



Open Archive Toulouse Archive Ouverte

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible

This is an author's version published in:

<http://oatao.univ-toulouse.fr/26432>

Official URL

<https://doi.org/10.1109/ICIP40778.2020.9191013>

To cite this version: El Mansouri, Oumaima and Basarab, Adrian and Figueiredo, Mario and Kouamé, Denis and Tourneret, Jean-Yves *Ultrasound and magnetic resonance image fusion using a patch-wise polynomial model*. (2020) In: IEEE International Conference on Image Processing (ICIP 2020), 25 October 2020 - 28 October 2020 (Abu Dhabi, United Arab Emirates).

Any correspondence concerning this service should be sent to the repository administrator: tech-oatao@listes-diff.inp-toulouse.fr

ULTRASOUND AND MAGNETIC RESONANCE IMAGE FUSION USING A PATCH-WISE POLYNOMIAL MODEL

O. El Mansouri^{(1),(2)} A. Basarab⁽²⁾ M. A. T. Figueiredo⁽³⁾ D. Kouamé⁽²⁾ J.-Y. Tourneret⁽¹⁾

⁽¹⁾ University of Toulouse, IRIT/INP-ENSEEIH/TéSA, 31071 Toulouse Cedex 7, France

⁽²⁾ University of Toulouse, IRIT, CNRS UMR 5505, Université Paul Sabatier, Toulouse, France

⁽³⁾ Instituto de Telecomunicações, Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisbon, Portugal

ABSTRACT

This paper introduces a novel algorithm for the fusion of magnetic resonance and ultrasound images, based on a patch-wise polynomial model relating the gray levels of the two imaging systems (called modalities). Starting from observation models adapted to each modality and exploiting a patch-wise polynomial model, the fusion problem is expressed as the minimization of a cost function including two data fidelity terms and two regularizations. This minimization is performed using a PALM-based algorithm, given its ability to handle non-linear and possibly nonconvex functions. The efficiency of the proposed method is evaluated on phantom data. The resulting fused image is shown to contain complementary information from both magnetic resonance (MR) and ultrasound (US) images, i.e., with a good contrast (as for the MR image) and a good spatial resolution (as for the US image).

Index Terms— Image fusion, magnetic resonance imaging, ultrasound imaging, super-resolution, despeckling, image enhancement, patch-based method.

1. INTRODUCTION

Magnetic resonance (MR) and ultrasound (US) images have been used intensively in many clinical diagnosis and guided surgery applications. While they both carry important information in assessing the condition of organs, they exploit different physical phenomena and thus have their own advantages and limitations. In particular, US imaging offers a good spatial resolution and high frame rate compared to MRI, at the cost of a very low signal to noise ratio (SNR), low contrast (depending on the central frequency of the probe), a presence of speckle noise and a reduced field of view. In contrast, MRI enables a wide field of view, with a good SNR, high contrast, but relatively low spatial resolution [1]. As a consequence of these complementary properties, MR and US images are commonly used jointly in various clinical applications. The objective of this paper is to propose a method to fuse the two

images in a single image in order to improve the diagnosis capacity of each modality.

Image fusion refers to assembling all the important information from multiple images and including them in fewer images or into a single image. Its purpose is not only to reduce the amount of data but also to build enhanced images that are more comprehensible and informative for human and machine insight [2]. Fusion of medical images is becoming very common for the study of a given pathology [3–5], and generally allows for a better medical decision in clinical studies. Medical images that are commonly fused include CT scans and positron emission tomography [6], or gammagraphy and US images [7]. However, to the best of our knowledge, the fusion of MR and US images, which is the purpose of this work, has been less addressed in the existing literature.

In our previous work on MR and US image fusion [8], we introduced a new algorithm performing both super-resolution of the MR image and despeckling of the US image. That algorithm was based on a polynomial function relating the US and MR images, accounting for the discrepancy between these two modalities. The coefficients of this polynomial were pre-estimated from the observed images. This paper further improves the polynomial relation between the two images by estimating the polynomial coefficients patch-wise, thus allowing for a better matching between the two images to be fused. Note that a similar idea was used in [9] for MRI images.

The paper is organized as follows. Section 2 presents the observation models, the patch-based polynomial function relating the US and MR images, and the optimization problem considered to fuse these images. The algorithm proposed to solve the fusion problem is detailed in Section 3. Simulation results are presented in Section 4. Conclusions and perspectives are finally reported in Section 5.

2. MAGNETIC RESONANCE AND ULTRASOUND IMAGE FUSION

2.1. Observation models

Denote as $\mathbf{y}_{\text{mr}} \in \mathbb{R}^M$ and $\mathbf{y}_{\text{us}} \in \mathbb{R}^N$ the registered MR and US images, with M and N the number of pixels in each

The authors would like to thank Fabien Vidal for providing the ultrasound and magnetic resonance data, as well as for the fruitful discussions about the clinical pertinence of the proposed algorithm.

image¹. This section introduces two observation models accounting for the low spatial resolution of MR images and the low SNR of US images. The low resolution of the MR image is modeled by a downsampling operation and a low pass filter [11], while an additive noise model is considered for the US B-mode image. Note that speckle is assumed to be a multiplicative noise, leading to additive perturbations when applying log-compression, which is classically considered before forming B-mode images. Furthermore, this work assumes that the speckle noise affecting B-mode images is distributed according to a log-Rayleigh distribution, as in [12, 13]. The two resulting observation models are

$$\begin{aligned} \mathbf{y}_{\text{us}} &= \mathbf{x}_{\text{us}} + \mathbf{n}_{\text{us}} \\ \mathbf{y}_{\text{mr}} &= \mathbf{S}\mathbf{H}\mathbf{x}_{\text{mr}} + \mathbf{n}_{\text{mr}}, \end{aligned} \quad (1)$$

where $\mathbf{y}_{\text{us}} \in \mathbb{R}^N$ is the observed B-mode US image, $\mathbf{x}_{\text{us}} \in \mathbb{R}^N$ is the noiseless US image, $\mathbf{n}_{\text{us}} \in \mathbb{R}^N$ is the log-Rayleigh speckle noise, $\mathbf{x}_{\text{mr}} \in \mathbb{R}^N$ is the high-resolution MR image, $\mathbf{y}_{\text{mr}} \in \mathbb{R}^M$ is the observed (low-resolution) MR image, and $\mathbf{n}_{\text{mr}} \in \mathbb{R}^N$ is an additive Gaussian noise. The matrix $\mathbf{H} \in \mathbb{R}^{N \times N}$ is the blurring matrix and $\mathbf{S} \in \mathbb{R}^{M \times N}$ (with $N = d^2M$) is a decimation operator with decimation factor d . Note that the decimation factor is such that \mathbf{x}_{us} and \mathbf{x}_{mr} have the same spatial sampling.

2.2. Patch-based polynomial model

The patch-based polynomial model proposed in this work (relating the gray levels of MR and US images) is motivated by the fact that US images highlight the interfaces between different anatomical structures with different acoustic impedances [14]. More precisely, the US image is expressed as a function of the MR image and its spatial gradient is computed in the direction of US wave propagation

$$\mathbf{x}_{\text{us}} = f(\mathbf{x}_{\text{mr}}, \nabla \mathbf{x}_{\text{mr}}^H \mathbf{u}), \quad (2)$$

where $f : \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}^N$ is unknown and $\nabla \mathbf{x}_{\text{mr}}^H \mathbf{u} \in \mathbb{R}^N$ contains in its i th line the inner product between the i th local gradient \mathbf{x}_{mr} and the US scan direction \mathbf{u} .

The function f was represented by a global polynomial in our previous work on image fusion [15], and in [14] for multimodal image registration. However, the relationship between MR and US images may depend on tissue acoustic and magnetic properties, and thus may change from one image region to another. Thus, considering a global polynomial model may lead to inaccurate gray level matching in specific image regions. To overcome this issue, this paper introduces a more general patch-based polynomial model, fitting independently low-order polynomial functions to each overlapping patch extracted from MR and US images. This patch-based polynomial model is defined as

¹The MR and US images are supposed to be registered with an algorithm such as [10] and the possible registration errors are ignored.

$$\mathbf{P}_p \mathbf{x}_{\text{us}} = f_p(\mathbf{P}_p \mathbf{x}_{\text{mr}}, \mathbf{P}_p \nabla \mathbf{x}_{\text{mr}}^H \mathbf{u}), \quad (3)$$

where $\mathbf{P}_p \in \mathbb{R}^{n \times N}$ is a binary operator that extracts the p th patch of size n from an image of size N . In the following, N_p will denote the total number of patches. Replacing f_p by a polynomial function, the relation between patches from the US and MR images becomes

$$\mathbf{P}_p \mathbf{x}_{\text{us}} = \sum_{l+k \leq d_p} c_{l,k,p} \mathbf{P}_p \mathbf{x}_{\text{mr}}^l \odot (\mathbf{P}_p \nabla \mathbf{x}_{\text{mr}}^H \mathbf{u})^k, \quad (4)$$

where $p = 1, \dots, N_p$ is the patch number, d_p and $c_{l,k,p}$ are the order and the coefficients of the polynomial function f_p corresponding to patch $\#p$, \odot is the Hadamard product (element by element multiplication) and the power operations applied to vectors are element-wise. In this paper, the final function f is obtained by averaging patch-wise polynomials, since each pixel of the image is contained in several overlapping patches. More precisely, the transformation of the i th pixel denoted as $f_i : \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}$ is the average of all the polynomials associated with the patches containing this pixel.

2.3. Cost function

Using the observation models in (1), the relationship between MR and US images defined in (3) and (4), and the ideas proposed in [15], this paper formulates image fusion as the following optimization problem:

$$\begin{aligned} \hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} & \underbrace{\frac{1}{2} \|\mathbf{y}_{\text{mr}} - \mathbf{S}\mathbf{H}\mathbf{x}\|^2}_{\text{MRI data fidelity}} + \underbrace{\tau_1 \|\nabla \mathbf{x}\|^2 + \tau_2 \|\nabla f(\mathbf{x}, \nabla \mathbf{x}^H \mathbf{u})\|^2}_{\text{regularization}} \\ & + \underbrace{\tau_3 \sum_{i=1}^N \left[\exp(y_{\text{us},i} - f_i(\mathbf{x}, \nabla \mathbf{x}^H \mathbf{u})) - \lambda(y_{\text{us},i} - f_i(\mathbf{x}, \nabla \mathbf{x}^H \mathbf{u})) \right]}_{\text{US data fidelity}}, \end{aligned} \quad (5)$$

where $\hat{\mathbf{x}}$ is the fused image, $y_{\text{us},i}$ is the i th pixel of \mathbf{y}_{us} and where τ_1, τ_2, τ_3 are hyperparameters balancing the weights of the MR and US data fidelity terms and regularizations. Note that, following [15], total variation was used to regularize the solution, thus promoting piecewise constant fused images both in the US and MR domains.

3. OPTIMIZATION

3.1. PALM algorithm for MR and US image fusion

The cost function in (5) is non-convex because of the presence of the polynomial functions f and f_i . Therefore, we investigate a solution based on the proximal alternating linearized minimization (PALM) algorithm [16]. In order to fit the general form of this algorithm, we propose the following parametrization:

$$\begin{aligned}
l(\mathbf{x}) &= \frac{1}{2} \|\mathbf{y}_{\text{mr}} - \mathbf{S}\mathbf{H}\mathbf{x}\|_2^2 + \tau_1 \|\nabla \mathbf{x}\|^2, \\
g(\mathbf{v}) &= \tau_3 \sum_i [\exp(y_{\text{us},i} - v_i) - \gamma(y_{\text{us},i} - v_i)] + \tau_2 \|\nabla \mathbf{v}\|^2, \\
H(\mathbf{x}, \mathbf{v}) &= \tau_4 \|\mathbf{v} - f(\mathbf{x}, \nabla \mathbf{x}^H \mathbf{u})\|^2,
\end{aligned}$$

where

$$\mathbf{v} = f(\mathbf{x}, \nabla \mathbf{x}^H \mathbf{u}).$$

This parametrization allows (5) to be rewritten as

$$\operatorname{argmin}_{\mathbf{x}, \mathbf{v}} l(\mathbf{x}) + g(\mathbf{v}) + H(\mathbf{x}, \mathbf{v}), \quad (6)$$

where l and g are related to the MRI and US images, and H ensures the coupling between the two modalities (whose importance is controlled by the hyperparameter τ_4).

The PALM algorithm iteratively minimizes the cost function in (6) with respect to \mathbf{x} and \mathbf{v} (the reader is invited to consult [16] for more details about PALM). Note that this cost function depends on the coefficients $c_{l,k,p}$ and degrees d_p of the different polynomials, which need to be estimated for each patch, as shown in the next subsection.

3.2. Estimation of the polynomial functions f_p

For a given degree d_p , the polynomial function f_p relating patches $\mathbf{P}_p \mathbf{x}_{\text{mr}}$ and $\mathbf{P}_p \mathbf{x}_{\text{us}}$ is defined by $(d_p + 1)(d_p + 2)/2$ coefficients assembled in the vector $\mathbf{c}_{d,p} = \{c_{k,l,p} \mid k + l \leq d_p\}$. To estimate these coefficients, we consider that the p th observed MR and US patches are related according to

$$\mathbf{P}_p \mathbf{y}_{\text{us}} = \sum_{k+l \leq 3} c_{k,l,p} \mathbf{P}_p \mathbf{y}_{\text{mr}}^l \odot (\mathbf{P}_p \nabla \mathbf{y}_{\text{mr}}^H \mathbf{u})^k + \epsilon_p,$$

or in a matrix form

$$\mathbf{P}_p \mathbf{y}_{\text{us}} = \mathbf{A}_{\text{mr},p} \mathbf{c}_{d,p} + \epsilon_p, \quad (7)$$

where $\mathbf{A}_{\text{mr},p}$ is a matrix whose elements are $\mathbf{P}_p \mathbf{y}_{\text{mr}}^l \odot (\mathbf{P}_p \nabla \mathbf{y}_{\text{mr}}^H \mathbf{u})^k$ for $l + k \leq d_p$, and ϵ_p is the measurement error. The least-squares estimator of $\mathbf{c}_{d,p}$ is defined by

$$\hat{\mathbf{c}}_{d,p} = \mathbf{A}_{\text{mr},p}^\dagger \mathbf{P}_p \mathbf{y}_{\text{us}}, \quad p = 1, \dots, N_p,$$

where $\mathbf{A}_{\text{mr},p}^\dagger = (\mathbf{A}_{\text{mr},p}^T \mathbf{A}_{\text{mr},p})^{-1} \mathbf{A}_{\text{mr},p}^T$ is the pseudo-inverse of the matrix $\mathbf{A}_{\text{mr},p}$.

In order to estimate the polynomial degree of the p th patch, we minimize the least square distance between $\mathbf{P}_p \mathbf{y}_{\text{mr}}$ and $\mathbf{P}_p \mathbf{y}_{\text{us}}$, i.e., solve the following problem

$$\operatorname{argmin}_{d_p} \|\mathbf{P}_p \mathbf{y}_{\text{us}} - f_p(\mathbf{P}_p \mathbf{y}_{\text{mr}}, \mathbf{P}_p \nabla \mathbf{y}_{\text{mr}}^H \mathbf{u})\|^2,$$

where we highlight that the polynomial degree d_p depends on the patch size. In the results provided in this paper, patches of size 30×30 were extracted from images containing 600×600 pixels, with an overlap of 25%. The degree of the polynomial relating the patches was constrained to $d_p \in \{1, \dots, 3\}$.

4. RESULTS AND DISCUSSION

The proposed MR-US image fusion algorithm was validated on experimental phantom data.² Figs. 1(a,b) show the observed MR and US images.

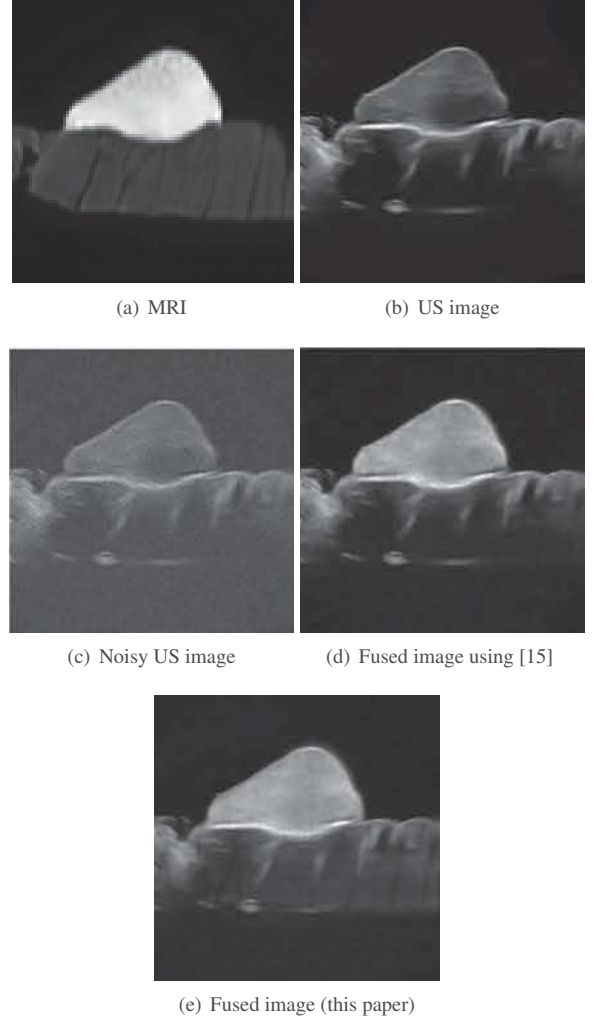


Fig. 1. Original MR image (200×200 pixels) and US image (600×600 pixels) and fusion results: (a) observed MRI, (b) original US image, (c) noisy US image, (d) fused image using a global polynomial model [14], (e) fused image with the proposed path-based polynomial model.

To mitigate the relatively good SNR obtained due to the phantom design, the US image was further degraded by log-Rayleigh noise as shown in Fig. 1(c). Figs. 1(a,b,c) highlight the differences in gray levels, spatial resolution, contrast, and noise between the two MR and US images. Three main structures can be observed in these images: a PVC phantom

²More details about the experimental model design and image acquisition can be found in [17].

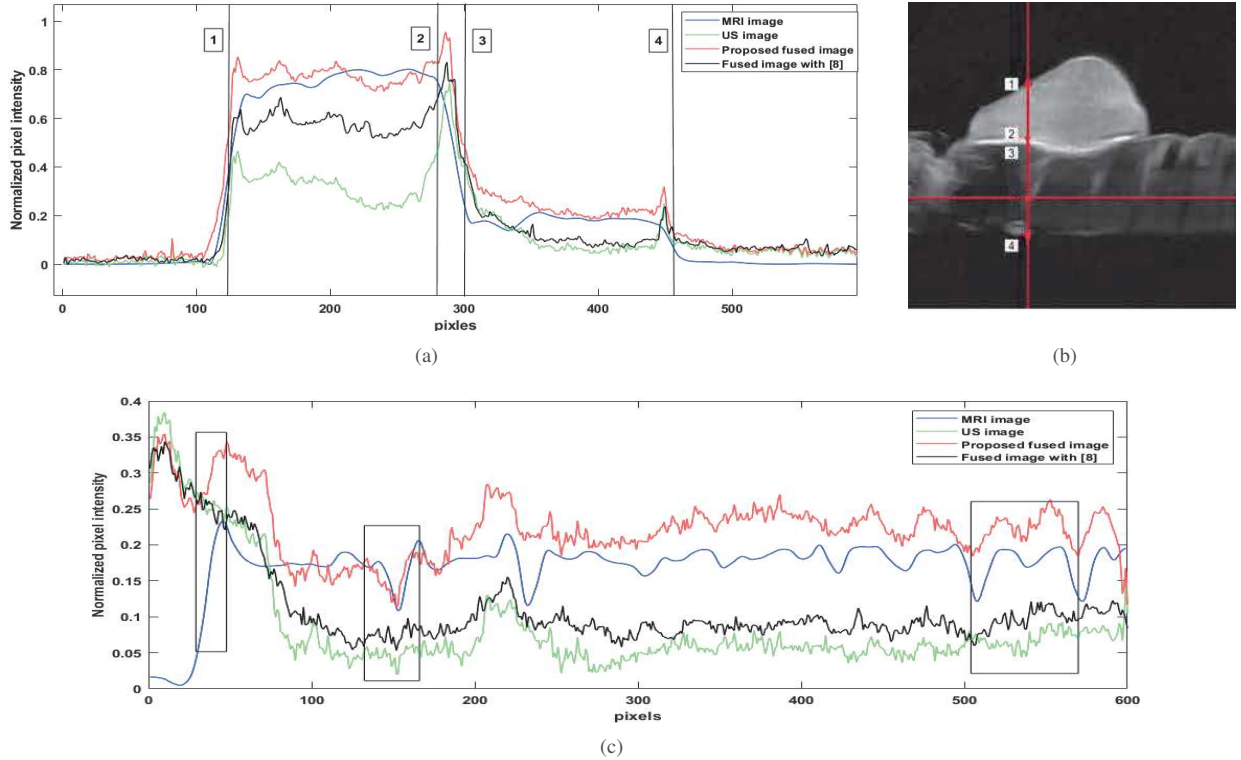


Fig. 2. (a) and (c) show normalized pixel intensities of extracted lines from the MR, US and fused images (using [15] and the proposed patch-based fusion). (a) shows the vertical straight lines in (b) whereas (c) displays the horizontal straight lines in (b).

(bright structure in the MR image), a piece of beef meat (gray structure in the MR image), and the glue used to attach them, only visible in the US image. Figs. 1(c,d) show the fused images obtained with the algorithm in [15] and the new proposed approach. Both fused images gather information from MR and US images (with a small preference to the proposed method): they provide a good contrast between the PVC and the beef tissue (similar to MRI), a good spatial resolution (similar to US) allowing small structures such as the glue to be distinguished, and good SNR. Moreover, the image obtained after fusion seems to carry more information than MRI, especially in the beef tissue.

Table 1. CNR results

CNR			
MRI	US	Fused image with [15]	Proposed
48.76 dB	20.64 dB	37.73 dB	41.72

In addition to visual inspection, the performance of the proposed patch-wise method was evaluated using two quantitative measures and compared to the global fusion method of [15]: 1) the contrast-to-noise ratio (CNR) [18] between the PVC and the beef meat, and 2) the slope between two neighboring structures as an indication of the spatial resolution [19]. As reported in Tables 1 and 2, the patch-wise approach offers a better compromise between MR and US im-

Table 2. Slope values at the interface between different regions of interest in the MR, US and fused images, corresponding to the vertical profile in Fig. 2.

Slope	MRI	US	Fused image with [15]	Proposed
#1	2.89	7.42	7.42	7.42
#2	-0.10	8.89	6.86	7.15
#3	3.57	5.47	4.61	5.24
#4	-1.35	-1.95	-2.05	-2.05

ages with a CNR close to that of the MRI and a slope close to that of the US image. Fig. 2 confirms these results, showing that the patch-wise fused image captures more details from the MRI than the global model-based fused image.

5. CONCLUSION

This paper studied a new approach for MR and US image fusion. The relation between MR and US images was modeled locally by low-order polynomial functions associated with the image patches. Interestingly, results obtained on a phantom show the advantage of using local polynomials associated with the image patches. A natural progression of this work is to combine the proposed fusion method with multimodal image registration in order to correct the registration errors and to further validate the algorithm on *in vivo* data.

6. REFERENCES

- [1] M. C. Mahoney and M. S. Newell, "Screening MR imaging versus screening ultrasound: pros and cons," *Magnetic Resonance Imaging Clinics*, vol. 21, no. 3, pp. 495–508, 2013.
- [2] M. Amin-Naji and A. Aghagolzadeh, "Multi-focus image fusion in DCT domain using variance and energy of Laplacian and correlation coefficient for visual sensor networks," *Journal of AI and Data Mining*, vol. 6, pp. 233–250, 2018.
- [3] J. Du, W. Li, K. Lu, and B. Xiao, "An overview of multi-modal medical image fusion," *Neurocomputing*, vol. 215, pp. 3–20, 2016.
- [4] A. P. James and B. V. Dasarathy, "Medical image fusion: A survey of the state of the art," *Information Fusion*, vol. 19, pp. 4–19, 2014.
- [5] A. Loza, D. Bull, N. Canagarajah, and A. Achim, "Non-Gaussian model-based fusion of noisy images in the wavelet domain," *Computer Vision and Image Understanding*, vol. 114, no. 1, pp. 54–65, 2010.
- [6] D. W. Townsend and T. Beyer, "A combined PET/CT scanner: the path to true image fusion," *The British journal of radiology*, vol. 75, pp. S24–S30, 2002.
- [7] S. Gautier, G. L. Besnerais, A. Mohammad-Djafari, and L. Blandine, "Data fusion in the field of non destructive testing," in *Maximum Entropy and Bayesian Methods*, New Mexico, U.S.A., 1996, pp. 311–316.
- [8] O. El Mansouri, F. Vidal, A. Basarab, P. Payoux, D. Kouamé, and J.-Y. Tournet, "Fusion of magnetic resonance and ultrasound images for endometriosis detection," *IEEE Trans. Image Process.*, vol. 29, pp. 5324–5335, 2020.
- [9] N. Eslahi and A. Foi, "Joint sparse recovery of misaligned multimodal images via adaptive local and non-local cross-modal regularization," in *Proc. IEEE Int. Workshop Comput. Adv. Multi-Sensor Adaptive Process. (CAMSAP)*, Guadeloupe, West Indies, 2019.
- [10] A. Yavariabdi and A. Bartoli, "Mapping and characterizing endometrial implants by registering 2D transvaginal ultrasound to 3D pelvic magnetic resonance images," *Computerized Medical Imaging and Graphics*, vol. 45, pp. 11–25, 2015.
- [11] F. Shi, J. Cheng, L. Wang, P.-T. Yap, and D. Shen, "LRTV: MR image super-resolution with low-rank and total variation regularizations," *IEEE Trans. Med. Imaging*, vol. 34, pp. 2459–2466, 2015.
- [12] D. Kundu and M. Z. Raqab, "Generalized Rayleigh distribution: different methods of estimations," *Computational statistics & data analysis*, vol. 49, no. 1, pp. 187–200, 2005.
- [13] T. Tuthill, R. Sperry, and K. Parker, "Deviations from Rayleigh statistics in ultrasonic speckle," *Ultrasonic imaging*, vol. 10, no. 2, pp. 81–89, 1988.
- [14] A. Roche, X. Pennec, G. Malandain, and N. Ayache, "Rigid registration of 3D ultrasound with MR images: a new approach combining intensity and gradient information," *IEEE Trans. Med. Imaging*, vol. 20, no. 10, pp. 1038–1049, 2001.
- [15] O. E. Mansouri, A. Basarab, F. Vidal, D. Kouamé, and J.-Y. Tournet, "Fusion of magnetic resonance and ultrasound images: A preliminary study on simulated data," in *Proc. IEEE Int. Symp. Biomed. Imaging (ISBI)*, Venice, Italy, 2019, pp. 1733–1736.
- [16] J. Bolte, S. Sabach, and M. Teboulle, "Proximal alternating linearized minimization or nonconvex and non-smooth problems," *Mathematical Programming*, vol. 146, no. 1-2, pp. 459–494, 2014.
- [17] F. Vidal, O. E. Mansouri, D. Kouamé, and A. Basarab, "On the design of a pelvic phantom for magnetic resonance and ultrasound image fusion," in *2019 IEEE International Ultrasonics Symposium (IUS)*, Glasgow, Scotland, 2019, pp. 2400–2403.
- [18] M. Welvaert and Y. Rosseel, "On the definition of signal-to-noise ratio and contrast-to-noise ratio for fMRI data," *PloS one*, vol. 8, no. 11, p. e77089, 2013.
- [19] R. Morin, A. Basarab, S. Bidon, and D. Kouamé, "Motion estimation-based image enhancement in ultrasound imaging," *Ultrasonics*, vol. 60, pp. 19–26, 2015.