Robust Real-Time-Constrained Estimation of Respiratory Motion for Interventional MRI on Mobile Organs

Sébastien Roujol, Jenny Benois-Pineau, Baudouin Denis de Senneville, Mario Ries, Bruno Quesson, Chrit Moonen

Robust real time constrained estimation of respiratory motion for interventional MRI on mobile organs.

Sébastien Roujol*, Jenny Benois-Pineau, *Member, IEEE,* Baudouin Denis de Senneville, Mario Ries, Bruno Quesson, Chrit Moonen

***Abstract*—Real time magnetic resonance (MR) imaging is a promising tool for image-guided interventions. For applications such as thermotherapy on moving organs, a fine image-based compensation of motion is required in real time to allow quantitative analysis, retro-control of the interventional device, or determination of the therapy endpoint. Since interventional procedures are usually restricted to a part of the organ/tissue under study, reduced FOV imaging represents a promising way to improve spatial and / or temporal resolution. However, it introduces new challenges for the target motion estimation since structures near the target may appear transiently due to the respiratory motion and the limited spatial coverage.**

In this paper, a new image based motion estimation method is proposed combining a global motion estimation with a novel optical flow approach extending the initial Horn & Schunck (H&S) method by an additional regularization term. This term integrates the displacement of physiological landmarks, which are obtained in a preparation step by pattern matching into the variational formulation of the optical flow problem. A smooth regulation of the constraint point influences is achieved using a spatial weighting function. The method was compared to the same registration pipeline employing the H&S approach. A first evaluation was performed on synthetic dataset where the accuracy of the motion estimated with the H&S method was improved by a factor of 2 using the proposed approach. An *in vivo* study was then realized on both the heart and the kidney of twelve volunteers. Compared to the H&S approach, a significant improvement ($p<0.05$) of the DICE similarity criterion computed between the reference and the registered organ positions was achieved.

***Index Terms*—Image registration, Motion analysis, Biomedical image processing, Magnetic resonance imaging.**

I. Introduction

REAL time MR imaging proved to be a promising candidate for guiding non- and mini-invasive surgical interventions [1]. The additional use of quantitative MR-measurements of tissue specific properties such as MR-relaxation times, or the local temperature, allows monitoring of the therapeutic progress and helps to determine the endpoint of the intervention. However, quantitative measurements require a motion correction to enable a direct comparison on a voxel-by-voxel basis between the dynamically acquired images. In the particular case of MR-guided thermal ablations, real time MR-thermometry is used to compute temperature maps inside the human body [2]. The temperature evolution can be used as an accurate and immediate prediction of tissue necrosis [3] and thus requires all temperature maps registered to a common position. Moreover, the whole registration process has to be realized within the interval of two MR-acquisitions to ensure the real time condition.

Typically, interventional procedures are usually restricted to a part of the organ/tissue under study making reduced field of view (FOV) imaging desirable. This would allow improvement of the spatial and / or temporal resolution in order to decrease partial volume effects (undesirable for quantitative analysis) and to increase imaging framerate (required to observe rapid phenomena). Several strategies have been proposed toward this direction by using saturation slabs [4], outer volume suppression [5] or interactive reduced FOV imaging [6]. However, a reduced FOV may introduce new challenges for the target motion estimation since structures may appear transiently due to the respiratory motion and the limited spatial coverage.

A variety of motion estimation algorithms [7], [8] have been suggested in the field of medical imaging. Optical flow algorithms [9] have been proposed for motion estimation on abdominal organs for MR-guided laser ablation [10]. Recently, an approach using an initial global motion estimation followed by an optical flow algorithm was developed for real time MR-thermometry in abdominal organs [11]. Optical flow algorithms allow estimating a velocity field assuming an intensity conservation during displacement, mathematically expressed by the optical flow equation (OFE):

$$I_x u + I_y v = -I_t. \quad (1)$$

Here u and v are the displacement vector components, and $I_{x,y,t}$ are the spatio-temporal partial derivatives of the image pixel intensity. However, a direct estimation by minimizing the deviation from OFE (equation 1) is an under-determined problem and an additional constraint is required. The method proposed by Horn and Schunck (referred to as H&S in this paper) introduces additional physical constraints enforcing the smoothness of the motion field [12]. They seek $u(x, y)$ and

This work was supported in part by Ligue Nationale Contre le Cancer, Conseil Régional d'Aquitaine, Diagnostic Molecular Imaging, Agence National de Recherche, Fondation InNaBioSanté and Philips Medical System.

M. Roujol*, Denis de Senneville, Ries, Quesson and Moonen are with the Laboratory for Molecular and Functional Imaging: from Physiology to Therapy FRE 3313 CNRS/ University of Bordeaux 2 - 146 rue Leo Saignat, Case 117, 33076 Bordeaux, France. e-mail: {s.roujol,m.ries,b.quesson,chrit.moonen,baudouin}@imf.u-bordeaux2.fr. M. Roujol and Mrs Benois-Pineau are with Laboratoire Bordelais de Recherche en Informatique, UMR 5800 CNRS/University of Bordeaux 1 - 351, cours de la Liberation, 33405 Talence, France e-mail: jenny.benois@labri.fr. M. Quesson is also with Centre de recherche cardio thoracique, Inserm U1045, University of Bordeaux 2 - 146 rue Leo Saignat, 33076 Bordeaux, France. *Asterisk indicates the corresponding author.*

$v(x, y)$ minimizing:

$$E(u, v) = \iint \left([I_x u + I_y v + I_t]^2 + \alpha^2 \left[\|\nabla u\|_2^2 + \|\nabla v\|_2^2 \right] \right) dxdy, \quad (2)$$

where α^2 is a weighting factor designed to link both intensity variation and motion field regularity and ∇u and ∇v are the spatial gradients of $u(x, y)$ and $v(x, y)$, respectively.

However, when reduced FOV imaging is employed, the assumption of energy conservation may be violated due to the potential presence of transient structures. In such conditions, the Horn & Schunck algorithm is expected to fail locally (near the transient structure) to recover the correct motion. This is known as one of the ill-posed problems in motion estimation, namely the occlusion problem [12]. In detail, if a structure is only present in the reference image, the algorithm might try to match the signal from this structure (in the reference image) with a different structure (in the image to be registered). To improve the robustness of the algorithm against this effect, one can increase the value of α^2 to increase the smoothness constraint of the motion field. However, this will reduce at the same time the ability of the algorithm to provide reliable motion estimation in presence of complex deformation.

Therefore, the first step of this study was to hypothesize that a combination of intensity and landmark registration could then be used to improve the robustness of the registration against transient structures. Such combinations have been proposed in the past such as in [13] where a large number of landmarks were employed for brain registration. This method was further extended to deal with a smaller number of points (manually defined) for interactive registration of medical images [14]. In these approaches, both intensity and landmark metrics were minimized sequentially inside an iterative minimization process. Several unified minimization framework were also proposed. In [15] the algorithm was designed to obtain a final motion field fitting each pre estimated landmark motion. In the presented application, due to the low SNR, a non negligible uncertainty of the landmark correspondence may be expected and would thus strongly affect such registration approach. Another solution proposed by Becciu *et al.* [16], attempt to use tags obtained from MR-tagged sequence in their registration framework for cardiac contraction assessment. Unfortunately, this sequence is generally unsuitable for interventional MRI, since images are tagged by regular lines where the signal has been removed. Recently, a variational approach, integrating segmented region motion, was proposed for large displacement estimation [17]. This method uses a linearized OFE deviation together with regularization terms which include the correspondence of region displacements in the image plane. Despite the interest of such an approach in general purpose video sequences, its application to MRI sequences is not straightforward due to the inherent difficulties of segmentation of frames into spatially coherent regions. Recently, a grid-based deformation model was proposed [18], but was also shown to be sensitive/limited by the landmark extraction process according to the authors.

In this paper, we propose a new real-time motion estimation method for MRI sequences which can operate in the case of reduced FOV imaging. Preliminary results of this approach can be found in [19]. The contribution of the paper is twofold. We first define the two stage interventional protocol for a robust choice of constraint points and then formulate the constrained optical flow estimation by introducing an additional regularization term in the H&S method. In our formulation, the introduction of a smooth weighting function allows for a local control of the influence of constraint points. Furthermore, in order to ensure the real-time requirement together with a short latency, all computationally intensive calculations were off-loaded to a dedicated graphics processing unit (GPU). The proposed algorithm is referred to as constrained motion estimation (CME) in the scope of this paper. It was compared with the same registration pipeline employing the H&S optical flow approach. Algorithm evaluations were conducted on both synthetic data and cardiac & kidney MR-images of healthy volunteers under free breathing conditions.

II. Material and Methods

The proposed CME algorithm is a two-step procedure (see Fig. 1). The first step consists in selecting the constraint points along the boundary of the organ in the reference image of the time series. The contour is then subsampled and refined based on the position of extracted feature points. In a second step the motion is estimated for each image as follows: a global translational motion estimation is performed and used to initialize a local estimation of the displacement of constraint points. Non-physiological constraint point displacements are automatically identified and corresponding constraint points are discarded. The displacements of the constraint points are then integrated into the constrained optical flow algorithm (using the global estimated motion as preconditioning) to obtain the final motion field. A detailed description of each algorithm step is now presented in the next sections.

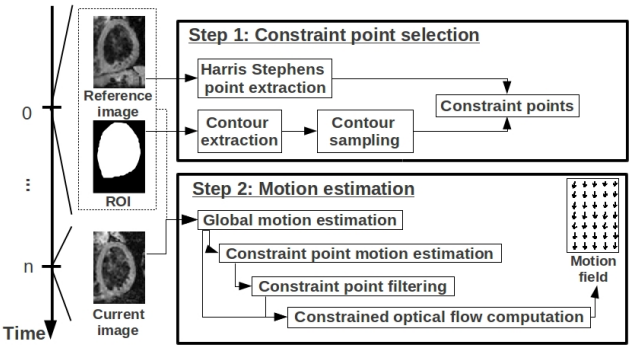

Fig. 1: General scheme of the algorithm. Prior to the intervention, constraint points are automatically extracted from a reference image (step 1). Then, during the procedure, the motion field is estimated for each frame (step 2).

A. Step 1: Constraint point selection

Anatomical points are localized and tracked over the time in order to guide and to constrain the motion estimation

of the target. To select them, anatomical structures such as organ boundaries, which remain present during the acquisition and follow the target, are suitable. For this, a region of interest (ROI) was manually set around the target of the reference image. The contour of the ROI was first extracted and then regularly sampled in space to obtain a set of N points surrounding the target. To allow a certain degree of freedom on the ROI drawing a refinement of the location of the N points was then performed by moving them a near feature point computed on the reference image. Due to the abundance of works in stereo matching and image retrieval, a large amount of feature point detectors has been tested and reported in literature [20]. The critical point is the stability of these methods with respect to affine transformations of image plane, lightening, scale variations and noise. In the case of MR images, the noise and deformable motion are the main factors. According to the evaluation in [20], the Harris-Stephens detector [21] appeared to provide a good compromise between robustness and computation time. The feature point detection is based on the following response function:

$$R(x,y) = Det(M_{x,y}) - k.Tr(M_{x,y})^2, \tag{3}$$

with

$$M_{x,y} = \sum_{i,j \in S} w_{i,j} \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix}_{i,j}, \tag{4}$$

where (x,y) denotes the spatial coordinates, Det denotes the determinant of a matrix, Tr is the trace of a matrix, w is a weighting factor (Gaussian kernel over the region S centered on (x,y)) and k is a sensitivity parameter. The value of k is generally chosen in the [0.04-0.15] interval. In our case, a value of 0.04 was empirically found suitable for our application. The response is positive in a corner region. Therefore, the feature point with the higher response in a small neighborhood (e.g. 3x3 pixels) of the contour point was selected as landmark point. The small neighborhood search was designed to prevent positioning a landmark on a different structure/organ that may depict a different motion from the targeted organ. If no feature point was present, the initial contour point was then selected.

B. Step 2: motion estimation algorithm

An optical flow based algorithm is more efficient when it is initialized near the global optimum. Therefore, to initialize it, a global motion estimation is first performed with a simple translational model. The translation parameters (horizontal and vertical) are estimated using a sign-gradient-descent with fixed step inspired by the Netravali-Robbins method [22]. The estimation is restricted to the ROI defined in section II-A.

Then, constraint point displacements are individually estimated (two translation parameters) using the global estimated displacement as initial estimate. This estimation is restricted to a small patch centered on each constraint point intersected with the initial ROI to allow a local refinement of the global displacement. We experimentally found that a patch size of 10×10 pixels was satisfactory using a given voxel size of 2-3 mm obtained with the employed MR-acquisitions.

To remove occasional non-physiological estimates, the following outliers rejection was added. The displacement vector (d_x, d_y) of a constraint point was supposed to follow a bivariate Gaussian distribution with independence of d_x and d_y coordinates. A constraint point was automatically rejected if at least one of its displacement components violated the marginal 3-sigma rule.

The idea of the presented approach is to constrain the H&S formulation by locally estimated displacements of feature points. Hence, we propose the following extension of the Horn & Schunck formulation with an additional regularization term:

$$E_c(u,v) = \iint \Big(\left[I_x u + I_y v + I_t\right]^2 + \alpha^2 \left[\|\nabla u\|_2^2 + \|\nabla v\|_2^2\right] + \lambda^2 \sum_{i=1}^{N} \left(\rho(d_i, R) \left[(u-u_i)^2 + (v-v_i)^2 \right] \right) \Big) dxdy, \tag{5}$$

where (u_i, v_i) are the horizontal and vertical components of the displacement estimated for the i^{th} constraint point. λ^2 is the regularization parameter that allows balancing between the initial behavior of the H&S algorithm and the constraint influence. ρ is a distance function, defined as

$$\rho(d,R) = \exp\left(-d^2/R^2\right) \tag{6}$$

where d represents the Euclidean distance between the pixel of coordinates (x,y) and the i^{th} constraint point, R is a bandwidth parameter. To minimize $E_c(u,v)$, we used the calculus of variation and obtained the following system:

$$\begin{cases} I_x^2 u + I_x I_y v = \alpha^2 \nabla^2 u - I_x I_t + \lambda^2 \sum_{i=1}^{N} \left(\rho(d_i,R) u_i\right) \\ I_x I_y u + I_y^2 v = \alpha^2 \nabla^2 v - I_y I_t + \lambda^2 \sum_{i=1}^{N} \left(\rho(d_i,R) v_i\right) \end{cases}. \tag{7}$$

Then, we used the approximation of the Laplacian, as suggested by Horn and Schunck, where $\nabla^2 u = \overline{u} - u$, with $\overline{u}$ the mean value of u in the neighborhood (3x3 pixels) of the estimated point [12]. Therefore, the system can be rewritten as:

$$\begin{cases} a_{11} u + a_{12} v = b_1 \\ a_{21} u + a_{22} v = b_2 \end{cases}, \tag{8}$$

with

$$\begin{cases} a_{11} = I_x^2 + \alpha^2 + \lambda^2 \sum_{i=1}^{N} \rho(d_i,R) \\ a_{12} = a_{21} = I_x I_y \\ a_{22} = I_y^2 + \alpha^2 + \lambda^2 \sum_{i=1}^{N} \rho(d_i,R) \\ b_1 = \alpha^2 \overline{u} - I_x I_t + \lambda^2 \sum_{i=1}^{N} \rho(d_i,R) u_i \\ b_2 = \alpha^2 \overline{v} - I_y I_t + \lambda^2 \sum_{i=1}^{N} \rho(d_i,R) v_i \end{cases}. \tag{9}$$

Finally, based on the Jacobi method, the system can be solved with the following iterative scheme:

$$u^{n+1} = \frac{b_1^n a_{22} - a_{12} b_2^n}{a_{11}a_{22} - a_{12}a_{21}} \quad , v^{n+1} = \frac{a_{11} b_2^n - a_{21} b_1^n}{a_{11}a_{22} - a_{12}a_{21}}. \quad (10)$$

C. Implementation

A registration to a reference frame has been preferred to a concatenation of frame to frame estimations since the latter generally lead to error accumulations, especially when high framerate is applied such as in abdominal imaging (10-15 Hz). In cardiac application, since the imaging framerate is generally limited to the cardiac frequency, the respiratory motion between two successive images generally depicts a large amplitude, devaluating the benefit of a frame to frame motion estimation. For both the H&S and the CME implementation, the multi-resolution approach from [23] was used, which refines the motion estimation algorithm from the 3^{rd} level of sub-resolution to the full resolution. We off-loaded the most time consuming task, i.e. the iterative numerical scheme of the optical flow, to a dedicated GPU.

In-vivo calibration of the employed algorithms is not an easy task since it depends on the criterion to be optimized. The calibration of the H&S method (α^2 value) has to deal with contradictory effects. A small regularization of the motion field smoothness is required to enable the estimation of complex motion and to have a globally reliable estimated motion. On the other hand, a high constraint on the motion field smoothness would reduce the registration artifact induced by intrusive structures but will at the same time limit the ability of the algorithm to estimate complex motion. This may thus deteriorate the registration in the entire organ and thus the value of such estimated motion in the perspective of quantitative analysis and clinical application. Therefore, we decided to calibrate the employed algorithms in a way to maintain optimal performance in the general case of full FOV imaging (without the presence of intrusive structures). As recently shown [24], a reliable in-vivo calibration of the H&S algorithm was obtained for a range of α^2 values between 0.1-0.5. A plateau was generally observed for these ranges of values and its lower bound was suggested as a good way to cope with variations of the breathing pattern (such as an amplitude variation or drift of the respiration pattern). Therefore, a α^2 value of 0.1 was employed for the H&S algorithm. We empirically found a near-optimal solution for the CME calibration by employing the following parameters: α^2=0.1, λ^2=0.1, N=20 and R^2=5. Note that a similar optimal configuration was obtained for the synthetic dataset experiment (see result part) except for the α^2 values that was higher due the lowest complexity of the synthesized motion.

The overall algorithm was implemented in C++ and evaluated on a dual processor (INTEL 3.1 GHz Penryn, two cores). The GPU implementation was based on the Compute Unified Device Architecture (CUDA) framework [25] using a NVIDIA GTX280 card.

D. Experimental setup

The proposed algorithm was evaluated on both synthetic and *in vivo* datasets:

1) Synthetic dataset experiment:

Data creation: A sequence of T(= 30) images was created. To simulate respiratory motion typically encountered on mobile organs, a periodical (period=6 frames) geometric transformation composed by a 2D translation (T_x,T_y) and scaling (S_x,S_y) was synthesized ($T_x = \{0, 0.5, 1, 1.5, 1, 0.5\}$ pixels,$T_y = \{0, 2.5, 5, 7.5, 5, 2.5\}$ pixels, $S_x = S_y = \{1, 1.03, 1.06, 1.09, 1.06, 1.03\}$). A signal-to-noise ratio (SNR_{dB}) of 1.3 was chosen to simulate a realistic acquisition (typically between 0.7 and 1.3). A structure appearing transiently in the lower part of the image was added in half of the images to simulate the effect potentially encountered with reduced FOV imaging.

Quality assessment of the motion estimation: Since the real motion ($D_{gt} = (u_{GT}, v_{GT})$) and the estimated motion ($D = (u, v)$) are available for each pixel in such synthetic dataset experiment, the measures commonly reported in the optical flow community such as the endpoint error (EE) and the angular error (AE) of the flow [9], [26] were computed, with:

$$EE = \sqrt{(u - u_{GT})^2 + (v - v_{GT})^2}, \quad (11)$$

$$AE = \cos^{-1} \frac{1 + u \times u_{GT} + v \times v_{GT}}{\sqrt{1 + u^2 + v^2}\sqrt{1 + u_{GT}^2 + v_{GT}^2}}. \quad (12)$$

In addition, to provide additional information about the smoothness of the estimated motion field, the harmonic energy of the estimated flow [27] was reported.

2) in vivo experiments:

In vivo experiments were conducted on the heart and the kidney of a total of 12 healthy volunteers. The MRI scans were performed under free breathing conditions using a 1.5 Tesla scanner (Philips Achieva/Intera, Best, The Netherlands). The volunteers were positioned head first in supine position. The MR sequences employed saturation slabs to obtain structures appearing transiently in the FOV as typically encountered when zoom imaging is used (these conditions are then referred to as "zoom imaging conditions" in the scope of this paper).

In-vivo study on the heart under zoom imaging conditions: Dynamic MRI was performed on the heart of six healthy volunteers. The acquisition sequence was ECG-gated (to observe the heart in the same cardiac phase) using a five element phased array cardiac coil. Five contiguous adjacent slices were acquired per cycle (200 cycles per scan), in short axis view, at end diastolic phase. A slice tracking technique [28] was used to compensate for respiratory motion in the third dimension. Blood signal reduction was obtained using saturation slabs positioned on each side of the imaging stack. The single shot EPI sequence employed the following parameters: FOV=260 $\times$ 260 mm^2, voxel size=2.7 $\times$ 2.7 $\times$ 7 mm^3, echo time=20 ms, repetition time=40 ms, SENSE acceleration factor=1.6 [29]. A saturation slab was positioned underneath the extreme position of the heart (corresponding to the position at maximum respiratory displacement) to simulate zoom imaging conditions.

In-vivo study on the kidney under zoom imaging conditions: 200 frames (single slice) in coronal orientation were acquired using a four element phased array body coil. A dual shot EPI sequence employed the following parameters:

FOV=200 × 400 mm^2, voxel size=2.3 × 2.3 × 6 mm^3, echo time=26 ms, repetition time=52 ms, flip angle=35°. Zoom imaging conditions were achieved using a saturation slab positioned on the top of the extreme position of the kidney.

Quality assessment of the motion estimation: In such conditions, the typical amplitude of both heart and kidney motions is about 8 pixels in imaging plane between two extreme images in the respiratory cycle. Since the true motion is unknown the quality assessment of the registration was analyzed by computing the DICE similarity coefficient (DSC) [30] between the position of the organ (ROI_t) in each registered frame (t) and its position in the reference frame (ROI_{ref}), as follows:

$$DSC(t) = \frac{2(ROI_{ref} \cap ROI_t)}{ROI_{ref} + ROI_t}, \quad (13)$$

Each ROI was obtained by manual segmentation. A value of 1 for DSC indicates an ideal registration (perfect ROI matching). The harmonic energy has been also computed and reported for all tested cases to assess the energy of the deformation fields.

III. Results

A. Synthetic dataset experiment

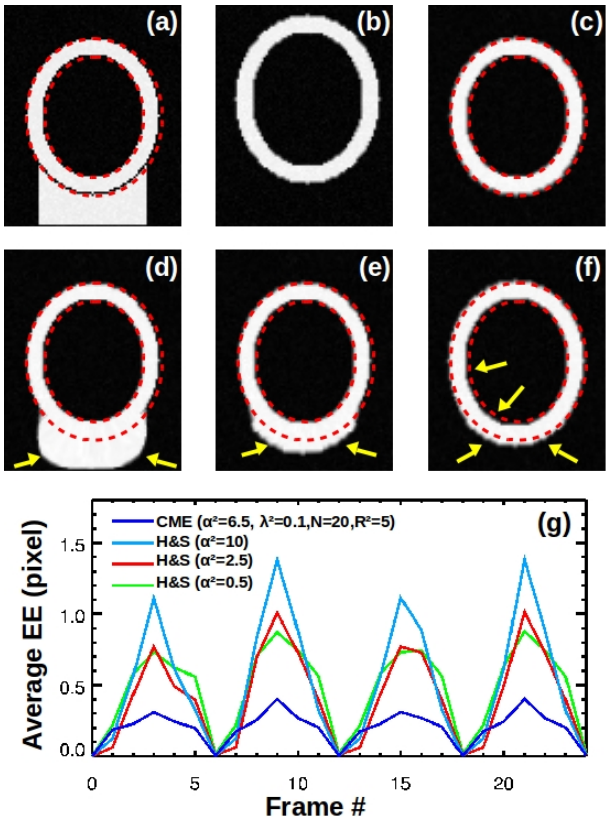

Fig. 2: Registration example obtained on the synthetic dataset experiment. The reference image (a) and the image to be registered (b) are displayed together with the registered images obtained with the CME (c) and the H&S approach using a value of α^2 of 0.5 (d), 2 (e) and 6.5 (f). The time evolution of the average endpoint error (EE) inside the target (between both dashed red ellipses) is reported in (g).

Example of synthetic images is shown in Fig. 2 where the reference image (2a, containing the underneath transient structure) and an image corresponding the maximum synthesized motion (2b, 4^{th} image of the cycle) are displayed. The contours (interior and exterior) of the target obtained from the reference image are reported in dashed red curves. The H&S approach was employed to register the image in (2b) using different α^2 values. When using a small value ($\alpha^2 = 0.5$), the registration is strongly influence by the underneath structure leading to a severe registration artifact (yellow arrows in 2d). By using an optimally calibrated value ($\alpha^2 = 2.5$, see Fig. 3a), the registration artifact was reduced but still present (yellow arrows in 2e). A higher value of $\alpha^2(= 10)$ is then reported in 2f where the registration artifact introduced by the transient structure appeared substantially reduced due to higher weight on the smoothness of the motion field. However, in this case a different registration artifact was observed due to the inability of the method to handle complex motion (here the scaling effect) as shown by the yellow arrows. The CME approach provided a reliable registration where the registered target perfectly matched the reference target contour. These results were confirmed in the plot of the averaged EE values over time (computed inside the target, between the two red dashed curves) in 2g. While the H&S approach periodically failed to recover the correction motion, the CME offered more robust performances over time.

In order to better characterize the potential gain and limitations of the compared methods, their performance in term of averaged error endpoint, harmonic energy and averaged angular error were precisely investigated in function of the employed parameters (Fig. 3). As previously observed in Fig. 2d, low α^2 values (3a, 3e and 3i) provided poor performance since the methods became very sensitive to the presence of the intrusive structure. On the contrary, high values limited the ability of the algorithm to estimate complex motion and also deteriorated the motion estimates as confirmed by the convergence of the harmonic energy toward a very small value. The averaged EE values obtained with an optimal α^2 calibration were O.42 for the H&S approach ($\alpha^2 = 2.5$) and 0.21 for the CME method (α^2=6.5) showing a reduction of the averaged EE by a factor of 2. The λ^2 value influence was then investigated ((3b, 3f and 3j). As expected, small λ^2 values tends to the H&S performance and high values tends to the extrapolation of the constraint point motions (and their associated uncertainty) leading in both cases to a deterioration of the registration. A good calibration of λ^2(=0.1, blue curve) provided a significant improvement of the motion estimation (3b, 3j). The influence of the constraint point number (N) was then evaluated (3c, 3g and 3k). Although an optimal value was reached around 20 points (3c, 3k, blue curve), the sensitivity of the CME to this parameter was limited. Finally the influence of the bandwidth (R^2) provided an optimal calibration for a value of 5 ((3d, (3h, (3i, blue curve).

B. *in vivo* experiments

Similar results were obtained in *in vivo* experiments in both the heart and the kidney of healthy volunteers. Registration

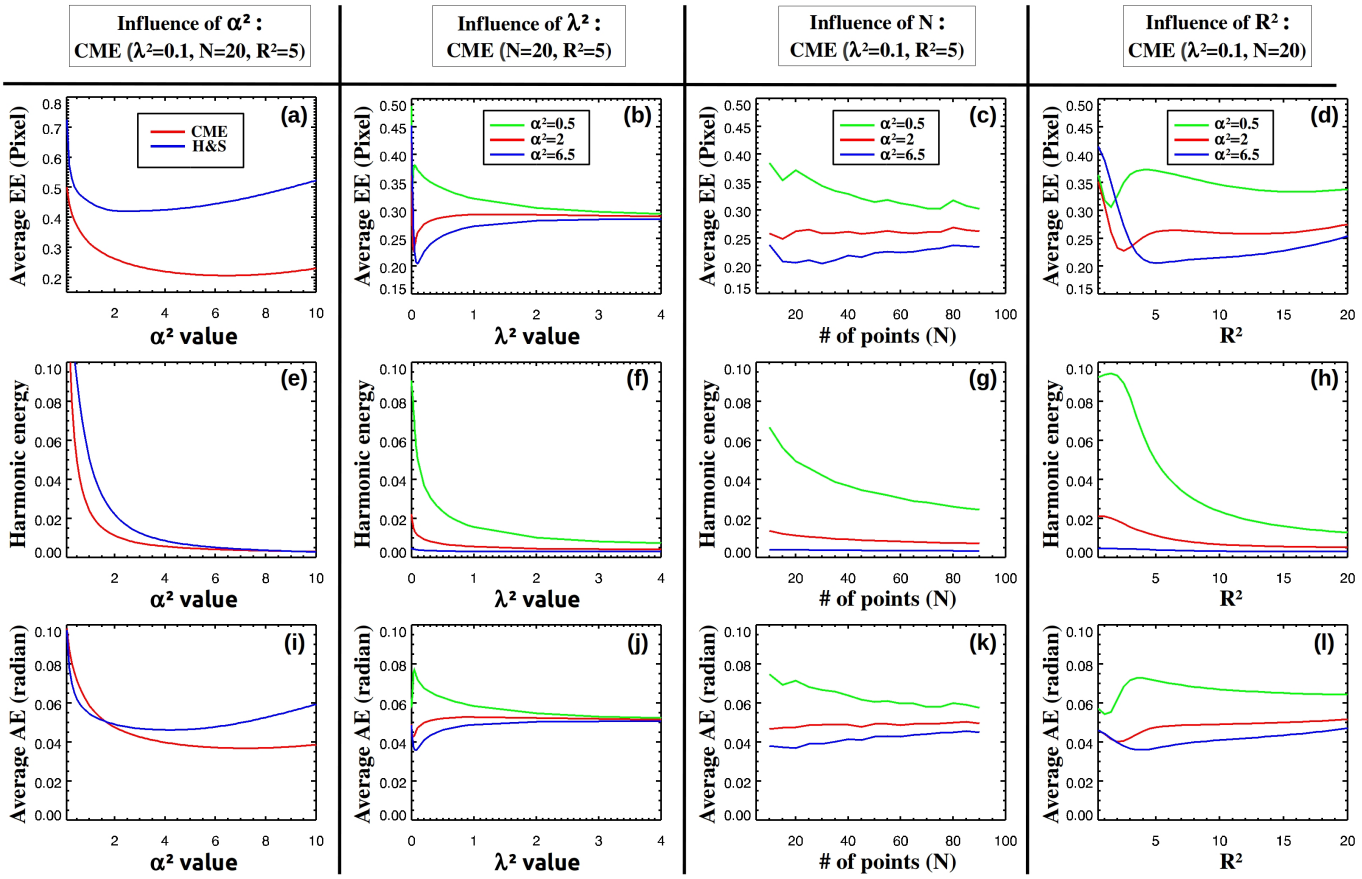

Fig. 3: Influence of the free parameter tuning on the H&S and the CME methods. The averaged error endpoint (EE), the harmonic energy (HE) and the average angular error (AE) are reported in the first, second and third lines of the figure respectively. Note that the averaged EE and AE computation were restricted to the target area (defined between the two red dashed ellipses in Fig. 2a). The influence of the motion field weighting (α^2) on both the H&S and the CME approaches are displayed in the first column. Then, other CME parameters were analyzed: the constraint point term weighting (λ^2, second column), the number of constraint points (N, third column) and the ray influence of the constraint point (R^2, forth column).

examples are shown for both organs in Fig. 4. By comparing the reference images (4a, 4i) with images acquired at different positions in the respiratory cycle (4b, 4j), one can observe some structures appearing transiently (see yellow arrows) due to the signal cancellation obtained from the saturation bands. In the cardiac images, the area with low signal intensity in the lower part corresponds to a saturation band that may allow for the reduction of the FOV without additional fold-over artifacts. The signal of the liver (below the heart) almost disappeared, due to the displacement induced by the respiration. In the abdominal images, two perturbations were observed in the top part of the kidney: the liver, above the kidney, partially disappeared in 4j and the intensity of the upper part of the kidney depicted a high variation. In both examples, while the registration obtained with the H&S approach (4c, 4k) was severely deteriorated in the regions near the transient structures (see red arrows), a reliable registration was obtained in the totality of the organs using the CME (4d,4l).

These findings were typical for the entire sequences as confirmed by the time evolutions of the DICE similarity coefficient and the harmonic energy, respectively shown in (4e,4m) and (4f,4n). Due to the respiratory cycle, the transient structures appeared periodically in the time series and the H&S approach periodically failed to recover a reliable motion estimate leading to low Dice similarity coefficient and elevated harmonic energy values. The CME clearly outperformed the H&S approach by providing a better overall registration and more stable performance.

Over the 12 volunteers, the averaged DICE similarity coefficient (central point inside the box) obtained with the H&S method has been significantly improved ($p<0.05$) using the proposed CME as shown in (4g and 4o. The minimal DICE similarity coefficient values were typically very low for certain frames using the H&S method, whereas the proposed CME allowed maintaining a better performance for all the frames (around 0.92 and 0.96 for the heart dataset and the kidney dataset, respectively)). As expected, higher harmonic energy values were obtained with the H&S method since the method periodically failed to recover the motion and also because a relative lower weight of the smoothness constraint term was employed since the same α^2 value (=0.1) was employed for both algorithms.

Constraint point filtering allowed the rejection of constraint points with non-physiological estimated displacement. In av-

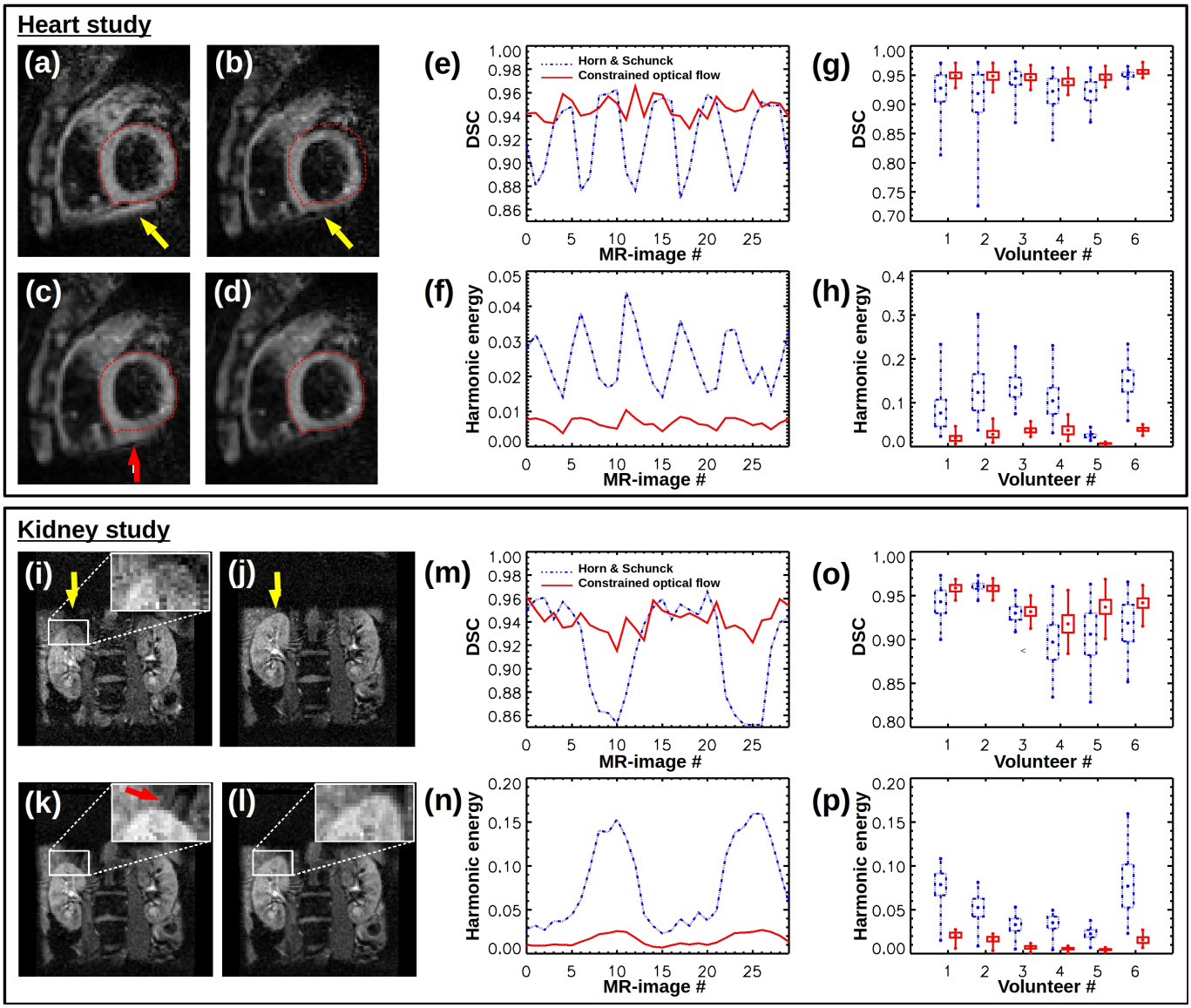

Fig. 4: Registration result obtained *in vivo* in the heart and the kidney of free breathing volunteers. Registration examples in the heart (volunteer #5) and in the kidney (volunteer #6) are given where the reference images (a,i), the images to be registered (b,j) and the registered image obtained with the H&S (c,k) and the CME (d,l) methods are displayed. The time evolution of both the DICE similarity coefficient (DSC) and the harmonic energy (HE) are shown for each registration example in (e,m) and (f,n), respectively. Finally, the DSC (g,o) and HE (h,p) are reported for each volunteer over time as Box and Whisker plot where the minimum (lower point), maximum (upper point), average (point inside the box) and standard deviation (box height) values are shown. While the H&S method was disturbed by the presence of transient structures, the CME provided robust and reliable registration performance for all the frames.

erage over the volunteers, less than 0.37 % and 2.67 % of the constraint points were rejected for the heart and the kidney dataset registration, respectively (with a maximum of two constraint points for a frame.

C. Real time benchmarking

Benchmarking was realized for each processing step for an image sequence of spatial resolution 128 × 128. The GPU based implementation allowed a significant reduction of the whole computation time. The total computation time of the proposed approach was evaluated to 22 ms (against 87.5 ms using a CPU only implementation) and was composed by: global motion estimation (5 ms), otion estimation of constraint points (10 ms), optical flow iterative scheme (7 ms) and image registration (0.01 ms). An acceleration factor of 10 was achieved for the computation time of the iterative numerical scheme of the optical flow algorithm (equation 10).

IV. Discussion

A. Performance of the proposed method

On the synthetic dataset, the proposed CME outperformed the H&S approach that failed to estimate the real motion in

presence of structures appearing transiently. The endpoint error of flow obtained with the optimal H&S calibration has been reduced by a factor of 2 using the proposed CME as shown in Fig. 3a.

Furthermore, the volunteer studies confirmed the *in vivo* feasibility of the CME in both the heart and the kidney. A reliable registration was obtained in the totality of the organ in all frames. Using the H&S method, similar performance were also achieved in frames with similar structures as in the reference image. However, a severe degradation of its performance was observed in presence of additional intrusive structures. Although this artifact could have been reduced by increasing the regularization of the motion field smoothness, this would have in the same time decrease its ability to estimate complex motion (as can be observed in Fig. 2f and 3a using a high value of α^2). As recently shown [24], such effect is generally observed with the H&S method by using α^2 values above the interval [0.1-0.5] with the employed in-vivo images.

B. Real time feasibility of the method

MR-guidance of interventional procedures relies on the instantaneous availability of the processed images. Therefore, this limits the available computation time. In addition, Denis de Senneville *et al.* demonstrated in [31] that large latencies have to be compensated with the help of accurate motion prediction. However, the performance of the prediction algorithm increases greatly with short latencies. Recently, in the particular case of a HIFU ablation on mobile organs, it was demonstrated that a latency inferior to 100 ms was required for the adjustment of the beam position in order to ensure an energy deposition similar to a static experiment [32]. Here, the GPU implementation offered an acceleration factor of 10 for the computation of the constrained optical flow which is in agreement with the published work [11], [33]. Significant higher acceleration factor would be expected using higher resolution as shown in [33] since it would benefit from a larger amount of data to process (higher occupations of each processor) and a lower relative overhead. Overall, the demonstrated CPU/GPU implementation allows the acceleration of the required processing time by a factor of four and thus ensures the real time conditions with a short low latency. Further reduction of the latency may be obtained by investigating the use of more complex optimization scheme aiming to improve the convergence speed of the algorithm.

C. Calibration of the CME

Contrary to the previous works using constraint points, a comprehensive formulation of the minimization problem was proposed. In the proposed approach, the confidence into the predetermined displacement of the selected constraint points can be freely adjusted with the regularization parameter λ^2. The quality of the obtained optical flow depends on the quality of the initial constraint point vectors, the number of constraints (N) and the bandwidth (R) of the ρ function parameters:

- For the motion estimation of the constraint points, only a translational model was considered as it was the most robust for small patch sizes surrounding constraint points. The optimal patch size in 128×128 MRI sequences was found to be 10×10 for our images.
- In order to control optical flow, the constraint points have to be placed near eventual occlusion (or problematic area). The manual choice of constraint points is not realistic during an interventional procedure, and we can only encourage the staff physician to approximately trace the contour of the ROI. Hence, the subsampling has to be sufficient in order to get a good coverage of problematic area. On the other hand, a too large number of constraint points will slow down the computational process and may degrade the registration performance by leading to a quasi interpolation of the constraint displacements and their associated uncertainties. Therefore, for the demonstrated application, N=20 was found to be near-optimal.
- The bandwidth R of the ρ function regulates the influence of remote points. The large bandwidth yields a quasi interpolation of constraint point displacements over the whole image. An optimal experimental value was R=$\sqrt{5}$.
- Outlier rejection for constraint point vectors was found particularly useful for small patch sizes where the estimation is more sensitive to out-of-plane motion, noise, etc.

D. Limitations, clinical perspectives and future works

Due to technical limitations of fast MR acquisition sequences, extensive 3D volume imaging on mobile organs is hard to achieve. The proposed technique has thus been evaluated in the 2D case. An extension of the method toward 3D motion correction may also be considered in future works and may contribute to better correct the motion in the third dimension by reducing out-of-plane motion artifacts. However, although the proposed algorithm can be easily extended to 3D, the main challenge would likely remain in the design of a reliable 3D MR-sequence. In such acquisition, the scan time increase would render the sequence more sensitive to intra-scan motion and fat/blood signal regrowth. Echo volumar imaging [34] may help to decrease the scan time of such 3D acquisition, however, its associated low resolution and its robustness against the latter artifacts should be carefully investigated.

A robust formulation of equation 5 using robust metrics such as Humber function or Lagrangian [35], especially in the first two terms of equation 5 should be investigated in order to account for both brightness variation and smoothness violation (motion discontinuities). Also, the integration of a median filter at each iteration step of the iterative minimization scheme has to be carefully investigated since significant improvements have been achieved in this direction [36]. In addition, the integration of additional regularization terms such as in [17] should be investigated. Although these approaches may improve the motion estimation quality, the balance between robustness, simplicity (with regard to the number of free parameters) and performance has to be carefully investigated in the perspective of a clinical use. The presented framework also opens great perspectives for integration of other motion information such as navigator or ultrasonic echoes.

The method shows a promising potential for clinical integration for two reasons. First, from its simplicity it only requires a small intervention of the staff physician (only for the mask drawing that only requires few seconds $\sim$5s) and few free parameters to be tuned a priori. Then, the employed values of these parameters were always identical for both heart and kidney studies demonstrating the non-necessity of a re-calibration of the parameters for each sequence and the robustness against the choice of the parameters.

Finally, although the feasibility of the method has been shown on healthy volunteers, its feasibility on other organs such as the liver or on patients with, for example, irregular cardiac motion remains to be investigated. In addition, the method has been tested in conditions of a non invasive procedure, its feasibility in the presence of invasive devices, such as a catheter, will have to be evaluated in future studies.

V. Conclusion

In this paper a new regularization constraint of the energy functional of the H&S method was presented. This approach represents a flexible solution to integrate constraint point displacements into the optical flow estimation. This extension has been demonstrated to render optical flow methods well suited to accurately estimate the motion for interventional MRI on mobile organs in presence of intrusive structures. Significant improvements were achieved compared to the Horn & Schunck approach. Finally, the use of parallel processing on affordable commodity graphics hardware demonstrates the feasibility of the algorithm in real time with very short latency required for interventional procedures.

References

[1] R. Edelman, J. Hesselink, M. Zlatkin, and J. Crues, *Clinical magnetic resonance imaging*. WB Saunders Philadelphia, 1996, vol. 2.

[2] V. Rieke and K. Butts Pauly, "Mr thermometry," *Journal of Magnetic Resonance Imaging*, vol. 27, no. 2, pp. 376–390, 2008.

[3] S. Sapareto and W. Dewey, "Thermal dose determination in cancer therapy," *Int. J. Rad. Onc. Biol. Phys.*, vol. 10, pp. 787–800, 1984.

[4] A. Haase, "Localization of unaffected spins in NMR imaging and spectroscopy (LOCUS spectroscopy)," *Magn Res Med*, vol. 3, no. 6, pp. 963–969, 1986.

[5] B. Wilm, J. Svensson, A. Henning, K. Pruessmann, P. Boesiger, and S. Kollias, "Reduced field-of-view MRI using outer volume suppression for spinal cord diffusion imaging," *Magn Res Med*, vol. 57, no. 3, pp. 625–630, 2007.

[6] T. Schaeffter, V. Raschea, P. Bornerta, and G. Mensb, "Interactive reduced FOV imaging for projection reconstruction and spiral acquisition," *Magn Res Imaging*, vol. 19, no. 5, pp. 677–684, 2001.

[7] J. Maintz and M. Viergever, "A survey of medical image registration," *Med Image Anal*, vol. 2, no. 1, pp. 1–36, 1998.

[8] T. Makela, P. Clarysse, O. Sipila, N. Pauna, Q. Pham, T. Katila, and I. Magnin, "A review of cardiac image registration methods," *IEEE Trans Med Imaging*, vol. 21, no. 9, pp. 1011–21, 2002.

[9] J. Barron, D. Fleet, and S. Beauchemin, "Performance of optical flow techniques," *International Journal of Computer Vision*, vol. 12, no. 1, pp. 43–77, Feb 1992.

[10] G. Zientara, P. Saiviroonporn, P. Morrison, M. Fried, S. Hushek, R. Kikinis, and F. Jolesz, "MRI monitoring of laser ablation using optical flow," *Journal of Magnetic Resonance Imaging*, vol. 8, no. 6, pp. 1306–1318, 1998.

[11] S. Roujol, M. Ries, B. Quesson, C. Moonen, and B. Denis de Senneville, "Real-time MR-thermometry and dosimetry for interventional guidance on abdominal organs," *Magn Res Med*, vol. 63, no. 4, pp. 1080–1087, 2010.

[12] B. Horn and B. Schunck, "Determining optical flow," *Artificial intelligence*, vol. 17, pp. 185–203, 1981.

[13] P. Cachier, J. Mangin, X. Pennec, D. Rivière, D. Papadopoulos-Orfanos, J. Régis, and N. Ayache, "Multisubject non-rigid registration of brain mri using intensity and geometric features," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2001*. Springer, 2001, pp. 734–742.

[14] A. Azar, C. Xu, X. Pennec, and N. Ayache, "An interactive hybrid non-rigid registration framework for 3d medical images," in *Biomedical Imaging: Nano to Macro, 2006. 3rd IEEE International Symposium on*. IEEE, 2006, pp. 824–827.

[15] S. Loncarič and T. Macan, "Point constrained optical flow for LV motion detection," in *SPIE Proceedings of Medical Imaging*, vol. 3978, 2000, pp. 521–529.

[16] A. Becciu, H. van Assen, L. Florack, S. Kozerke, V. Roode, and B. ter Haar Romeny, "A multi-scale feature based optic flow method for 3D cardiac motion estimation," in *SSVN*, 2009.

[17] T. Brox, C. Bregler, and J. Malik, "Large displacement optical flow," *CVPR*, vol. 0, pp. 41–48, 2009.

[18] A. Sotiras, Y. Ou, B. Glocker, C. Davatzikos, and N. Paragios, "Simultaneous geometric-iconic registration," *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, pp. 676–683, 2010.

[19] S. Roujol, J. Benois-Pineau, B. Denis de Senneville, B. Quesson, M. Ries, and C. Moonen, "Real time constrained motion estimation for ECG-gated cardiac MRI," in *ICIP*, 2010, pp. 757–760.

[20] K. Mikolajczyk and C. Schmid, "Scale and affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 60, no. 1, pp. 63–86, 2004.

[21] C. Harris and M. Stephens, "A combined corner and edge detector," in *In Proceedings of The Fourth Alvey Vision Conference*, 1988, pp. 147–151.

[22] A. M. Tekalp, *Digital video processing*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1995.

[23] I. Pratikakis, C. Barillot, P. Hellier, and E. Mémin, "Robust multiscale deformable registration of 3D ultrasound images," *International Journal of Image and Graphics*, vol. 3, no. 4, pp. 547–566, 2003.

[24] S. Roujol, M. Ries, C. Moonen, and S. Denis, "Automatic nonrigid calibration of image registration for real time mr-guided hifu ablations of mobile organs." *IEEE transactions on medical imaging*, vol. 30, no. 10, p. 1737, 2011.

[25] *NVIDIA CUDA Compute Unified Device Architecture - Programming Guide*, 2007.

[26] S. Baker, D. Scharstein, J. Lewis, S. Roth, M. Black, and R. Szeliski, "A database and evaluation methodology for optical flow," *International Journal of Computer Vision*, vol. 92, no. 1, pp. 1–31, 2011.

[27] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Diffeomorphic demons using itks finite difference solver hierarchy," *The Insight Journal*, 2007.

[28] M. Ries, B. Denis de Senneville, S. Roujol, S. Hey, G. Maclair, M. Kohler, B. Quesson, and C. Moonen, "Three dimensional motion compensation for real-time MRI guided focused ultrasound treatment of abdomainal organs," in *ISTU*, vol. 1215, 2010, pp. 239–42.

[29] K. Pruessmann, M. Weiger, M. Scheidegger, and P. Boesiger, "SENSE: Sensitivity encoding for fast MRI," *Magn Res Med*, vol. 42, no. 5, pp. 952–62, Nov 1999.

[30] L. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[31] B. Denis de Senneville, C. Mougenot, and C. Moonen, "Real time adaptive methods for treatment of mobile organs by MRI controlled high intensity focused ultrasound," *Magn Res Med*, vol. 57, no. 2, pp. 319–30, 2007.

[32] M. Ries, B. Denis de Senneville, S. Roujol, and C. Moonen, "Latency compensation for real-time 3D HIFU beam-steering on moving targets," in *ISMRM*, 2010.

[33] B. Denis de Senneville, K. Noe, M. Ries, M. Pedersen, C. Moonen, and T. Sorensen, "An optimised multi-baseline approach for on-line MR-temperature monitoring on commodity graphics hardware," 2008, pp. 1513–1516.

[34] C. Rabrait, P. Ciuciu, A. Ribes, C. Poupon, P. Le Roux, G. Dehaine-Lambertz, D. Le Bihan, and F. Lethimonnier, "High temporal resolution functional mri using parallel echo volumar imaging," *Journal of Magnetic Resonance Imaging*, vol. 27, no. 4, pp. 744–753, 2008.

[35] M. Black and P. Anandan, "A framework for the robust estimation of optical flow," in *Computer Vision, 1993. Proceedings., Fourth International Conference on*. IEEE, 1993, pp. 231–236.

[36] D. Sun, S. Roth, and M. Black, "Secrets of optical flow estimation and their principles," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 2432–2439.