



**HAL**  
open science

## Fusion of 3D real-time echocardiography and cine MRI using a saliency analysis

Angélica Atehortúa, Mireille Garreau, Antoine Simon, Erwan Donal, Mathieu Lederlin, Eduardo Romero

► **To cite this version:**

Angélica Atehortúa, Mireille Garreau, Antoine Simon, Erwan Donal, Mathieu Lederlin, et al.. Fusion of 3D real-time echocardiography and cine MRI using a saliency analysis. *International Journal of Computer Assisted Radiology and Surgery*, 2020, 15 (2), pp.277-285. 10.1007/s11548-019-02087-w . hal-02364096

**HAL Id: hal-02364096**

**<https://univ-rennes.hal.science/hal-02364096v1>**

Submitted on 11 Dec 2019

**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.

# 1 Fusion of 3D real-time echocardiography and cine 2 MRI using a saliency analysis

3 **Angélica Atehortúa · Mireille Garreau ·**  
4 **Antoine Simon · Erwan Donal · Mathieu**  
5 **Lederlin · Eduardo Romero**

6  
7 Received: date / Accepted: date

## 8 **Abstract**

### 9 **Purpose:**

10 This paper presents a novel 3D multimodal registration strategy to fuse 3D  
11 real-time echocardiography images with cardiac cine MRI images. This align-  
12 ment is performed in a saliency space, designed to maximize similarity between  
13 the two imaging modalities. This fusion improves the quality of the available  
14 information.

### 15 **Methods:**

16 The method performs in two steps: temporal and spatial registrations. A tem-  
17 poral alignment is firstly achieved by non-linearly matching pairs of correspon-  
18 dences between the two modalities using a dynamic time warping. A temporal  
19 registration is then carried out by applying non-rigid transformations in a  
20 common saliency space where normalized cross correlation between temporal  
21 pairs of salient volumes is maximized.

### 22 **Results:**

23 The alignment performance was evaluated with a set of 18 subjects, 3 with  
24 cardiomyopathies and 15 healthy, by computing the Dice score and Haus-  
25 dorff distance with respect to manual delineations of the left ventricle cavity  
26 in both modalities. A Dice score and Hausdorff distance of  $0.86 \pm 0.04$  and  
27  $13.61 \pm 3.86$  mm respectively were obtained. In addition, the deformation field

---

A. Atehortúa (orcidID 0000-0002-6192-1757)  
Univ Rennes, CHU Rennes, Inserm, LTSI - UMR 1099, F-35000 Rennes, France  
Tel.: +33 02 23 23 62 20  
E-mail: amatehortual@univ-rennes1.fr

M. Garreau · A. Simon (orcidID 0000-0001-6023-6427) · E. Donal (orcidID 0000-0002-9083-  
1582) · M. Lederlin (orcidID 0000-0003-1388-5277)  
Univ Rennes, CHU Rennes, Inserm, LTSI - UMR 1099, F-35000 Rennes, France  
E-mail: {mireille.garreau,antoine.simon,erwan.donal,mathieu.lederlin}@univ-rennes1.fr

A. Atehortúa · E. Romero (orcidID 0000-0003-2088-2509)  
Universidad Nacional de Colombia, Bogotá, Colombia  
E-mail: edromero@unal.edu.co

was estimated by quantifying its foldings, obtaining a 98% of regularity in the deformation field.

### Conclusions:

The 3D multimodal registration strategy presented is performed in a saliency space. Unlike state of the art methods, the presented one takes advantage of the temporal information of the heart to construct this common space, ending up with two well aligned modalities and regular deformation fields. This preliminary study was evaluated on heterogeneous data composed by two different datasets, healthy and pathological cases, showing similar performances in both cases. Future work will focus on testing the presented strategy in a larger dataset with a balanced number of classes.

**Keywords** Image registration · spatio-temporal saliency · cardiac cine MRI images · echocardiography

## 1 Introduction

Cardiac imaging plays an important role in diagnosis and treatment planning of prevalent heart diseases. Non-invasive procedures, specifically 3D real-time echocardiography (3DUS+t) and cine cardiac MRI (cMRI), provide structural and functional information of the heart. While 3DUS+t is a first option for cardiac evaluation and follow-up, cMRI is used when a diagnosis doubt persists after the ultrasound examination [9]. In pathologies such as pulmonary arterial hypertension, ischaemic heart disease, arrhythmogenic ventricular cardiomyopathy and atrial arrhythmias [24,18,2], the information obtained with these two modalities results to be complementary. Under these considerations, multimodal image fusion should provide a more accurate and reliable examination by combining morphological (cMRI) and functional (3DUS+t) assessments [18,24]. This fusion could facilitate anatomo-functional correlations of cardiac abnormalities that lead to design new diagnostic and follow-up markers [6,17,2] to detect early impairments and subtle changes in the cardiac function. Consequently, a better therapy follow-up could be performed to improve patient categorization, personalized treatment and survival.

Image fusion is generally carried out by a registration process [12]. However, a useful registration between 3DUS+t and MRI images is a complex and difficult process, basically because the represented information is captured using very different physical sensors. Furthermore, 3DUS+t images have a lower signal-to-noise ratio (SNR) than 2DUS+t images. There are also several factors that increase a poor definition of the left ventricular chamber and walls during the acquisition in echocardiography, such as anatomic particularities (obesity, asthenic habitus) and other related information, such as acquired disease that affects the acoustic window.

Few works in the state of the art have addressed the challenge of registering 3D echocardiography and cine MRI images. These approaches can be divided in two groups: intensity-based registration [10,28,13] and surfaced-based registration [15,20]. In the first group, works map MRI and 3DUS modalities to a

71 common space using only spatial information. Some of them transformed both  
 72 modalities into their corresponding phase maps using Fourier transform [28].  
 73 However, in some cases, this phase representation is very similar to the orig-  
 74 inal image and may lead to erroneous registrations. Other works [10,13,27]  
 75 align the modalities in the intensity space, where the registration is prone to  
 76 error due to the high data heterogeneity and the noisy US images. For the  
 77 second group, surfaced-based registration, cardiac structures are segmented  
 78 in both modalities and then, a surface extracted from one modality is non  
 79 rigidly-registered to the other surface modality. In this group, strategies based  
 80 on the registration of common landmarks in both modalities are also taken  
 81 into account. Nevertheless, these approaches are highly sensitive to the seg-  
 82 mentation and precise setting of landmarks, due to the inter- and intra-shape  
 83 variability of the heart, its complex dynamic and the image quality.

84 Unlike these works, which are based only on spatial registration or surface  
 85 deformations, this paper proposes a novel 3D multimodal registration frame-  
 86 work by combining saliency and temporal information, which is applied to  
 87 cMRI and 3DUS+t images. In this work, a non rigid deformation is performed  
 88 to cope with local differences between images, since that heart is intrinsically  
 89 deformable, even if intra-patient cases are considered. Although a preliminary  
 90 work of this study has been previously published as a conference paper [30],  
 91 this paper is substantially different from that work in that it incorporates an  
 92 initial registration to perform a more robust alignment between both modal-  
 93 ities and it includes more experimental validation results and the use of quan-  
 94 titative metrics for the evaluation.

95 The paper is organized as follows: In Section 2 the registration framework  
 96 is described. Experiments and results are provided in Section 3. Finally, con-  
 97 clusions are presented in Section 4.

## 98 2 Methodology

99 Each image modality (cMRI noted as  $MR$  and 3DUS+t noted as  $US$ ) is rep-  
 100 resented as a sequence  $I(x, y, z, t)$  of images, where  $(x, y, z) \in \Omega_I$  corresponds  
 101 to the spatial coordinates in the domain of the image ( $\Omega_I$ ) and  $t \in \{1, T\}$   
 102 represents a temporal position in the sequence, being  $T$  the number of frames  
 103 per cardiac cycle for each modality, MR and US, which can be different, i.e.,  
 104  $T^{mr}$  and  $T^{us}$ . This work proposes a 3D registration strategy to map each voxel  
 105 of the moving modality to its corresponding voxel in the reference modality.  
 106 The registration process is decoupled into two temporal and spatial steps of  
 107 alignment, that are hereafter described.

### 108 2.1 Temporal alignment

109 Firstly, a temporal alignment sets the frame correspondences between both  
 110 image sequences. For doing so, two sub-steps are performed: i) an image-

111 based cardiac dynamic descriptor and ii) a nonlinear mapping between both  
 112 modalities, using the method proposed by Betancur et al [4].

### 113 2.1.1 Image-based cardiac dynamic descriptor.

114 A  $\Phi(t; I)$  curve that describes events of the cardiac cycle consistently [19,  
 115 27,4] is constructed using all volumes or image sequence  $I_t$  of the cardiac  
 116 cycle. This process is independently performed for each image sequence  $I_t$ , by  
 117 computing the similarity between the volume at the end of the diastole (ED)  
 118 with respect to each of the remaining volumes of the cycle, Eq. 1. The ED  
 119 frame is arbitrarily set to the first frame of the sequence,  $I_{t=1}$ . The classical  
 120 normalized cross correlation is used as similarity measure since this measure  
 121 is less sensitive to different noisy conditions.

$$122 \quad \Phi(t; I_t) = \left\langle \frac{I_{t=1} - \bar{I}_{t=1}}{\|I_{t=1} - \bar{I}_{t=1}\|}, \frac{I_t - \bar{I}_t}{\|I_t - \bar{I}_t\|} \right\rangle \quad \forall t \in \{1, T\} \quad (1)$$

123 being  $\bar{I}_t$  the average intensity value of the image  $I_t$ .

### 124 2.1.2 Nonlinear mapping between the images in the reduced space.

125 Due to the non-linear nature of the cardiac function, a dynamic time warping  
 126 (DTW) [3] is applied to map each point of the curve  $\Phi^{mr}(t)$  to its correspond-  
 127 ing point of the curve  $\Phi^{us}(t)$ , being  $\Phi(t)$  the curve described in equation 1. As  
 128 a result of this mapping, a set of aligned temporal pairs  $A = \{(t_i^{us}, t_j^{mr})\}$  is  
 129 obtained for each subject. Despite that several US volumes could be matched  
 130 with the same MR volume or the other way around, this does not pose any  
 131 issue because this temporal alignment matches volumes that are close in time  
 132 and from the same cardiac phase.

## 133 2.2 Spatial alignment

134 A main contribution of this work is to spatially register the image sequences by  
 135 using spatio-temporal information of the cardiac function in the 3D domain.  
 136 For doing such registration, two sub-steps are performed: i) mapping the image  
 137 sequence to a common saliency space and ii) setting customized registration  
 138 metric in the saliency space.

### 139 2.2.1 Mapping the image sequence to a saliency space.

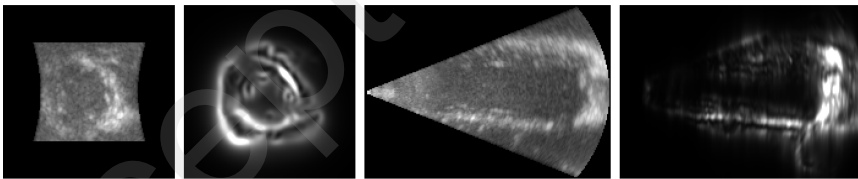
140 Multimodal cardiac images are statistically dependent by the fact of being cap-  
 141 tured from the very same object, i.e., the heart, and such dependency could be  
 142 used to fuse them. The human visual system (HVS) is able to extract relevant  
 143 information by decomposing the scene into low level representations which are  
 144 integrated and encoded as visual information [22]. Under such hypotheses, the

145 salient volume is built upon a multi-scale region-based decomposition which  
 146 preserves the most correlated information and is obtained by applying a classic  
 147 Gaussian pyramidal filtering [5]. For every image sequence, five scales  $s$  are  
 148 taken, namely the original scale, a finer level and three incrementally coarser  
 149 levels of the original. The finer level provides more details of the heart, while  
 150 the coarser levels provide a global localization of the heart. For each scale  $s$ ,  
 151 regions  $r$  are extracted by using the simple linear iterative clustering (SLIC)  
 152 algorithm [1]. The Cartesian coordinates of these regions are stored as a grid  
 153  $G$  since every processing is from here on performed exclusively within this  
 154 grid. Afterwards, each image is filtered for each scale  $s$  and for each instant  
 155 of the time  $t$  by applying steerable filters [7]. These filters correspond to an  
 156 oriented first derivative of a 3D Gaussian with a standard deviation set to 3,  
 157 due to the noise in cMRI and 3DUS images can be modeled by Gaussian or  
 158 Laplacian distribution [29]. The result of this bank of filters is an image  $SV_t$   
 159 for each scale  $s$  where their structural features, such as intensity, orientation  
 160 and edges, are enhanced, while it reduces noise.

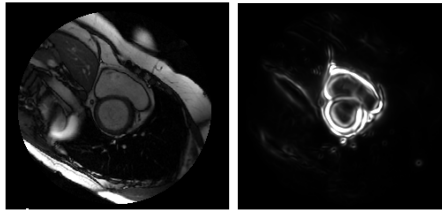
161 Saliency is defined by a consistent difference between regional  $r$  features  
 162 and their surroundings during the time, Eq. 2. Independently of each scale  $s$ ,  
 163 the salient volume at time  $t$  is estimated by comparing it with respect to the  
 164 remaining frames. This ensures temporal coherence of the salient region, even  
 165 in scenes with a higher level of complexity, such as low frame rate and noisier  
 166 sequences [11,26].

$$167 \quad \text{saliency}(t, r, s) = \sum_{\tau \in T} |SV_t(r, s) - SV_\tau(r, s)|^2 \quad \forall r \in G \quad (2)$$

168 The final salient volume  $S$  is obtained by averaging saliency for every scale,  
 169 previously set to the original size either by interpolation or decimation. This set  
 170 of salient volumes constitutes the saliency space. Figures 1 and 2 show exam-  
 171 ples of saliency measured in echocardiography and magnetic resonance imaging  
 172 respectively. Observe how the myocardium is highlighted in both modalities.  
 173 This representation provides a common space where both image modalities  
 174 can be spatially aligned.



**Fig. 1** Salient images in echocardiography. Panels from left to right: original image in axial view, salient image in axial view, original image in coronal view and salient image in coronal view.

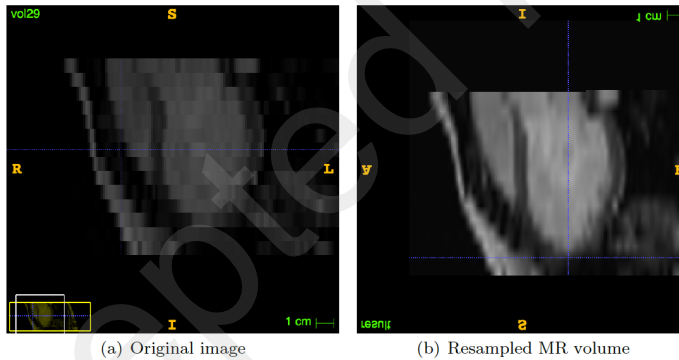


**Fig. 2** Saliency image in short axis cine MRI. Left panels shows the original image and the right panel illustrates its respective saliency.

### 175 2.2.2 Setting customized registration metric in the saliency space.

176 Since the field of view is not the same in both modalities, a coarse manual  
 177 delineation of the epicardium of the left ventricle (LV) at the end of the diastole  
 178 (ED) is obtained to delimit a common region of interest (ROI). Herein, this LV  
 179 ROI is used for the whole cardiac cycle, under the hypothesis that myocardium  
 180 contracts towards the center of the LV and such ROI encloses all the cardiac  
 181 structures of interest [16].

182 The intensity image sequence  $I_t$ , the LV ROI and the saliency image  $S$   
 183 are oriented to the same Right-Anterior-Inferior radiological coordinates  $RAI$ ,  
 184 defining a common  $x, y, z$  order, using the ITK-snap tool [25]. This orientation  
 185 only describes the mapping from the voxel coordinate system to the phys-  
 186 ical coordinate system. As 3DUS+t modality has a higher spatial resolution  
 187 than cMRI, the oriented LV ROI, the salient volume and the intensity im-  
 188 age sequence from cMRI modality are linearly resampled to obtain the same  
 189 dimensions and voxel size of the echocardiographic image. Figure 3 shows a  
 190 resampled MRI image for a healthy subject in the longitudinal direction.



**Fig. 3** Example of a resampled MRI volume in the longitudinal direction.

191 Afterwards, a spatial registration is performed in the saliency space de-  
 192 limited by the LV ROI. For each temporal pair  $(t_i^{us}, t_j^{mr}) \in A$  obtained as  
 193 described in section 2.1, a mapping-salient function  $\mathcal{F}_t$  sets each voxel in the

194 cMRI image to its corresponding voxel in the 3DUS domain at the instant  
 195 of time  $\hat{t} = (t_i^{us}, t_j^{mr})$ . The mapping-salient function for each temporal pair is  
 196 then defined as:

$$197 \quad \mathcal{F}_{\hat{t}}(S_{t_i}^{us}, S_{t_j}^{mr}) = \underset{T_x}{argmax} \quad sim(S_{t_i}^{us}, S_{t_j}^{mr}(T_x); T_x) \quad (3)$$

198 where  $T_x$  is a transformation model that maximizes the similarity (*sim*)  
 199 between both salient volumes,  $S_{t_i}^{us}$  and  $S_{t_j}^{mr}$ . In this work,  $T_x$  is a non-rigid  
 200 transformation applied to the salient volumes via interpolation with a B-spline  
 201 of order 3 that maximizes the normalized cross correlation. This similarity  
 202 measure is used since it is invariant to affine intensity distortions [8]. Herein,  
 203 an adaptive stochastic gradient descent [14] was used as optimizer. In order to  
 204 achieve an accurate alignment, this non-rigid registration was initialized by an  
 205 Eulerian transformation [21] that rigidly aligns the LV ROI centroids of both  
 206 modalities.

207 Once the mapping-salient function  $\mathcal{F}_{\hat{t}}$  is obtained for each temporal pair  
 208  $\hat{t}$ , it is applied to the original image sequence cMRI,  $I_{t_j}^{mr}(\mathcal{F}_{\hat{t}})$ , mapping it to  
 209 the 3D echocardiography image domain  $I_{t_i}^{us}$ . This process is applied for all  
 210 the aligned temporal pairs and in consequence, the whole cMRI sequence is  
 211 covered.

## 212 3 Experiments and results

### 213 3.1 Image Data

214 Two datasets of pairs of cMRI and 3DUS+t data were used: healthy subjects  
 215 and pathological subjects.

#### 216 3.1.1 Healthy group.

217 This group comprises 15 healthy subjects [23]. cMRI and 3DUS+t acquisitions  
 218 were obtained within an interval of  $3.5 \pm 3.3$  days. Acquisition features for  
 219 cMRI and 3DUS+t are, respectively: temporal resolution of 30 and  $14.93 \pm$   
 220  $3.03$  frames per second; slice thickness of 8 mm and  $0.71 \pm 0.09$  mm; and in-  
 221 plane resolution voxel size of  $1.22 \pm 0.05$  mm and  $0.84 \pm 0.16$  mm. Delineated  
 222 volumes of the left ventricle cavity are provided with each modality for all the  
 223 cardiac cycle.

#### 224 3.1.2 Pathological group.

225 It is composed of 3 subjects with dilated and hypertrophic cardiomyopathy,  
 226 who followed an acquisition protocol comprising both cMRI and 3DUS+t. This  
 227 dataset was acquired at CHU-Pontchaillou in Rennes, France. This study was  
 228 part of a systematic database review conducted following the Declaration of  
 229 Helsinki and approved under the CNIL (National Commission on Informatics



and Liberty of France). Acquisition features for cine MRI and 3DUS+t were, respectively: temporal resolution of  $28.33 \pm 5.77$  and  $22.33 \pm 14.98$  frames per second; slice thickness of 8 mm and  $0.74 \pm 0.11$  mm; and in-plane resolution voxel size of  $1.62 \pm 0.1231$  mm and  $1.08 \pm 0.15$  mm. Delineated volumes of the left ventricle cavity are manually provided in each modality at the end of the diastole and end of the systole.

The method performance was evaluated for each step, temporal and spatial alignment, as follows.

### 3.2 Evaluation of the temporal alignment

The temporal alignment was validated for each image modality and dataset by computing the Mean Absolute Error (MAE) between the time point corresponding to the end of the systole (ES), which was manually set by an experienced observer, and the time point with the minimum value of the curve  $\Phi$  obtained in Section 2.1.1. This MAE measures the distance in number of frames between the ES position estimated by the method and the position set by the expert. The MAE for the healthy group was  $2.73 \pm 1.67$  frames and  $0.93 \pm 0.70$  frames for cMRI and 3DUS+t, respectively. While the MAE for the pathological group was  $1.0 \pm 1.0$  frames and  $2.33 \pm 1.53$  frames for cMRI and 3DUS+t, respectively.

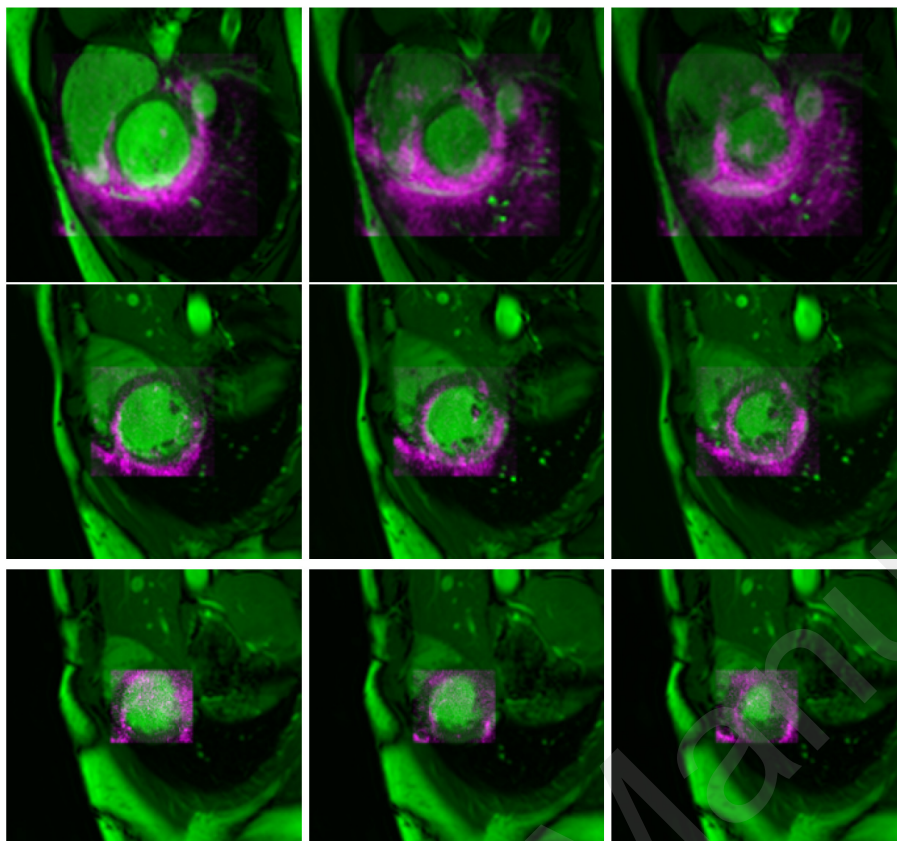
### 3.3 Evaluation of the spatial alignment

The results differ according to the temporal duration and the phases that were considered in the saliency estimation. Different tests have been conducted, along the heart cycle providing the results presented in the paper and along each cardiac period (diastole and systole) providing some differences. These second results showed a slight difference in accuracy about the 5%. In this sense, the complete cardiac cycle is used to compute the saliency in each modality.

Qualitative results of the spatial alignment are illustrated for one pathological case in Fig. 4, from the base to the apex (in rows) at four instants of time presented in the columns (End of the diastole ED, end of the systole ES, intermediate time between the ED and ES, and intermediate time between the ES and the final frame). Observe the good spatial alignment between cMRI (green color) and 3DUS+t (magenta color) volumes using the saliency descriptor and the registration framework, in particular for the basal, basal-mid and mid views, i.e., the myocardium is good overlapped in both modalities.

To quantitatively evaluate the registration performance, aligned delineations of the LV cavity from both modalities were compared at the ED and ES by computing the Dice Score and Hausdorff distance.

Table 1 and Fig. 5 show the spatial registration performance of the presented strategy, when it is evaluated in terms of the Dice score and Hausdorff



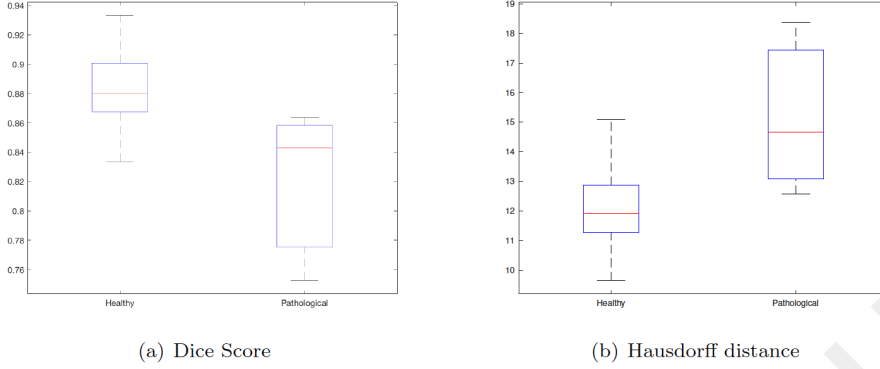
**Fig. 4** Illustration of a spatial alignment for one pathological case. MRI (green color) superimposed to the 3DUS (magenta color). Rows correspond to the volume levels, from top to bottom: Basal, mid and apex-mid. Columns correspond to the instants of time, from left to right: end of the diastole ED, end of the systole ES and intermediate time between ED and ES.

270 distance. The method achieves accurate correspondences between both modalities  
271 with an average Dice score higher than 0.8, a level that could be clinically  
272 acceptable. However, a registration method can be useless if there are too  
273 many foldings in the deformation field, even for cases of good match. Hence,  
274 regularity of the deformation field, for the whole dataset, was evaluated by  
275 computing the number of voxels for which the determinant of the Jacobian is  
276 non-positive. This analysis showed only  $0.2 \pm 0.3\%$  of the voxels exhibited such  
277 values. In addition, registration with the elastix [31] software corresponded to  
278 the nonrigid B-spline method [35,34], which has been described to robustly  
279 operate in similar problems [32,33].

280 In addition, the benefit of this method has been evaluated by computing  
281 the Dice Score and the Hausdorff distance at each step of the method to see

**Table 1** Spatial registration performance, for healthy and pathological groups, by computing the Dice score and the Hausdorff distance

Measure	Healthy	Pathological
Dice Score	$0.88 \pm 0.03$	$0.83 \pm 0.08$
Hausdorff distance (mm)	$12.02 \pm 1.36$	$15.20 \pm 2.93$

**Fig. 5** Boxplot that shows the minimum and maximum values for a) Dice Score and b) Hausdorff distance computed in the healthy and pathological datasets

282 the contribution of the registration step in the saliency space. The results are  
 283 presented in Table 2. Columns show the results about the type of transforma-  
 284 tions applied in the saliency space. The first column represents the comparison  
 285 of the LV surfaces without performing any transformation, i.e, the original LV  
 286 surfaces compared between them. The remaining columns are then results of  
 287 the initial, affine and non-rigid transformations applied to the salient volumes.  
 288 Note how, the Dice Score value passed from 0.40 to 0.76 after applying the  
 289 initial rigid transformation and then, the registration performance improved  
 290 to 0.88, when the final stage, non-rigid registration, was performed. The same  
 291 observations can be made for the Hausdorff distance, in which a value of  
 292 13.01 mm was obtained after applying the initial registration, while a value of  
 293 12.02 mm was estimated at the final stage.

**Table 2** Spatial registration performance in each transformation: without any transformation (without), initial registration and final registration (affine and non-rigid), using the healthy dataset

Measure	Transformation			
	Without	Initial	Affine	Nonrigid
Dice Score	$0.40 \pm 0.13$	$0.76 \pm 0.7$	$0.80 \pm 0.05$	$0.88 \pm 0.03$
Hausdorff distance (mm)	$15.30 \pm 13.40$	$13.01 \pm 0.13$	$12.89 \pm 0.13$	$12.02 \pm 1.36$

**Table 3** Spatial registration performance in the intensity space, for healthy and pathological groups, by computing the Dice score and the Hausdorff distance

Measure	Healthy	Pathological
Dice Score	0.77±0.05	0.79 ± 0.08
Hausdorff distance (mm)	19.34 ± 3.67	19.32 ± 4.04

294 Moreover, two additional experiments to emphasize the advantage of the  
 295 method were performed. The first experiment consisted of aligning the inten-  
 296 sity images using the same process computed for the salient volumes. The  
 297 results are presented in Table 3, showing that the Dice Score measure was  
 298 of 0.77 for 15 healthy cases and 0.79 for pathological cases, in comparison  
 299 with the presented method, 0.88 and 0.83 for healthy and pathological sub-  
 300 jects, respectively. In a similar way, the Hausdorff distance results were of  
 301 19.34 mm and 19.32 mm for healthy and pathological subjects, while 12.02 m  
 302 and 15.20 m for the presented strategy. These results highlight the advantage  
 303 of using the saliency space to align the modalities. The second experiment con-  
 304 sisted of adding Gaussian noise (with variance of 0.02) to the intensity images  
 305 and after applying the registration in the saliency space. **For this experiment,**  
 306 **a subset composed by five randomly selected healthy subjects was tested, re-**  
 307 **porting a lower value of the Dice score and Hausdorff distance about 10% with**  
 308 **respect to the results shown in Table 1.** These results suggest the saliency in  
 309 echocardiography is affected by the noise, due to the fact that cardiac struc-  
 310 tures can not be well recognized by fuzzy edges. This could be mitigated by  
 311 performing a robust 3D denoising algorithm that improves the image quality  
 312 in echocardiography.

313 As mentioned in the introduction, this paper is substantially different to the  
 314 work published in [30], since a more detailed evaluation was performed and  
 315 a larger dataset was used. **In addition,** the Dice score was computed for the  
 316 resulting aligned LV surfaces from pathological subjects using the registration  
 317 strategy described in that published work. **A Dice score of 0.78 was reported**  
 318 **in that published work, while the presented method outperforms such work**  
 319 **with a Dice score of 0.83, about 7%..** The advantages from this new work is  
 320 the use of a 3D saliency, which is applied to each modality and the initial reg-  
 321 istration. Unlike the previous work, which was performed using a 2D approach  
 322 and without any initial alignment.

## 323 4 Conclusions

324 A novel 3D multimodal registration method was proposed. This strategy in-  
 325 cludes the temporal information of the heart to obtain a common space of  
 326 reference based on a saliency measure, where both modalities are aligned.  
 327 The method has been tested on 15 healthy cases and 3 patients with car-  
 328 diomyopathies. The evaluation has been realized by a visual and qualitative

way and by a quantitative strategy using Dice Score and 3D Hausdorff measures. Results have shown an interest of this method based on registration in saliency space by comparison to a registration approach based only on intensity. In addition, the results show an improvement when the initial rigid registration step results are compared with the results obtained at the final stage by using the saliency. Results have shown a good alignment between cMRI and 3DUS+t, even if the method has to be evaluated in a quantitative way on a large database. Future research will focus on improving the 3DUS image quality to obtain more reliable region of interest of the left ventricle. In addition, a rigid transformation between the modalities in the saliency space will be explored to study the effects of the heart deformation by our algorithm. The presented strategy will be tested with a larger dataset of healthy and pathological subjects and will be compared with other methods of the literature.

### Acknowledgments

This work was supported by Colciencias-Colombia, Grant No. 647 (2015 call for National PhD studies), and Région Bretagne in the framework of the Investissement d'Avenir Program through Labex CAMI (ANR-11-LABX-0004).

### Conflict of Interest

The authors declare that they have no conflict of interest.

### References

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: Slic superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelligence* 34(11), 2274-2282 (2012)
2. Badesch, D.B., Champion, H.C., Sanchez, M.A.G., Hoepfer, M.M., Loyd, J.E., Manes, A., McGoon, M., Naeije, R., Olschewski, H., Oudiz, R.J.: Diagnosis and assessment of pulmonary arterial hypertension. *Journal of the American College of Cardiology* 54(1 Supplement), S55-S66 (2009)
3. Berndt, D.J., Clifford, J.: Using dynamic time warping to find patterns in time series. In: *KDD workshop*. vol. 10, pp. 359-370. Seattle, WA (1994)
4. Betancur, J., Simon, A., Langella, B., Leclercq, C., Hernández, A., Garreau, M.: Synchronization and registration of cine magnetic resonance and dynamic computed tomography images of the heart. *IEEE journal of biomedical and health informatics* 20(5), 1369-1376 (2016)
5. Burt, P., Adelson, E.: The laplacian pyramid as a compact image code. *IEEE Transactions on communications* 31(4), 532-540 (1983)
6. Dahlem, K., Michels, G., Kobe, C., Bunck, A.C., Ten Freyhaus, H., Pfister, R.: Diagnosis of cardiac transthyretin amyloidosis based on multimodality imaging. *Clinical Research in Cardiology* 106(6), 471-473 (2017)
7. Freeman, W.T., Adelson, E.H.: The design and use of steerable filters. *IEEE Transactions on Pattern Analysis & Machine Intelligence* (9), 891-906 (1991)

- 370 8. Grewenig, S., Zimmer, S., Weickert, J.: Rotationally invariant similarity measures for  
371 nonlocal image denoising. *Journal of Visual Communication and Image Representation*  
372 22(2), 117-130 (2011)
- 373 9. Haugaa, K.H., Basso, C., Badano, L.P., Bucciarelli-Ducci, C., Cardim, N., Gaemperli,  
374 O., Galderisi, M., Habib, G., Knuuti, J., Lancellotti, P.: Comprehensive multi-modality  
375 imaging approach in arrhythmogenic cardiomyopathy expert consensus document of the  
376 european association of cardiovascular imaging. *European Heart Journal-Cardiovascular*  
377 *Imaging* 18(3), 237-253 (2017)
- 378 10. Huang, X., Hill, N.A., Ren, J., Guiraudon, G., Boughner, D., Peters, T.M.: Dynamic 3d  
379 ultrasound and mr image registration of the beating heart. In: *International Conference*  
380 *on Medical Image Computing and Computer-Assisted Intervention*. pp. 171-178. Springer  
381 (2005)
- 382 11. Itti, L.: Quantifying the contribution of low-level saliency to human eye movements in  
383 dynamic scenes. *Visual Cognition* 12(6), 1093-1123 (2005)
- 384 12. James, A.P., Dasarathy, B.V.: Medical image fusion: A survey of the state of the art.  
385 *Information Fusion* 19, 4-19 (2014)
- 386 13. Kiss, G., Thorstensen, A., Amundsen, B., Claus, P., D'hooge, J., Torp, H.: Fusion of 3d  
387 echocardiographic and cardiac magnetic resonance volumes. In: *2012 IEEE International*  
388 *Ultrasonics Symposium*. pp. 126-129. IEEE (2012)
- 389 14. Klein, S., Pluim, J.P., Staring, M., Viergever, M.A.: Adaptive stochastic gradient descent  
390 optimisation for image registration. *International journal of computer vision* 81(3), 227  
391 (2009)
- 392 15. Ma, Y.L., Penney, G.P., Rinaldi, C.A., Cooklin, M., Razavi, R., Rhode, K.S.: Echocar-  
393 diography to magnetic resonance image registration for use in image-guided cardiac  
394 catheterization procedures. *Physics in medicine and biology* 54(16), 5039 (2009)
- 395 16. Mizuguchi, Y., Oishi, Y., Miyoshi, H., Iuchi, A., Nagase, N., Oki, T.: The functional role  
396 of longitudinal, circumferential, and radial myocardial deformation for regulating the early  
397 impairment of left ventricular contraction and relaxation in patients with cardiovascular  
398 risk factors: a study with two-dimensional strain imaging. *Journal of the American Society*  
399 *of Echocardiography* 21(10), 1138-1144 (2008)
- 400 17. Nagueh, S.F., Chang, S.M., Nabi, F., Shah, D.J., Estep, J.D.: Imaging to diagnose and  
401 manage patients in heart failure with reduced ejection fraction. *Circulation: Cardiovascular*  
402 *Imaging* 10(4), e005615 (2017)
- 403 18. Olsen, F.J., Bertelsen, L., de Knegt, M.C., Christensen, T.E., Vejlstrop, N., Svendsen,  
404 J.H., Jensen, J.S., Biering-Sorensen, T.: Multimodality cardiac imaging for the assessment  
405 of left atrial function and the association with atrial arrhythmias. *Circulation: Cardiovas-*  
406 *cular Imaging* 9(10), e004947 (2016)
- 407 19. Perperidis, D., Mohiaddin, R.H., Rueckert, D.: Spatio-temporal free-form registration  
408 of cardiac mr image sequences. *Medical image analysis* 9(5), 441-456 (2005)
- 409 20. Puyol-Anton, E., Sinclair, M., Gerber, B., Amzulescu, M.S., Langet, H., De Craene, M.,  
410 Aljabar, P., Piro, P., King, A.P.: A multimodal spatiotemporal cardiac motion atlas from  
411 mr and ultrasound data. *Medical image analysis* 40, 96-110 (2017)
- 412 21. Szasz, O.: On products of summability methods. *Proceedings of the American Mathe-*  
413 *matical Society* 3(2), 257-263 (1952)
- 414 22. Thorpe, S., Fize, D., Marlot, C.: Speed of processing in the human visual system. *nature*  
415 381(6582), 520 (1996)
- 416 23. Tobon-Gomez, C., De Craene, M., Mcleod, K., Tautz, L., Shi, W., Hennemuth, A.,  
417 Prakosa, A., Wang, H., Carr-White, G., Kapetanakis, S.: Benchmarking framework for  
418 myocardial tracking and deformation algorithms: An open access database. *Medical image*  
419 *analysis* 17(6), 632-648 (2013)
- 420 24. Valsangiacomo Buechel, E.R., Mertens, L.L.: Imaging the right heart: the use of inte-  
421 grated multimodality imaging. *European heart journal* 33(8), 949-960 (2012)
- 422 25. Yushkevich, P.A., Piven, J., Hazlett, H.C., Smith, R.G., Ho, S., Gee, J.C., Gerig, G.:  
423 User-guided 3d active contour segmentation of anatomical structures: significantly im-  
424 proved efficiency and reliability. *Neuroimage* 31(3), 1116-1128 (2006)
- 425 26. Zhai, Y., Shah, M.: Visual attention detection in video sequences using spatiotemporal  
426 cues. In: *Proceedings of the 14th ACM international conference on Multimedia*. pp. 815-  
427 824. ACM (2006)

- 428 27. Zhang, W., Noble, J.A., Brady, J.M.: Spatio-temporal registration of real time 3d ul-  
429 trasound to cardiovascular mr sequences. In: International Conference on Medical Image  
430 Computing and Computer-Assisted Intervention. pp. 343-350. Springer (2007)
- 431 28. Zhang, W., Noble, J.A., Brady, J.M.: Adaptive non-rigid registration of real time 3d  
432 ultrasound to cardiovascular mr images. In: Biennial International Conference on Infor-  
433 mation Processing in Medical Imaging. pp. 50-61. Springer (2007)
- 434 29. Zhao, N., Basarab, A., Kouamé, D., & Tourneret, J. Y. Joint segmentation and decon-  
435 volution of ultrasound images using a hierarchical Bayesian model based on generalized  
436 Gaussian priors. *IEEE transactions on Image Processing*, 25(8), 3736-3750. (2016).
- 437 30. Atehortúa, A., Garreau, M., & Romero, E. Fusion of 4D echocardiography and cine  
438 cardiac magnetic resonance volumes using a salient spatio-temporal analysis. In 13th In-  
439 ternational Conference on Medical Information Processing and Analysis (Vol. 10572, p.  
440 105721A). International Society for Optics and Photonics. (2017)
- 441 31. Klein, S., Staring, M., Murphy, K., Viergever, MA & Pluim, JPW. elastix: A toolbox for  
442 intensity-based medical image registration. *IEEE Trans Med Imag.* pp. 196–205. (2010)
- 443 32. Kanai, T., Kadoya, N., Ito, K., Onozato, Y., Cho, S. Y., Kishi, K., & Jingu, K. Eval-  
444 uation of accuracy of B-spline transformation-based deformable image registration with  
445 different parameter settings for thoracic images. *Journal of radiation research*, 55(6), 1163-  
446 1170. (2014)
- 447 33. Ou, Y., Akbari, H., Bilello, M., Da, X., & Davatzikos, C. Comparative evaluation of  
448 registration algorithms in different brain databases with varying difficulty: results and  
449 insights. *IEEE transactions on medical imaging*, 33(10), 2039-2065. (2014)
- 450 34. Klein, S., Staring, M., & Pluim, J. P. Evaluation of optimization methods for nonrigid  
451 medical image registration using mutual information and B-splines. *IEEE transactions on*  
452 *image processing*, 16(12), 2879-2890. (2007)
- 453 35. Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L., Leach, M. O., & Hawkes, D. J.  
454 Nonrigid registration using free-form deformations: application to breast MR images. *IEEE*  
455 *transactions on medical imaging*, 18(8), 712-721. (1999)