



HAL
open science

An improved optical flow tracking technique for real-time MR-guided beam therapies in moving organs

Cornel Zachiu, Nicolas Papadakis, Mario Ries, Chrit Moonen, Baudouin Denis de Senneville

► To cite this version:

Cornel Zachiu, Nicolas Papadakis, Mario Ries, Chrit Moonen, Baudouin Denis de Senneville. An improved optical flow tracking technique for real-time MR-guided beam therapies in moving organs. *Physics in Medicine and Biology*, 2015. hal-01212632

HAL Id: hal-01212632

<https://hal.science/hal-01212632v1>

Submitted on 7 Oct 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.

An improved optical flow tracking technique for real-time MR-guided **beam therapies** in moving organs

C Zachiu¹, N Papadakis², M Ries¹, C Moonen¹ and B Denis de Senneville^{1,2}

¹ Imaging Division, UMC Utrecht, Heidelberglaan 100, 3584 CX, Utrecht, Netherlands

² “Institut de Mathématiques de Bordeaux”, Université Bordeaux, 351 Cours de la Libération, 33405 TALENCE Cedex, France

E-mail: C.Zachiu@umcutrecht.nl, nicolas.papadakis@math.u-bordeaux1.fr, M.Ries@umcutrecht.nl, C.Moonen@umcutrecht.nl, b.desenneville@umcutrecht.nl

Abstract. Magnetic resonance (MR) guided **high intensity focused ultrasound (HIFU) and external beam radiotherapy (EBRT) interventions, to which we shall refer to as beam therapies/interventions**, are promising techniques for the non-invasive ablation of tumors in abdominal organs. Therapeutic energy delivery in these areas becomes, however, challenging due to the continuous displacement of the organs with respiration. Previous studies have addressed this problem by coupling high-framerate MR-imaging with a **tracking** technique based on the algorithm proposed by Horn&Schunck (H&S), which was chosen due to its fast convergence rate and highly parallelizable numerical scheme. Such characteristics were shown to be indispensable for the real-time guidance of **beam** therapies. In its original form, however, the algorithm is sensitive to local gray-level intensity variations not attributed to motion such as those that occur, for example, in the proximity of pulsating arteries.

In this study, an improved motion estimation strategy which reduces the impact of such effects is proposed. Displacements are estimated through the minimization of a variation of the H&S functional for which the quadratic data fidelity term was replaced with a term based on the linear L^1 norm, resulting in what we have called an L^2 - L^1 functional.

The proposed method was tested in the liver and kidney of two healthy volunteers under free-breathing conditions, on a data set comprised of 3000 images equally divided between the volunteers. Results have shown that, compared to the existing approaches, our method demonstrates a greater robustness to local gray-level intensity variations introduced by arterial pulsations. Additionally, the computational time required by our implementation make it compatible with the work-flow of real-time **MR-guided beam interventions**.

To the best of our knowledge this study was the first to analyze the behavior of an L^1 -based optical flow functional in an applicative context: real-time MR-guidance of **beam** therapies in moving organs.

Keywords: Motion analysis, Real-time MR-guidance, EBRT, HIFU.

Submitted to: *Phys. Med. Biol.*

1. Introduction

MR guided high intensity focused ultrasound (MRg-HIFU) together with MR guided external beam radiotherapy (MRg-RT) are attractive techniques for treating tumors deep inside the human body (Cline et al. 1992, Hynynen et al. 1996, Mutic & Dempsey 2014, Crijs et al. 2012). The non-invasive nature of such interventions opens up additional treatment options for patient groups affected by tumors that are inaccessible to current surgical equipment (Baskar et al. 2012, Foley et al. 2013, Tempany et al. 2011). In particular, MRg-HIFU has the potential to ablate primary and metastatic tumors in a single session (Illing et al. 2005). It has already met success in treating several conditions such as symptomatic uterine fibroids, palliative treatment of bone metastases, prostate cancer, breast fibroadenoma and a number of functional neuropathic disorders (Foley et al. 2013). Although external beam radiotherapy (EBRT) is already a standard approach for treating pathologies in numerous locations inside the human body (Delaney et al. 2005), **at the time of this study there is just a single clinically-available device which makes use of real-time MR guidance, i.e. the ViewRay system** (Mutic & Dempsey 2014). Nevertheless, there are great perspectives for MRg-RT clinical interventions with the development of the first **MR-Linac** prototypes (Crijs et al. 2012, Raaymakers et al. 2009, Crijs et al. 2011). One of the challenges that might be encountered during **EBRT and HIFU interventions, to which for the purpose of this study we shall alternatively refer to as beam therapies/interventions**, is the necessity of a tracking technique for the targeted tissue. When therapy is conducted in the upper abdomen or in the thorax, energy delivery is hampered by physiological motion, which can be coarsely categorized in the following sub-types: respiratory motion, cardiac motion, long term motion (e.g. peristalsis) and spontaneous motion (e.g. coughing, twitching) (Zachiu et al. 2015). In the present work the focus will mainly be on respiratory motion, which, during energy deliveries, was identified to be the most problematic (Booth 2002, Langen & Jones 2001, Goitein 2004). In order to prevent unnecessary damage to healthy tissues, therapeutic energy delivery needs to be correlated with the underlying displacements (Langen & Jones 2001, Ries et al. 2010). For example, a gating strategy could be employed in which the beam is turned on and off, depending on whether the targeted area is in a predefined fixed location or not (Crijs et al. 2011, Ohara et al. 1989, Okada et al. 2006). An alternative would be to continuously re-adjust the focus of the beam according to the motion pattern exhibited by the target tissue (Ries et al. 2010). Taking motion into account during energy deliveries will also prevent the therapeutic energy from being diffused along the motion trajectory, which in turn would lead to the under-treatment of the treated tumorous tissue (Langen & Jones 2001, Ries et al. 2010).

An additional challenge is the need for a real-time therapy efficiency monitoring strategy. For example, during a MRg-HIFU intervention, the interaction of the ultrasonic beam with the target tissue and the organs at risk (OAR) is observed through a real-time temperature feedback loop. Therefore, high-framerate MR imaging has to be run in

parallel with the energy delivery. However, the time varying geometric distortions induced by the moving target might end up degrading the temperature measurements (Denis de Senneville et al. 2011).

Thus, a real-time motion compensation strategy is necessary both for gating and/or steering the energy beam and, in the particular case of a HIFU intervention, correcting temperature maps. Note that the quality of the latter has a direct impact on the thermal dose measurements, which is the main indicator on whether the endpoint of a HIFU therapy has been reached or not (Ries et al. 2010, Roujol et al. 2010).

Motion compensation for EBRT was historically achieved by several approaches including the increase of treatment margins, establishing a relation between external and internal motion through surrogate devices usually situated outside the patients body or tracking the target motion directly by implanting MV/kV detectable fiducial markers. Increasing tumor margins, however, often leads to an increased unnecessary toxicity to otherwise healthy tissues. Surrogate external devices were shown to aid in target tracking for several particular cases (Seppenwoolde et al. 2007), however, for some sites, it becomes difficult to establish a correlation with target motion, since the relationship between external and internal motion may vary during treatment (Feng et al. 2009). While the depiction of target motion directly by tracking implanted fiducial markers may be a more reliable solution than external surrogates (Minn et al. 2009), such an approach is usually unable to handle deformations and involves a certain degree of invasiveness. All these approaches and several additional methods, together with their advantages and disadvantages are discussed in great detail in (Keall et al. 2006). An image-guided intervention, thus, provides a more attractive solution since tracking can be performed in a completely non-invasive manner. In particular, MR imaging can provide real-time target visualization with high soft tissue contrast. Thus, in anticipation to the future release of the MR-Linac, recent research aims at integrating MR-based estimated displacements in the computation of the delivered radiation dose and based on the results, corrections can be made and the interventional plan re-optimized (Crijns et al. 2011, Guckenberger et al. 2012).

Designing a real-time motion estimation strategy generates, however, additional challenges. The configuration and convergence of the target tracking algorithm must not exceed the sampling time of the image acquisition process. Also, a temporal margin has to be taken into account for the image acquisition itself and for several additional operations (such as thermometry, energy dose reconstruction, etc.). In this paper the focus will mainly be on the target tracking algorithm itself. It has already been shown in (Ries et al. 2010) and (Roujol et al. 2010) that, for respiratory motion compensation, the algorithm proposed by Horn & Schunck (Horn & Schunck 1981) is compatible with the work-flow and real-time requirements of MRg-HIFU interventions in the abdomen. The optical flow formulation of Horn & Schunck, initially proposed in the context of estimating motion in video sequences in 1981, assumes that pixels conserve their intensity along their trajectory, to which a spatial regularity constraint of the estimated motion is added. This can be mathematically expressed with the following functional

involving two quadratic terms:

$$E_{L^2L^2}(u, v) = \iint_{\Omega} (I_x u + I_y v + I_t)^2 + \alpha^2 (\|\nabla u\|_2^2 + \|\nabla v\|_2^2) \, dx dy \quad (1)$$

where $\Omega \subseteq \mathbb{R}^2$ is the image domain, u and v are the components of the 2D displacement vectors and α is a user defined weighting factor designed to link the data fidelity term (first term of the integral in Eq. (1)) and the regularity of the estimated motion field (second term of the integral in Eq. (1)). In the data fidelity term, assuming pixel gray-level intensity conservation between successive images, $I_{x,y}$ and I_t are the spatial and respectively temporal partial derivatives of the image pixel intensity. The regularization term is given by $\|\nabla u\|_2^2 = u_x^2 + u_y^2$ and $\|\nabla v\|_2^2 = v_x^2 + v_y^2$, with u_x , u_y , v_x and v_y being the partial spatial derivatives of u and v respectively. For the remainder of this paper $E_{L^2L^2}(u, v)$ will be referred to as the L^2 - L^2 functional. The estimated displacements are given by the minimizers of $E_{L^2L^2}(u, v)$ with respect to u and v .

Motion estimation based on the L^2 - L^2 functional has several properties that make it attractive for respiratory motion compensation during real-time MR-guided **beam** interventions:

- It requires tuning only one parameter, namely, the regularization parameter α . In practice, the value of α needs to be optimized only once for a particular contrast weighting, after which the same one can be used for an unlimited number of patients. Its value needs to be re-optimized only if the MR-sequence is significantly changed (Roujol et al. 2011).
- The numerical scheme employed for the minimization of $E_{L^2L^2}(u, v)$ has a pixel-wise nature which facilitates parallel processing. This, in turn, reduces the computational time of the algorithm, which is advantageous for applications with short processing latency requirements.

In general, respiratory motion compensation during MR-guided **EBRT or HIFU** interventions requires high-framerate imaging ($> 10\text{Hz}$)(Ries et al. 2010, Roujol et al. 2010, Roujol et al. 2011). At such sampling frequencies, depending on the image contrast weighting, arterial pulsations and peristaltic contractions may become apparent, and manifest themselves as local intensity variations between successive images. Consequently, pixel gray-level intensity is no longer conserved. This will have a direct impact on the quality of the motion estimates, since the assumption made for the construction of the data fidelity term in the H&S functional is locally violated. It is expected, however, that in such cases, these effects are partially reduced by the smoothness constraint imposed on the estimated motion by the regularization term. For example, if tracking is performed for the liver, even if the registration process might be influenced by the pulsating hepatic arteries, the displacements estimated at the edges of the liver will be propagated inwards and reduce the effects caused by the local gray-level intensity variations. This, however, might become problematic when the hepatic tissue appears in the images with a low gray-level intensity, and the registration process

has to rely mostly on the information provided by the blood vessels. Such is the case, for example, of an MRg-HIFU intervention for which, during energy delivery, real-time thermometry usually provides T_2^* -weighted images. Under such a contrast weighting, the hepatic tissue appears dark, while the blood vessels are hyper-intense, thus it is expected that arterial pulsations will have a large impact on the motion estimates. To the authors knowledge such effects have not been previously quantified in the context of image guided **beam** interventions.

In the current work we propose solving the motion estimation problem via the following functional:

$$E_{L^2L^1}(u, v) = \iint_{\Omega} |I_x u + I_y v + I_t| + \beta^2 (\|\nabla u\|_2^2 + \|\nabla v\|_2^2) \, dx dy \quad (2)$$

Given that the data fidelity term of the above functional is no longer quadratic, it is expected that the estimated motion will be less prone to errors caused by local gray-level intensity perturbations in cases such as the one described above. For the remainder of this paper $E_{L^2L^1}(u, v)$ will be referred to as the L^2 - L^1 functional.

In general, quantifying the improvement in the quality of the estimates provided by a new optical flow functional over the ones provided by the original model is challenging since:

- The new functional might compensate for the drawbacks of the original model only in a limited number of scenarios, having a weaker performance otherwise.
- The input parameters of the compared motion estimation models must be set in such a way that the difference in the quality of the estimates is not biased by a sub-optimal configuration.
- In order to keep the intervention non-invasive, *in-vivo* quantification of the quality of the estimates is made difficult due to the lack of a reliable gold standard.

Thus, in the current study, we additionally propose an experimental setup designed to:

- Quantify the effect of arterial pulsations on the motion estimates obtained using the L^2 - L^2 model.
- Quantify the improvement in the quality of the estimates provided by the proposed L^2 - L^1 functional over the ones provided by the L^2 - L^2 criterion.

2. Method description

2.1. Proposed motion estimation strategy based on the L^2 - L^1 functional

For the original L^2 - L^2 model, minimization with respect to u and v is achieved by solving the Euler-Lagrange equations. However, the proposed L^2 - L^1 functional is no longer differentiable, thus it cannot be minimized by using the same approach. Since $E_{L^2L^1}(u, v)$ is a convex functional, we propose its minimization via the primal-dual algorithm (Chambolle & Pock 2011). Special attention was paid to the configuration

parameters of the numerical scheme in order to ensure a fast convergence of the algorithm under various conditions in terms of noise/observed organ displacement amplitudes. The implementation of the algorithm was included as a part of a larger motion estimation library that is freely accessible[†].

2.1.1. Implemented optimization scheme We propose the minimization of the L^2 - L^1 functional via the primal-dual algorithm. Details concerning the primal-dual algorithm, its adaptation to the L^2 - L^1 criterion and the stability of the associated numerical scheme are found in Appendix A.

2.1.2. Coarse-to-fine scheme For both the L^2 - L^2 and the L^2 - L^1 models the data fidelity term is obtained through a linearization by Taylor approximation of the pixel intensity conservation constraint. By ignoring the higher order terms of the expansion, displacements that are larger than the size of one pixel cannot be estimated. In order to overcome this limitation, we adopted an approach similar to the one described in (Brox et al. 2004) : A coarse-to-fine strategy was carried out, which iterated the registration algorithm from a 4-fold downsampled image step by step to the original image resolution. In addition, an iterative refinement of the motion estimates was performed within each resolution. This implies running the algorithm several times at the same resolution, initializing the motion fields at the current run of the algorithm with the motion fields that resulted during the previous run. In this manner, the stability of the numerical scheme is improved and at the same time a better quality of the estimates is obtained.

2.1.3. Convergence criterion for the numerical scheme It was considered that the numerical scheme in Eq. (A.6) converged when the average variation of the motion magnitude from one iteration to the next was smaller than 10^{-3} pixels.

2.2. Experimental validation

All images (I) were registered to a common reference image (I_{ref}). Image acquisition and processing has been performed in 2D. This is due to the fact that, on currently available MR-systems, the acquisition of a 3D MR image with sufficiently good quality coupled with a registration procedure becomes timewise challenging considering the low latency requirements of a real-time MR-guided **beam** intervention.

The displacements estimated using the L^2 - L^1 model were compared in terms of both accuracy and precision to the ones estimated by the existing techniques based on the L^2 - L^2 functional.

2.2.1. MR imaging protocol Dynamic MR imaging was performed under free-breathing conditions on the abdomen of two healthy volunteers. For each volunteer a total number of 1500 dynamics have been acquired over a duration of ~ 2 min, with an

[†] The full library can be downloaded from <http://bsenneville.free.fr/RealTITracker/>.

imaging frame-rate of 12 images/s. The MR sequence was a single-shot gradient recalled echo employing the following parameters: repetition time (TR)=80 ms, echo time (TE)=37 ms, bandwidth in readout direction = 1250 Hz, flip angle=20°, field of view (FOV)=40×40 cm², matrix=160×160, in-plane voxel size=2.5 × 2.5mm², slice thickness=7 mm, using a 12 element phased array body coil. Each dynamic consisted in a single slice acquired in the coronal plane. Acquisition has been conducted on a Philips Achieva 1.5 T (Philips Healthcare, Best, The Netherlands).

2.2.2. Implementation of the motion estimation strategy based on the L^2 - L^2 functional

Given that the expression in Eq. (1) is differentiable, its minimizers were found by solving the associated Euler-Lagrange equations. These, in turn, generate a linear system which was solved using the Gauss-Seidel method. In order to make a proper comparison with the estimates provided by the L^2 - L^1 functional, the same coarse-to-fine scheme and convergence criterion as described in sections 2.1.2 and respectively 2.1.3 were employed.

2.2.3. Quantifying the quality of the estimated motion

Optical flow algorithms are intrinsically sensitive to gray-level intensity variations not attributed to motion. When the intervention is performed for organs in the upper abdomen, such variations might occur due to arterial pulsations:

- In the vicinity of the targeted region. This could happen, for example, when motion estimation is performed in the kidney, case in which the estimated displacements might be affected by the pulsations of the renal aorta.
- Within the targeted region. For example, when the intervention is performed in the liver, the estimated motion may be influenced by the pulsations of the hepatic arteries.

To assess the quality of the estimated motion, we propose a criterion based on the pixel-wise endpoint error (EE):

$$EE(x, y) = \sqrt{(u(x, y) - u_{gold}(x, y))^2 + (v(x, y) - v_{gold}(x, y))^2} \quad (3)$$

where (x, y) indicates the location on the image grid, (u, v) is the estimated motion and (u_{gold}, v_{gold}) is the gold standard motion. The spatial and temporal distribution of the EE was analyzed over two manually defined regions of interest (ROI) encompassing the kidney and the liver respectively. The procedure employed for obtaining u_{gold} and v_{gold} will be detailed in the following paragraphs.

To assess and compare the quality of the motion estimates provided by the L^2 - L^2 and the L^2 - L^1 models we propose three test benches:

Reliability test #1 The purpose of this reliability test was to quantify the impact of arterial pulsations on the respiratory displacements estimated by the two compared

functionals. Fig. 1 summarizes the processes and their succession implied by the current test bench. For the creation of the gold standard, the initial image input stream was temporally sub-sampled by retrospectively selecting solely the images that were acquired during the systolic cardiac phase. The selection procedure will be detailed in section 2.2.4. This operation will generate a stream of images for which apparent arterial pulsations are minimized, however, the image sampling frequency is reduced from 12 images/s (as in the light green columns) to ~ 1 image/s (as in the light blue column), which is the approximate frequency of the cardiac cycle. Displacements were then estimated on the resulting images using both the L^2-L^2 and the L^2-L^1 functional. Since arterial pulsations are minimal, we can assume that their influence on the motion estimates is negligible. Following the registration procedure, positional updates with a frequency of ~ 1 image/s are obtained. The resulting motion fields were then upsampled through linear interpolation back to the sampling frequency of the original image sequence. In this test, the resulting interpolated motion fields played the role of gold standard in the computation of the EE.

Subsequently, registration was carried out on the original set of images, in which the arterial pulsations are still present, using both the existing and the proposed motion estimation criterion. Finally, the EE was computed between the displacements estimated on the original set of images and the previously created gold standard.

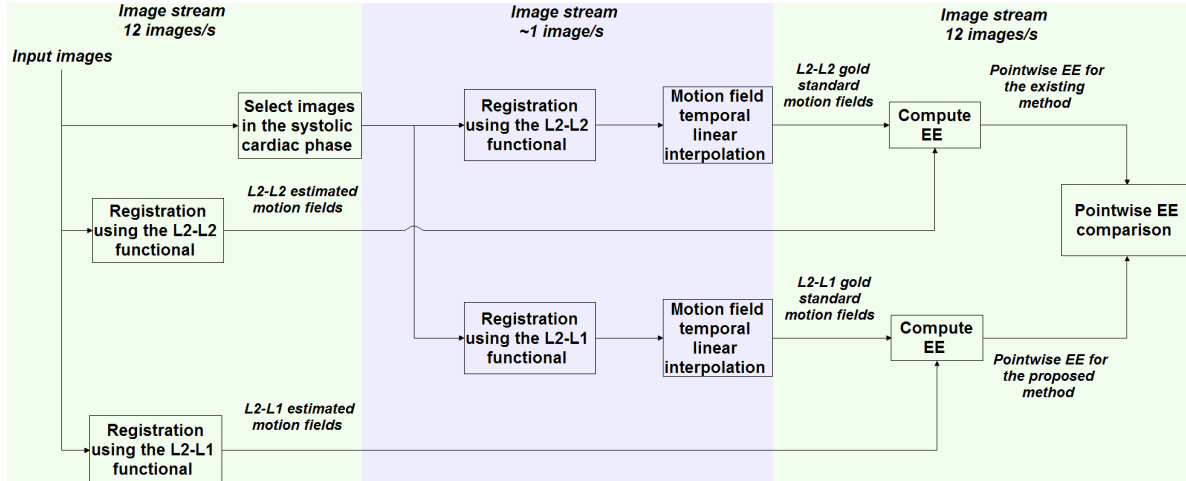


Figure 1: Diagram that summarizes the processes and their succession for test bench #1.

Reliability test #2 This test was designed as a “sub-set” of reliability test #1; i.e. it operated on a more refined dataset in order to emphasize the extent of the potential benefits brought by the proposed L^2-L^1 criterion that might not have been apparent during test #1. For this, in a first step, both apparent arterial pulsations and respiratory displacements were reduced through a succession of retrospective image selection procedures. Details on these procedures will be discussed in section 2.2.4.

Only images acquired in the systolic cardiac phase and at full exhalation were further considered for registration (see Fig. 2). Note that the initial selection with respect to the cardiac phase will reduce the sampling frequency of the original image stream from 12 images/s (as in the light green column) to ~ 1 image/s (as in the light blue column), while the subsequent selection according to the respiratory phase will further reduce it to ~ 0.2 images/s (as in the light red column), which is the approximate frequency of the respiratory cycle. The resulting images were then separately registered to two reference images, both acquired at full exhalation but at different stages of the cardiac cycle. Registration was performed using both the both the L^2-L^2 and the L^2-L^1 functional. Since the organs of interest are apparently immobilized, the gold standard for this test was chosen equal to zero. Thus, the EE coincides with the magnitude of the estimated motion vectors.

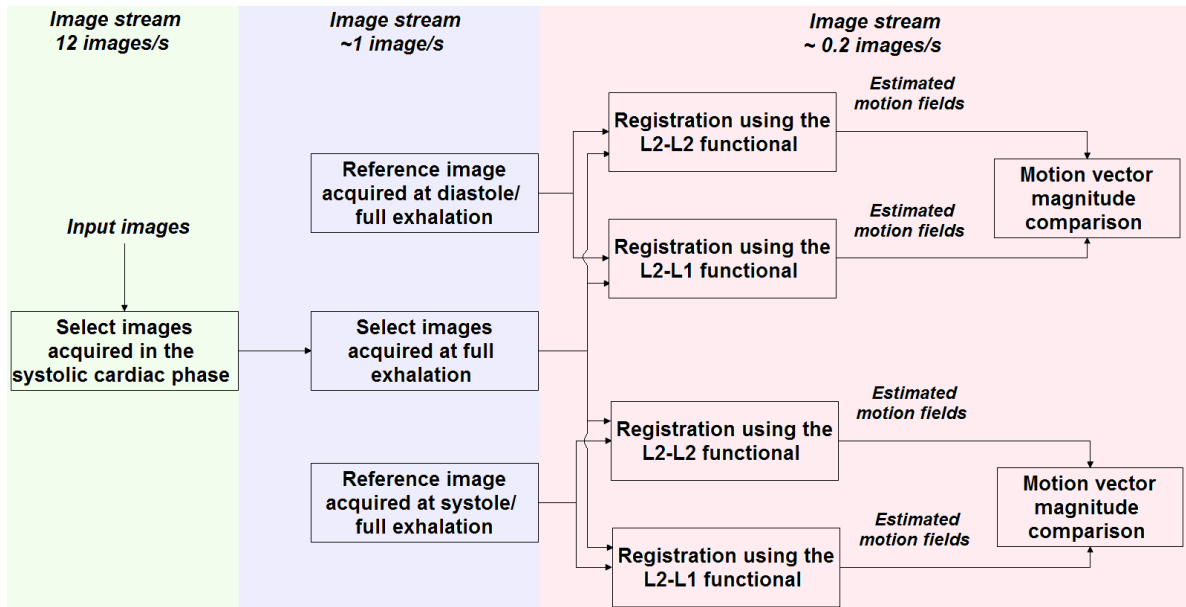


Figure 2: Diagram that summarizes the processes and their succession for test bench #2.

Reliability test #3 The purpose of this test was to establish the robustness to noise of the L^2-L^1 model compared to the L^2-L^2 . Fig. 3 illustrates a diagram summarizing this test. Precision of the estimated organ displacements was analyzed for both the proposed and the existing motion models under various SNR conditions as follows. In a first step, for each of the volunteers all the images acquired in the systolic cardiac phase and at full exhalation were retrospectively selected and stored in a buffer. An image with a high SNR was created by temporally averaging all images stored inside the buffer. Subsequently, increasing levels of Rician noise were added to the resulting noise-free image. Registration was then performed on pairs of images affected by the same level, but resulting from different realizations of the noise. The spatial average of

the EE was used as a quality quantification criterion for the motion estimates. The u_{gold} and v_{gold} for the computation of the EE were set to zero, since the registration operated on copies of the same image with the difference that they are affected by independent realizations of the same level of noise. The image sampling frequency in each processing block is indicated at the top of the colored columns in Fig. 3.

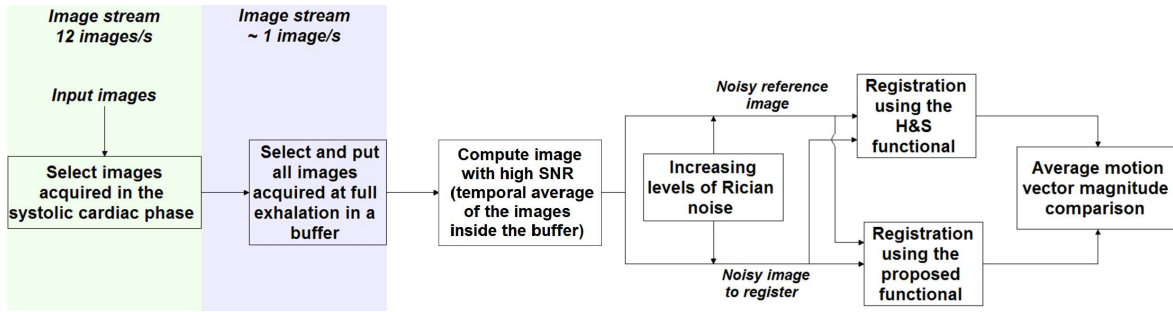


Figure 3: Diagram that summarizes the processes and their succession for test bench #3.

2.2.4. Retrospective image selection procedures All three test benches described in section 2.2.3 imply selective processing of images depending on the specific time at which they were acquired during the respiratory and/or respectively the cardiac cycle. For this, decision criteria have to be established for the retrospective selection of images acquired at a particular respiratory and/or cardiac phase.

Image selection with respect to the cardiac phase The slice position of the dynamic MR images acquired on the volunteers was set such that the renal aorta was clearly visible. A ROI encompassing the latter was defined and the averaged gray-level intensity inside the region was analyzed over time. The zeroes of the first order derivative of the obtained time-curve were used to identify the moments at which the signal inside the ROI was either minimal or maximal. This allowed deciding whether the images were acquired during the systolic or respectively diastolic cardiac phase.

Image selection with respect to the respiratory phase Initially, images acquired during the systolic cardiac phase were selected using the procedure described above. Registration was then performed on the resulting image set using the existing L^2 - L^2 motion estimation criterion. Since the previously performed image selection minimizes the intensity variations originating from arterial pulsations, their impact on the estimated motion is expected to be negligible. A ROI encompassing the liver was defined and the averaged displacements inside the region were analyzed over time. The zeroes of the first order derivative of the obtained time-curve provided the minima and the maxima of the average displacement, which were used to decide whether the images were acquired at the apex of expiration or respectively inspiration.

2.2.5. Calibration of the compared functionals Both the L^2 - L^2 and the L^2 - L^1 functionals require the fixation of a weighting parameter that links the data fidelity term and the regularization term (see Eq. (1) and (2)). During the registration process the image gray-level intensities were normalized such that they lie between 0 and 1. Since in this interval $\|x\|_1 \geq \|x\|_2^2$ (where x is an n -dimensional vector) it is expected that $\beta \geq \alpha$ for the functionals to provide similar motion estimates.

In order to reduce the impact of arterial pulsations during the calibration process, images were selected with respect to the cardiac phase in which they were acquired using the procedure detailed in section 2.2.4. Using solely the images acquired during the systolic phase, the weighting parameters were fixated as follows:

Calibration of α for the $E_{L^2L^2}$ functional Two landmarks, one in the liver and one in the kidney, were manually tracked over the resulting set of images. Subsequently, registration of the images was performed using the L^2 - L^2 motion model, while varying α between 0 and 1 with a fixed increment of 1/20. The estimated landmark displacements were compared in terms of EE to the ones resulting from the manual tracking. The value of α which provided the smallest time-averaged EE was established as optimal.

Calibration of β for the $E_{L^2L^1}$ functional After optimizing the value of α for the L^2 - L^2 functional, β was chosen such that the displacements estimated at the previously established landmarks by the L^2 - L^1 model are as close as possible to the ones estimated using L^2 - L^2 . For this purpose, registration was initially performed using the L^2 - L^2 functional with the established optimal value for the parameter α . Subsequently, the displacements were estimated using the L^2 - L^1 model while varying the value of β between 0 and 1 with an increment of 1/20. The value of β which minimized the temporally averaged EE between the displacements estimated using the L^2 - L^2 and respectively the L^2 - L^1 functional was established as optimal.

2.2.6. Estimation of the true motion induced by arterial pulsations It is expected that the motion estimated in the vicinity of arteries is also influenced by tissues being displaced by the periodic increasing and decreasing vessel diameter with the cardiac cycle. The extent of these displacements has been evaluated as follows. A set of 8 multi-slice MR images have been acquired on the abdomen of the same volunteers as in section 2.2.1 with the following protocol: repetition time (TR)=1714 ms, echo time (TE)=55 ms, flip angle=90°, field of view (FOV)=40×32×6 cm³, slice thickness=5 mm, matrix=268×268×10, with inflow suppression of blood. The images were acquired using respiratory and cardiac gating such that all images reflect the same respiratory phase, but at different stages of the cardiac cycle. The diameter of the two most visible hepatic arteries was then manually measured at the different stages of the cardiac cycle.

2.2.7. Hardware Implementation of the motion estimation algorithms was performed on an Intel 3.2 GHz i7 workstation (8 cores) with 16 GB of RAM. All computationally

intensive calculations were offloaded to a dedicated NVIDIA Tesla C2075 graphics processing unit (GPU) with 6 GB of DRAM.

2.2.8. Implementation A multi-threaded C++ implementation was performed for both the Horn&Schunck and the primal-dual algorithm, employed for the minimization of the L^2 - L^2 and respectively the L^2 - L^1 functional. A total number of 8 threads was found to provide optimal acceleration for the convergence of the numerical schemes. Above 8, the overhead introduced by thread synchronization would prevent further acceleration. An implementation of the algorithms on a GPU using the Compute Unified Device Architecture (CUDA) was also performed.

3. Results

Initially, in section 3.1, we analyze the performance of the image selection techniques described in section 2.2.4. Section 3.2 reports the values of the regularization parameters α and β resulting from the optimization procedures described in section 2.2.5. Both image selection and the calibration of the regularization parameters subsequently allowed analyzing the performance of the two compared functionals against the three proposed test benches (see section 2.2.3). The outcome of this analysis is reported in section 3.3. In section 3.4, we illustrate the measured tissue displacements due to the periodic contraction and dilatation of the arteries. Finally, section 3.5 reports the computational time of the numerical scheme implemented for the minimization of the proposed L^2 - L^1 functional.

3.1. Retrospective image selection with respect to the cardiac and respiratory phase

3.1.1. Image selection with respect to the cardiac phase Fig. 4a and 4b illustrate two dynamics acquired on volunteer #1 during the systolic and respectively diastolic cardiac phase. The image acquisition protocol is described in section 2.2.1. In order to do an automatic selection of the images according to the cardiac phase in which they were acquired, the average gray-level intensity inside a ROI (depicted by the red dotted line in Fig. 4) placed on the renal aorta was used. It can be seen that during systole (Fig. 4a), the average gray level inside the ROI is lower than during diastole (Fig. 4b). Fig. 4c displays for volunteer #1, the average gray level inside the ROI for the first 20 s of the study. The zeroes of the first order derivative of the displayed curve indicate the position of its minima (red triangles) and maxima (blue circles), which, in turn, identify the cardiac phase in which the images were acquired: systole or respectively diastole. The white and the blue ROIs depicted in Fig. 4a and 4b identify the liver and respectively the kidney in the two images.

3.1.2. Image selection with respect to the respiratory phase Initially, apparent arterial pulsations were minimized using the procedure described in section 2.2.4. The average

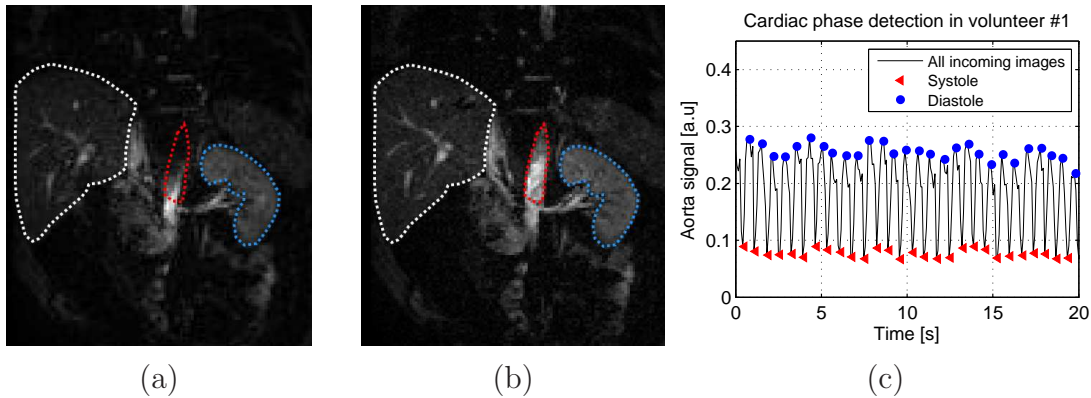


Figure 4: Employed strategy for retrospectively selecting images with respect to the cardiac phase in which they were acquired, exemplified for volunteer #1: (a): Image acquired during systole; (b): Image acquired during diastole; (c): Curve which displays the average gray-level intensity inside a ROI encompassing the renal aorta, delimited by the red dotted line in (a) and (b) during the first 20 s of the study. The red triangles and the blue circles indicate the locations of the minima and respectively the maxima of the curve, as identified by the zeroes of its first order derivative.

displacement inside a ROI encompassing the liver (depicted by the white dotted line in Fig. 5a) was then estimated over time using the L^2 - L^2 functional. The resulting curve is displayed in Fig. 5b. The zeroes of the first order derivative of this signal were used to find its minima and maxima, which, in turn, will identify the respiratory phase in which the image was acquired: full inhalation or full exhalation respectively.

3.2. Calibration of the input parameters

3.2.1. Calibration of the α parameter for the L^2 - L^2 functional The calibration of the parameter α which links the data fidelity term and the regularization term in the L^2 - L^2 functional was performed as described in section 2.2.5. The two white arrows in Fig. 6a indicate the manually tracked landmarks in volunteer #1. Fig. 6b displays as a function of α the time averaged EE between the manually tracked displacements and the ones estimated using the L^2 - L^2 criterion. In this case, the minimum is attained for $\alpha = 0.3$. This value is in good correspondence with previous reportings (Roujol et al. 2011, Denis de Senneville et al. 2015). Fig. 6c depicts the trajectory of the landmark inside the liver of volunteer #1 that resulted from the manual tracking (red dashed line) and respectively from the estimation using the L^2 - L^2 criterion for various values of α . It can be noted that for $\alpha = 0.3$ the best approximation of the manually tracked trajectory is attained. After optimizing the value of α for the second volunteer, the same value was found to generate the lowest EE. This is also reflected by the bar graph illustrated in Fig. 6d, where for both of the volunteers, the values of the time-averaged EE were plotted for several values of α .

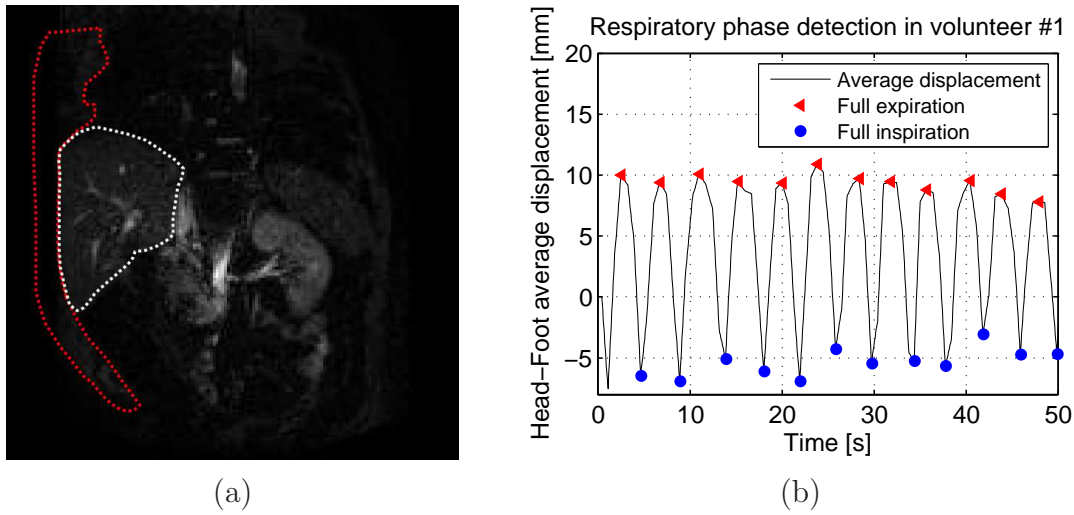


Figure 5: Employed strategy for selecting images with respect to the respiratory phase in which they were acquired, exemplified for volunteer #1. (a): Image used as reference for estimating the average displacement of the liver for images acquired in the systolic cardiac phase. (b): Curve displaying the evolution of the average displacement of the pixels inside a ROI encompassing the liver depicted by the white dotted line in figure (a). The minima and the maxima of the curve indicate the respiratory phase in which the images were acquired. The red ROI identifies the location of the abdominal wall.

3.2.2. Calibration of the β parameter for the L^2 - L^1 functional The value of β in the L^2 - L^1 functional was optimized using the procedure described in section 2.2.5. The optimization process was performed in terms of the EE between the displacements estimated at the previously established anatomical landmarks using the L^2 - L^2 and respectively the L^2 - L^1 criteria. Registration operated on images for which apparent cardiac activity has been minimized. Fig.7a illustrates the temporally averaged EE between the displacements estimated by the two criteria as a function of β at the landmark inside the liver of volunteer #1. For $\beta = 0.6$, the L^2 - L^1 functional estimated respiratory induced displacements that are most similar to the ones estimated by the L^2 - L^2 criterion. Fig.7b depicts under the shape of a bar graph the time-averaged EE between the displacements estimated by the L^2 - L^1 functional and the manually tracked trajectory in the liver of both volunteers for several values of β . Even though the relative differences in terms of EE as a function of β are not as evident as during the calibration of α , a value of $\beta = 0.6$ was found to be optimal for both volunteers. Fig.8a depicts the spatial distribution of the temporally averaged EE between the respiratory induced displacements estimated in the liver and kidney by the two compared criteria. During the registration process, parameters α and β were set at the previously found optimal values. The boxplots displayed in Fig.8b reflect the statistical distribution of the previously computed EE. It is observed that for both the liver and the kidney, errors remain under 0.8 mm for both of the volunteers. This indicates that a proper

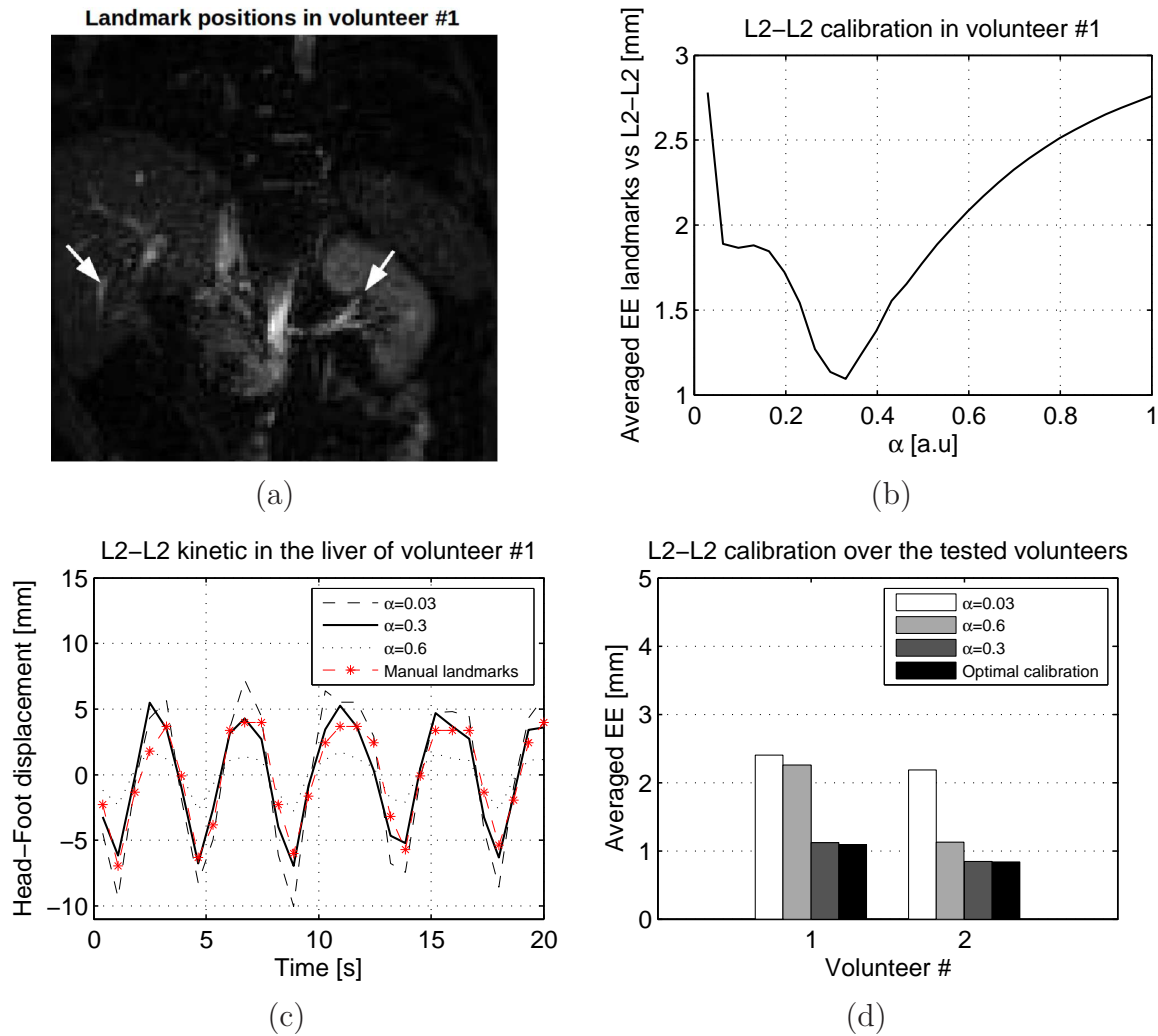


Figure 6: Calibration of the α parameter for the L^2-L^2 functional. (a): The reference image upon which the landmarks are indicated by white arrows; (b): Time averaged EE between the displacements resulting from manually tracking the landmarks and the displacements estimated with the L^2-L^2 functional as a function of α ; (c): Trajectory of the tracked landmarks. The red dashed line corresponds to the manually tracked trajectory while the black lines are the trajectories estimated by the L^2-L^2 criterion for various values of α ; (d): Time averaged EE for both the volunteers for several values of α . The results displayed in (a), (b) and (c) were obtained for volunteer #1.

calibration of the value of β was attained. Additionally, the relatively small overall EE is an indicator that when arterial pulsations are minimized, there is a good match between the motion estimates returned by the two compared functionals, when α and β are properly calibrated. Each boxplot was constructed in the following fashion: the lower limit of the box corresponds to the first quartile, the upper limit of the box is the third quartile, the median is represented by the red horizontal line in the middle of the box, the lower whisker corresponds to the 5th percentile and the upper whisker to the

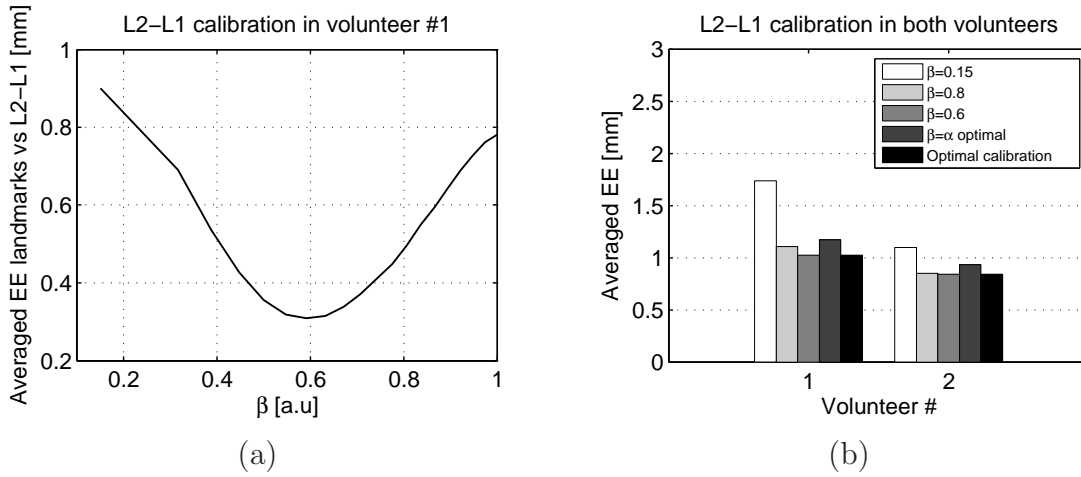


Figure 7: Calibration of the β parameter in the L^2-L^1 functional. (a): The temporally averaged EE between the motion estimates obtained using the L^2-L^2 and respectively the L^2-L^1 criteria at the landmark in the liver of volunteer #1 (see Fig. 6a) as a function of β . (b): Bar graph illustrating the EE between the L^2-L^1 estimates and the manually tracked trajectory for a landmark in the liver of both volunteers for various values of β . Estimation has been conducted only on images acquired during the systolic cardiac phase in order to minimize the influence of arterial pulsations.

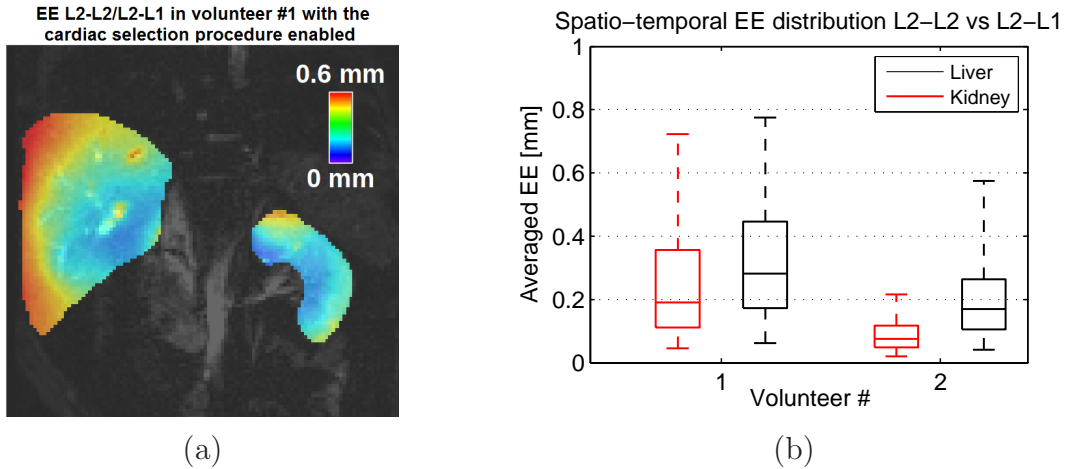


Figure 8: (a): Spatial distribution of the temporally averaged EE, obtained for volunteer #1, between the respiratory induced displacements estimated with the L^2-L^2 and respectively the L^2-L^1 criteria in the liver and kidney for the optimal values of α and β . (b): The EE from displayed in (a) put under the shape of a boxplot for both volunteers. The relatively small values of the EE is an indicator that the value of β was properly calibrated.

95th percentile.

3.3. Performance comparison between the L^2 - L^2 and the L^2 - L^1 criteria in the presence of arterial pulsations

3.3.1. Reliability test #1 The quality of the estimated motion in the presence of arterial pulsations was initially quantified using the procedure described in section 2.2.3. Fig.9a and 9b depict, for volunteer #1, the time averaged EE between the displacements estimated by the two motion estimation criteria and their corresponding gold standard. The construction of the gold standard for each criterion is detailed in section 2.2.3. It is observed that errors tend to be overall larger when estimation is performed with the existing L^2 - L^2 functional. In particular, error peaks occur in the proximity of the hepatic arteries (indicated by the white arrow in Fig.9a) and also in the superior part of the kidney closest to the renal aorta. For the L^2 - L^1 criterion the estimation errors do not manifest any clear tendency in the liver. In the kidney, however, errors tend to increase in the area closest to the renal aorta. Fig.9c displays the absolute difference between the EE illustrated in Fig.9a and 9b. This difference further emphasizes the spatial locations in which the two functionals provide motion estimates of different quality in the particular case of reliability test #1. Fig.9d and 9e illustrate, under the shape of boxplots, the previously computed EE in the liver and kidney of both volunteers. While there is no notable difference in the kidney between the two criteria, there is an obvious improvement in the liver when the estimation is performed using the L^2 - L^1 functional. The maximal EE was reduced in the latter from ~ 2.5 to ~ 1.5 mm, while the median of the EE decreased from ~ 2.1 mm to ~ 1.3 mm. Fig.9f and 9g display the trajectory of the landmark placed in the liver of volunteer #1 (indicated by a white arrow in figure 6a). The red dashed line depicts the trajectory of the landmark given by the gold standard of the two criteria. The black lines correspond to the trajectory of the landmark estimated by the L^2 - L^2 and respectively the L^2 - L^1 functional. For both criteria there are fluctuations of the landmark position during the positive peak of the curves. Since the landmark was chosen close to one of the hepatic arteries we assume that the source of the fluctuations in the estimated displacement is the local intensity variations due to the pulsations. It can be observed, however, that the amplitude of the fluctuations is much smaller for the L^2 - L^1 criterion than for the L^2 - L^2 .

3.3.2. Reliability test #2 This test was designed to evaluate the behavior of the existing and the proposed motion estimation models when registering images for which apparent respiratory motion is minimized, with only arterial pulsations being present. Details concerning this test bench were discussed in section 2.2.3. All images acquired in the systolic cardiac phase at full expiration were separately registered, using both the motion estimation criteria, to two reference images: one acquired at systole and full expiration, the other at diastole and full expiration. Table 1 reports the mean and the standard deviation of the temporally averaged EE between the displacements estimated by the two functionals and their corresponding gold standard (which for this test was established to be zero - see section 2.2.3). The results were obtained by analyzing the data from

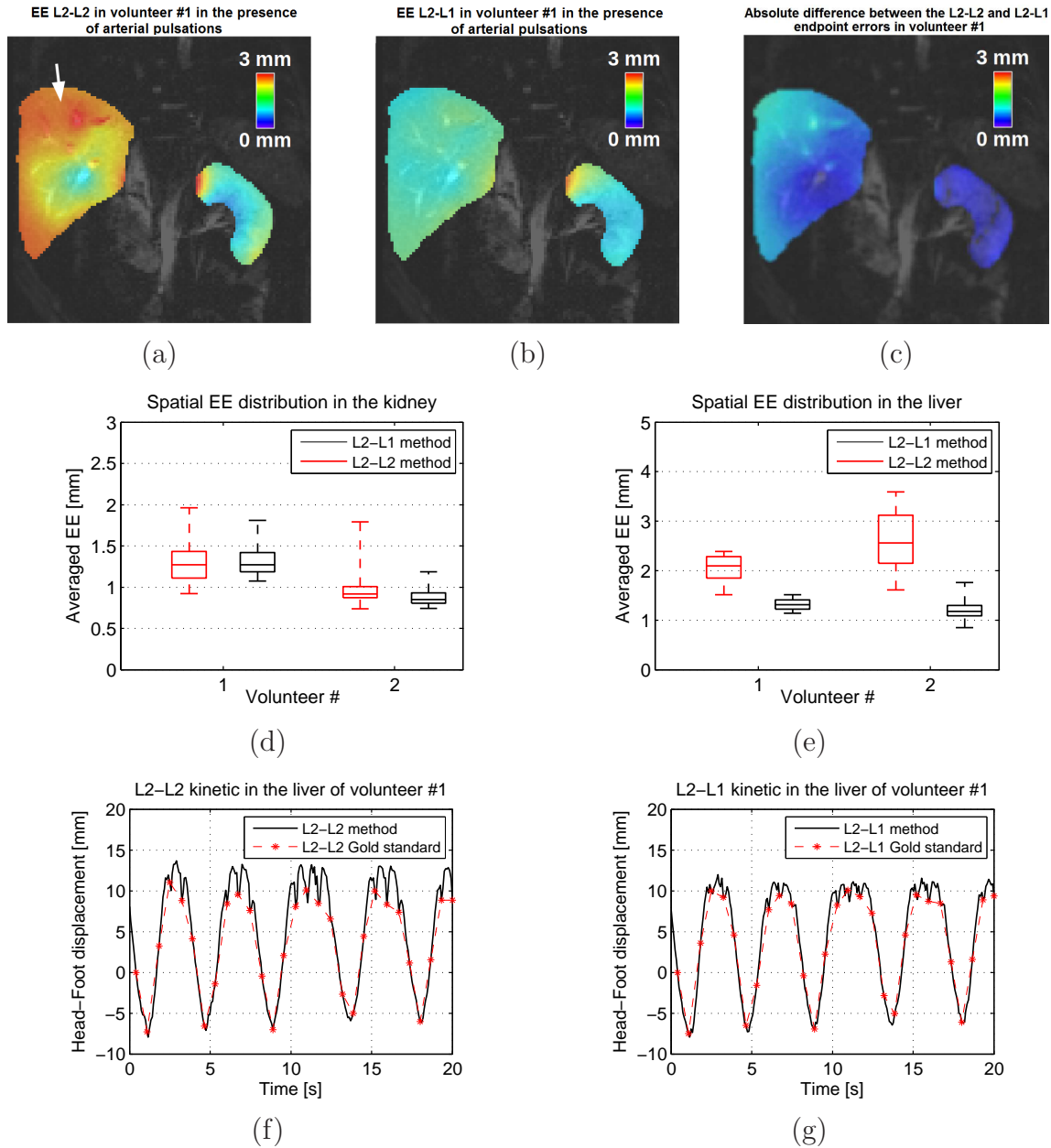


Figure 9: Performances of the L^2-L^2 and the L^2-L^1 criteria obtained during reliability test #1. (a), (b): Temporally averaged EE, for volunteer #1, between the displacements estimated with the L^2-L^2 and respectively L^2-L^1 functional and their corresponding gold standards; (c): Absolute difference between the EE displayed in (a) and (b); (d), (e): The EE in (a) and (b) put under the shape of a boxplot separately for the liver and the kidney of both volunteers; (f), (g): Trajectory of a landmark placed close to an artery in the liver of volunteer #1 estimated by the L^2-L^2 and respectively the L^2-L^1 criterion. The red dashed line is the trajectory given by the corresponding gold standard. Note that the fluctuations in the estimated landmark position are lower when the proposed L^2-L^1 functional is used.

both volunteers, but separately for the liver and the kidney. It can be observed that for either of the organs there is no remarkable difference in EE when the reference image is acquired at systole - full exhalation, regardless of the functional which is used for estimation. However, when the reference image is in the diastole - full exhalation phase, both the average and the standard deviation of the EE were reduced with more than 30% by using the proposed L^2-L^1 model instead of the existing L^2-L^2 , for both the liver and the kidney. Note that while no evident improvement could be observed for the latter in test #1, the current test, which is based on a more restricted subset of images, emphasizes the benefit of using the L^2-L^1 functional for estimating kidney motion.

Cardiac phase in which the reference image was acquired	EE in [mm] for the kidney		EE in [mm] for the liver	
	L^2-L^2	L^2-L^1	L^2-L^2	L^2-L^1
Systole	0.6 ± 0.5	0.6 ± 0.5	0.7 ± 0.7	0.7 ± 0.6
Diastole	1.2 ± 1.2	0.9 ± 0.7	1.8 ± 1.2	1.0 ± 0.8

Table 1: Mean and standard deviation of the EE obtained during the reliability test #2 by the L^2-L^2 and respectively the L^2-L^1 functionals. The motion estimation procedure operated on images acquired at systole - full exhalation that were independently registered to two reference images acquired at full exhalation, but at the two extreme cardiac phases. Analysis was performed on both volunteers, but separately for the liver and kidney.

3.3.3. Reliability test #3 This test had the purpose to evaluate the robustness to noise of the existing and the proposed motion estimation criteria. The evaluation procedure is detailed in section 2.2.3. Fig.10a illustrates an image acquired on volunteer #1 during the systole - full exhalation phase both in its noisy and temporally filtered form. Filtering was performed as described in section 2.2.3 and resulted in an SNR increase from 12.8 in the liver and 20 in the kidney to 20 in the liver and 29.5 in the kidney. Two copies of the filtered image were registered to one another using both motion estimation criteria, while gradually decreasing their SNR by applying the same level but independent realizations of Rician noise. Fig.10b depicts the spatially averaged EE between the resulting estimates and their corresponding gold standard (see section 2.2.3) as a function of the SNR of the images. The analysis has been performed separately for the liver and kidney. Results show that a moderate penalty of 0.1 mm arises for $\text{SNR} > 5$ from the use of the L^2-L^1 metric instead of the L^2-L^2 .

3.4. Estimation of the true motion due to arterial pulsations

Figure 11 displays for volunteer #2 a slice from three of the images acquired using the protocol described in section 2.2.6, with focus on the liver. The images illustrated in

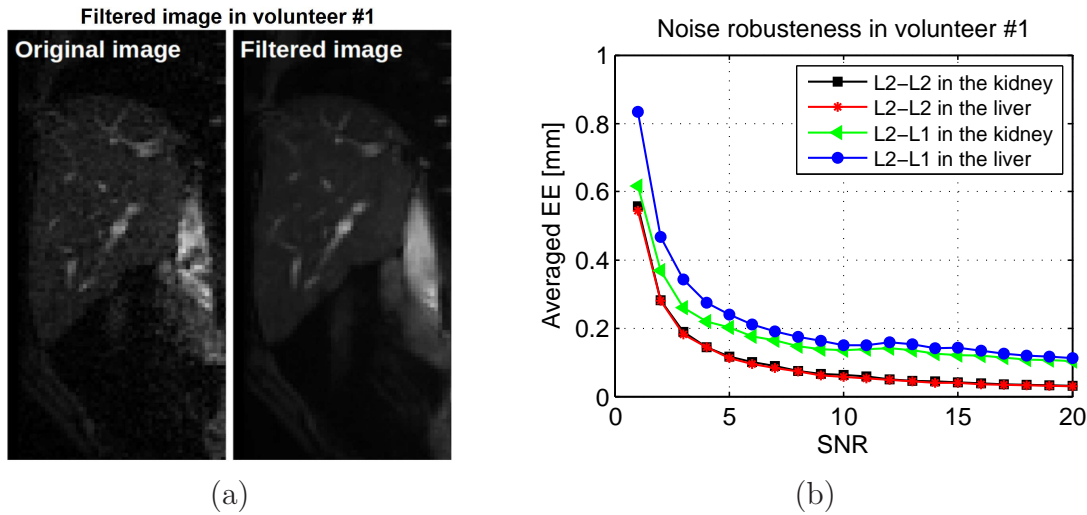


Figure 10: Robustness to noise of the proposed and existing motion estimation criteria. (a): Noisy ($\text{SNR}_{\text{liver}}=12.8$, $\text{SNR}_{\text{kidney}}=20$) and filtered version ($\text{SNR}_{\text{liver}}=20$, $\text{SNR}_{\text{kidney}}=29.5$) of the image used as reference in reliability test #3 for volunteer #1. Only the liver is displayed in the current figure. (b): Spatially averaged EE between the displacements estimated by each of the criteria and their corresponding gold standard as a function of the SNR of the registered images. The analysis has been conducted separately for the liver and the kidney.

Fig. 11a, 11b and 11c have been acquired at mid-diastole, 600 ms after the mid-diastole and respectively 1000 ms after the mid-diastole. The arteries indicated by the white arrows had their radius manually measured in cardiac phases distributed over the entire cardiac cycle. The results were as follows. The radius of artery #1 measured 1.78 ± 0.07 mm (range 1.7 - 1.9 mm), while artery #2 measured 2.32 ± 0.07 mm (range 2.2 - 2.4 mm).

3.5. Computational time of the primal-dual algorithm

The computational time of the primal-dual algorithm used for the minimization of the proposed L^2-L^1 functional was measured for a total of 3000 registrations, i.e. for all the images acquired on the two volunteers. Our C++ implementation converged in 25 ± 1.4 ms per image, with computational times ranging from ~ 14 ms to ~ 29 ms. The GPU implementation did not provide further acceleration.

4. Discussion

In recent years, L_1 -based variational methods have become increasingly popular for addressing the issue of motion estimation in the video processing domain (Wedel et al. 2009, Rakêt et al. 2011, d'Angelo et al. 2011). However, the benefit of such approaches in medical imaging has yet to be shown. To our knowledge, this study is

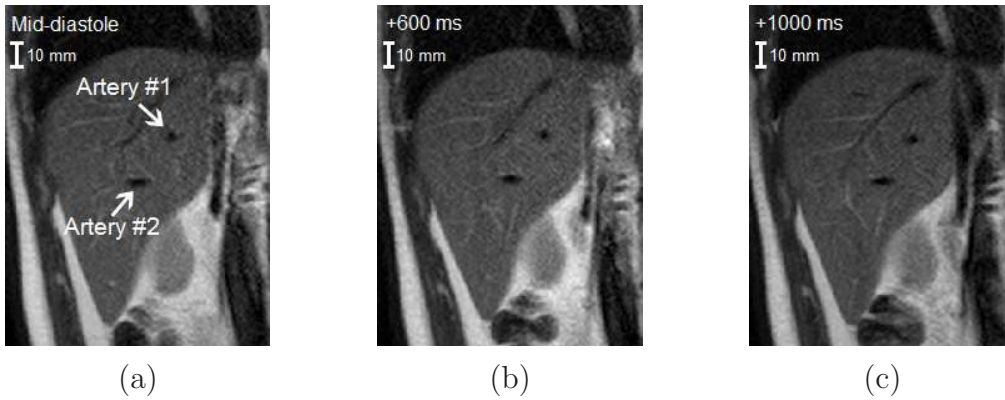


Figure 11: Slice from three of the images used to measure the artery radius in volunteer #2 over the cardiac cycle, with focus on the liver. (a): Image acquired at mid-diastole; (b): Image acquired 600 ms after the mid-diastole; (c): Image acquired 1000 ms after mid-diastole. Measurements have been performed for the arteries indicated by the white arrows.

the first to analyze the behavior of these techniques in an applicative context: real-time **MR-guidance of beam** therapies in mobile organs.

In the current study we proposed a motion estimation method that allows real-time tracking of mobile organs which, compared to the existing approaches, has an increased robustness to local gray-level intensity variations not attributed to motion. In order to assess the performance of the proposed approach with respect to previous methods, we have additionally developed a series of tests benches operating on *in-vivo* data sets.

Previous studies have shown that the variational approach for motion estimation proposed by Horn&Schunck has good potential for therapy guidance during MRg-HIFU interventions in moving organs (Ries et al. 2010, Roujol et al. 2010, Roujol et al. 2011). We hypothesize that the success of the method is due to the good correspondence between the motion model assumed by the H&S functional and the underlying physical reality. First, the data fidelity term of the functional (see Eq. 1) was built on the assumption that the moving image has the same, possibly displaced, content as the reference image. Since the amplitude of the displacements exhibited by the abdominal organs is limited, as long as the FOV covered by the images is properly chosen, this assumption is, in general, satisfied. Second, the regularization term of the functional constrains the estimated motion fields to be spatially continuous. This again is a justified assumption since abdominal organs are incompressible and elastic, thus shearing effects will, in general, not occur. For these reasons, the H&S model was used as a starting point for the construction of the motion estimation functional proposed by the current study.

There are however scenarios in which the H&S model becomes less reliable. For example, real-time thermometry during MRg-HIFU interventions is usually based on T_2^* -weighted images. For such a contrast, the blood might appear more intense than the tracked

organs (see, for example, Fig. 4a). Since the H&S functional relies on structural information (which provides non-zero spatio-temporal gradients in the data fidelity term) in order to solve the motion estimation problem, blood vessels will have a higher influence on the outcome of the registration process. Given the fast imaging required for the real-time thermometry, there is a high chance that the reference and the moving image are acquired at different phases of the cardiac cycle. This leads to arteries having different diameters, and possibly gray-level intensities, in the two images, which violates the pixel intensity conservation hypothesis used for the construction of the data fidelity term. Since the blood vessels have a high influence on the estimated displacement fields, this will lead to erroneous results. In this study, we address this issue by employing a version of the H&S functional for which the quadratic data fidelity term is replaced with an L_1 -based term, which is linear. In this manner, the motion estimates become more robust to gray-level intensity variations not attributed to motion.

In their excursion due to respiration, the liver and kidneys slide on the thoracic and respectively abdominal walls which, by comparison, manifest a reduced amount of motion. The arising shearing effects result in two perturbations disturbing the L^2 - L^2 functional: one on the data fidelity term and one on the regularization term. The effect on the data fidelity term of local gray-level intensity variations not attributed to motion was reduced using the proposed L^2 - L^1 functional. We anticipate that the systematic moderate differences between L^2 - L^2 and L^2 - L^1 estimates in the vicinity of the thoracic and abdominal walls (see, for example, the red ROI in Fig.5a) may be in favor of the L^2 - L^1 criterion. However, since the walls are adjacent to the liver and kidneys, the specific impact on the motion regularization term, for its part, may not be negligible, and will need to be addressed in future studies. Due to its quadratic nature, the regularization term will cause the estimated motion fields to be continuous. For this reason, the transition from areas within the liver and kidneys to the thoracic and abdominal walls will be smooth. This implies that false motion will be estimated for the latter due to the propagation of the motion vectors from inside the organs. Inversely, the minimally moving walls will cause a decrease in the magnitude of the motion vectors in adjacent areas from the liver and kidneys. Shearing motion is usually addressed for video sequences by replacing the quadratic regularization term with a linear term (e.g. an L^1 norm). Such techniques are however not straightforward in a clinical context. In particular, a penalty may arise within regions depicting low apparent anatomical structures. In such areas, the low spatio-temporal gradients will cause the data fidelity term to become close to zero. The quadratic regularization would normally propagate the displacements estimated in regions with higher gradients towards these areas. By replacing the quadratic regularization with a linear term the extent to which the motion is propagated will be reduced and thus, in regions with low apparent anatomical structures, the motion estimation problem might become ill-posed. Specific experimental benches will thus be mandatory to assess the benefit of such approaches *in-vivo*. It is, however, important to note that, as long as the assumptions made by the H&S motion model hold, the proposed and the existing motion estimation

methods provide similar motion estimates (as shown in Fig. 8). Thus, in such a case, it becomes irrelevant which of the methods is chosen to address the motion estimation problem.

A challenging task when estimating displacements for *in-vivo* targets during non-invasive beam interventions is the acquisition of a reliable gold standard. This was of particular interest for reliability test #1 (see section 2.2.3). At a particular point in the processing chain dedicated to obtaining the gold standard displacement fields, a temporal upsampling procedure was required. This was achieved by the means of a linear interpolation procedure which is, in general, prone to errors. We have evaluated the extent of these errors on an analytic breathing motion pattern (details are found in Appendix B). A theoretical maximal error of 0.25 ± 0.2 mm has been estimated for a peak-to-peak motion amplitude of 20 mm, a TR of 80 ms, and the typical breathing and cardiac frequency of 0.2 Hz and 1 Hz, respectively (Seppenwoolde et al. 2002). Given the amplitude of liver and kidney excursions during normal respiration of 11 ± 3 mm and respectively 11 ± 4 mm (Langen & Jones 2001) such errors were deemed to be negligible. However, if the images are acquired with a lower temporal resolution, the interpolation errors are expected to increase. It is worth noting that the resulting interpolation errors are, to a certain degree, sensitive to the simulated breathing motion pattern (see Eq. B.1) and to the injected model parameters such as the breathing and the cardiac frequency. Nevertheless, the true target motion together with the breathing and cardiac frequency are not expected to vary from the simulated values to such an extent that the interpolation errors exceed sub-millimeter values.

In the current study we have proposed three test benches that aid in comparing the performances of the proposed L^2 - L^1 and respectively the existing L^2 - L^2 functionals in the presence of local gray-level intensity variations not attributed to motion. In two of the tests, the proposed criterion proved to be more robust to such effects than the existing one. In reliability test #1 respiratory motion was estimated over a series of images containing gray-level intensity variations due to arterial pulsations. By using the L^2 - L^1 functional instead of L^2 - L^2 , while no notable improvement was observed in the kidney, the temporally averaged EE was reduced for almost all pixels inside the liver. Reliability test #2 registers images acquired at the same cardiac and respiratory phase to reference images acquired at the same respiratory phase but at the two extreme phases of the cardiac cycle. This process generates for each image two motion fields for which the magnitude of the contained motion vectors should be close to zero, since we have seen in section 3.4 that the true motion due to arterial pulsations is minimal. For this test, the proposed motion estimation criterion decreased the mean and the standard deviation of estimated motion compared to the existing criterion in both the liver and the kidney when the moving and the reference image were acquired at different phases of the cardiac cycle. Reliability test #3 was designed to test the robustness to noise of the two compared functionals. In this scenario, there are two contributions that impact the gray level intensity variations in the data fidelity term of $E_{L^2L^2}$ and $E_{L^2L^1}$: Spatial variations along anatomical structures present in the FOV, as well as spatio-temporal

variations induced by the simulated noise. For an SNR larger than 1, the contribution of anatomical structures exceeds the noise contribution and the quadratic impact of the data fidelity term is an advantage in this case. Nevertheless, since the SNR of the images used for registration is 12.8 in the liver and 20 in the kidney, thus on the right side of the curve depicted by Fig. 10, the reduction of the impact of pulsating arteries using the proposed L^2 - L^1 criterion is by far more important than the moderate penalty arising from the reduced robustness to noise. Also note from the graph in Fig. 10b that, as the SNR drops below 5, the quality of the motion estimates degrades exponentially. This result is in good correspondence with previous reportings (Zachiu et al. 2015). Thus, beyond this threshold, the estimated displacements become unreliable. However, this is not the case for the images used in the current study, since the SNR in the regions of interest lie well above this threshold.

The original Horn & Schunck functional requires a single parameter α as input, the value of which controls the smoothness of the estimated motion fields. This makes it a particularly attractive solution for motion estimation during **EBRT and HIFU** interventions, since it implies a minimal amount of intervention from the clinician delivering the therapy. Additionally, once the value of α has been optimized for a particular MR contrast weighting, the same value can practically be used for any number of patients as long as the parameters of the sequence remain unchanged. This is also beneficial for the real-time requirements of **the aforementioned** therapies since α does not have to be re-optimized over the duration of the intervention, thus reducing the overall time delay introduced by the registration process. The same is true for the proposed L^2 - L^1 criterion for which parameter β plays the same role as α in the L^2 - L^2 functional.

Real-time MR-guided **beam** interventions usually exploit the MR-image information for a direct retroactive control of the energy delivery device, thus it is necessary to distinguish two key concepts: temporal resolution (i.e. the time interval between two updates/observations) and information latency (i.e. the delay between the actual time of displacement and the availability of motion information). This latency is composed by the sum of the remaining acquisition time after echo-formation, the required data transport time and the image processing time. The data presented in the scope of this paper has been acquired with 12 images/s and then post-processed off-line. This was necessary since we had to compare the new proposed L^2 - L^1 approach with established gold standards (and the L^2 - L^2 approach) and to sort images into artefacted/non-artefacted categories. In a real image guided therapy scenario, however, the algorithm would run in real-time with processing on-the-fly. The implementation of the L^2 - L^1 algorithm potentially allows average frame rates of ~ 40 images/s (in practice this is usually limited by the capabilities of the MR scanner) with a typical end-to-end processing latency (from the beginning of the MR-acquisition of the slice to the final output of the motion-vector fields) of under 100 ms. Both are well within the requirements for real-time guidance (see for example (Ries et al. 2010)).

Both the L^2 - L^2 and the L^2 - L^1 optical flow algorithms allow the voxel-wise estimation

of displacement vectors with sub-voxel precision. While the $2.5 \times 2.5 \times 7\text{mm}^3$ voxel size might appear rather large compared to the typical $1 \times 1 \times 7\text{mm}^3$ HIFU focus size, previous studies have shown that therapy guidance can still be performed with acceptable errors even at such image resolutions (Ries et al. 2010, Roujol et al. 2010).

A limiting factor of the current study is the lack of validation of the proposed method under realistic **beam therapy** scenarios. However, we have shown that, as long as the assumptions made by the H&S algorithm are not violated, the proposed L^2 - L^1 motion estimation method has a similar performance to the existing ones based on the L^2 - L^2 functional, both in the quality of motion estimates (see Fig. 8) and computational time. Moreover, it was proven that, in the eventuality that the pulsating arteries become the dominant structure in the images used for therapy guidance, our method provides motion estimates of superior quality, while maintaining the short latencies required by the real-time nature of the discussed interventions. Thus, given that the L^2 - L^2 method was already validated under clinical conditions in previous studies for both MRg-HIFU (Ries et al. 2010, Roujol et al. 2010) and MRg-RT (Stemkens et al. 2015) interventions, the proposed L^2 - L^1 method can be easily integrated in the work-flow of such therapies and provide reliable therapy guidance capabilities with increased robustness to arterial pulsations.

The presented work proposes a respiratory motion correction strategy with the effects of cardiac motion/pulsations as a source of artifacts. However, there are additional types of physiological motion might occur during the lengthy **EBRT and HIFU** interventions, such as spontaneous motion (e.g. coughing or twitching of the patient) and long term drifts (e.g. due to peristaltic activity or slow changes in the respiration baseline). A frame-work for slow 3D organ motion was presented in a previous study, see (Zachiu et al. 2015), which has been validated in-vivo for MRg-HIFU interventions. Additionally, an adaptation of the framework to the work-flow of an MRg-RT intervention is also discussed. Spontaneous motion is a difficult problem due to its infrequent nature and very short timescale. One way to deal with spontaneous motion would be to detect it as fast as possible and to cut the power of the interventional device rapidly and to subsequently to reassess the situation. The respiratory motion correction presented here is fully compatible with the previously published approaches for slow 3D motion and potentially even spontaneous motion.

The proposed motion estimation framework is limited in the fact that it takes into consideration exclusively in-plane displacements. While the trajectory of the kidney and the lower part of the liver could be approximated in a first order through a linear shift, the trajectory of the upper liver is a curve in a 3D space, making it difficult to contain in a 2D slice with static geometry. Dynamic 3D imaging would be ideal, however, on current systems, the acquisition of 3D images coupled with a motion estimation procedure becomes challenging if the temporal resolution required by real-time **beam interventions** is to be maintained. One solution is to acquire additional information in the third dimension, such as navigator echoes, and dynamically adapt the slice position/geometry with respect to the through-plane displacements provided

by the navigators (Ries et al. 2010, Feinberg et al. 2010, Köhler et al. 2011). A different full-3D approach, proposed in (Glitzner et al. 2015), consists in dynamically acquiring low-resolution 3D volumes of the target area and its surroundings and register the images, by the means of the H&S algorithm, to a reference volume acquired at the beginning of the intervention. It was shown that such an approach might accelerate the overall image-based therapy guidance process, while maintaining the quality of the motion estimates within acceptable limits. While the method was proven to have only near-real-time capabilities, it shows good potential for future investigations towards real-time guidance of **EBRT and HIFU** interventions. Additionally, 3D trajectories may be estimated from 2D MRI using one or several volumetric scans obtained before the intervention, as shown in (Brix et al. 2014, Arnold et al. 2011) and (Stemkens et al. 2015). Note, however, that some of these approaches have not yet underwent pre-clinical or clinical validation.

The 2D real-time motion estimation strategy proposed in this study is not limited to improving the robustness of the motion estimates specifically to arterial pulsations. It also offers good perspectives for improving, in both 2D and 3D, the robustness of the estimates to local gray-level intensity variations originating from other sources, such as: peristaltic contractions, incomplete blood suppression for interventions in the heart or local contrast variations due to the thermal build-up during HIFU ablations.

5. Conclusion

The presented study addresses target tracking for MR-guided **EBRT and HIFU** interventions in the abdomen through an improved real-time optical flow-based motion estimation strategy. The proposed approach based on the L^2 - L^1 optical flow functional was shown to be more resilient to local gray level intensity variations not attributed to motion compared to previous approaches which employed the original H&S functional. Our method was shown to be potentially beneficial for interventions in the liver and kidneys, where local gray level intensity variations introduced by pulsating arteries were shown to have a lesser impact on the estimated motion compared to the existing approaches. Additionally, the sub-second temporal resolution of the proposed method renders it compatible with high duty cycle MR-guided **beam** interventions for treating tumors in the abdomen under free-breathing conditions.

Acknowledgments

This work was supported by the Dutch Technology Foundation (STW) (project OnTrack #12813) and in part by the European Research Council (project ERC-2010-AdG-20100317, Sound Pharma) and ITEA 2 (project SoRTS). The authors would also like to thank Charles Dossal (Institut de Mathématiques de Bordeaux (IMB), Université de Bordeaux, France) for his many helpful suggestions. The authors gratefully acknowledge the two anonymous referees whose remarks and suggestions have helped us greatly in

improving the quality of the manuscript.

Appendix A. The primal-dual algorithm

Appendix A.1. Theoretical background

Let $F : X \rightarrow [0; +\infty[$ and $G : X \rightarrow [0; +\infty[$ be two proper, convex and lower semi-continuous functions and K a linear operator with $\|K\| = L$. We seek the solution of the problem:

$$\min_{x \in X} F(Kx) + G(x) \quad (\text{A.1})$$

Solving the problem described by Eq. (A.1) is equivalent to finding the solution of the associated dual problem (Chambolle & Pock 2011):

$$\max_{y \in Y} -(-G^*(-K^*y) + F^*(y)) \quad (\text{A.2})$$

where Y is the dual space of X , $F^* : Y \rightarrow [0; +\infty[$ and $G^* : Y \rightarrow [0; +\infty[$ are the convex conjugates of F and G respectively and K^* is the adjoint operator of K .

Based on Eq. (A.1) and Eq. (A.2) the primal-dual problem can be defined:

$$\min_{x \in X} \max_{y \in Y} \langle Kx, y \rangle + G(x) - F^*(y) \quad (\text{A.3})$$

the solving of which returns the solution of both the primal and its associated dual problem. The application $\langle Kx, y \rangle$ is the inner product between Kx and y .

The problem in Eq. (A.3) admits a unique solution given by the fixed point of the following iterative scheme:

$$\begin{cases} y_{n+1} = \text{prox}_{\sigma F^*}(y_n + \sigma Kx_n) \\ x_{n+1} = \text{prox}_{\tau G}(x_n + \tau K^*y_{n+1}) \end{cases} \quad (\text{A.4})$$

for any σ and τ such that $\sigma\tau L^2 < 1$ and with $x_0 = 0$ and $y_0 = 0$. $\text{prox}_H(z)$ is the proximal operator of function H in point z . Additional technical details regarding the primal-dual algorithm applied for optical flow estimation can be found in (Chambolle & Pock 2011).

Appendix A.2. Minimization of the L^2 - L^1 functional via the primal-dual algorithm

The L^2 - L^1 functional was minimized via the primal-dual algorithm, detailed in Appendix A.1, with:

$$\begin{cases} F(\nabla w) = \beta^2 \|\nabla w\|_2^2 \\ G(w) = \|(\nabla I)^T \cdot w + I_t\|_1 \end{cases} \quad (\text{A.5})$$

where $\nabla I = (I_x, I_y)^T$ is the 2D spatial gradient of the image being registered, I_t is the difference between the reference image and the image to be registered (i.e. I_t is

the temporal gradient), $w = (u, v)^T$ is the 2D displacement vector, F is the L_2 norm multiplied by β^2 and applied on the gradient of w and G is the L_1 norm applied on the data fidelity term. The numerical scheme of the primal-dual algorithm used for the minimization of $E_{L_2L_1}$ consists in the following:

- (i) Initialize: $w_0^* = 0$ and $w_0 = 0$.
- (ii) For $n \geq 1$ perform the following iterations:

$$\begin{cases} w_{n+1}^* &= \text{prox}_{\sigma F^*}(w_n^* + \sigma \nabla w_n) \\ w_{n+1} &= \text{prox}_{\tau G}(w_n - \tau(-\text{div}(w_{n+1}^*))) \end{cases} \quad (\text{A.6})$$

where w^* is the dual variable of w and div is the divergence operator. The remainder of the terms are explained in the following paragraphs. Intuitively, the algorithm consists in four steps:

- (i) A gradient ascent with a fixed step σ performed in the dual space.
- (ii) A projection of the result obtained in step 1 on σF^* via the proximal operator, where F^* is the convex conjugate of F .
- (iii) A gradient descent with a fixed step τ performed in the primal space.
- (iv) A projection of the result obtained in step 3 on G via the proximal operator.

The proximal operator of a functional H in a point $y \in \mathbb{R}^n$ is given by:

$$\text{prox}_H(y) = \underset{x \in \mathbb{R}^n}{\text{argmin}} \frac{\|x - y\|^2}{2} + H(x) \quad (\text{A.7})$$

It can be shown that the proximal operator of σF^* in a point w^* is given by:

$$\text{prox}_{\sigma F^*}(w^*) = \frac{2\beta^2}{2\beta^2 + \sigma} w^* \quad (\text{A.8})$$

The proximal operator of τG is computed as (see (Chambolle & Pock 2011)):

$$\text{prox}_{\tau G}(w) = \begin{cases} w + \tau \nabla I & \text{if } \rho(w) < -\tau |\nabla I|^2 \\ w - \tau \nabla I & \text{if } \rho(w) > \tau |\nabla I|^2 \\ -\rho(w) \frac{\nabla I}{|\nabla I|^2} & \text{if } |\rho(w)| \leq \tau |\nabla I|^2 \end{cases} \quad (\text{A.9})$$

where $\rho(w) = (\nabla I)^T w + I_t$. Note that the proximal operators in Eq. (A.8) and (A.9) are computed on a pixel-by-pixel basis.

Appendix A.3. Stability of the primal-dual numerical scheme

It was proven in (Chambolle & Pock 2011) that the numerical scheme described by Eq. (A.6) will converge only if $\tau\sigma\|\nabla\|^2 < 1$, where $\|\nabla\|$ is the norm of the gradient operator. According to (Chambolle 2004) for 2D images $\|\nabla\|^2 \leq 8$. Thus, τ and σ need to be chosen such that $\tau\sigma < 1/8$. An additional condition for the convergence

of the numerical scheme is that τ itself should be smaller than $\|\nabla\|^2$. Therefore, the parameters τ and σ were established as follows: τ was chosen as close as possible to $1/8$ such that the numerical scheme converges, after which σ was selected as the largest value that fulfills $\tau\sigma < 1/8$.

Appendix B. Theoretical analysis of the precision of the reliability test #1

In the reliability test #1 (detailed in section 2.2.3), a temporally under-sampled signal (imaging was retrospectively synchronized to the cardiac activity) was upsampled at the original temporal resolution to build a gold standard data set. To quantify the inherent theoretic uncertainty (denoted by ε) arising from this process, we first define a generic analytic breathing motion pattern as follows:

$$s(t) = A \cdot \sin^4\left(\frac{\pi}{T_r} \cdot t\right) \quad (\text{B.1})$$

where $s(t)$ is the actual displacement amplitude at instant t , A the peak-to-peak displacement, and T_r the period of the breathing activity. Previous studies have shown that such a pattern closely approximates the actual realistic breathing motion (Seppenwoolde et al. 2002, Lujan et al. 2003, Killoran et al. 2011).

ε can be mathematically expressed as follows:

$$\varepsilon = \frac{\int_0^{+\infty} [s(t) - S(t)] \cdot \delta_{TR}(t) dt}{\int_0^{+\infty} \delta_{TR}(t) dt} \quad (\text{B.2})$$

where $S(t)$ is the signal reconstructed from the temporally under-sampled data, and δ_{TR} the impulse train corresponding to the original temporal resolution of the acquired MR-images (the time period between impulses is equal to the repetition time TR of employed MR-imaging sequence). $S(t)$ can be mathematically expressed as follows:

$$S(t) = (s(t) \cdot \delta_{T_c}(t)) * \Lambda_{T_c}(t) \quad (\text{B.3})$$

where T_c the period of the cardiac cycle, $\delta_{T_c}(t)$ is the impulse train corresponding to the repetition of the cardiac cycle, and Λ_T is a tent function used in the linear interpolation required by the reconstruction of the temporally under-sampled signal. Λ_T can be expressed as follows:

$$\Lambda_T(t) = \Pi_T(t) * \Pi_T(t) \quad (\text{B.4})$$

where T is the sampling distance, and Π_T is a box function of width T .

References

- Arnold, P., Preiswerk, F., Fasel, B., Salomir, R., SCheffler, K. & Cattin, P. (2011). 3D organ motion prediction for MR-guided high intensity focused ultrasound, *Proc. MICCAI* **14**(2): 623–630.

- Baskar, R., Lee, K. A., Yeo, R. & Yeoh, K. (2012). Cancer and radiation therapy: Current advances and future directions, *Int J Med Sci* **9**(3): 193–199.
- Booth, J. (2002). *Modelling the impact of treatment uncertainties in radiotherapy*, PhD thesis, University of Adelaide.
- Brix, L., Ringgaard, S., Sorensen, T. S. & Poulsen, P. R. (2014). Three-dimensional liver motion tracking using real-time two-dimensional MRI, *Med. Phys.* **41**(4): 042302:1–10.
- Brox, T., Bruhn, A., Papenberg, N. & Weickert, J. (2004). High accuracy optical flow estimation based on a theory for warping, *Computer Vision-ECCV 2004* pp. 25–36.
- Chambolle, A. (2004). An algorithm for total variation minimization and applications, *J Math Imaging Vis* **20**(1-2): 89–77.
- Chambolle, A. & Pock, T. (2011). A first-order primal-dual algorithm for convex problems with applications to imaging, *J Math Imaging Vis* **40**(1): 120–145.
- Cline, H., Schenck, J., Hynynen, K., Watkins, R., Souza, S. & Jolesz, F. (1992). MR-guided focused ultrasound surgery, *J Comput Assist Tomogr* **16**: 956–965.
- Crijns, S., Kok, J., Lagendijk, J. & Raaymakers, B. (2011). Towards MRI-gated linear accelerator control: gating on an MRI accelerator, *Phys. Med. Biol.* **56**(15): 4815–4825.
- Crijns, S. P. M., Raaymakers, B. W. & Lagendijk, J. J. W. (2012). Proof of concept of MRI-guided tracked radiation delivery: tracking one dimensional motion, *Phys. Med. Biol.* **57**(23): 7863–7872.
- d’Angelo, E., Paratte, J., Puy, G. & Vandergheynst, P. (2011). Fast TV-L1 optical flow for interactivity, *In: IEEE International Conference on Image Processing (ICIP11)*, pp. 1925–1928.
- Delaney, G., Jacob, S., Featherstone, C. & Barton, M. (2005). The role of radiotherapy in cancer treatment, *Cancer* **104**(6): 1129–1137.
- Denis de Senneville, B., El Hamidi, A. & Moonen, C. (2015). A direct PCA-based approach for real-time description of physiological organ deformations, *IEEE Transaction on Medical Imaging* **34**(4): 974–982.
- Denis de Senneville, B., Ries, M., Maclair, G. & Moonen, C. (2011). MR-guided thermotherapy of abdominal organs using a robust PCA-based motion descriptor, *IEEE Transaction on Medical Imaging* **30**(11): 1987–1995.
- Feinberg, D. A., Giese, D., Bongers, D. A., Ramanna, S., Zaitsev, M., Markl, M. & M., G. (2010). Hybrid ultrasound MRI for improved cardiac imaging and real-time respiration control, *Magn. Reson. Med.* **63**(2): 2290–296.
- Feng, M., Balter, J. M., Normolle, D., Adusumilli, S., Cao, Y., Chenevert, T. L. & Ben-Josef, E. (2009). Characterization of pancreatic tumor motion using cine MRI: surrogates for tumor position should be used with caution, *Int. J. Radiation Oncology Biol. Phys.* **74**(3): 884 – 91.
- Foley, J., Eames, M., Snell, J., Hananel, A., Kassell, N. & Aubry, J.-F. (2013). Image-guided focused ultrasound: state of the technology and the challenges that lie ahead, *Imaging Med.* **5**(4): 357–370.
- Glitzner, M., Denis de Senneville, B., Lagendijk, J., Raaymakers, B. & Crijns, S. (2015). Real-time 3D anatomy estimation from undersampled MR acquisitions, *Med. Phys.* **42**(6): 3320–3320.
- Goitein, M. (2004). Organ and tumor motion: An overview, *Seminars in Radiation Oncology* **14**: 2–9.
- Guckenberger, M., Richter, A., Boda-Heggemann, J. & Lohr, F. (2012). Motion compensation in radiotherapy, *Crit Rev Biomed Eng.* **40**(3): 187–197.
- Horn, B. & Schunck, B. (1981). Determining optical flow, *Artificial Intelligence* **17**: 185–203.
- Hynynen, K., Freund, W., Cline, H., Chung, A., Watkins, R., Vetro, J. & Jolesz, F. (1996). A clinical, noninvasive, MR imaging-monitored ultrasound surgery method, *Radiographics* **16**: 185–195.
- Illing, R., Kennedy, J., ter Haar, G., Protheroe, A., Friend, P., Gleeson, F., Cranston, D., Phillips, R. & Middleton, M. (2005). The safety and feasibility of extracorporeal high-intensity focused ultrasound (hifu) for the treatment of liver and kidney tumours in a western population, *Br J Cancer* **93**(8): 890–895.
- Keall, P. J., Mageras, G. S., Balter, J. M., Emery, R. S., Forster, K. M., Jiang, S. B., Kapatoes, J. M.,

- Low, D. A., Murphy, M. J., Murray, B. R., Ramsey, C. R., Van Herk, M. B., Vedam, S. S., Wong, J. W. & Yorke, E. (2006). The management of respiratory motion in radiation oncology report of aapm task group 76a, *Med. Phys.* **33**(10): 3874 – 3900.
- Killoran, J. H., Gerbaudo, V. H., Mamede, M., Ionascu, D., Park, S. J. & Berbeco, R. (2011). Motion artifacts occurring at the lung/diaphragm interface using 4D CT attenuation correction of 4D PET scans, *J. Appl. Clin. Med. Phys.* **12**(4): 261–274.
- Köhler, M. O., Denis de Senneville, B., Quesson, B., Moonen, C. T. & M., R. (2011). Spectrally selective pencil-beam navigator for motion compensation of MR-guided high-intensity focused ultrasound therapy of abdominal organs, *Magn. Reson. Med.* **66**(1): 102–111.
- Langen, K. & Jones, D. (2001). Organ motion and its management, *Int J Radiation Oncology Biol Phys* **50**(1): 265–278.
- Lujan, A. E., Balter, J. M. & Haken, R. K. (2003). A method for incorporating organ motion due to breathing into 3D dose calculations in the liver: sensitivity to variations in motion, *Med. Phys.* **30**(10): 2643–2649.
- Minn, A. Y., Schellenberg, D., Maxim, P., Suh, Y., McKenna, S., Cox, B., Dieterich, S., Xing, L., Graves, E., Goodman, K. A., Chang, D. & Koong, A. C. (2009). Pancreatic tumor motion on a single planning 4d-ct does not correlate with intrafraction tumor motion during treatment, *Am. J. Clin. Oncol.* **32**(4): 364 – 8.
- Mutic, S. & Dempsey, J. F. (2014). The viewray system: magnetic resonance-guided and controlled radiotherapy, *Semin Radiat Oncol.* **24**(3): 196–199.
- Ohara, K., Okumura, T., Akisada, M., Inada, T., Mori, T., Yokota, H. & Calaguas, M. (1989). Irradiation synchronized with respiratory gate, *Int J Radiation Oncology Biol Phys* **17**(4): 853–857.
- Okada, A., Murakami, T., Mikami, K., Onishi, H., Tanigawa, N., Marukawa, T. & Nakamura, H. (2006). A case of hepatocellular carcinoma treated by mr-guided focused ultrasound ablation with respiratory gating, *Magn Reson Med Sci.* **5**(3): 167–171.
- Raaymakers, B., Lagendijk, J., Overweg, J., Kok, J., Raaijmakers, A., Kerkhof, E., van der Put, R., Meijnsing, I., Crijns, S., Benedosso, F., van Vulpen, M., de Graaff, C., Allen, J. & Brown, K. (2009). Integrating a 1.5 T MRI scanner with a 6 MV accelerator: proof of concept, *Phys. Med. Biol.* **54**(12): 229–237.
- Rakêt, L., Roholm, L., Nielsen, M. & Lauze, F. (2011). TV-L1 optical flow for vector values images, *Energy minimization methods in computer vision and pattern recognition*, Vol. 6819 of *Lecture notes in computer science*, Springer Berlin Heidelberg, pp. 329–343.
- Ries, M., Denis de Senneville, B., Roujol, S., Berber, Y., Quesson, B. & Moonen, C. (2010). Real-time 3D target tracking in MRI guided focused ultrasound ablations in moving tissues, *Magnetic Resonance in Medicine* **64**: 1704–1712.
- Roujol, S., Ries, M., Moonen, C. T. W. & Denis de Senneville, B. (2011). Automatic non-rigid calibration of image registration for real time MR-guided HIFU ablations of mobile organs, *IEEE Transaction on Medical Imaging* **30**(10): 1737–1745.
- Roujol, S., Ries, M., Quesson, B., Moonen, C. & Denis de Senneville, B. (2010). Real-time MR-thermometry and dosimetry for interventional guidance on abdominal organs, *Magnetic Resonance in Medicine* **63**: 1080–1087.
- Seppenwoolde, Y., Berbeco, R. I., Nishioka, S., Shirato, H. & Heijmen, B. (2007). Accuracy of tumor motion compensation algorithm from a robotic respiratory tracking system: a simulation study, *Med. Phys.* **34**(7): 2774 – 84.
- Seppenwoolde, Y., Shirato, H., Kitamura, K., Shimizu, S., van Herk, M., Lebesque, J. V. & Miyasaka, K. (2002). Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy, *Int. J. Radiation Oncology Biol. Phys.* **53**(4): 822–834.
- Stemkens, B., Tijssen, R. H. N., Denis de Senneville, B., Heerkens, H. D., van Vulpen, M., Lagendijk, J. J. W. & van den Berg, C. A. T. (2015). Optimizing 4-dimensional magnetic resonance imaging data sampling for respiratory motion analysis of pancreating tumors, *Int. J. Radiation Oncology*

Biol. Phys. **91**(3): 571–578.

Tempany, C., McDannold, N., Hynynen, K. & Jolesz, F. (2011). Focused ultrasound surgery in oncology: overview and principles, *Radiology* **36**(5): 39–56.

Wedel, A., Pock, T., Zach, C., Bischof, H. & Cremers, D. (2009). An improved algorithm for TV-L1 optical flow, *Statistical and geometrical approaches to visual motion analysis*, Vol. 5604 of *Lecture notes in computer science*, Springer Berlin Heidelberg, pp. 23–45.

Zachiu, C., Denis de Senneville, B., Moonen, C. & Ries, M. (2015). A framework for the correction of slow physiological drifts during MR-guided HIFU therapies: Proof of concept, *Med. Phys.* **42**(7): 4137–4148.