JOINT REGISTRATION AND FUSION OF 3D MAGNETIC RESONANCE AND 2D ULTRASOUND IMAGES FOR ENDOMETRIOSIS SURGERY

Youssra El Bennioui[(1)], *Fabien Vidal*[(3)], *Adrian Basarab*[(2)], *Jean-Yves Tourneret*[(1)]

[(1)]Institut de Recherche en Informatique de Toulouse, CNRS UMR 5505, Université de Toulouse / TéSA Laboratory, Toulouse, France

[(2)]Univ Lyon, Université Claude Bernard Lyon 1, CNRS, Inserm, CREATIS UMR 5220, U1294, France

(3) Department of gynecologic surgery Clinique Croix Du Sud Ramsey Santé, France

ABSTRACT

This paper investigates a general framework for the registration of 3D magnetic resonance (MR) and 2D ultrasound (US) images. This framework is divided into a rigid slice-to-volume 3D-2D MR/US registration and a 2D-2D US/MRI fusion algorithm to generate an image having a better resolution than the MR image and a better contrast than the US image. The accuracy of the joint registration and fusion method is analyzed by means of quantitative and qualitative tests conducted on experimental phantom and realistic synthetic data generated from an in vivo MRI volume, with a specific attention to endometriosis treatment.

Index Terms— Registration, Image fusion, Magnetic Resonance Imaging, Ultrasound Imaging, Endometriosis.

1. INTRODUCTION

Endometriosis is a complex disease that occurs when the tisue that normally lines the uterus grows outside of it, and bleeds cyclically as it is under the influence of the menstrual cycle. It is very often located in the ovaries, the intestine or the bladder or, more rarely, on other internal organs. The average time to detect endometriosis is about 7 years during which patients suffer from physical and psychological consequences, such as debilitating chronic pelvic pain and infertility.

Endometriosis diagnosis is based on two medical imaging modalities, namely ultrasound (US) and magnetic resonance (MR) images. The main treatment consists of laparoscopic surgery that requires a precise localization of the endometrial implant and its depth of infiltration. Therefore, US and MRI play a crucial and complementary role on endometriosis examination. Transvaginal and transrectal US scanners use high-frequency probes, resulting in high spatial resolution images that allow a precise localization and characterization of small lesions. In particular, US images allow lesion depth of infiltration to be evaluated. However, US images are characterized by a low signal to noise ratio and a limited field of view. On the contrary, MR images provide a wide field of view of the patient anatomy with a good SNR ratio, but small lesions cannot be detected because of its limited spatial resolution close to 1 to 2 mm. Based on the previous clinical rationale, a US-MR image fusion algorithm was recently proposed to estimate a hybrid image gathering the advantages of both modalities in the context of endometriosis diagnosis [1]. The fused image was shown to have a spatial resolution comparable to the US image and SNR/contrast close to the MR image. Our final objective is to use this fused image during surgery using augmented reality. However, the 2D-2D fusion algorithm of [1] requires registered 2D US and MR images. In practical applications, US and MRI examinations are performed separately, resulting in non aligned 2D US images and 3D MRI volumes. A 3D-2D registration step is thus required before fusion, leading to a 2D US image aligned with an oblique MR slice extracted from the 3D volume.

Existing registration approaches for MRI and US can be classified into two categories, namely, intensity-based [2] and feature-based methods [3]. Intensity-based methods use cost functions commonly based on the sum of squared distances [4], the mutual information [5], or the cross-correlation [6]. These methods are automatic and use the entire image without any post-processing. Feature-based methods extract features from the image such as surfaces, volumes, and contours and use them to construct similarity measures between images [3]. Most of the current state-of-the-art algorithms in the medical field focus on 3D/3D [7], [8] or 2D/2D registration, *i.e.*, on the alignment of two volumes or two images. An iconic slice-to-volume registration was also studied in [9] in the context of image guided surgeries. The idea is that low-resolution single photon emission computed tomography can be registered with a high-resolution MRI volume, which could be subsequently fused with live-time interventional MRI. A 3D mutual information registration method is used for the first step, and a robust slice to volume registration algorithm with special features for the latter. A robust patch-based correlation ratio was investigated in [10] for the registration of 3D US and MR images. More recently, deep learning has been suggested as a tool to build efficient registration methods, e.g., learning appropriate features from the images using training data [11].

The main objective of this work is to propose a rigid slice to volume registration algorithm inspired by [12], and to use it as a prior step for US/MR image fusion. The algorithm of [12] was pioneer in using discrete methods to solve the challenging slice-to-volume registration task, giving promis-

ing results when compared to a continuous approach. The results can be further refined using a continuous method, generating more accurate estimations. The fusion algorithm following that registration step is based on two image formation models highlighting the advantages and drawbacks of each modality, as originally proposed in [1]. Both qualitative and quantitative results show the interest of fusing MR and US images. The resulting fused images, benefiting from the 3D-2D registration step, is better resolved than the MR slice and less degraded by noise compared to the US image.

The remainder of the paper is organized as follows. Section 2 presents the 3D-2D registration algorithm based on a discrete graph-based formulation to solve rigid slice-to-volume registration. The 2D-2D fusion algorithm between US and MR images is summarized in Section 3. Experiments are presented and discussed in Section 4. Conclusions and perspectives are finally reported in Section 5.

2. US-MRI SLICE-TO-VOLUME REGISTRATION

The slice-to-volume registration algorithm proposed herein is based on the method in [12]. A 2D US image is registered to a 3D MR volume, by searching the best matching MR slice, not necessarily following one of the three orthogonal directions of acquisition. This slice is defined by three rotation angles and three translation parameters. Its resemblance to the US image is based on a similarity measure, which can have an important impact on the registration results. Note that the proposed registration approach uses a discrete optimization within a Markov random field framework.

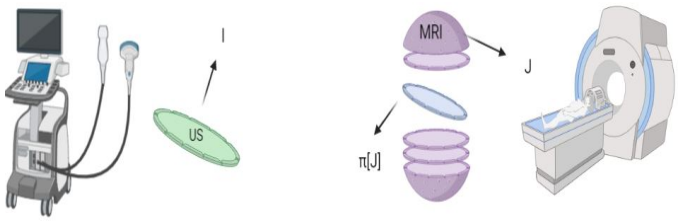


Fig. 1: Schematic view of the proposed registration algorithm. I represents a 2D US image, related by a geometrical transformation π to a best-matching MRI slice $\pi[J]$.

Given an MR volume J and a 2D US image I, the goal of the method is to estimate the rigid transformation π defined by $\pi = (r_x, r_y, r_z, t_x, t_y, t_z)$, containing the rotation angles r and the translations t in the three spatial dimensions denoted by x, y and z, that better aligns I with a slice from J. This is achieved by solving the following optimization problem:

$$\widehat{\pi} = \underset{\pi}{\operatorname{argmin}}\, M(I, \pi[J]), \tag{1}$$

where $\pi[J]$ is a slice extracted from J specified by the rigid transformation π, as illustrated in Fig. 1 and M is the matching criteria that defines the dissimilarity between the 2D image I and the slice $\pi[J]$, which is inversely proportional to a similarity measure. Each parameter of the transformation π will be constrained to a finite set and optimized in order to satisfy (1). More precisely, starting from an initial rigid transformation $\pi_0 = (r_x^0, r_y^0, r_z^0, t_x^0, t_y^0, t_z^0)$, the space of solutions is explored by sampling discrete variations of π_0 to determine the transformation associated with the slice $\pi[J]$ best matching the image I through the similarity measure M. For a maximum size ω_i and a quantization factor k_i, the variations of the variable v_i associated with one component of π are constrained to be in $[0, \pm\frac{\omega_i}{k_i}, \pm\frac{2\omega_i}{k_i}, ..., \pm\frac{k_i\omega_i}{k_i}]$. After each iteration, the maximum size ω_i is decreased by a factor α_i to allow a finer exploration of the search space. The total number of resulting values of v_i at iteration i is $l_i = 2k_i + 1$. Note that 0 is also included because the current parameter value can be preserved. For example, when the rotation variable r_x is considered, for $\omega_0 = 0.4$ and $k_0 = 2$, the search space for v_0 is $[r_x^0, r_x^0 \pm 0.2, r_x^0 \pm 0.4]$. Since the number of possible solutions is exponential, it is not possible with a reasonable time to explore all the possible values of π. Instead, as explained in [12], only variations for all pairs of variables can be considered leading to an optimization algorithm referred to as Fast-PD [13]. Fast-PD is a discrete optimization algorithm based on principles from linear programming and primal dual strategies, which generalizes the α-expansion method [14]. Using the final discrete solution as an initialization, a continuous optimization strategy based on a simplex method is finally run to further improve the result.

3. A 2D/2D US-MR FUSION MODEL

Once the registration between the MR and US images has been performed, the MR slice $\pi[J]$ that best matches the US image I has been extracted. However, the dissimilarities between the two kinds of images, *e.g.*, the differences in resolution, contrast and noise still make it difficult to fully take advantage of the properties of each modality for endometriosis surgery. Our idea is to fuse MR and US images in order to summarize the important information from both modalities and to produce a more informative image that will be superimposed to the video stream collected during surgery using augmented reality. Since MR and US images have different properties, the following two observations models can be used (see [1] for more details):

$$\begin{aligned} \pi[\mathbf{J}] &= \boldsymbol{S}\boldsymbol{H}\boldsymbol{x}_{\mathbf{m}} + \boldsymbol{n}_{\mathbf{m}}, \\ \mathbf{I} &= \boldsymbol{x}_{\mathbf{u}} + \boldsymbol{n}_{\mathbf{u}}, \end{aligned} \tag{2}$$

where $\boldsymbol{x}_m \in \mathbb{R}^N$ is the non-observable high-resolution vectorized MR image, $\pi[J] \in \mathbb{R}^M$ is the low-resolution observed MR image, $\boldsymbol{n}_m \in \mathbb{R}^N$ is an independent identically distributed (i.i.d.) additive white Gaussian noise with variance σ_m^2, $\boldsymbol{H} \in \mathbb{R}^{N\times N}$ is a block circulant with circulant blocks matrix modelling the blurring effect of the MRI point

spread function (PSF) with circulant boundary conditions, $\boldsymbol{S} \in \mathbb{R}^{M \times N}$ (with $N = d_m^2$) is a decimation operator with a decimation factor d. On the other hand, $\boldsymbol{I} \in \mathbb{R}^N$ is the vectorized observed B-mode US image, $\boldsymbol{x}_u \in \mathbb{R}^N$ is the vectorized noise-free US image and $n_u \in \mathbb{R}^N$ is an i.i.d. additive log-Rayleigh noise sequence with localization parameter γ. Following [8] [1], a polynomial function is used to link the two images:

$$\boldsymbol{x}_u = f\left(\boldsymbol{x}_m, \nabla \boldsymbol{x}_m^H \boldsymbol{u}\right), \tag{3}$$

where $f : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}^N$ is an unknown polynomial function of the image $\boldsymbol{x}_m$, its gradient, and the scan direction $\boldsymbol{u}$. The polynomial function will be denoted as $g(\boldsymbol{x}, \boldsymbol{u}) = f\left(\boldsymbol{x}, \nabla \boldsymbol{x}^H \boldsymbol{u}\right)$ to shorten equations. This leads to the following optimization problem:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \underbrace{\frac{1}{2}\|\pi[\mathbf{J}] - \mathbf{SHx}\|^2}_{\text{MRI data fidelity}} + \underbrace{\tau_1\|\nabla \mathbf{x}\|^2 + \tau_3\|\nabla g(\mathbf{x}, \boldsymbol{u})\|^2}_{\text{regularization}} + \underbrace{\tau_2 \sum_{i=1}^{N} \Big[\exp(I_i - g_i(\boldsymbol{x}, \boldsymbol{u})) - \gamma(I_i - g_i(\boldsymbol{x}, \boldsymbol{u}))\Big]}_{\text{US data fidelity}}$$

that can be solved using the proximal alternating linearized minimization (PALM) [1].

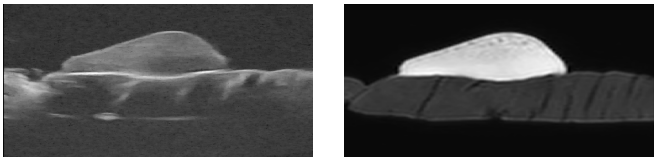
4. EXPERIMENTS

4.1. Phantom data

This section evaluates the proposed joint registration/fusion method combining a rigid 3D-2D registration and an MR and US slice fusion on a phantom data. The experimental phantom was designed to mimic uterus and endometrium responses to MR and US imaging. It was made of a beef steak on top of which was stuck a polyvinyl alcohol (PVA) phantom, using cyanoacrylate instant glue. The beef meat is composed of muscular tissues and its echogenicity and response to MR are similar to those of uterus tissue. The PVA phantom has roughly the same echogeneicity as the beef meat, but has different magnetic properties resulting in high contrast in the MR image. From this viewpoint, its properties are similar to those of endometrium. Finally, the glue between the two structures is visible on US images due to their high resolution and absent in MR image because of its limited resolution, mimicking the depth of penetration information, a crucial element for the surgery. MRI acquisitions were performed using a 3T clinical imaging system (Philips Achieva dStream, Inserm/UPS UMR 1214, ToNIC Technical plateform, Toulouse, France). Axial fat-suppressed T1-weighted sequences (multishot mode; 4 mm slice thickness; voxel matrix $4 \times 1 \times 4$ mm) and axial, sagittal and coronal T2-weighted sequences (multishot mode; 2 mm slice thickness; voxel matrix $0.8 \times 2 \times 2$ mm) were acquired. For US image acquisition, the phantom was immersed in a bucket full of water. US examination was performed using a Voluson S10 system (General Electrics, USA). All images were acquired with a 10-MHz linear array transducer [15].

The sizes of the acquired images are (600×600) for the US, and $(320 \times 320 \times 90)$ for the MRI volume. The field of view of the MR image is wider than that of the US image. Therefore, the MR volume was manually cropped to $(100 \times 100 \times 90)$ to ensure similar fields of view for the two modalities. Finally, a despeckling of the US image and a bicubic interpolation of the MR image were performed to ensure the same pixel size in MR and US images.

For the 3D-2D registration task, the ground truth was manually constructed by searching into the MR volume the best slice matching the 2D US image acquired while scanning the phantom. Fig. 2 shows the MR slice manually registered to the considered US image. The registration algorithm proposed in this work was then tested using different initialization settings, representing MR slices whose positions varied from small to more significant perturbations from the manual ground truth.



(a) US image (b) Ground truth MR

Fig. 2: Pair of MR slice and US image corresponding to manual 3D-2D registration.

The intensity-based metric used for this example is the sum of squared differences (SSD) defined as follows (for two vectorized images $\boldsymbol{X}$ and $\boldsymbol{Y}$):

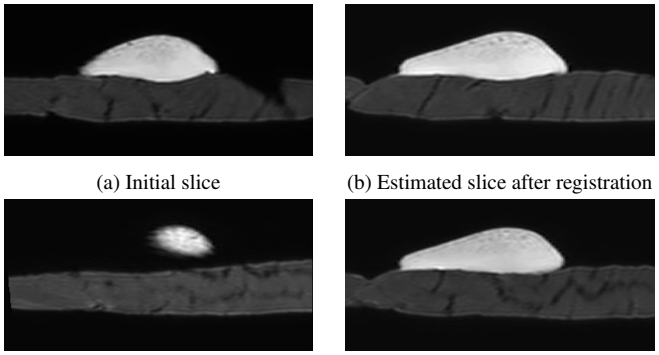
$$\text{SSD}(\boldsymbol{X}, \boldsymbol{Y}) = \sum_{i \in \Omega} [X(i) - Y(i)]^2, \tag{4}$$

where Ω is the region of interest (here the full image). This measure is not always adapted to two different image modalities such as MR and US images. However, it is simple and deserves to be considered as a reference.

The performance of the registration method was evaluated qualitatively through visual inspection of the registered image pairs, and quantitatively using the root mean square error (RMSE) defined as: RMSE = $\sqrt{\frac{1}{N}\|\hat{\boldsymbol{\pi}}[J] - \boldsymbol{x}_{\text{true}}\|_2^2}$ where $\boldsymbol{x}_{\text{true}}$ is the MR slice extracted manually to best match the US image, and $\hat{\boldsymbol{\pi}}[J]$ is the MR slice automatically estimated by the registration algorithm. Moreover, the performance of the proposed algorithm was also evaluated after the fusion step, using the contrast-to-noise ratio (CNR) and the profile slope between two structures computed on the fused image. For two patches extracted from two different structures (PVA phantom and beef steak in this case), the CNR is defined as

CNR $= \frac{|\mu_i - \mu_j|}{\sqrt{\sigma_i^2 + \sigma_j^2}}$ where $\mu_i, \mu_j, \sigma_i^2, \sigma_j^2$ are the means and standard deviations of two blocks of pixels.

Figs. 3 (a) and (c) show two initial slices given as input to the registration algorithm, corresponding to close and far locations from the ground truth. The registration results obtained for these initializations are displayed in Figs. 3 (b) and (d), showing that the estimated MR images are close to the ground truth of Fig. 2 (b) in both cases.



(a) Initial slice (b) Estimated slice after registration

(c) Initial slice (d) Estimated slice after registration

Fig. 3: Image registration results obtained using an initialization slice with a small then an important deviation from the manual ground truth.

The following quantitative analysis focuses on the first set of images presented in Fig. 3. The RMSE between the ground truth and estimated MR images is $7e^{-4}$, which is very promising. This result shows that using SSD as a similarity metric can be sufficient for this experimental case, due to the similarity between the images acquired on the phantom.

As explained previously, the second step of the proposed algorithm consists of fusing the extracted MR slice (after registration) and the US image. Fig. 4 shows the fused image, which has a good contrast, comparable to the MRI, and a good spatial resolution similar to the US image. Moreover, the fused image can differentiate neighbouring structures and highlight small structures as the glue, contrary to the MR image. For example, the part between the PVA phantom and the glue is not distinguishable in the MRI, while it is clearly visible in the US and fused images. These results are confirmed in Table 1 presenting the CNR values and the slopes computed at the frontier between the steak and the glue. Note that the slopes are commonly used as an indicator of spatial resolution in US imaging.

	MRI	US	Fused image
CNR	54.21 dB	32.37 dB	43.17 dB
Slope	0.3×10^{-2}	1.5×10^{-2}	1.8×10^{-2}

Table 1: CNR and slope values after fusion.

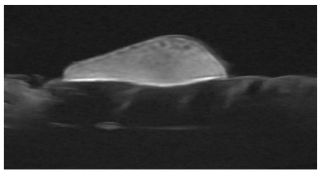


Fig. 4: Fused image

4.2. Synthetic data from real MR acquisition

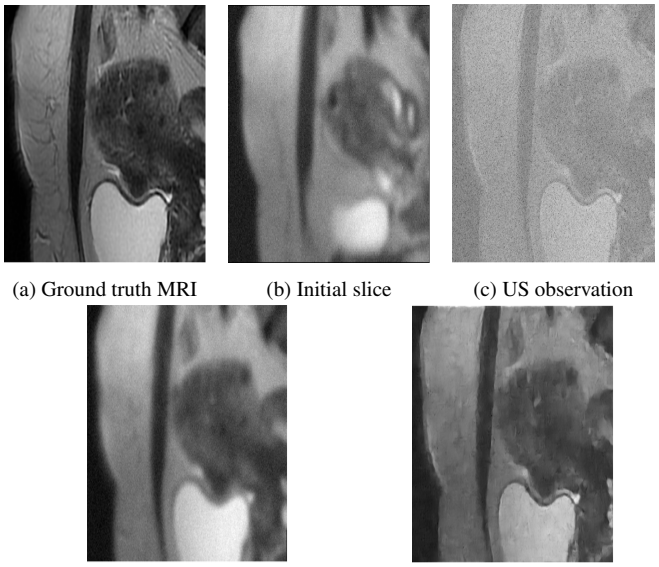
The joint multimodal registration and fusion task is challenging in most medical applications, including the one targeted by this study, i.e., endometriosis detection and localization. The simulations presented in this section have been obtained using a real high resolution MR image that has been degraded to generate an image close to that obtained for endometriosis surgery. The 3D high resolution MR volume corresponds to a real pelvic MRI capturing the uterus, bladder and endometriosis lesions. A blurred and noisy 3D MRI is then generated from this high-resolution MR volume. More precisely, the HR volume was contaminated by an additive white Gaussian noise (SNR = 21.5 dB), and the blurring kernel was a 2D Gaussian filter of standard deviation $\sigma^2 = 4$. The ground truth high-resolution MR image is shown in Fig. 5 (a), whereas the initial slice extracted from the degraded volume and used as an input to the registration algorithm is displayed in Fig. 5 (b). A third order polynomial (see [1] for details) was used to generate the clean US images from the corresponding clean high-resolution MR image, together with log-Rayleigh additive noise, yielding the image displayed in Fig. 5 (c) (SNR= 8 dB).
The intensity-based metric used in this section is the mutual information. It is a measure of image matching that is often used for two different image modalities. It quantifies the "amount of information" provided by one random variable $\boldsymbol{X}$ given that the other random variable $\boldsymbol{Y}$ is observed and is defined as:

$$\mathrm{MI}(\boldsymbol{X}, \boldsymbol{Y}) = \sum_{x \in \Omega_{\boldsymbol{X}}} \sum_{y \in \Omega_{\boldsymbol{Y}}} p(x,y) \log \left[\frac{p(x,y)}{p(x)p(y)} \right], \quad (5)$$

where $\Omega_{\boldsymbol{X}}$ and $\Omega_{\boldsymbol{Y}}$ are the regions of interest for the two images $\boldsymbol{X}$ and $\boldsymbol{Y}$.

Given the initial slice in Fig. 5 (b), Fig. 5(d) shows the estimated registered image, with an RMSE between the ground truth and estimated MR images equal to RMSE = 0.02. The fused image obtained using the two registered US and MR images is displayed in Fig. 5(e). This image provides a good compromise between the US and MR data. Specifically, the fused image is less affected by US speckle and MRI blur, provides well-defined contours and good contrast compared to the native MR and US images.

In addition to the visual inspection of the different images, CNR was used to evaluate the contrast between two different



(a) Ground truth MRI (b) Initial slice (c) US observation

(d) Extracted slice after registration (e) Fused image

Fig. 5: (a) True high resolution MR image. (b) Initial MRI slice (MR low-resolution and blurred image). (c) US image (polynomial function of the noiseless MRI with additive log-Rayleigh noise). (d) Registration result. (e) Fused image.

structures of the images. The two regions considered here are extracted from the uterus and the bladder. The CNR values for the MR and the US images are respectively 41 dB and 19.17 dB. The final fused image had a CNR of 39.01 dB, which clearly demonstrates that the fusion process improves the contrast in the images compared to the US image.

5. CONCLUSION

This paper introduced a new framework for rigid slice to volume 3D-2D registration and 2D-2D fusion for MR and US images. The performance of the proposed joint registration/fusion algorithm was evaluated on experimental and realistic synthetic data having characteristics in agreement with endometriosis, leading to promising results. An important perspective of this work is to introduce non rigid registration to further improve the results. A long term objective of this work is to merge the fused image with the video stream collected during laparoscopy, allowing a safer decision-making and therefore a more precise endometriosis surgery.

6. REFERENCES

[1] O. El Mansouri et al., "Fusion of Magnetic Resonance and Ultrasound images for endometriosis detection," *IEEE Trans. Image Process.*, vol. 29, 2020.

[2] F. Maes et al., "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, 1997.

[3] A. Yavariabdi et al., "Mapping and characterizing endometrial implants by registering 2D transvaginal ultrasound to 3D pelvic magnetic resonance images," *Comput. Med. Imaging Graph.*, vol. 45, 2015.

[4] M. Fogtmann et al., "A unified approach to diffusion direction sensitive slice registration and 3-D DTI reconstruction from moving fetal brain anatomy," *IEEE Trans. Med. Imag.*, vol. 33, no. 2, 2014.

[5] J. P. Pluim et al., "Mutual-information-based registration of medical images: a survey," *IEEE Trans. Med. Imag.*, vol. 22, no. 8, 2003.

[6] J. N. Sarvaiya et al., "Image registration by template matching using normalized cross-correlation," in *Proc. Int. Conf. on Advances in Computing, Control, and Telecommunication Technol.*, Bangalore, India, 2009.

[7] A. Atehortúa et al., "Fusion of 3D real-time echocardiography and cine mri using a saliency analysis," *Int. J Comput Assist Radiol Surg*, vol. 15, no. 2, 2020.

[8] A. Roche et al., "Rigid registration of 3D ultrasound with MR images: a new approach combining intensity and gradient information," *IEEE Trans. Med. Imag.*, vol. 20, no. 10, 2001.

[9] F. Baowei et al., "Registration and fusion of SPECT, high-resolution MRI, and interventional MRI for thermal ablation of prostate cancer," *IEEE Trans. Nuclear Science*, vol. 51, no. 1, 2004.

[10] H. Rivaz et al., "Automatic deformable MR-Ultrasound registration for image-guided neurosurgery," *IEEE Trans. Med. Imag.*, vol. 34, no. 2, 2015.

[11] G. Haskins et al., "Deep learning in medical image registration: a survey," *Machine Vision and Applications*, vol. 31, no. 1-2, 2020.

[12] R. Porchetto et al., "Rigid slice-to-volume medical image registration through Markov random fields," *CoRR*, vol. abs/1608.05562, 2016.

[13] N. Komodakis et al., "Fast, approximately optimal solutions for single and dynamic MRFs*," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Minneapolis, MN, USA, 2007.

[14] N. Komodakis et al., "Performance vs computational efficiency for optimizing single and dynamic MRFs: Setting the state of the art with primal-dual strategies," *Comput. Vis. and Image Understanding*, vol. 112, 2008.

[15] F. Vidal et al., "On the design of a pelvic phantom for Magnetic Resonance and Ultrasound image fusion," in *In Proc. IEEE IUS Symp.*, Glasgow, UK, 2019.