the physiologic constraints. Due to the lack of (real-time 3D) data being eligible to be considered as ground-truth, this scenario has not been considered in this work. A potential data source could be retrospectively binned 4D MRI data.

Possibly, the accepted error could be evaluated per site or per subject with pre-treatment assessments: high-resolution data could be acquired and gradually

downsampled until remarkable sites exceed an error threshold. These assessment could take place under worst-case treatment conditions, e.g. during rectal or bladder filling in pelvic cases, in order to test the error for an extreme scenario. The reward would be a highly accelerated MRI acquisition, without relying on the further introduction of signal models.

## 5. Conclusion

This investigation showed, that deformation estimation on coarsely resolved images yields congruent results when compared to more highly resolved images. The examined insensitivity of nonrigid image registration to spatial downsampling can be used as a facility to speed up MRI acquisitions, specifically for real-time therapy guidance. Careful optimization between acceptable registration error and necessary imaging speed will however be necessary, since this trade-off cannot be made in an intuitive way.

## 6. Acknowledgements

The authors thank the ITEA (project 12026, SoRTS), the European Research Council (project ERC-2010-AdG-20100317, Sound Pharma) and Elekta AB (Stockholm, Sweden) for funding.

## References

- Baker S, Scharstein D, Lewis J P, Roth S, Black M J & Szeliski R 2011 A database and evaluation methodology for optical flow *Int. J. Comput. Vis.* **92**(1), 1–31.
- Barron J L, Fleet D J & Beauchemin S S 1994 Performance of optical flow techniques *Int. J. Comput. Vis.* **12**(1), 43–77.
- Berbeco R I, Nishioka S, Shirato H & Jiang S B 2006 Residual motion of lung tumors in end-of-inhale respiratory gated radiotherapy based on external surrogates. *Med. Phys.* **33**(11), 4149–4156.
- Blaimer M, Breuer F, Mueller M, Heidemann R M, Griswold M A & Jakob P M 2004 SMASH, SENSE, PILS, GRAPPA: how to choose the optimal method. *Top. Magn. Reson. Imaging* **15**(4), 223–236.
- Bruhn A, Weickert J & Schnörr C 2005 Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods *Int. J. Comput. Vis.* **61**(3), 1–21.
- de Senneville B D, Ries M, Bartels L W & Moonen C T W 2011 in T Kahn & H Busse, eds, ‘Interv. Magn. Reson. Imaging’ Springer Berlin Heidelberg pp. 349–366.
- Hore P 1983 Solvent suppression in fourier transform nuclear magnetic resonance *J. Magn. Reson.* **55**(2), 283–300.
- Horn B K & Schunck B G 1981 Determining optical flow *Artif. Intell.* **17**(1-3), 185–203.
- Legendijk J J W, Raaymakers B W & van Vulpen M 2014 The Magnetic Resonance Imaging-Linac System *Semin. Radiat. Oncol.* **24**(3), 207–209.
- Lustig M, Donoho D & Pauly J M 2007 Sparse MRI: The application of compressed sensing for rapid MR imaging *Magn. Reson. Med.* **58**(6), 1182–1195.
- Oliveira F P & Tavares J a M R 2012 Medical image registration: a review *Comput. Methods Biomech. Biomed. Engin.* (April 2015), 1–21.

- Østergaard Noe K, De Senneville B D, Elstrøm U V, Tanderup K & Sørensen T S 2008 Acceleration and validation of optical flow based deformable registration for image-guided radiotherapy. *Acta Oncol.* **47**(7), 1286–1293.
- Raaymakers B W, Lagendijk J J W, Overweg J, Kok J G M, Raaijmakers a J E, Kerkhof E M, van der Put R W, Meijsing I, Crijs S P M, Benedosso F, van Vulpen M, de Graaff C H W, Allen J & Brown K J 2009 Integrating a 1.5 T MRI scanner with a 6 MV accelerator: proof of concept. *Phys. Med. Biol.* **54**(12), N229–N237.
- Roujol S, Ries M, Moonen C & Denis De Senneville B 2011 Automatic nonrigid calibration of image registration for real time MR-guided HIFU ablations of mobile organs *IEEE Trans. Med. Imaging* **30**(10), 1737–1745.
- Roujol S, Ries M, Quesson B, Moonen C & De Senneville B D 2010 Real-time MR-thermometry and dosimetry for interventional guidance on abdominal organs *Magn. Reson. Med.* **63**(4), 1080–1087.