

An Adaptive Non-Local-Means Filter for Real-Time MR-Thermometry

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

▶ **To cite this version:**

Cornel Zachiu, Mario Ries, Chrit Moonen, Baudouin Denis de Senneville. An Adaptive Non-Local-Means Filter for Real-Time MR-Thermometry. IEEE Transactions on Medical Imaging, 2017, 36 (4), pp.904-916. 10.1109/TMI.2016.2627221 . hal-01461332

HAL Id: hal-01461332
https://hal.science/hal-01461332

Submitted on 21 Dec 2018

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 Adaptive Non-Local-Means Filter for Real-Time MR-Thermometry

Cornel Zachiu, Mario Ries, Chrit Moonen, and Baudouin Denis de Senneville

***Abstract*—Proton resonance frequency shift-based magnetic resonance thermometry is a currently used technique for monitoring temperature during targeted thermal therapies. However, in order to provide temperature updates with very short latency times, fast MR acquisition schemes are usually employed, which in turn might lead to noisy temperature measurements. This will, in general, have a direct impact on therapy control and endpoint detection. In the current study we address this problem through an improved non-local filtering technique applied on the temperature images. Compared to previous non-local filtering methods, the proposed approach takes into account not only spatial information, but also exploits temporal redundancies. The method is fully automatic and designed to improve the precision of the temperature measurements while at the same time maintaining output accuracy. Additionally, the implementation was optimized in order to ensure real-time availability of the temperature measurements while having a minimal impact on latency. The method was validated in three complementary experiments: a simulation, an *ex-vivo* and an *in-vivo* study. Compared to the original non-local means filter and two other previously employed temperature filtering methods, the proposed approach shows considerable improvement in both accuracy and precision of the filtered data. Together with the low computational demands of the numerical scheme, the proposed filtering technique shows great potential for improving temperature measurements during real-time MR thermometry dedicated to targeted thermal therapies.**

***Index Terms*—Image denoising, MR-thermometry, Real-time system.**

I. Introduction

Magnetic resonance (MR) thermometry using proton resonance frequency shifts (PRFS) is one of the currently used mechanisms for temperature monitoring during thermal ablations of pathologies inside the human body [1]–[3]. However, such interventions typically require a good spatial coverage of the heated region and/or temperature updates with a high temporal resolution [4], [5]. For this reason, fast MR-imaging sequences are in general necessary, which might lead to an acquisition having a low signal-to-noise ratio (SNR), resulting in noisy temperature measurements [6], [7]. This is expected to have a direct impact on therapy control and endpoint detection, which might lead to the over or under-treatment of the pathology. In the current study we address this issue through a non-local image filtering technique as a means to improve the SNR of the PRFS-based temperature measurements.

Cornel Zachiu, Mario Ries, Baudouin Denis de Senneville and Chrit Moonen are with the Imaging Division, UMC Utrecht, Utrecht, Netherlands, (e-mail: {c.zachiu,m.ries,c.moonen,b.desenneville}@umcutrecht.nl).

Baudouin Denis de Senneville is also with the "Institut de Mathématiques de Bordeaux" (IMB), UMR 5251 CNRS/University of Bordeaux, F-33400 Talence, France.

Manuscript received ...; revised ...

Image denoising through non-local means (NLM) was originally proposed by Buades *et al.* in the context of SNR restoration for digital images and video sequences altered by white Gaussian noise [8]. A pixel denoised with the NLM method is basically a weighted average of all the pixels in the noisy image. The weights themselves are proportional to the gray-level intensity similarity between the local neighborhoods of the pixel under study and the pixel contributing to the average: the more similar the two neighborhoods are, the larger the weight, hence the greater the impact a pixel has on the resulting denoised value. Since the most important contributions in denoising a pixel are provided by pixels that are most similar to it, the NLM filter is more prone to preserve high-frequency image content such as edges and/or small structures, compared to other denoising techniques [8]. However, the fact that the NLM method is based on spatial similarities is also the source of one of its major drawbacks. In case a particular structure/pixel lacks self-similarity in an image, it might end-up being insufficiently/poorly denoised. Such is the case, for example, of the focal point and its surroundings in the temperature images acquired during targeted thermal ablations. The focal point in particular has a high chance of being unique in the image. In its attempt to denoise the temperature at the focal point, the NLM filter, due to its functioning principle, will perform a weighted average between dissimilar pixels, which might result in a blurred/biased measurement. In addition, since the filter performs on the quantification of the image (i.e the color domain) and not on its sampling support (i.e the pixel domain), the filtering results fully depend on the image content. Therefore, contrary to other existing filtering approaches (such as convolution-based or (in)finite-impulse response techniques), a theoretical characterization of the overall filtered results becomes, from a mathematical point of view, a challenging task. It is thus difficult to evaluate/estimate the inherent resulting bias on the filtered data. Although this consideration is of minor interest for digital photography or video sequences, it becomes of major importance in the field of medical imaging, for which both the accuracy and precision of the quantitative image information are important.

Previous studies have addressed the issue of filtering PRFS-thermometry images through so-called predictive methods. For example, Roujol *et al.* [9] proposed a temperature filtering/prediction technique relying on the model-based Kalman filter [10]. The output of the filter was a linear combination of the measured noisy data and a temperature evolution model (the bioheat transfer equation (BHTE)). While showing great filtering capabilities for high-frame-rate PRFS thermometry ($>$10 images/s), the performance of the method is expected to deteriorate as the temperature update frequency decreases.

Nevertheless, an NLM filter should be seen as complementary to such an approach rather than a replacement. Depending on the application, one might out-perform the other.
In this work we propose an improved NLM filtering method for dynamic PRFS thermometry during targeted thermal therapies. Performance analysis of the method was conducted using simulations, *ex-vivo* data acquired during a radiofrequency (RF) thermal ablation experiment on a calf liver and *in-vivo* data acquired on a porcine kidney during an MR-guided high intensity focused ultrasound (MRg-HIFU) ablative sonication. The contribution of this study is fourfold:

1) We propose denoising PRFS temperature maps using a variation of the NLM filter which during the restoration process takes into account not only spatial information, but also exploits temporal redundancies. More precisely, each new incoming temperature map is also filtered based on information available from previous measurements. This is expected to improve the resulting filtered data compared to the original approach, especially in the focal point and its surroundings, since more reliable temperature information in this area is available in the temporal dimension rather than in the spatial one.
2) The proposed spatio-temporal NLM filter is further enhanced by dynamically adapting its filtration strength, through an optimization process, such that the accuracy of the resulting temperature measurements is maximized (i.e. the blur/bias on the filtered temperature is minimized).
3) The performance of the NLM filter can usually be improved if prior knowledge on the noise distribution in the image is available [8], [11], [12]. Therefore, in the current study, we propose a fully automatic method that estimates 'on-the-fly' the amount of noise altering the temperature measurements.
4) An implementation with real-time capabilities was performed for the proposed methods in order to provide improved temperature measurements while having a minimal impact on latency. For a smooth MR temperature monitoring work-flow, the computational time of the methods has to lie well beneath the time interval between two temperature measurements.

II. Method Description

Initially, section II-A provides a brief mathematical description of the basic NLM filter together with the manner in which it can be applied for denoising MR thermal maps. Section II-B details the proposed improvements to the NLM filter which are intended to compensate for the shortcomings of the original version in the context of dynamic MR-thermometry. The section also includes a description of the proposed method which allows dynamic characterization of the noise distribution in the temperature measurements. Finally, section II-C describes the protocol used to evaluate the performance of the proposed variations to the original NLM filter and that of the noise characterization method, together with implementation details of the proposed methods.

A. The non-local means filter applied to thermal maps

The NLM filter proposed in this study used as a starting point the formulation suggested by Coupé *et al.* in [11]. Compared to the original version of Buades [8], Coupé *et al.* proposed an improved approach in which the filtration strength is computed in a deterministic way rather than empirically. While their study focused on 3D diagnostic brain images, the current technical limitations of MR systems and the computational time of NLM filters, notorious for their high processing latency, make real-time 3D PRFS-thermometry paired with a non-local filtering method a challenging task. For this reason, the current study was conducted solely for 2D temperature maps. According to the formulation proposed by Coupé *et al.*, the filtered temperature $T_f(\vec{r_i})$ at pixel $\vec{r_i} = (x_i, y_i)$ is given by the weighted average of all the noisy measurements in the temperature image T_m (principle which is graphically summarized in Fig. 1):

$$T_f\left(\vec{r_i}\right) = \sum_{j \in \Omega} w\left(\vec{r_i}, \vec{r_j}\right) T_m\left(\vec{r_j}\right) \tag{1}$$

with:

$$w\left(\vec{r_i}, \vec{r_j}\right) = \frac{1}{Z_i} e^{-\frac{\|T_m(N_i) - T_m(N_j)\|^2}{2\beta\hat{\sigma}(\vec{r_i})^2 |N_i|}} \tag{2}$$

where Ω is the image grid, Z_i is a normalization constant ensuring that $\sum_j w\left(\vec{r_i}, \vec{r_j}\right) = 1$, $|N_i|$ is the size of a local neighborhood around $\vec{r_i}$, $\hat{\sigma}(\vec{r_i})$ is an estimation of the local standard deviation of the noise and β is a constant that needs to be manually tuned. As mentioned in [11], for Gaussian noise, if the estimation of $\hat{\sigma}$ is correct, the value of β is theoretically close to 1. However, the automatic dynamic adjustment of β will be the keystone of the adaptive NLM strategy proposed in this study. Additionally, note that, while in theory the filtered value is a weighted average of all the noisy measurements in the image, in practice the contributing pixels are limited to a smaller window around the filtered pixel.

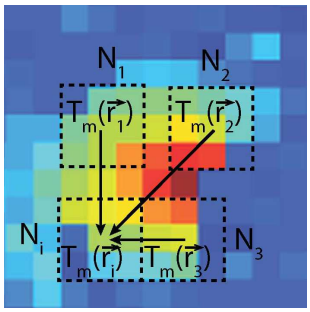


Fig. 1: Noisy temperature map acquired during a targeted thermal therapy upon which the principle of the NLM filter is illustrated: The denoised value of the pixel at $\vec{r_i}$ is the weighted average of the pixels at $\vec{r_j}$, with the weights computed based on the similarity between their neighborhoods N_i and N_j. For simplicity, the process is illustrated only for three values of j, rather than the entire image domain Ω.

B. Improvements of the non-local means filter

1) Extension of the NLM filter to dynamic MRI: Instead of relying solely on spatial information in order to filter the noisy temperature maps, in this work we propose an extension of the NLM technique also to include the temporal dimension. In effect, a filtered pixel is the weighted average of the noisy measurements not only from the current dynamic, but also from the previous N dynamics. This results in what we have called the NLM-(2D+t) filter, having the following mathematical formulation:

$$T_f\left(\vec{r_i},t\right)=\sum_{n=0}^{N-1}\sum_{j\in\Omega}w\left(\vec{r_i},\vec{r_j},t-n\right)T_m\left(\vec{r_j},t-n\right) \tag{3}$$

with:

$$w\left(\vec{r_i},\vec{r_j},t-n\right)=\frac{1}{Z_i}e^{-\frac{\|T_m(N_i,t)-T_m(N_j,t-n)\|^2}{2\beta\hat{\sigma}(\vec{r_i})^2|N_i|}} \tag{4}$$

where t is the temporal index of the current dynamic.

2) Adaptive NLM strategy: Due to its underlying principle (i.e. averaging in the color domain), the NLM filter will unavoidably introduce a certain amount of blurring in the filtered data, which might lead to inaccurate temperature measurements. The filtering strength of the NLM is mainly controlled by the denominator in the exponential defined by Eq. (2). In particular, the β factor plays the most important role in this regard, since the other parameters are usually fixed. For example, while $\beta=0$ implies that no filtering is performed, a value of β equal to 1 results in filtering performed assuming a standard deviation equal to σ. In this work, we propose an optimization scheme for β such that for every pixel of each new incoming temperature map, the filter ensures a maximal output accuracy.

Filter accuracy was evaluated on a pixel-by-pixel basis by computing, for each new incoming image, the temperature bias $\epsilon\left(\vec{r}\right)$ over a temporal window. At instant t, the bias at pixel of coordinates $\vec{r}$ was calculated as follows:

$$\epsilon\left(\vec{r}\right)=\frac{1}{N}\sum_{n=0}^{N-1}\left[T_m(\vec{r},t-n)-T_f(\vec{r},t-n)\right] \tag{5}$$

where $T_m(\vec{r},t)$ is the measured (noisy) temperature at location $\vec{r}$ and instant t and $T_f(\vec{r})$ is the filtered output. For ideally filtered thermal maps, $\epsilon\left(\vec{r}\right)$ is a Gaussian centered random variable of standard deviation:

$$\sigma_e(\vec{r})=\frac{1}{N}\sqrt{\sum_{n=0}^{N-1}\hat{\sigma}(\vec{r},t-n)^2} \tag{6}$$

In such an ideal case, 95% of the realizations of $\epsilon\left(\vec{r}\right)$ comply with the following inequality:

$$\left|\epsilon\left(\vec{r}\right)\right|\leq 2\sigma_e(\vec{r}) \tag{7}$$

The value $2\sigma_e(\vec{r})$ could thus be taken as an upper bound for $\left|\epsilon\left(\vec{r}\right)\right|$ in order to warrant the output accuracy. For each temperature map, an exhaustive enumeration of β was performed between 0 and 1 with an increment of $\frac{1}{10}$. The resulting filtered data and the associated error maps $\epsilon\left(\vec{r}\right)$ were stored in a map data structure, for which the entry key was the corresponding tested β value. The optimal value(s) for $\beta\left(\vec{r}\right)$ was(were) obtained for each pixel individually as follows:

$$\hat{\beta}\left(\vec{r}\right)=\underset{\{\beta,|\epsilon(\vec{r})|\leq 2\sigma_e(\vec{r})\}}{\arg\max}\left|\epsilon\left(\vec{r}\right)\right| \tag{8}$$

The final filtered temperature map was reconstructed on a pixel-by-pixel basis using the previously stored data corresponding to the values of $\hat{\beta}\left(\vec{r}\right)$. Intuitively, the proposed adaptive NLM method can be summarized as follows: Search for the value of $\beta\left(\vec{r}\right)$ which leads to a maximal amount of filtering (denoted by the "$\arg\max$" in Eq. (8)), while minimizing accuracy losses (imposed by the constraint illustrated in Eq. (7)). For the remainder of the manuscript we will refer to this filter as the adaptive NLM-(2D+t), abbreviated as ANLM-(2D+t).

3) Dynamic characterization of the temperature noise distribution: As illustrated by Eq. (2), the NLM filter requires prior knowledge of the noise distribution in the image. Please note that for a zero-mean Gaussian noise, as assumed in the current study, the noise distribution is fully characterized by its standard deviation. For this reason, throughout this manuscript we will use the terms of noise distribution and standard deviation interchangeably. A number of methods have already been proposed to estimate the noise distribution directly from the image itself. However, for the particular case of temperature images, the estimation process might be hampered by several factors such as pixels with low magnitude/signal or phase artifacts due to susceptibility/temporal drift of the magnetic field [13]–[15]. Since the magnitude component of the MR-signal is less prone to the such artifacts, we decided to use the latter to estimate the noise distribution in the temperature maps. This was achieved by first estimating the standard deviation of the noise in the phase image by making use of the formula proposed by Conturo in [13]:

$$\sigma(\phi(\vec{r}))=\frac{\sigma(M)}{|M(\vec{r})|} \tag{9}$$

where ϕ and M are the phase and the magnitude of the complex MR-signal and $\vec{r}$ is the spatial location. Based on the noise distribution estimated in the phase image, the standard deviation of the noise in the temperature maps was then computed based on the following relationship between phase and temperature:

$$\sigma(T(\vec{r}))=\sqrt{2}\frac{\sigma(\phi(\vec{r}))}{k} \tag{10}$$

where k is the PRFS-thermometry constant [3].

The standard deviation of the noise in the magnitude image can be estimated through several methods, depending on whether the noise is assumed to be homogenous [11], [16] or spatially variant [12]. In the current study, since the noise was assumed to be homogeneously distributed over the field of view, the standard deviation of the noise in the magnitude image was estimated via the wavelet-based approach proposed in [16]:

$$\sigma_{\text{est}}(M)=\frac{\text{median}|y_i|}{0.6745} \tag{11}$$

where $\sigma_{\text{est}}(M)$ is the estimated value of the noise standard deviation in the magnitude image and y_i are the high frequency

coefficients of the first level wavelet decomposition of the image. The method was chosen due to the good compromise between computational time versus quality of the estimation it has shown during our experiments. Note that since the standard deviation of the noise in the magnitude image is not expected to vary significantly over time, the estimation in Eq. (11) can be carried-out only for one image in the thermometry sequence.

For the rest of the manuscript, unless specified otherwise, σ refers to the standard deviation of the noise in the temperature maps.

C. Experimental setup

1) Performance assessment of the proposed filtering methods: The following filtering techniques were compared in the context of dynamic MR-thermometry during a targeted thermal therapy:

- An infinite-impulse-response (IIR) temporal low-pass Cauer filter of the 5^{th} order. The pass-band and stop-band were adjusted to result in an overall temporal resolution divided by a factor 2.
- A 2D spatial Gaussian filter (kernel size equal to 3×3, σ=0.5).
- The 2D spatial NLM filter proposed by Coupé *et al.* in [11]. For the remainder of the manuscript we will refer to this filter as the NLM-2D.
- The proposed NLM-(2D+t) filter.
- The proposed ANLM-(2D+t) filter.

The first two methods were selected as they represent two previously employed strategies for real-time dynamic MRI-thermometry [17]–[19]. It is expected that the two will provide a similar SNR improvement by a factor $\sqrt{2}$, since each of them imply a reduction of the resolution by a factor 2, one in the temporal and the other in the spatial domain. The other 3 methods have the purpose to evidence the improvements provided by the proposed NLM filtering methods compared to the existing one. For all the NLM filters, the patch sizes $|N_i|$ and $|N_j|$ were chosen equal to 25 (5×5). Additionally, each pixel was filtered using the contribution of the noisy measurements situated within a 7×7 encompassing window. The size of the latter and also the patch size were chosen as to provide the best compromise between quality of the filtered data and computational time. Following the same reasoning, the temporal window N (see Eq. (3)) was chosen equal to 6. Since the main concern was output accuracy, the performance of the tested filters was assessed quantitatively by evaluating the time averaged absolute difference between the output of the filters and a gold standard. The construction of the latter will be further detailed in future sections. The performance assessment was conducted in two different regions:

- In a region of interest covering 7×7 pixels, centered on the focal point position, over the entire duration of the thermal energy deposition. In this manner, the assessment is performed within an area excluding background pixels which are not of great interest.
- At the focal point itself (i.e in a single pixel located at the focal point position).

The performance of the filters was evaluated in three complementary studies:

a) Simulation study: Temperature maps resulting from a HIFU energy delivery were simulated using the BHTE model. The simulation extended over a duration of 100 s and 250 W of acoustic power was applied in a single point during the [10 s; 60 s] time interval. The spatial and temporal evolution of the temperature resulting from the acoustic energy delivery was given by the solution of the BHTE, solved using finite elements on a $1 \times 1\mathrm{mm}^2$ lattice and with a temporal resolution of 1 s. The simulated temperature maps were subsequently corrupted by additive white Gaussian noise with increasing standard deviation. The latter was varied between 0°C and 4°C with a regular sampling step of 0.5°C. The noisy temperature maps were then filtered using the approaches enumerated in section II-C1 and their performance was assessed based on the metric described therein. The gold standard was the simulated, unaltered temperature maps. Note that the assessment was averaged over 10 experiments with different realizations of noise, in order to reduce random impact on the overall results.

b) Ex-vivo study: A heating *ex-vivo* experiment was conducted on a calf liver. The heating itself was the result of delivering 20 W of RF power via a bipolar electrode, for a duration of ∼50 s. The energy delivery was carried-out under MR-guidance with MR-thermometry being performed before, during and after heating the tissue. The MR-acquisition was performed via a dual-shot gradient recalled echo-planar imaging (EPI) sequence with TE = 13 ms, TR = 70 ms, N_{slices} = 1, image size 128×128 and a 1.5×1.5×5 mm^3 voxel size, resulting in an image being acquired every 140 ms (∼7.15 Hz). The intense and quick heating induced by the RF energy delivery lead to notable local gray-level intensity drops in the magnitude image in the close vicinity of the heated region.

The purpose of this particular experiment was to analyze the impact of a noise distribution in the temperature maps with a spatio-temporally varying standard deviation. When heating tissues, relaxation times get longer and the magnitude of the MR-signal changes. According to Eq. (9) and (10) this leads to local spatio-temporal variations in the standard deviation of the noise in the MR-temperature maps. This effect might end-up having an impact on the quality of the filtered data provided by the proposed methods.

A high-SNR dataset was derived from the original sequence by a centered sliding temporal averaging applied on both the magnitude and the phase components of the MR signal, followed by a temporal subsampling, such that the output data had an equivalent temporal frequency of ∼1 Hz. This operation lead to an average SNR of the magnitude image of more than 40 (SNR computation was restricted to the liver). The MR temperature maps were then recalculated based on the averaged phase images and set as a goldstandard for the current experiment. In order to evaluate the performance of the compared denoising methods in this particular experiment, a noisy dataset was constructed based on the temporally averaged magnitude and phase images. The latter were used to recalculate the real and imaginary channels of the MR signal, upon which a Gaussian noise with a standard deviation of 10% of the maximum intensity in the magnitude image was

then added. The noisy real and imaginary parts were then recombined to compute the noisy magnitude and phase images, after which a set of noisy temperature maps was constructed, based on the noisy phase images. The denoising methods were then applied on the resulting temperature maps and the filtered data was compared in terms of accuracy against the goldstandard.

c) In-vivo study: An MRg-HIFU shot was performed *in-vivo* on a porcine kidney, which consisted in the delivery of 250 W of acoustic power over a duration of 50 s. The animal was under general anesthesia, with the targeted area only exhibiting minimal motion. Thus no correction of motion related temperature errors was required. The MR-thermometry acquisition protocol consisted of a multi-slice image, with each slice being acquired via a slice-selective single-shot echo planar imaging sequence employing the following parameters: TE=25 ms, TR=1143 ms, flip angle=35°, FOV=142.5×285 mm^2, bandwidth$_{\text{read}}$ = 2078 Hz and voxel size=$3 \times 3 \times 6\text{mm}^3$. Each dynamic included a total of 9 coronal slices, symmetrically distributed around the focal spot. This would ensure the coverage of the entire kidney and part of the near-field, in order to detect any secondary heating spots. A total of 175 dynamics were acquired over a duration of 200 s, providing temperature information during pre-sonication, the energy delivery itself and respectively post-sonication. The temperature measurements provided by the middle slice (which included the primary focal spot) were filtered using the approaches enumerated in section II-C1 and their performances were compared based on the metric described therein.

The gold standard was obtained as follows: A pixelwise temporal centered sliding average was performed on the temperature data provided by the middle slice with a span extending over six temperature maps followed by a 6-fold temporal subsampling. This increased the SNR by a factor 2.45 and resulted in a dynamic sequence with a temporal resolution equal to 0.15 Hz. Note that the original noisy image sequence was also temporally subsampled such that it's frequency matches the one of the gold standard. Subsequently, a Mann-Whitney test was carried out in order to study whether the differences between the noisy and the filtered data are statistically significant. A significance threshold of $p = 0.05$ was used.

The animal experiment was conducted in agreement with the European law on animal experimentation and in compliance with the institution's rules for animal care and use.

2) Performance analysis of the noise estimation algorithm: A high-SNR image was initially computed by temporally averaging all the corresponding magnitude images acquired with the protocol described in section II-C1c. After normalizing the resulting image to the $[0, 1]$ interval, increasing levels of Gaussian noise were added to it, progressively lowering its SNR. The standard deviation of the added noise was varied between 0.005 and 0.06 with an increment of 0.005. Note that, according to Eq. (9) and (10), this leads to a standard deviation in the temperature maps that varies between $\sim 0.5°C$ and $\sim 4.5°C$. For each noise level, the standard deviation was then estimated in the magnitude images using the algorithm specified by section II-B3. The estimated value was then compared to the true standard deviation in terms of the percentage relative error (denoted by δ_σ):

$$\delta_\sigma = 100 \times \left| 1 - \frac{\hat{\sigma}}{\sigma} \right| \tag{12}$$

where $\hat{\sigma}$ and σ are the estimated and respectively the true value of the noise standard deviation added to the magnitude images. Subsequently, the percentage relative estimation errors versus the SNR of the image was analyzed. Note that the estimation errors were averaged over 1000 realizations of each noise level. Note that assessment and processing was performed exclusively for the slice containing the primary focal spot.

3) Implementation: A C++ multi-threaded implementation (8 threads) was performed for all filtering methods enumerated in section II-C1 on an Intel 3.2 GHz i7 workstation (8 cores) with 16 GB of RAM. In the scope of this study, the filtering process was restricted to a region of interest around the focal point. The size of the region was chosen in accordance with the spatial extent of the heating. This was decided both for computational time considerations and, in the scope of this study, for the lack of interest in the pixels outside the heated area.

III. Results

Section III-A illustrates the performance during the simulation study (see section II-C1a for details) of the filtering methods enumerated in II-C1 including: an example of the spatial distribution of the accuracy errors introduced by each of the methods, the performance of the filters for increasing noise levels and the sensitivity of the proposed ANLM-(2D+t) method to errors of the noise estimation algorithm. Section III-B displays the results obtained during the *ex-vivo* experiment described in section II-C1b. Of interest here were the accuracy errors introduced by the compared methods in the presence of heating-induced spatio-temporal SNR variations in the MR magnitude images. Finally, section III-C illustrates the performance of the filters for data obtained during the *in-vivo* experiment (detailed in section II-C1c) together with an assessment of the noise estimation algorithm (see sections II-B3 and II-C2) and the computational requirements of the proposed methods.

A. Simulation study

1) Performance assessment of the proposed NLM-(2D+t) and ANLM-(2D+t) filtering methods: Fig. 2 illustrates the temperature bias introduced by the compared filtering methods (enumerated in section II-C1), aiming to reflect the accuracy of the resulting filtered data. Fig. 2(a) shows the temperature spatial distribution 40 s into the heating simulation detailed in section II-C1a. This played the role of gold standard in comparing the performance of the filtering methods. Fig. 2(b) displays the gold standard temperature map altered by a Gaussian noise with a 2°C standard deviation, while Fig. 2(c)-(g) display the filtered temperature measurements provided by each approach. The resulting temporally averaged absolute errors are illustrated in Fig.2(h)-(m). From a visual inspection it already becomes apparent that, in the focal point and its

close proximity, the NLM-(2D+t) and respectively the ANLM-(2D+t) methods provide the lowest errors, observation that will be further confirmed by subsequent results. Fig. 3 and 4 compare in terms of accuracy the proposed ANLM-(2D+t) against the IIR and respectively the Gaussian filter. Fig. 3 in particular illustrates the temperature evolution in the focal point over the duration of the simulation. Fig. 3(a) traces the gold standard temperature evolution in the focal point (black curve) overlapped with its noisy version (red curve). The noise standard deviation in this experiment was set to $3°$C. The filtered data together with the gold standard is displayed in Fig. 3(b). It can be observed that the intrinsic temporal lag of the IIR filter, which becomes evident especially in the proximity of the starting and end point of the energy delivery, leads to a consistent amount of bias in the filtered data. A consistent amount of bias can also be observed for the Gaussian filter, which by construction has a blurring effect on the filtered measurements. A visual analysis of the traces in Fig. 3(b) would indicate that the ANLM-(2D+t) filter (blue curve) provides the most accurate results. This is confirmed in Fig. 4, where the mean absolute error of the filtered temperature versus the simulated noise standard deviation is plotted. Note that analysis was performed for the same filters as in Fig. 3, but separately for the focal point (Fig. 4(a)) and the area surrounding the focal point (Fig. 4(b)). While the IIR (red line) and the spatial Gaussian (green line) filters manifest similar tendencies in terms of accuracy of the filtered data, the ANLM-(2D+t) outperforms both for all the analyzed noise levels, especially in the focus point for low noise levels. The black line corresponds to the mean absolute error of the noisy data. As a side-note, the comparable performance of the IIR and the spatial Gaussian filter can be explained by their similar approach in improving the SNR of the noisy image. They both reduce the information resolution by a factor of 2 (see section II-C1), one however operates in the temporal while the other in the spatial domain, in order to improve the SNR of the image.

Fig. 5 shows the accuracy of the temperature filtered using the ANLM-(2D+t) (blue line) compared to the original NLM (red line) and to the non-adaptive NLM-(2D+t) (green line). The analysis was conducted for several noise levels in both the focal point (Fig. 5(a)) and the area around the focal point (Fig. 5(b)). Note that the blue curves corresponding to the ANLM-(2D+t) filter are identically the same as the ones as in Fig. 4(a) and 4(b) respectively. Except for very low noise levels ($< 0.5°C$), where all three filters have a similar performance, the original NLM is constantly outperformed by the other two filters. As already anticipated from Fig. 2(l) and 2(m), the NLM-(2D+t) and the ANLM-(2D+t) have a similar performance, particularly if the errors are averaged over the area surrounding the focal point or if the noise levels are low. However, for a noise standard deviation larger than $2°$C, the ANLM-(2D+t) clearly outperforms the NLM-(2D+t) when analysis is conducted in the focal point, especially in the rightmost part of the graph in Fig. 5(a). Note that the black line corresponds to the errors in the noisy data. A noteworthy fact that can be observed from both Fig. 4 and 5 is that the ANLM-(2D+t) filter does not yield higher errors than the noisy data for any of the analyzed noise levels. The rest of the filtering methods, on the other hand, at least in some particular cases, actually end-up altering the temperature measurements even further compared to when no filtering is performed.

2) Robustness of the ANLM-(2D+t) to calibration errors: As illustrated by Eq. (2) and (4), the NLM, NLM-(2D+t) and the ANLM-(2D+t) filters require the noise standard deviation as an input parameter, which has to be estimated (see section II-B3). Fig. 6 displays for the ANLM-(2D+t) in particular its sensitivity to noise estimation errors. The figure illustrates the mean absolute error of the filtered temperature as a function of the percentage relative estimation error of the noise standard deviation. The analysis was conducted in both the focal point (red curve) and the area around the focal point (blue curve) for temperature maps altered by a noise with a standard deviation of $1°$C. It can be observed that an under-estimation of the noise, while the accuracy of the filtered data is still improved, precision is being hampered, indicated by the fact that the difference between the filtered and the noisy temperature becomes less evident. Note, however, that the filter still remains reliable for noise estimation errors of up to 50% in either direction.

B. Ex-vivo study

Fig. 7 displays the performance of the analyzed filtering methods following the MR-guided *ex-vivo* RF thermal ablation study conducted on a calf liver (described in II-C1b). Fig. 7(a) in particular, illustrates one of the temporally averaged high-SNR MR magnitude images, calculated prior to starting the energy delivery (see section II-C1b for details). For better visibility, the treated area (identified by the red dashed square) has been magnified, while at the same time indicating through blue arrows several pixels that were of interest during the *ex-vivo* study. Points (1) and (2), for example, correspond to the approximate position of the tips of the RF-emitting needles, while points (3) and (4) were selected such that the former would be in close proximity to one of the focal spots while the latter is completely outside the heated area. The precise purpose of points (3) and (4) will be detailed in later paragraphs. Fig. 7(b) and 7(c) display the gold standard temperature distribution and its noisy version, 80 s into the experiment. Fig. 7(d)-7(h) showcase the effect each of the filtering methods have on the noisy temperature map from Fig. 7(c), while Fig.7(i)-7(n) illustrate the bias introduced in the filtered data by each of the approaches. Consistent with the observations made during the simulation study is the fact that both the NLM-(2D+t) and the ANLM-(2D+t) outperform the other filtering methods, with the ANLM-(2D+t) performing slightly better than the NLM-(2D+t), especially in the proximity of the focal spots. This is further confirmed by Fig. 8 where the temperature evolution during the *ex-vivo* experiment is traced for point (3) in the absence (Fig. 8(a)) and in the presence (Fig. 8(b)) of a filtering method. It can be observed from the latter that compared to the ANLM-(2D+t), the Gaussian and the IIR filter have indeed a tendency to introduce additional bias in the filtered data.

What sets the *ex-vivo* experiment apart from the simulation

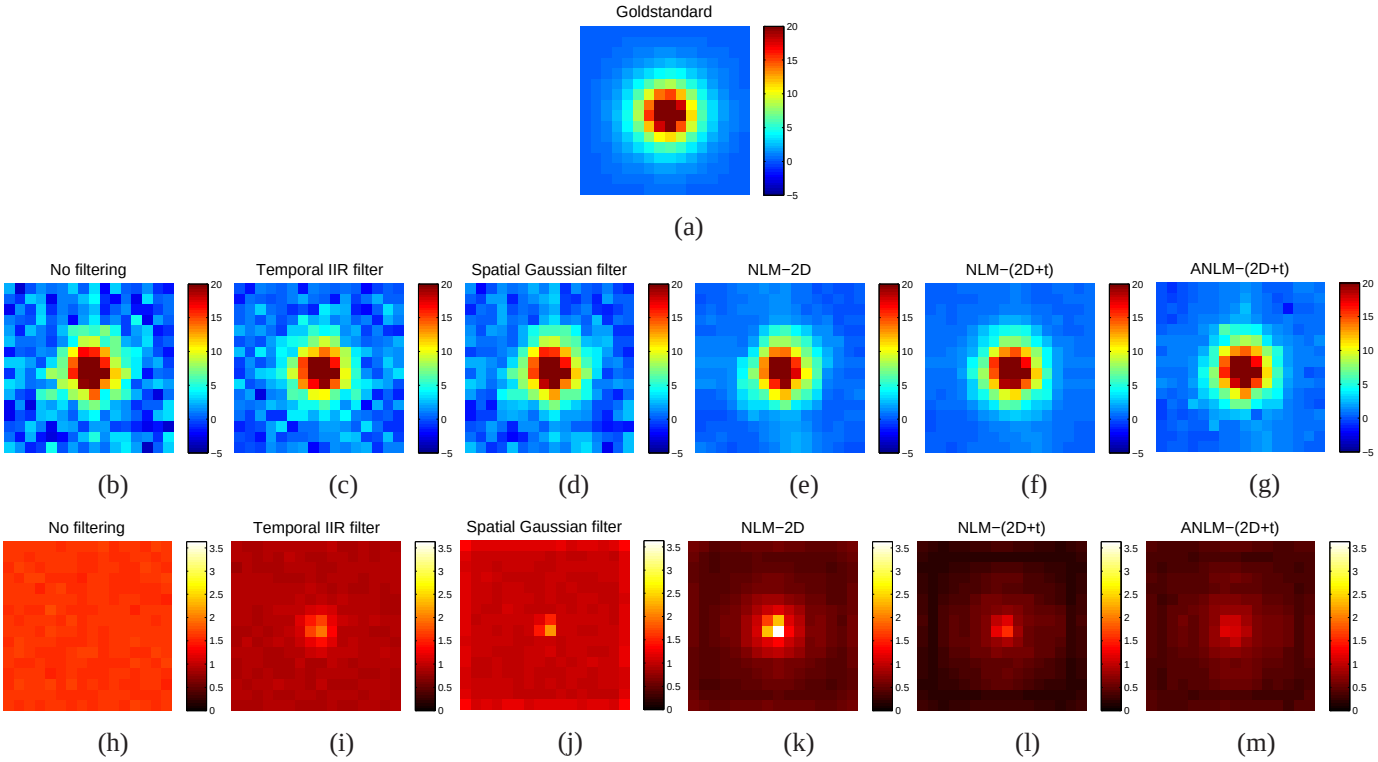


Fig. 2: Example of thermometry results obtained during the simulation study. (a): The gold standard noise-free temperature map; (b): Corresponding noisy map (standard deviation $2°C$); (c)-(g): Filtered temperature maps; (h)-(m): Spatial distribution of the temporally averaged absolute errors for each filtering method.

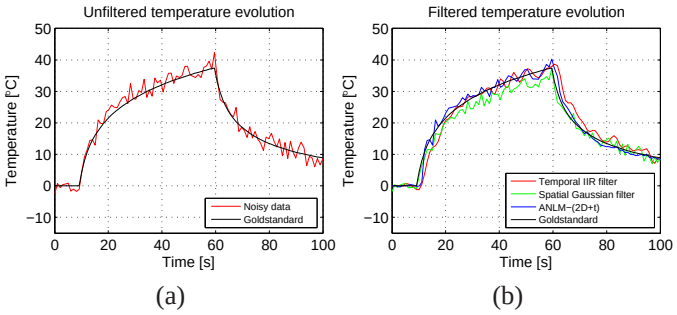


Fig. 3: Temperature time curve obtained in a single pixel located at the focal point. (a): The noisy time temperature evolution (red curve, standard deviation $3°C$). (b): Filtered temperature values obtained using the proposed ANLM(2D+t) method (blue curve), the spatial Gaussian filter (green curve) and the temporal IIR filter (red curve). The gold standard temperature illustrated in (a) and (b) by the black curve.

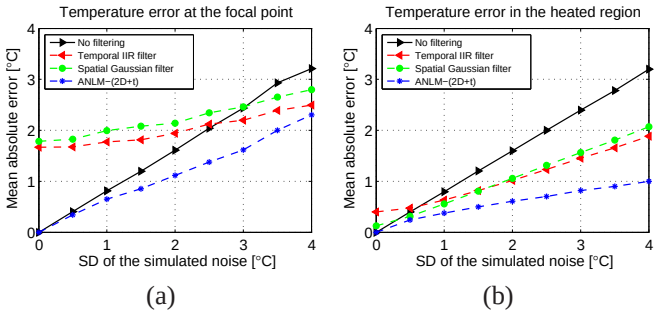


Fig. 4: Accuracy of the ANLM-(2D+t) filter (blue line) versus the IIR (red line) and respectively spatial Gaussian (green line) filter. The two figures illustrate the mean absolute errors as a function of the simulated noise standard deviation (a): In the focal point; (b): In the surrounding area of the focal point.

study is the fact that the temperature maps in the *ex-vivo* experiment are altered by a noise with a spatially variant standard deviation, which was constant during the simulation. The variation of the noise standard deviation stems mainly from the signal drops induced in the MR magnitude images by the heating process, factor which, for the sake of simplicity, was not taken into account during the simulation study. The local signal drops lead to an increase in the temperature standard deviation according to Eq. (9) and (10). Fig. 9 displays the temporal evolution of the gray-level intensity at points (3) and (4) respectively. It can be noted that, as the temperature at point (3) increases, the signal intensity decreases, while at point (4), which was chosen outside the heated area, the signal intensity remains more or less constant. Nevertheless, even in such conditions the proposed filtering methods outperform the existing ones, as shown in Fig. 7 and 8.

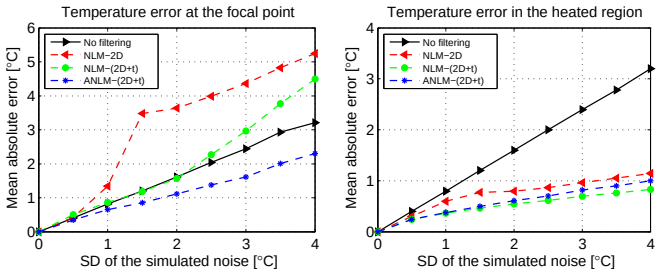


Fig. 5: Accuracy of the ANLM-(2D+t) filter (blue line) versus the original NLM (red line) and the NLM-(2D+t) (green line) filter. The mean absolute temperature errors are reported for various noise levels (a): In the focal point; (b): In the area surrounding the focal point.

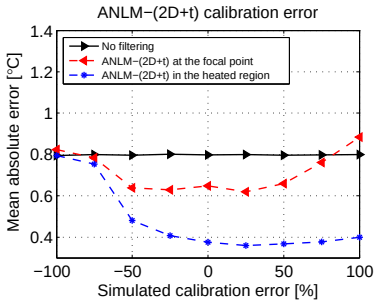


Fig. 6: Robustness of the proposed ANLM-(2D+t) filter against potential calibration errors for a nominal noise standard deviation of $1°C$. The mean absolute temperature errors are reported for various relative noise estimation errors.

C. In-vivo study

1) Performance assessment of the proposed NLM-(2D+t) and ANLM-(2D+t) filtering methods: Fig. 10 illustrates the accuracy of each of the compared filtering methods (see section II-C1) when applied on the *in-vivo* dataset. Details concerning the *in-vivo* experiment can be found in section II-C1c. Fig. 10(a) and 10(b) correspond to the gold standard temperature map and its noisy version, respectively, after 50 s of ultrasonic energy delivery. Fig. 10(c)-(g) display the filtered temperature map using each of the compared methods. The accuracy of the filters is shown in Fig. 10(h)-(m), where the spatial distribution of the temporally averaged absolute errors is illustrated. Just as in the simulation and the *ex-vivo* experiments, a visual analysis of the results indicates that the NLM-(2D+t) and the ANLM-(2D+t) outperform the other approaches. Fig. 11(a) and 11(b) illustrate the temperature evolution in the focal point over the duration of the *in-vivo* experiment. The black and the red curves in Fig. 11(a) correspond to the gold standard and to the noisy temperature measurements, respectively. The filtered values are displayed in Fig. 11(b), where the red, green and blue curves correspond to the IIR, spatial Gaussian and the ANLM-(2D+t) filter, respectively. The black curve again represents the gold standard temperature. Similar to the results provided by the simulation and the *ex-vivo* experiment, the temporal lag of the IIR filter introduces consistent accuracy errors in the filtered data. The spatial Gaussian also biases the temperature measurements due to its intrinsic blurring effect. The ANLM-(2D+t) seems to provide the highest fidelity to the gold standard data.

2) Statistical analysis of the filtering errors: The pixel-wise errors illustrated in Fig. 10(h)-(m) were all pooled in a set, separately for each filtering method, and displayed as boxplots in Fig. 12. The analysis was performed both in the focal point (Fig. 12(a)) and in a 7×7 area around the focal point (Fig. 12(b)). It can be observed that, compared to the other filters, the ANLM-(2D+t) and the NLM-(2D+t) provide the best performance in terms of both accuracy and precision. A marginal improvement in the focal point can be observed for the ANLM-(2D+t) over the NLM-(2D+t). Table I reports the *p*-values provided by a Mann-Whitney test applied on the data illustrated in Fig. 12. The test indicates whether the improvement or alteration induced by the filters, compared to the case when no filtering is applied, is statistically significant or not. The null hypothesis was that there are no statistically significant differences between the errors found in the noisy and respectively the filtered data. When analysis is conducted in the area around the focal point (right column), of interest are the entries associated to the NLM-(2D+t) and the ANLM-(2D+t) filters. Their *p*-values are beneath the threshold, confirming that the improvement due to filtration is statistically significant. In the focal point itself (left column), however, the differences between the aforementioned filters and the noisy data are not statistically significant. For the IIR and spatial Gaussian, even though the *p*-value confirms the alternate hypothesis, it is obvious from Fig. 12 that, at least in the focal point, they actually end up deteriorating rather than improving the data.

Filtering method	*p*-value at the focal point	*p*-value in the heated region
Temporal IIR filter	0.0155	4.9×10^{-5}
Spatial Gaussian filter	2.2×10^{-6}	3.7×10^{-12}
NLM-2D	0.26	1.3×10^{-18}
NLM-(2D+t)	0.8	9.8×10^{-6}
ANLM-(2D+t)	0.3	2.7×10^{-5}

TABLE I: Statistical significance of the temperature error distributions obtained for each filtering method tested during the *in-vivo* experiment, as compared to the scenario when no filtering is performed.

3) Computational performance of the ANLM-(2D+t) algorithm: In terms of computational performance, our multi-threaded C++ implementation of the ANLM-(2D+t) filter required 311 ± 10 ms per image, with processing times ranging from 297 ms to a maximum of 385 ms. Note that the reported values were obtained following a total of 1000 executions of the algorithm.

4) Assessment of the noise estimation algorithm: Figure 13 illustrates the percentage relative noise estimation errors versus the SNR of the noisy magnitude images. Details concerning the images and the computation of the noise estimation errors can be found in section II-C2. Noteworthy is the fact that as the SNR of the image increases, the noise estimation errors also have a tendency to increase, exceeding 50% for an SNR

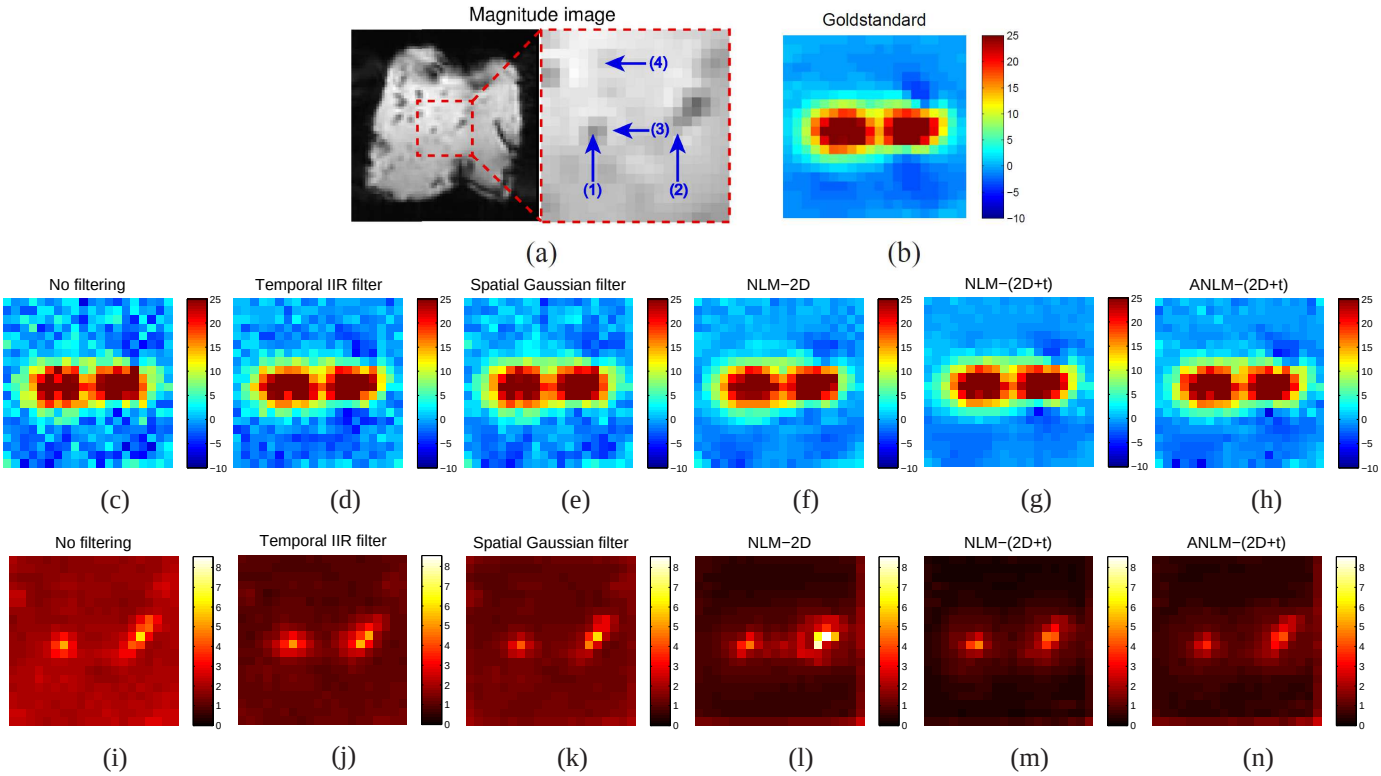


Fig. 7: Performance of the analyzed filtering methods following the *ex-vivo* study. (a): High SNR MR magnitude computed prior to the energy delivery together with a magnification of the treated area (red dashed square) upon which several points of interest are indicated by blue arrows; (b), (c): The gold standard temperature map together with its noisy version; (d)-(h): Filtered temperature maps; (i)-(n): Spatial distribution of the temporally averaged absolute errors for each filtering method.

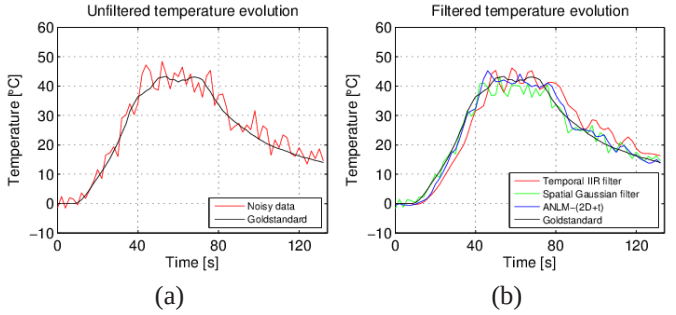


Fig. 8: Temperature time curve at point (3) from Fig. 7(a). (a): The noisy time temperature evolution (red curve). (b): Filtered temperature values obtained using the proposed ANLM(2D+t) method (blue curve), the spatial Gaussian filter (green curve) and the temporal IIR filter (red curve), respectively. The gold standard temperature illustrated in (a) and (b) by the black curve.

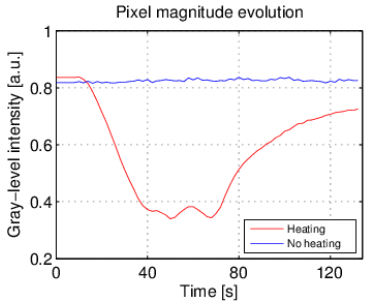


Fig. 9: The effects of heating on the magnitude of the MR signal during the *ex-vivo* experiment: Temporal evolution of the signal intensity at points (3) (red curve) and (4) (blue curve) from Fig. 7(a).

of 30.

Concerning computational demands, the algorithm required on average 16 ms per image, evaluated from a total of over 10000 estimations.

IV. Discussion

This study proposes an improved non-local filtering method for real-time dynamic PRFS thermometry. The technique was designed using the NLM formulation proposed by Coupé *et al.* in [11] as a starting point, in addition to which, temporal information was also exploited in order to improve SNR. Moreover, an adaptation technique is proposed such that, the amount of filtration is dynamically optimized as to maximize output accuracy. Compared to previously employed filtration methods for real-time thermometry, such as the IIR [17] and the spatial Gaussian [18], [19], the proposed technique has

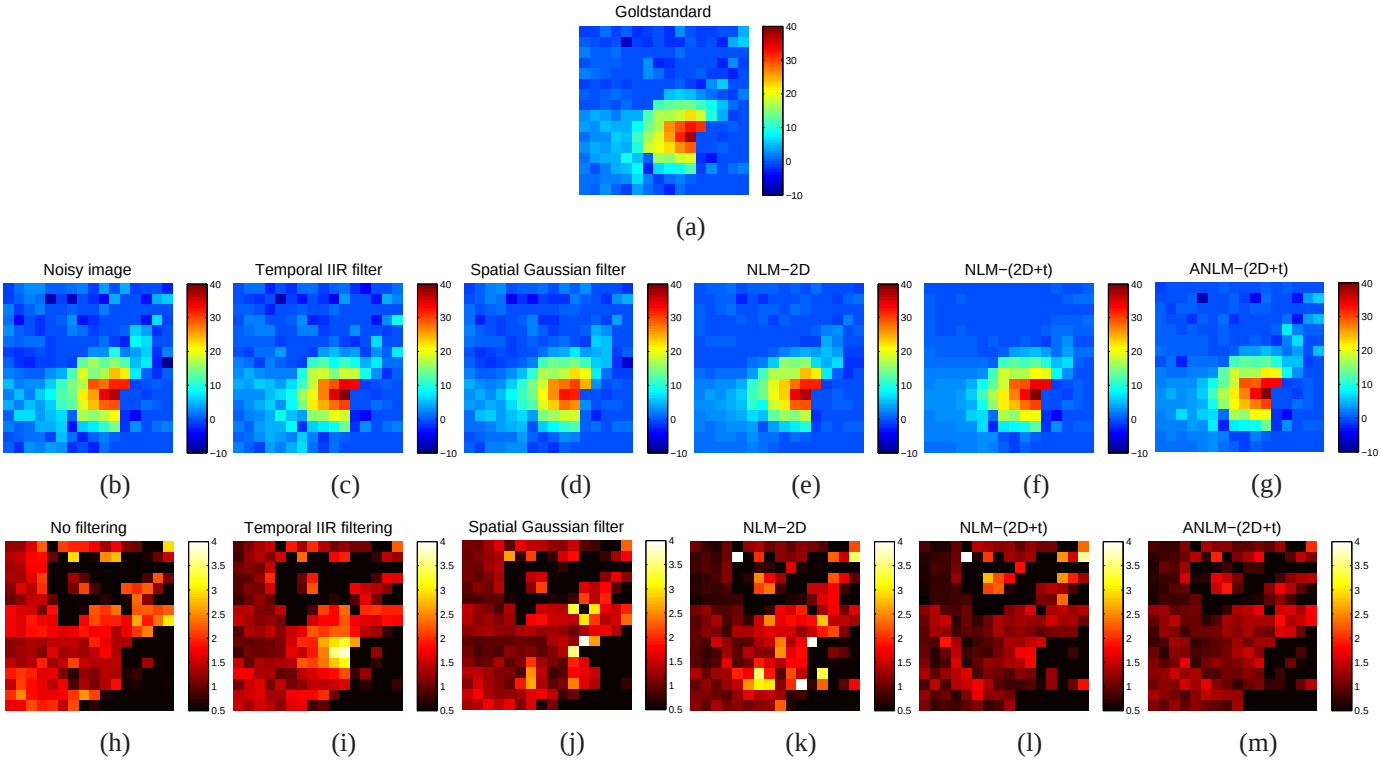


Fig. 10: Accuracy of the compared filtering methods on the *in-vivo* porcine dataset. (a): The gold standard temperature map; (b): Noisy temperature map; (c)-(g): Filtered temperature maps as provided by the compared methods; (h)-(m): Spatial distribution of the temporally averaged absolute errors for each filtering method.

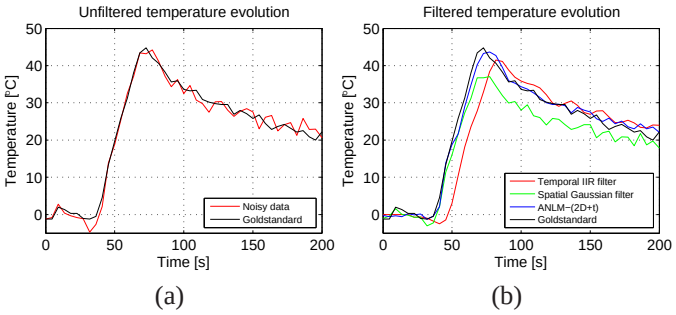


Fig. 11: Temperature time curve obtained in the focal point during the *in-vivo* experiment. (a): The noisy time temperature evolution (red curve). (b): Filtered temperature values obtained using the proposed ANLM-(2D+t) method (blue curve), the spatial Gaussian filter (green curve) and the temporal IIR filter (red curve). The gold standard temperature is illustrated in (a) and (b) by the black curve.

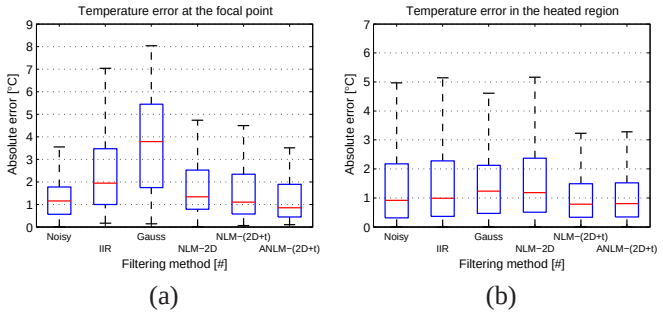


Fig. 12: Statistical analysis of the absolute errors introduced by each filtering method for the *in-vivo* experiment (a): In the focal point; (b): In the area surrounding the focal point.

shown improved performance in all tested scenarios. However, this was to be expected since, while the IIR and the Gaussian filters exploit either temporal or spatial information, respectively, the proposed NLM-(2D+t) and ANLM-(2D+t) exploit simultaneously both sources of information.

As specified throughout this work, the spatial Gauss filter and the IIR were included in the study due to their prior use for denoising temperature measurements during real-time MR-thermometry. While the exclusion of other filtering methods can be seen as somewhat of a limitation, a criterion had to be established when selecting the filtering methods to include. This is due to the fact that there is an entire branch of digital image processing dedicated to denoising methods, proposing a vast number of algorithms. Thus, it seemed a justified choice to only select methods that have been previously used in the context of MR-thermometry and allow the analysis of other filtering methods to be the object of future studies.

Performance comparison between the analyzed filtering methods was performed in terms of the temporally averaged absolute error, which quantifies the accuracy errors in the

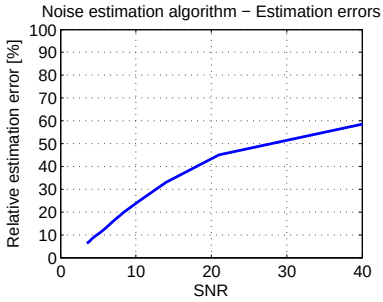


Fig. 13: Percentage relative noise estimation errors versus the SNR of the magnitude images.

(un)filtered temperature measurements. Depending on the application, alternate criteria could be used such as, for example, a thermal dose-based metric. However, the main purpose of this study was to propose a novel denoising method for real-time MR-thermometry in general. Thermal dose becomes a relevant factor only for applications involving temperatures above $\sim 43°$C [20], which, for example, is rarely the case for near-field heating surveillance or low-power test-shots. Moreover, some approaches dedicated to quantifying thermal damage rely on detecting a certain temperature threshold rather than on thermal dose [21], [22]. In such cases of interest are the precision and accuracy of the temperature measurements themselves. Thus, the impact of noise and filtering on other factors such as the thermal dose is a topic in itself and can be the subject of future studies, dedicated to more specific applications.

The original version of the NLM relies on the redundancies found in an image in order to improve its SNR. Since usually the temperature in the focal point is unique, it is natural to expect the underlying filtration errors to be large. This aspect is reflected very well throughout the majority of the results. This drawback is compensated, however, by using information regarding the temperature in focal point from the previous dynamics, as is the case of the NLM-(2D+t) and the ANLM-(2D+t). Additionally, the NLM filter recovers the signal in a pixel as the weighted average of pixels with similar neighborhoods. This will unavoidably introduce blurring effects, that lead to inaccuracies/biasing of the filtered data. The ANLM-(2D+t) circumvents this issue by optimizing the weights of the pixels used for signal recovery in such a way that accuracy of the filtered temperature is maximized. In the extreme case, when noise levels are low, a pixel might even be left unfiltered.

One of the key points in comparing the performance of the different filtering techniques in the current study was the construction/selection of a proper gold standard. This is a particularly challenging task during *in-vivo* experiments, since obtaining a completely noise-free temperature map is hardly feasible. In the current work temperature maps with only improved SNR were adopted as gold standard. Nevertheless, this was sufficient to quantify the performance of the proposed filtering methods and to show the statistical significance of the improvement over the existing ones (see Fig. 12 and Table I).

Since temporal data is used in the filtration process, image sampling/acquisition frequency becomes an important issue. In the current work, the simulation and the *ex-vivo* studies were carried out at a framerate of 1 Hz, while in the *in-vivo* experiment the sampling frequency was 0.15 Hz. Low MRI framerates become problematic especially for pure temporal filters such as the IIR evaluated in this study. As it can also be observed from Fig. 3, 8 and 11, the intrinsic temporal lag and the low temporal resolution will lead to considerable errors in the results provided by the IIR. It is expected that for low MRI framerates, due to low temporal redundancies, the proposed NLM-(2D+t) and the ANLM-(2D+t) will also suffer a decrease in performance. However, as seen in the results corresponding to the *in-vivo* experiments, the proposed filtering methods remain reliable even at frequencies as low as 0.15 Hz. This is an important aspect since, for example, several pre-clinical MRg-HIFU ablation protocols in moving organs involve respiratory gating, resulting in imaging frequencies of 0.2 - 0.4 Hz [23], [24].

Another important aspect that needs to be taken into consideration when exploiting temporal information in the context of NLM filters is motion. In theory, when denoising a pixel via the NLM filter, the search for redundancies extends over the entire image. In such a case, motion is unlikely to have an impact on filter performance. In practice, however, due to computational considerations, a pixel is usually filtered based only on its neighbors within a particular search window (which is most of the times significantly smaller than the image size). Since the latter is usually fixed in space, motion might lead to similar/identical pixels exiting/entering the search window. Naturally, this is expected to affect filter performance. However, during dynamic MR-thermometry, the acquired images are often registered/aligned to a reference position in order to allow a proper analysis of the temperature evolution in each pixel (which is particularly important in applications where the thermal dose is of interest). Thus, when applied in such a context, motion is not expected to have an impact on the performance of the proposed NLM-(2D+t) and ANLM-(2D+t) filters.

One of the inputs required by the filtering methods proposed in this study is the standard deviation of the noise altering the images. The details concerning its estimation can be found in section II-B3. While the noise estimator provided good results for low-SNR images, the performance of the algorithm starts to deteriorate as the images become less noisy (see Figure 13). We hypothesize that this is due to the fact that the anatomical structures start to become more apparent in the MR-images and the gray-level transitions from one to the other start being interpreted as noise. However, as it can be observed from Figure 6, the proposed ANLM-(2D+t) filter remains reliable even for relative noise estimation errors of 50%. The curve in Figure 13 indicates that such errors only occur for an SNR higher than 30, which for the magnitude images provided by real-time thermometry is rarely the case. For example, given the MR acquisition sequence employed during the *in-vivo* experiment, a noise standard deviation of $1°C$ in the temperature images would lead, according to Eq. (9) and Eq. (10), to an SNR in the magnitude images equal to 10. This translates to a percentage relative noise estimation error of $\sim 20\%$ (see Fig. 13), value which is within

acceptable margins, as previously noted. Any value of the noise standard deviation in the temperature images higher than $1°C$ will translate to an even lower SNR of the magnitude images, further reducing the uncertainty on the estimated noise distribution. Moreover, for low noise levels ($< 1°C$), while the relative estimation errors may be large, the absolute errors will be small. Thus, we anticipate that for a high SNR of the magnitude images (> 30), noise estimation errors of up to one order of magnitude will still have only a minimal impact on the performance of the filter.

In the current study, the noise altering the temperature maps was assumed to follow a Gaussian distribution. In theory, however, the noise distribution is somewhat more complex. PRFS temperature maps depend linearly on the difference between a current phase image and a reference acquired prior to heating. Thus, it is safe to assume that the noise distribution in the phase images and the temperature maps follow the same statistics. The MR phase image is computed as the inverse tangent of the ratio between the imaginary and the real part of the complex MR-signal. While the two channels are indeed subjected to Gaussian noise [13], the resulting expression for the noise distribution in the phase image itself becomes fairly complicated. It was shown, however, by Gudbjartsson and Patz in [25] that in practice, as long as the SNR of the magnitude image remains above 3, the noise distribution in the phase images is approximately Gaussian. For MR-thermometry acquisition sequences, this is almost always the case in the heated region, since a sequence that provides a magnitude image with an SNR beneath 3 is practically unusable to begin with. Thus, our assumption that the noise in the temperature images follows a Gaussian distribution is justified. In addition, estimation of the noise standard deviation in the temperature maps was computed based on the result of a noise estimation algorithm applied on the magnitude image. The algorithm is built on the assumption that the noise follows a Gaussian distribution. Theoretically, the noise distribution in the MR magnitude image is Rician, however, it was shown in previous studies that for SNR values higher than 3, it is well approximated by a Gaussian distribution [25], [26]. Nevertheless, these SNR limits need to be taken into consideration when denoising MR-temperature maps via the proposed methods, since beyond these boundaries no guarantees can be made on the quality of the filtered data. In case an application involves MR images with a severely low SNR, alternate methods may have to be considered.

Throughout the experiments performed in the current work, the noise distribution in the MR magnitude images (based on which the noise in the temperature measurements was calculated via Eq. (9) and (10)) was assumed to be spatially invariant. However, this hypothesis is likely to be violated in case MR-thermometry is performed using parallel imaging techniques such as SENSE or GRAPPA [12]. Depending on how the MR-signal received by the multi-coil system is re-combined/reconstructed, the noise standard deviation might end-up varying across the MR-magnitude image. Since the noise estimation method employed in our study operates on a global scale, this may lead to the proposed filtering methods to provide sub-optimal results. Thus, in case parallel acquisition is used, the global noise estimation method should be replaced with an approach that estimates the noise standard deviation in the MR-magnitude image on a pixel-by-pixel basis.

As we have shown, the proposed NLM filtering methods can be used on their own to improve the SNR of temperature maps. However, they also show great potential to be used in conjunction with other methods previously employed to improve/enhance PRFS thermometry. For example, as mentioned earlier in the manuscript, the output of the Kalman predictor-based filter proposed by Roujol *et al.* in [9] is a linear combination of the measured noisy data and a temperature evolution model, which is obviously noise-free. The resulting filtered data, thus, unavoidably contains a certain amount of noise. We anticipate that the approach can be further enhanced by coupling the output of the Kalman filter with the NLM methods proposed in the current study. Another instance in which the proposed filters could aid is in the context of the studies conducted by Todd *et al.* in [27]. They propose a method to accelerate MR-image acquisition, with the goal of rendering real-time 3D PRFS-thermometry feasible. While results have shown the great potential of the method, some limitations were encountered due to noisy measurements affecting the precision and accuracy of the resulting temperature data, which prevented further acceleration of the acquisition process. We again speculate that used in conjunction with one of the NLM filters illustrated in our study, the method could be improved.

An important aspect that needs to be taken into consideration is computational latency, i.e. the delay between the actual time of the measurement and the availability of the information. This includes the remaining acquisition time after the passage of the *k*-space center (equal to half the echo time), the data transport duration and the image processing time (which in the current study is the computational time of the ANLM-(2D+t) filter). Considering the MR-sequence and the hardware employed for the *in-vivo* study, half the echo time together with the data transport delay amount to $\sim$22.5 ms. With an average computational time of $\sim$300 ms for the ANLM-(2D+t) filter, the total latency reaches $\sim$322.5 ms per slice. In order to ensure a smooth temperature monitoring work-flow, this value must lie well beneath the time interval between two MR acquisitions. This constraint is fulfilled with a considerable margin for the application considered in this study, since respiratory gated MR-thermometry usually provides temperature updates with a frequency of 0.2 - 0.4 Hz (corresponding to the respiratory frequency). Note that the ANLM-(2D+t) filter can be even further accelerated by visiting the values of β (see Eq. (4)) in parallel, contrary to the exhaustive search performed in the current work. We anticipate that this would lead to an average processing time of the filter beneath 30 ms. Under such conditions the total latency would amount to $\sim$52.5 ms, allowing high-frame-rate MR-imaging. Additionally, an implementation of the ANLM-(2D+t) on a graphical processing unit is expected to provide even shorter computational times.

V. Conclusion

In this work we propose an improved non-local filtering technique for PRFS-based temperature measurements during real-time MR-guided targeted thermal therapies. Accurate and precise temperature maps are highly beneficial for proper therapy control and endpoint detection, potentially avoiding under or over-treatment of the targeted area. Compared to the original NLM filter and two other filtering methods previously employed during MR-thermometry, the proposed approaches have shown a considerable increase in both precision and accuracy. This together with their low computational demands, prove that the proposed filters are potentially beneficial as a means to improve temperature measurements during PRFS-based real-time MR thermometry dedicated to targeted thermal therapies.

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 gratefully acknowledge Pierrick Coupé (Laboratoire Bordelais de la Recherche Informatique (LaBRI), Université de Bordeaux, France) and Charles Deledalle (Institut de Mathématiques de Bordeaux (IMB), Université de Bordeaux, France) for the many helpful discussions and suggestions. Finally, we would like to thank the three anonymous reviewers whose remarks and suggestions have helped us greatly in improving the quality of the manuscript.

References

[1] Y. Ishihara, A. Calderon, H. Watanabe, K. Okamoto, Y. Suzuki, K. Kuroda, and Y. Suzuki, "A precise and fast temperature mapping using water proton chemical shift," *Magn Reson Med*, vol. 34, no. 6, pp. 814–823, 1995.

[2] J. De Poorter, D. D. Y. De Wagter, C., T. C., F. Stahlberg, and E. Achten, "Noninvasive MRI thermometry with the proton resondance frequency (PRF) method: in vivo results in human muscle," *Magn Reson Med*, vol. 33, no. 1, pp. 74–81, 1995.

[3] V. Rieke and K. B. Pauly, "MR-thermometry," *J Magn Reson Imaging*, vol. 27, no. 2, pp. 376–390, 2008.

[4] C. S. Mei, L. P. Panych, J. Yuan, N. J. McDannold, L. H. Treat, Y. Jing, and B. Madore, "Combining two-dimensional spatially selective RF excitation, parallel imaging, and UNFOLD for acceleraed MR thermometry imaging," *Magn Reson Med*, vol. 66, no. 1, pp. 112–122, 2011.

[5] J. Yuan, C. S. Mei, P. L. P., M. N. J., and B. Madore, "Towards fast and accurate temperature mapping with proton resonance frequency-based MR thermometry," *Quant Imaging Med Surg*, vol. 2, no. 1, pp. 21–32, 2012.

[6] Q. X. Yang, D. R. J., B. J. Dardzinski, B. W. Arnold, and M. B. Smith, "Multiple echo frequency-domain image contrast: improved signal-to-noise ratio and T2(T2*) weighting," *Magn Reson Med*, vol. 41, no. 2, pp. 423–428, 1999.

[7] D. Erdogmus, L. E. G., R. Yan, P. J. C., and F. J. R., "Asymptotic SNR-performance of some image combination techniques for phased-array MRI," *J Sign Proc*, vol. 84, no. 6, pp. 997–1003, 2004.

[8] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *J Multiscale Model Simul*, vol. 4, no. 2, pp. 490–530, 2005.

[9] S. Roujol, B. Denis de Senneville, S. Hey, C. Moonen, and M. Ries, "Robust Adaptive Extended Kalman filtering for real time MR-thermometry guided HIFU interventions," *IEEE Transaction on Medical Imaging*, vol. 31, no. 3, pp. 533–542, 2012.

[10] R. Kalman, "A new approach to linear filtering and prediction problems," *J of Basic Engineering*, vol. 82, no. 2, pp. 35–45, 1960.

[11] P. Coupe, P. Yger, P. Hellier, C. Kervrann, and C. Barillot, "An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images," *IEEE Trans Med Imaging*, vol. 27, no. 4, pp. 425–441, 2008.

[12] J. V. Manjon, P. Coupe, L. Marti-Bonmati, D. L. Collins, and M. Robles, "Adaptive non-local means denoising of MR images with spatially varying noise levels," *J Magn Reson Imaging*, vol. 31, no. 1, pp. 192–203, 2010.

[13] T. E. Conturo and G. D. Smith, "Signal-to-noise in phase angle reconstruction: Dynamic range extension using phase reference offsets," *Magn Reson Med*, vol. 15, no. 3, pp. 420–437, 1990.

[14] S. J. Riederer, "Recent technical advances in MR imaging in the abdomen," *J Magn Reson Imaging*, vol. 6, no. 5, pp. 822–832, 1996.

[15] R. D. Peters and R. M. Henkelman, "Proton-resonance frequency shift MR thermometry is affected by changes in the electrical conductivity of tissue," *Magn Reson Med*, vol. 43, no. 1, pp. 62–71, 2000.

[16] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

[17] M. Ries, B. Denis de Senneville, S. Roujol, Y. Berber, B. Quesson, and C. Moonen, "Real-time 3D target tracking in MRI guided focused ultrasound ablations in moving tissues," *Magnetic Resonance in Medicine*, vol. 64, pp. 1704–1712, 2010.

[18] B. Denis de Senneville, P. Desbarats, B. Quesson, and M. C. T. W., "Real-time artifact corrections for quantitative MR temperature mapping," *Journal of WSCG*, vol. 11, no. 1, 2003.

[19] B. Denis de Senneville, A. El Hamidi, and C. Moonen, "A direct PCA-based approach for real-time description of physiological organ deformations," *IEEE Transaction on Medical Imaging*, vol. 34, no. 4, pp. 974–982, 2015.

[20] S. Separeto and W. Dewey, "Thermal dose determination in cancer therapy," *Int J Radiat Oncol Biol Phys*, vol. 10, no. 6, pp. 787–800, 1984.

[21] T. J. Dubinsky, C. Cuevas, M. K. Dighe, O. Kolokythas, and J. H. Hwang, "High-intensity focused ultrasound: current potential and oncologic applications," *Am J Roentgenol*, vol. 190, no. 1, pp. 191–199, 2008.

[22] A. Kolandaivelu, M. M. Zviman, V. Castro, A. C. Lardo, R. D. Berger, and H. R. Halperin, "Noninvasive assessment of tissue heating during cardiac radiofrequency ablation using MRI thermography," *Circ Arrhythm Electrophysiol*, vol. 3, no. 5, pp. 521–529, 2010.

[23] V. Auboiroux, L. Petrusca, M. Viallon, A. Muller, S. Terraz, R. Breuget, X. Montet, C. D. Becker, and R. Salomir, "Respiratory-gated MRgHIFU in upper abdomen using an MR-compatible in-bore digital camera," *Biomed Res Int*, vol. 2014, pp. 1–9, 2014.

[24] J. W. Wijlemans, M. de Greef, G. Schubert, L. W. Bartels, C. T. Moonen, M. A. van den Bosch, and M. Ries, "A clinically feasible treatment protocol for magnetic resonance-guided high-intensity focused ultrasound ablation in the liver," *Invest Radiol*, vol. 50, no. 1, pp. 24–31, 2015.

[25] H. Gudbjartsson and S. Patz, "The Rician distribution of noisy MRI data," *Magn Reson Med*, vol. 34, no. 6, pp. 910–914, 1995.

[26] R. D. Nowak, "Wavelet-based Rician noise removal for magnetic resonance imaging," *IEEE Trans Imag Proc*, vol. 8, no. 10, pp. 1408–1419, 1999.

[27] N. Todd, J. Prakesh, H. Odeen, J. de Bever, A. Payne, P. Yalavarthy, and D. L. Parker, "Toward real-time availability of 3D temperature maps created with temporally constrained reconstruction," *Magn Reson Med*, vol. 71, no. 4, pp. 1394–1404, 2014.