A non-rigid registration approach for quantifying myocardial contraction in tagged MRI using generalized information measures.

Nicolas Rougon, Caroline Petitjean, Françoise Prêteux, Philippe Cluzel, Philippe Grenier

A non-rigid registration approach for quantifying myocardial contraction in tagged MRI using generalized information measures

Nicolas Rougon [a,*], Caroline Petitjean [a], Françoise Prêteux [a], Philippe Cluzel [b], Philippe Grenier [b]

[a] *ARTEMIS Project Unit, GET/INT, 9 Rue Charles Fourier, 91011 Evry, France*
[b] *Service de Radiologie, Groupe Hospitalier Pitié-Salpêtrière, 47-83 Bvd de l'Hôpital, 75651 Paris Cedex 13, France*

Abstract

We address the problem of quantitatively assessing myocardial function from tagged MRI sequences. We develop a two-step method comprising (i) a motion estimation step using a novel variational non-rigid registration technique based on generalized information measures, and (ii) a measurement step, yielding local and segmental deformation parameters over the whole myocardium. Experiments on healthy and pathological data demonstrate that this method delivers, within a reasonable computation time and in a fully unsupervised way, reliable measurements for normal subjects and quantitative pathology-specific information. Beyond cardiac MRI, this work redefines the foundations of variational non-rigid registration for information-theoretic similarity criteria with potential interest in multimodal medical imaging.

Keywords: Cardiac motion analysis; Myocardial function; MR tagging; Non-rigid registration; Generalized information measures

1. Introduction

Cardiovascular diseases are the leading cause of mortality in industrialized countries, and therefore a major public health challenge. In this framework, magnetic resonance imaging (MRI) is recognized as a key modality for dynamically imaging the heart anatomy and function. It provides a valuable investigation tool for early diagnosis and clinical/therapeutical follow-up, capable of delivering, in single examination, all the necessary information for assessing a large variety of cardiovascular pathologies, including those related to morphology and flow rate of coronary arteries, cardiac function, and myocardial perfusion/viability. In particular, the relevance of MRI for identifying and localizing acute and chronic myocardial ischemia through the detection of kinematic abnormalities has been established (Lipton et al., 2002). Accordingly, abundant research efforts have been devoted to cardiac motion recovery and deformation assessment from MR image sequences (Duncan and Ayache, 2000; Frangi et al., 2001, 2002).

The reference MR modality for imaging myocardial strain is tagged MRI (Zerhouni et al., 1988; Axel and Dougherty, 1989). Measuring myocardial deformations from tagged MR sequences relies on estimating a dense displacement field consistent with the motion of the structured tagging pattern (Kerwin et al., 2001; Amini and Prince, 2001; McVeigh and Öztürk, 2001). Classical approaches comprise: (i) differential optical flow-based methods coupled with intensity correction schemes (Prince and McVeigh, 1992; Denney, 1994; Denney

* Corresponding author. Tel.: +33 1 60 76 46 44; fax: +33 1 60 76 43 81.
E-mail address: nicolas.rougon@int-evry.fr (N. Rougon).
URL: http://www.artemis.int-evry.fr (N. Rougon).

and Prince, 1994; Gupta and Prince, 1995; Gupta et al., 1997; Dougherty et al., 1999); (ii) phase-based optical flow methods such as HARP (Osman et al., 1999; Osman and Prince, 2000; Amini and Prince, 2001); (iii) segmentation-based approaches combining tag segmentation, sparse motion estimation along tags, and dense motion interpolation over the image domain (Kumar and Goldgof, 1994; Guttman et al., 1994; Denney and Prince, 1995; Kraitchman et al., 1995; Young et al., 1995; Park et al., 1996; Declerck, 1997; Guttman et al., 1997; Amini et al., 1998; Denney, 1999; Huang et al., 1999; Kerwin and Prince, 1999; Clarysse et al., 2000; Haber et al., 2000; Öztürk and McVeigh, 2000; Amini et al., 2001). However, their clinical applicability is questioned by intrinsic limitations: differential optical flow fails for large displacements due to the local nature of constraint equations; phase-based optical flow is limited to 1-D tags and proves to be sensitive to artifacts; tag segmentation often requires supervision to deal with tag contrast attenuation over time due to T1 demagnetization (tag fading out), and tag destruction along myocardial boundaries induced by blood flows within the cardiac chambers.

Recently, intensity-based non-rigid registration (NRR) methods (Maintz and Viergever, 1998; Mäkelä et al., 2002; Pluim et al., 2003) have been suggested independently by Petitjean et al. (2002) and Chandrashekara et al. (2002) as robust, segmentation-free techniques to obtain dense myocardial motion estimates in a non-supervised way. Here, the use of information-theoretic similarity criteria allows for efficiently dealing with tag fading out without specifying a demagnetization model. Based on Rueckert's method (Rueckert et al., 1999), initially designed for aligning breast MR images, the approach in Chandrashekara et al. (2002), Perperidis et al. (2003) and Rao et al. (2003) delivers parametric motion estimates by maximizing normalized mutual information (Studholme et al., 1999) over a space of free form deformations. The hereby presented approach relies on the Ali-Silvey class of generalized information measures for which a generic variational optimization framework is developed (Petitjean et al., 2002; Petitjean, 2003c). It includes mutual information and normalized mutual information as subcases, and encompasses both non-parametric and parametric motion spaces. It is, therefore, applicable to model-free (Petitjean et al., 2003a) as well as compact model-based motion estimation using a dedicated statistical cardiac motion atlas (Petitjean et al., 2004).

1.1. Outline of the method

The proposed approach makes use of tagged MRI sequences acquired at multiple slice levels under short-axis or/and long-axis views, and of cine MRI sequences jointly acquired at the same slice levels in identical conditions. Tagged MR data provide access to intra-myocardial motion information, whereas cine MR data allow for dynamically assessing myocardium anatomy. Datasets are processed either integrally in a frame-to-frame fashion, yielding dynamic quantitative analysis during systole, or partially by restricting to end-diastolic (ED) and end-systolic (ES) frames to directly derive systolic deformation parameters. The latter mode is relevant in a clinical routine framework where fast quantitative assessment of cardiac function is desirable. It is enabled by the ability of the motion estimation and tracking methods to deal with large displacements and non-linear contrast variations.

The measurement of myocardial deformations is achieved in three steps:

(1) *dense non-parametric motion estimation* over the whole image domain from tagged MR data;
(2) *dynamic myocardium segmentation* via non-rigid alignment of cine MR data, and non-rigid propagation against the resulting transforms of an initial myocardium segmentation derived from the ED cine MR image;
(3) *computation of local and segmental deformation parameters* over the myocardium from displacement estimates.

Non-rigid motion estimation and anatomical alignment are both achieved via non-parametric NRR based on generalized information measures, which is, therefore, the core of the approach. Myocardium segmentation at ED can be achieved by any technique of choice (Suri, 2000). In the current implementation, we have favored mathematical morphology.

1.2. Non-rigid registration using statistical similarity measures

Statistical similarity measures underlie efficient techniques for matching mono-modal data with significant local gray level variations or multi-modal data. In these cases, the relationship between signals is generally complex and not explicitly known. The correlation ratio (Roche et al., 1998) provides an optimal answer for functionally dependent signals (Roche et al., 1999). The weaker assumption of statistically dependent data leads to using information theoretic measures, such as joint entropy (Collignon et al., 1995) and mutual information (MI) (Viola and Wells, 1997; Maes et al., 1997), which have motivated an abundant literature (Pluim et al., 2003).

Maximizing these measures is most often performed over low-dimensional parametric transform spaces. Optimization is achieved via direct techniques using empirical histogram-based (Maes et al., 1997; Pluim, 2000) or parametric (Leventon and Crimson, 1998)

probability estimates, or via gradient descent methods relying on analytic (Rueckert et al., 1999) or non-parametric Parzen probability estimates (Viola and Wells, 1997). Moreover, Parzen estimation provides a path for performing variational optimization over non-parametric transform spaces (Hermosillo et al., 2001; Hermosillo, 2002; D'Agostino et al., 2003), leading to well-posed models (Faugeras and Hermosillo, 2001).

Despite good performances, MI is sensitive to image overlap and can be trapped into local minima when dealing with large deformations (Pluim, 2000). This is illustrated in Fig. 1 for MI-based NRR between ED and ES cine MR images. To overcome these limitations, several solutions have been proposed and investigated for parametric transforms: building probabilistic priors via supervised learning (Leventon and Crimson, 1998); incorporating spatial information by weighting MI by a gradient-dependent term (Pluim et al., 2000); using alternative entropic criteria such as normalized information measures based on Shannon entropy (Maes et al., 1997; Studholme et al., 1999), or non-Shannon measures such as f-informations (Pluim, 2000; He et al., 2003).

This paper elaborates along the latter direction by bringing forward the Ali-Silvey class of generalized information measures as a superset of robust statistical similarity criteria which extend the previously proposed Shannon measures. It comprises f-informations as well as novel normalized generalized information measures which are introduced. Using Parzen estimation, we identify closed-form expressions for their gradient flows over non-parametric and parametric transform spaces. New results relative to previously proposed measures are readily derived as subcases, such as the flows for (i) MI for parametric transforms, (ii) normalized Shannon information measures for arbitrary transforms, and (iii) I_α-information (Pluim, 2000) and Renyi information (He et al., 2003) for arbitrary transforms.

1.3. Structure of the paper

The paper is organized as follows. Section 2 introduces the Ali-Silvey class of generalized information measures as an extension of the classical Shannon framework. In Section 3, a generic variational NRR approach based on these measures is developed, yielding simple and computationally efficient algorithms. Applied to tagged MR data, they deliver dense non-parametric displacement estimates in a robust and fully unsupervised way. Applied to cine MR data, they allow for dynamically segmenting the myocardium by warping

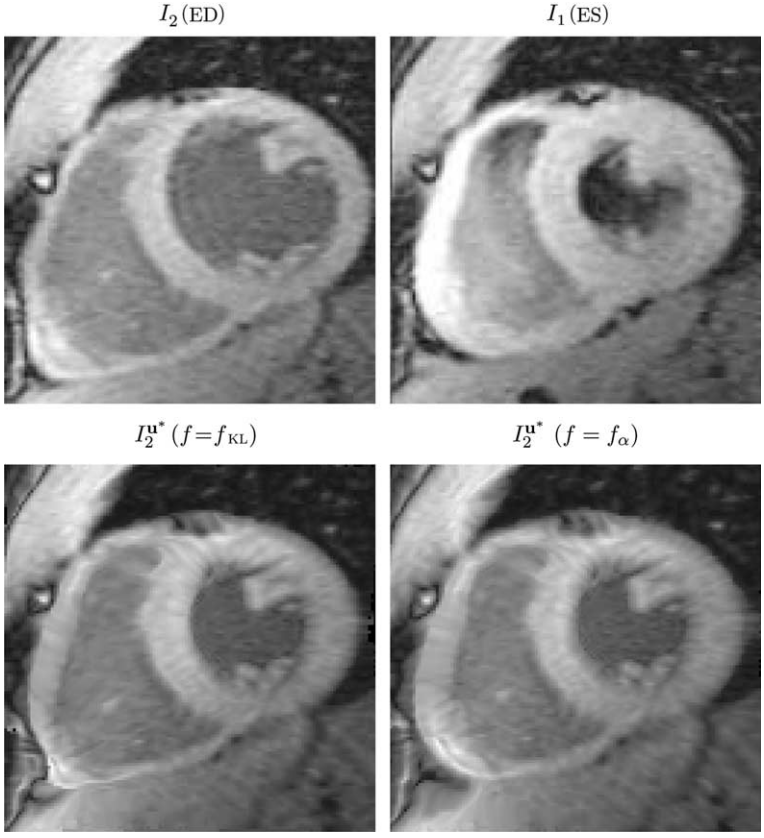

Fig. 1. Influence of the information metrics f on NRR between ED and ES cine MR images. The f-information similarity criterion $\mathscr{S}(\mathbf{u}) = -I_f(I_1, I_2^{\mathbf{u}})$ has been used.

an initial segmentation map. Section 4 deals with the computation of local and segmental deformation parameters over the segmented myocardium from estimated displacements. Section 5 details evaluation on simulated tagged MR data, and demonstrates performances for assessing cardiac function on healthy and pathological data. A discussion on the benefits and possible extensions of the proposed approach is given in Section 6 before concluding.

2. Generalized information measures

Let X and Y denote continuous random variables over a state space Ω with pdf p^X and p^Y, respectively. We also consider the joint variable (X,Y) over Ω^2 with pdf $p^{X,Y}$.

2.1. The Shannon framework

The Shannon entropy $H_S(X)$ of the random variable X, defined as

$$H_S(X) := -\int_\Omega p^X(x)\log p^X(x)\,\mathrm{d}x$$

measures its uncertainty or randomness. The Shannon entropy $H_S(X,Y)$ of the joint variable (X,Y) is defined similarly. When X and Y are independent, the following additivity property holds: $H_S(X,Y) = H_S(X) + H_S(Y)$.

Information gain between random variables is measured via relative entropy and mutual information (MI). The relative entropy, or Kullback–Leibler (KL) divergence, is the non-symmetric measure $D_{KL}(X\,\|\,Y)$ such that

$$D_{KL}(X\|Y) := \int_\Omega p^X(x)\log\frac{p^X(x)}{p^Y(x)}\,\mathrm{d}x.$$

MI is a symmetric entropic measure, denoted by $I(X,Y)$, which quantifies the reduction of the uncertainty of Y given X (and vice versa)

$$I(X,Y) := H_S(X) + H_S(Y) - H_S(X,Y).$$

Equivalently, MI is the relative entropy $D_{KL}((X,Y)\,\|\,X\times Y)$ between the joint variable (X,Y) and the product distribution $(X\times Y)$

$$I(X,Y) = \int_{\Omega^2} p^{X,Y}(x,y)\log\frac{p^{X,Y}(x,y)}{p^X(x)p^Y(y)}\,\mathrm{d}x\,\mathrm{d}y.$$

MI, thus, measures the statistical distance between X and Y and the reference case where these variables are assumed to be independent.

2.1.1. Normalized information measures

Alternative information gain measures, with enhanced robustness to lacunar joint statistics, have been proposed in the literature of image registration (Hajnal et al., 2001). They comprise normalized measures, such as the entropy correlation coefficient (ECC) (Maes et al., 1997)

$$\mathrm{ECC}(X,Y) := 2\frac{I(X,Y)}{H_S(X)+H_S(Y)} \tag{1}$$

and the normalized mutual information (NMI) (Studholme et al., 1999)

$$\mathrm{NMI}(X,Y) := \frac{H_S(X)+H_S(Y)}{H_S(X,Y)}. \tag{2}$$

Another information gain measure, denoted by Z, has also been proposed in (Maes et al. (1997)) without given terminology

$$Z(X,Y) := H_S(X,Y) - I(X,Y). \tag{3}$$

Since $I(X,Y)$ quantifies the information gain resulting from observing X and Y jointly compared to observing them separately, and $H_S(X,Y)$ is the information of the joint variable (X,Y), Z measures the information exclusively contained in X and Y when observed jointly (Fig. 2). We, thus, refer to Z as *exclusive information*.[1]

2.1.2. Limitations of the Shannon framework

It is known from information theory that MI leads to optimal statistical decision when uncertainty measurement is constrained to be additive[2] and no assumption is made about data statistics. MI provides, therefore, a simple and universal answer for statistically comparing arbitrary data. However, blind statistical decision may be unrobust to consistently deal with complex joint and/or marginal statistics which largely deviate from the Gaussian case.[3] Overcoming this intrinsic limitation requires relaxing the additivity hypothesis and allowing some prior statistical information to be accounted for.

Based on this idea, extensions of the Shannon framework have been proposed, where statistical prior knowledge is incorporated by specifying some given, non-additive uncertainty measurement model for the class of data under consideration. In this approach, Shannon entropy is a limiting case of more general entropies which fall into two categories: *integral entropies* and *non-integral entropies*. They induce integral and non-integral divergences from which generalized information gain measures are defined, according to the same design principles as for the Shannon case (Vajda, 1989).

[1] In Section 3.1.3, we show that exclusive information belongs to the same class of Shannon information gain measures as ECC and NMI. For this reason, and with an abuse of terminology in the case of Z, they will be uniformly referred to as normalized information measures.

[2] i.e., there exists a formal entropy H such that: $H(X,Y) = H(X) + H(Y)$ for independent random variables X and Y.

[3] In the framework of image registration, complex joint statistics can arise from large deformations or from highly multimodal matching problems (e.g., MR-US registration).

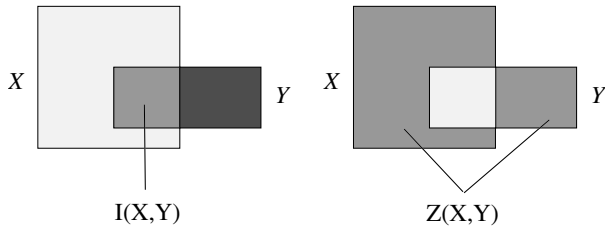

Fig. 2. Mutual information $I(X,Y)$ and exclusive information $Z(X,Y)$ of 2 random variables X and Y.

2.2. Generalized information measures

Entropy is basically defined for probability measures. In this setting, the entropy of a given measure explicitly depends on a reference measure. Hence, the terminology of *relative entropy*. In what follows, f (resp. ψ) will denote a continuous convex (resp. monotonic) function over $\mathbb{R}^+$.

Definition 1 (f-entropy). The integral (resp. non-integral) entropy $H_{f,\nu}(\mu)$ (resp. $H_{\psi,\nu}(\mu)$) of a probability measure μ w.r.t. the reference measure ν is

$$H_{f,\nu}(\mu) = -\int f\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right)\mathrm{d}\nu,$$

$$H_{\psi,\nu}(\mu) = -\log\psi^{-1}\left(\int \frac{\mathrm{d}\mu}{\mathrm{d}\nu}\psi\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right)\mathrm{d}\nu\right).$$

These definitions can be specialized to continuous random variables by letting $\mu \equiv X$ and ν be the natural Borel measure, and by using Radon–Nikodym derivatives $\frac{\mathrm{d}\mu}{\mathrm{d}x} \equiv p^X(x)$ and $\frac{\mathrm{d}\nu}{\mathrm{d}x} \equiv 1$. The Kullback metrics $f_{\mathrm{KL}}(x) := x\log x$ or $\psi_{\mathrm{KL}}(x) := \log x$ leads to Shannon entropy which has both an integral and a non-integral form. Many other entropies, corresponding to specific choices for the information metrics f or ψ, have been proposed, allowing to tune the measurement of uncertainty to specific statistical contexts (Vajda, 1989).

Similarity between probability measures is formalized via the notion of *divergence*, which derives from relative entropy. Several paths to defining divergence are possible (Vajda, 1989). Comparing probability measures individually (rather than comparing mixtures) leads to f-divergence.

Definition 2 (f-divergence). An f-divergence D_f is formally defined as: $D_f(\mu \parallel \nu) = -H_{f,\nu}(\mu)$.

Definition 2 delimits the *Ali-Silvey class* of divergences, which is divided into integral and non-integral f-divergences. Letting μ and ν be random variables X and Y such that $\frac{\mathrm{d}\mu}{\mathrm{d}x} \equiv p^X(x)$ and $\frac{\mathrm{d}\nu}{\mathrm{d}x} \equiv p^Y(x)$ yield the definitions for an integral f-divergence $D_f(X\|Y)$ and a non-integral f-divergence $D_\psi(X\|Y)$ of X w.r.t. Y. Choosing $f = f_{\mathrm{KL}}$ or $\psi = \psi_{\mathrm{KL}}$ yields KL divergence. Setting $f = f_\alpha$ with $f_\alpha(x) := \frac{x^\alpha - \alpha x + \alpha - 1}{\alpha(\alpha-1)}$ $(\alpha \neq \{0,1\})$ leads to the I_α-divergence I_α (Vajda, 1989) which verifies: $\lim_{\alpha\to 1} I_\alpha = D_{\mathrm{KL}}$. When $\psi = \psi_r$ with $\psi_r(x) := x^{r-1}$ $(r > 0, r \neq 1)$, one obtains the Renyi divergence I_r of order r which verifies: $\lim_{r\to 1} I_r = D_{\mathrm{KL}}$.

The notion of *information gain* induced by simultaneously observing two probability measures compared to their separate occurrence is tightly related to divergence. It results from quantifying the information content of the joint measure using f-entropy w.r.t. the product measure corresponding to independence. This leads to f-information.

Definition 3 (f-information). An f-information I_f is formally defined as: $I_f(\mu,\nu) = D_f((\mu,\nu) \parallel \mu\times\nu)$.

f-information measures the reduction of the uncertainty of μ given ν (and vice versa) w.r.t. the information metrics f. One distinguishes between integral and non-integral f-informations. Setting $\frac{\mathrm{d}\mu}{\mathrm{d}x\mathrm{d}y} \equiv p^{X,Y}(x,y)$ and $\frac{\mathrm{d}\nu}{\mathrm{d}x\mathrm{d}y} \equiv p^X(x)p^Y(y)$ lead to specialized definitions of f-information for random variables. Choosing $f = f_{\mathrm{KL}}$ yields MI, simply denoted by I.

2.2.1. Duality properties

Duality properties arise when studying relationships between integral and non-integral f-divergences. The so-called *mirror function* of function f, denoted by $\breve{f}$ and defined as

$$\breve{f}(x) := xf\left(\frac{1}{x}\right) \tag{4}$$

is then the key notion.[4] The terminology is motivated by the idempotence of the transform $f \to \breve{f}$, i.e., $\breve{\breve{f}} = f$. From Definition 1, it is seen that, barring a one-to-one mapping, integral and non-integral f-divergences are related by the following connection identity:

$$\psi(x) = -\breve{f}\left(\frac{1}{x}\right). \tag{5}$$

Eq. (5) provides the foundation for a unified axiomatic framework, based on homeomorphic idempotent information metrics, which integrates the various instances of f-divergences proposed in the literature (Vajda, 1989).

2.2.2. Normalized generalized information measures

Entropy correlation coefficient, normalized mutual information and exclusive information can be extended to generalized information measures in a consistent fashion with the Shannon case (Petitjean et al., 2002). To this end, Definitions 1–3 are first rewritten using the identities $H_{\mathrm{S}}(X,Y) = H_{\mathrm{S}}(X) + H_{\mathrm{S}}(Y) - I(X,Y)$ and

[4] The mirror function is also involved in the pseudo-symmetry property satisfied by f-divergences: $D_f(\mu \parallel \nu) = D_{\breve{f}}(\nu \parallel \mu)$. The latter states that permuting the reference measure when quantifying similarity between probability measures requires switching from the primal to the mirror information metrics.

$H_S(X) = I(X,X)$. Switching from f_{KL} to arbitrary f in the resulting expressions then leads to the following definitions:

- generalized entropy correlation coefficient

$$\mathrm{ECC}_f(\mu,\nu) := 2\frac{I_f(\mu,\nu)}{I_f(\mu,\mu)+I_f(\nu,\nu)}, \tag{6}$$

- generalized normalized mutual information

$$\mathrm{NMI}_f(\mu,\nu) := \frac{I_f(\mu,\mu)+I_f(\nu,\nu)}{I_f(\mu,\mu)+I_f(\nu,\nu)-I_f(\mu,\nu)}, \tag{7}$$

- exclusive f-information

$$Z_f(\mu,\nu) := I_f(\mu,\mu)+I_f(\nu,\nu)-2I_f(\mu,\nu). \tag{8}$$

By construction, a uniform convergence to the Shannon framework is verified: $\lim_{f\to f_{KL}}\mathrm{ECC}_f = \mathrm{ECC}$, $\lim_{f\to f_{KL}}\mathrm{NMI}_f = \mathrm{NMI}$, and $\lim_{f\to f_{KL}} Z_f = Z$.

3. Variational non-rigid registration using generalized information measures

Let I_1 and I_2 be images defined over a compact domain $D \subset \mathbb{R}^n$ with values in an interval Ω. Registering I_2 onto I_1 consists in finding a transformation ϕ: $D \to D$ within some function space $\mathcal{T}$ so that $I_2 \circ \phi$ is similar to I_1 according to a predefined criterion. This amounts to finding a displacement field $\mathbf{u} := \phi - \mathbb{I}\mathrm{d} \in \mathcal{T}$ such that $I_2^{\mathbf{u}}(\mathbf{x}) := I_2(\mathbf{x}+\mathbf{u}(\mathbf{x}))$ is similar to $I_1(\mathbf{x})$ for any $\mathbf{x} \in D$.

This problem is ill-posed and must be regularized. A classical approach defines a solution as a minimizer $\mathbf{u}^*$ of a functional $\mathcal{J}(\mathbf{u}) := \mathcal{S}(\mathbf{u}) + \lambda\mathcal{R}(\mathbf{u})$ over the space $\mathcal{T}$. $\mathcal{S}$ is a similarity functional which quantifies the discrepancy between I_1 and $I_2^{\mathbf{u}}$. In this paper, $\mathcal{S}$ is defined as a generalized information measure. Depending on its extremality properties, this yields the possible choices: $\mathcal{S}(\mathbf{u}) := -I_f(I_1, I_2^{\mathbf{u}})$, $\mathcal{S}(\mathbf{u}) := -\mathrm{ECC}_f(I_1, I_2^{\mathbf{u}})$, $\mathcal{S}(\mathbf{u}) := -\mathrm{NMI}_f(I_1, I_2^{\mathbf{u}})$ and $\mathcal{S}(\mathbf{u}) := Z_f(I_1, I_2^{\mathbf{u}})$. $\mathcal{R}$ is a stabilizing functional ensuring almost-everywhere smooth solutions (see Section 3.2). The regularization parameter $\lambda > 0$ controls the trade-off between data consistency and smoothing.

Minimizing $\mathcal{J}$ in a variational way is achieved by integrating, via some gradient descent technique, the associated Euler–Lagrange equations, obtained by zeroing its Gâteaux derivative. Defining the latter depends on the modeling context chosen for $\mathbf{u}$:

- *Non-parametric motion models*: letting $\mathcal{T}$ be a space of regular mappings leads to a dense minimization problem: $\mathbf{u}^* = \arg\min_{\mathbf{u}\in\mathcal{T}} \mathcal{J}(\mathbf{u})$. It yields a flow over D: $\partial_t \mathbf{u}(\mathbf{x}) = -\partial_{\mathbf{u}}\mathcal{J}(\mathbf{u}(\mathbf{x}))$ where $\partial_{\mathbf{u}}\mathcal{J}$ is the 1st variation of $\mathcal{J}$ over $\mathcal{T}$.
- *Parametric motion models:* when $\mathbf{u} := \mathbf{u}^{\Theta}$ lies in a space of smooth mappings depending on a parameter $\Theta \in \mathbb{R}^p$, $\mathbf{u}^*$ is searched as a solution $\mathbf{u}^{\Theta^*}$ of a sparse minimization: $\Theta^* = \arg\min_\Theta \mathcal{J}(\mathbf{u}^\Theta)$. This yields a flow over $\mathbb{R}^p$: $\partial_t\Theta = -\partial_\Theta\mathcal{J}(\mathbf{u}^\Theta)$, where $\partial_\Theta\mathcal{J}$ is the 1st variation of $\mathcal{J}$. Here, we restrict to linear models for which $\mathbf{u}^\Theta(\mathbf{x}) = \mathbf{B}(x)\Theta$ and $\mathbf{B}(\mathbf{x})$ is a matrix encoding a local basis[5] of $\mathcal{T}$.

In each setting, the key point relies on computing the flows $\partial_{\mathbf{u}}\mathcal{S}(\mathbf{u})$ or $\partial_\Theta\mathcal{S}(\mathbf{u}^\Theta)$ for f-informations. For the targeted application, we have focused on non-parametric motion models, for which computations are presented in detail. For the sake of completeness, the results for linear parametric motions, whose derivation follows the same lines, are given directly in a second step. They have been applied to atlas-based myocardial motion recovery in tagged MRI by means of statistical parametric motion models (Petitjean et al., 2004).

3.1. Gradient flow of generalized information measures

3.1.1. Integral f-information measures

Referring to Section 2, an integral f-information functional between reference image I_1 and deformed image $I_2^{\mathbf{u}}$ is

$$\begin{aligned}\mathcal{S}(\mathbf{u}) &= -I_f(I_1, I_2^{\mathbf{u}}) \\ &= -\int_{\Omega^2} p^{I_1}(i_1)p^{I_2^{\mathbf{u}}}(i_2)f(\rho^{\mathbf{u}}(i_1,i_2))\,\mathrm{d}i_1\,\mathrm{d}i_2,\end{aligned} \tag{9}$$

where we have set $\rho^{\mathbf{u}}(i_1,i_2) := \frac{p^{I_1,I_2^{\mathbf{u}}}(i_1,i_2)}{p^{I_1}(i_1)p^{I_2^{\mathbf{u}}}(i_2)}$. Here, $p^{I_1}(i_1)$ (resp. $p^{I_2^{\mathbf{u}}}(i_2)$) denotes the probability that a pixel in image I_1 (resp. $I_2^{\mathbf{u}}$) has intensity i_1 (resp. i_2), and $p^{I_1,I_2^{\mathbf{u}}}(i_1,i_2)$ is the probability that corresponding pixels $\mathbf{x}$ in image I_1 and $\mathbf{x}+\mathbf{u}(\mathbf{x})$ in image $I_2^{\mathbf{u}}$ have intensities i_1 and i_2, respectively.

Following Hermosillo et al. (2001) and Hermosillo (2002), the first variation $\partial_{\mathbf{u}}\mathcal{S}(\mathbf{u})$ of $\mathcal{S}(\mathbf{u})$ is obtained by computing the derivative $\frac{\partial\mathcal{S}(\mathbf{u}+\varepsilon\mathbf{h})}{\partial\varepsilon}$ of the perturbed cri-

[5] Important instances within this class are affine and B-splines models. In the 2-D case, the matrix $\mathbf{B}(\mathbf{x})$ has the generic block form: $\mathbf{B}(\mathbf{x}) := \begin{pmatrix}\mathbf{M}(\mathbf{x}) & \mathbf{0}\\ \mathbf{0} & \mathbf{M}(\mathbf{x})\end{pmatrix}$ where $\mathbf{M}(\mathbf{x})$ is a row vector depending on $\mathbf{x} := (x,y)$. For affine models, Θ is a 6-dimensional vector and $\mathbf{M}(\mathbf{x}) = (1\ x\ y)$. For B-spline models (Rueckert et al., 1999), Θ is a $(2n_xn_y)$-dimensional vector gathering the x- then y-coordinates of a $(n_x \times n_y)$ rectangular mesh of raster-ordered control points. Letting $\tilde{x} := \frac{x}{n_x}$ and $\tilde{y} := \frac{y}{n_y}$, and setting $i := \lfloor\tilde{x}\rfloor - 1$, $j := \lfloor\tilde{y}\rfloor - 1$, $u := \tilde{x} - \lfloor\tilde{x}\rfloor$ and $v := \tilde{y} - \lfloor\tilde{y}\rfloor$, $\mathbf{M}(\mathbf{x})$ is the (n_xn_y)-dimensional vector with components $M_{mn_x+m} := B_k(u)B_l(v)\delta_{i+k,j+l}(m,n)$ $((k,l) \in [0\ldots3]\times[0\ldots3]$, $(m,n) \in [1\ldots n_x]\times[1\ldots n_y])$. Here, B_k represents the kth basis function of the cubic B-spline and $\delta_{i,j}(m,n)$ denotes the standard Kronecker symbol which is 1 iff. $(i=m)\wedge(j=n)$, and 0 otherwise. The extension to the 3D case is straightforward.

terion $\mathscr{S}(\mathbf{u}+\varepsilon\mathbf{h})$, $\mathbf{h}\in\mathscr{T}$ at $\varepsilon=0$. The latter involves the derivatives $\frac{\partial p^{I_1,I_2^{\mathbf{u}+\varepsilon\mathbf{h}}}}{\partial\varepsilon}$ and $\frac{\partial p^{I_2^{\mathbf{u}+\varepsilon\mathbf{h}}}}{\partial\varepsilon}$ which are expressed under closed form by approximating $p^{I_1,I_2^{\mathbf{u}}}$ (resp. $p^{I_2^{\mathbf{u}}}$) with a global Parzen estimator with 2-D (resp. 1-D) kernel $\mathscr{K}_2$ (resp. $\mathscr{K}_1$) over the region of overlap $\underline{D}:=D\cap\phi(D)$ of both images:

$$p^{I_1,I_2^{\mathbf{u}}}(i_1,i_2)\approx\frac{1}{|\underline{D}|}\int_{\underline{D}}\mathscr{K}_2\big(I_1(\mathbf{x})-i_1,I_2^{\mathbf{u}}(\mathbf{x})-i_2\big)\mathrm{d}\mathbf{x},\tag{10}$$

$$p^{I_2^{\mathbf{u}}}(i_2)\approx\frac{1}{|\underline{D}|}\int_{\underline{D}}\mathscr{K}_1(I_2^{\mathbf{u}}(\mathbf{x})-i_2)\mathrm{d}\mathbf{x},\tag{11}$$

where $|\underline{D}|$ is the volume of $\underline{D}$ (Viola and Wells, 1997; Hermosillo et al., 2001). It is then shown in Appendix A that

$$\partial_{\mathbf{u}}\mathscr{S}(\mathbf{u})=-\partial_{\mathbf{u}}I_f(I_1,I_2^{\mathbf{u}})=-V_1^{I_1,I_2^{\mathbf{u}}}(\mathbf{x})\nabla I_2^{\mathbf{u}}(\mathbf{x}),\tag{12}$$

$$V_1^{I_1,I_2^{\mathbf{u}}}(\mathbf{x}):=K\Bigg\{\left[\mathscr{K}_2\bigstar\frac{\partial L_1^{\mathbf{u}}}{\partial i_2}\right]\big(I_1(\mathbf{x}),I_2^{\mathbf{u}}(\mathbf{x})\big)+E_{I_1}\left(\left[\mathscr{K}_1\bigstar\frac{\partial L_2^{\mathbf{u}|i_1}}{\partial i_2}\right](I_2^{\mathbf{u}}(\mathbf{x}))\right)\Bigg\},\tag{13}$$

where $E_{I_1}(\cdot)$ denotes expectation w.r.t. the random variable I_1 and the symbol $\bigstar$ is the 2-D (resp. 1-D) convolution operator over the intensity space Ω^2 (resp. Ω). The symbols $K(I_1,I_2^{\mathbf{u}})$, $L_1^{\mathbf{u}}$ and $L_2^{\mathbf{u}}$ are defined in Table 1 and $L_2^{\mathbf{u}|i_1}(\cdot):=L_2^{\mathbf{u}}(i_1,\cdot)$ for any given intensity value $i_1\in\Omega$. One has

$$\frac{\partial L_2^{\mathbf{u}|i_1}}{\partial i_2}=-\rho^{\mathbf{u}}\frac{\partial L_1^{\mathbf{u}}}{\partial i_2},\tag{14}$$

$$\frac{\partial L_1^{\mathbf{u}}}{\partial i_2}=\rho^{\mathbf{u}}f''(\rho^{\mathbf{u}})\left[\frac{\partial\log p^{I_1,I_2^{\mathbf{u}}}}{\partial i_2}-\frac{\partial\log p^{I_2^{\mathbf{u}}}}{\partial i_2}\right].\tag{15}$$

The flow (12) is along $\nabla I_2^{\mathbf{u}}$. Thus, constant intensity regions remain unaltered and are deformed only through interactions along their boundaries. Its magnitude is controlled both by contrast and luminance statistics, encoded in $\|\nabla I_2^{\mathbf{u}}\|$ and $V_1^{I_1,I_2^{\mathbf{u}}}$, respectively. From (13), it is seen that $V_1^{I_1,I_2^{\mathbf{u}}}$ involves regularized i_2-derivatives (defined via linear filtering against Parzen kernels) of the terms $L_i^{\mathbf{u}}$ $(i=1,2)$. The latter are functions of the variable $\rho^{\mathbf{u}}$ which measures the dependence between luminance distributions, and are interpreted as follows. Deforming the floating image $I_2^{\mathbf{u}}$ during the registration process causes the information between I_1 and $I_2^{\mathbf{u}}$ to vary, inducing a statistical force. Quantifying these variations w.r.t. the information metrics f requires defining an observed variable and a reference measure. Two choices are possible, yielding respective contributions to the statistical force:

1. The observed variable is the joint distribution $(I_1,I_2^{\mathbf{u}})$: in this case, the reference measure is chosen as the product distribution $I_1\times I_2^{\mathbf{u}}$. This leads to considering the variable $\rho^{\mathbf{u}}$. The variation of information w.r.t. the primal metrics f is then the flux $f'(\rho^{\mathbf{u}})$, i.e., $L_1^{\mathbf{u}}$.
2. The observed variable is the marginal distribution $I_2^{\mathbf{u}}$: comparison with I_1 requires generating hypotheses on possible realizations i_1 of I_1. Sampling I_1 allows to define the product observation $I_1\times I_2^{\mathbf{u}}$ whose information can be quantified w.r.t. to the joint measure $(I_1,I_2^{\mathbf{u}})$. This leads to considering the variable $\frac{1}{\rho^{\mathbf{u}}}$. The appropriate information metrics is then the dual metrics $\breve{f}$, and the variation of information w.r.t. $\breve{f}$ is the flux $\breve{f}'(\frac{1}{\rho^{\mathbf{u}}})$, i.e., $L_2^{\mathbf{u}|i_1}$. The process is repeated for every possible realization of I_1. Finally, the average value of information variation, corresponding to expectation w.r.t. I_1, is retained.

Eq. (12) generalizes to arbitrary integral f-informations the result derived in (Hermosillo et al., 2001; Hermosillo, 2002) for MI, which corresponds to $f=f_{\mathrm{KL}}$. In this case, the expectation in (13) vanishes and

$$\frac{\partial L_1^{\mathbf{u}}}{\partial i_2}=\frac{\partial\log p^{I_1,I_2^{\mathbf{u}}}}{\partial i_2}-\frac{\partial\log p^{I_2^{\mathbf{u}}}}{\partial i_2}.\tag{16}$$

The term $L_2^{\mathbf{u}}$ is, thus, informationless in the Shannon framework. In contrast, using $f\neq f_{\mathrm{KL}}$ allows to exploit the information content of $I_2^{\mathbf{u}}$ in a more extensive fashion.

Table 1
Analytical structure of the variational derivative of integral f-informations

$L_1^{\mathbf{u}}$	$f'(\rho^{\mathbf{u}})$
$L_2^{\mathbf{u}}$	$f(\rho^{\mathbf{u}})-\rho^{\mathbf{u}}f'(\rho^{\mathbf{u}})=\breve{f}'\left(\frac{1}{\rho^{\mathbf{u}}}\right)$
$L_3^{\mathbf{u}}$	$2f\left(\frac{1}{p^{I_2^{\mathbf{u}}}}\right)p^{I_2^{\mathbf{u}}}-f'\left(\frac{1}{p^{I_2^{\mathbf{u}}}}\right)$
$K(I_1,I_2^{\mathbf{u}})$	$\frac{1}{\lvert\underline{D}\rvert}$

3.1.2. Non-integral f-information measures

A non-integral f-information similarity functional is

$$\mathscr{S}(\mathbf{u})=-I_\psi(I_1,I_2^{\mathbf{u}})=-\log\psi^{-1}\big(B_\psi(I_1,I_2^{\mathbf{u}})\big),\tag{17}$$

$$B_\psi(I_1,I_2^{\mathbf{u}})=\int_{\Omega^2}p^{I_1,I_2^{\mathbf{u}}}(i_1,i_2)\psi(\rho^{\mathbf{u}}(i_1,i_2))\,\mathrm{d}i_1\,\mathrm{d}i_2.\tag{18}$$

Its variational derivative $\partial_{\mathbf{u}}\mathscr{S}(\mathbf{u})$ is derived along the same lines as above (Petitjean et al., 2002; Petitjean, 2003c). One obtains the same expressions (12) and (13) as for integral f-information, with specific definitions for the symbols $K(I_1,I_2^{\mathbf{u}})$, $L_1^{\mathbf{u}}$ and $L_2^{\mathbf{u}}$ given in Table 2. One has

Table 2
Analytical structure of the variational derivative of non-integral f-informations

$L_1^{\mathbf{u}}$	$\psi(\rho^{\mathbf{u}}) + \rho^{\mathbf{u}}\psi'(\rho^{\mathbf{u}})$
$L_2^{\mathbf{u}}$	$-(\rho^{\mathbf{u}})^2\psi'(\rho^{\mathbf{u}})$
$L_3^{\mathbf{u}}$	$\psi\left(\frac{1}{p^{I_2^{\mathbf{u}}}}\right) - \frac{1}{p^{I_2^{\mathbf{u}}}}\psi'\left(\frac{1}{p^{I_2^{\mathbf{u}}}}\right)$
$K(I_1, I_2^{\mathbf{u}})$	$\frac{1}{\lvert\underline{D}\rvert}\frac{(\psi^{-1})'(B_\psi(I_1,I_2^{\mathbf{u}}))}{\psi^{-1}(B_\psi(I_1,I_2^{\mathbf{u}}))}$

By defining $f(x) := -x\psi(x)$ according to the connection identity (5), the symbols $L_i^{\mathbf{u}}$ $(i = 1, 2, 3)$ rewrite as in Table 1. It follows that the variational derivatives of integral and non-integral f-informations are dual.

$$\frac{\partial L_2^{\mathbf{u}|i_1}}{\partial i_2} = -\rho^{\mathbf{u}}\frac{\partial L_1^{\mathbf{u}}}{\partial i_2}, \tag{19}$$

$$\frac{\partial L_1^{\mathbf{u}}}{\partial i_2} = \rho^{\mathbf{u}}[2\psi'(\rho^{\mathbf{u}}) + \rho^{\mathbf{u}}\psi''(\rho^{\mathbf{u}})]\left[\frac{\partial \log p^{I_1,I_2^{\mathbf{u}}}}{\partial i_2} - \frac{\partial \log p^{I_2^{\mathbf{u}}}}{\partial i_2}\right]. \tag{20}$$

Setting $\psi = \psi_{\mathrm{KL}}$ yields the MI flow (Hermosillo et al., 2001). For $\psi = \psi_r$, one obtains the Renyi information flow.

Using the identity (5) relating integral and non-integral f-divergences, it is straightforward to show that the defining equations for $L_1^{\mathbf{u}}$ and $L_2^{\mathbf{u}}$ provided in Tables 1 and 2, and, consequently, their derivatives expressed by Eqs. (14) and (15) and (19) and (20), are dual. It follows that the gradient flows for integral and non-integral informations are dual.

3.1.3. Normalized generalized information measures

Following the perturbative approach of Section 3.1.1, the variational derivative of the similarity functionals $\mathscr{S}(\mathbf{u}) = -\mathrm{ECC}_f(I_1, I_2^{\mathbf{u}})$, $\mathscr{S}(\mathbf{u}) = -\mathrm{NMI}_f(I_1, I_2^{\mathbf{u}})$ and $\mathscr{S}(\mathbf{u}) = Z_f(I_1, I_2^{\mathbf{u}})$ induced by normalized generalized information measures are shown to be of the following generic form (Petitjean et al., 2002; Petitjean, 2003c):

$$\partial_{\mathbf{u}}\mathscr{S}(\mathbf{u}) = -w_0^{I_1,I_2^{\mathbf{u}}}\left[w_1^{I_1,I_2^{\mathbf{u}}}\partial_{\mathbf{u}} I_f(I_1, I_2^{\mathbf{u}}) - w_2^{I_1,I_2^{\mathbf{u}}}\partial_{\mathbf{u}} I_f(I_2^{\mathbf{u}}, I_2^{\mathbf{u}})\right], \tag{21}$$

where the weighting coefficients $w_i^{I_1,I_2^{\mathbf{u}}}$ $(i \in \{0,1,2\})$ are defined in Table 3. The variational derivative $\partial_{\mathbf{u}} I_f(I_1, I_2^{\mathbf{u}})$ of f-information is given by (12) and (13). The variational derivative $\partial_{\mathbf{u}} I_f(I_2^{\mathbf{u}}, I_2^{\mathbf{u}})$ of self f-information is easily computed. The result is the following:

$$\partial_{\mathbf{u}} I_f(I_2^{\mathbf{u}}, I_2^{\mathbf{u}}) = V_2^{I_2^{\mathbf{u}}}(\mathbf{x})\nabla I_2^{\mathbf{u}}(\mathbf{x}), \tag{22}$$

$$V_2^{I_2^{\mathbf{u}}}(\mathbf{x}) := K\left\{\left[\mathscr{K}_1 \bigstar \frac{\partial L_3^{\mathbf{u}}}{\partial i_2}\right](I_2^{\mathbf{u}}(\mathbf{x}))\right\}, \tag{23}$$

where $K(I_1, I_2^{\mathbf{u}})$ and $L_3^{\mathbf{u}}$ are defined in Table 1 for the integral case where

$$\frac{\partial L_3^{\mathbf{u}}}{\partial i_2} = \left[2f\left(\frac{1}{p^{I_2^{\mathbf{u}}}}\right) - \frac{2}{p^{I_2^{\mathbf{u}}}}f'\left(\frac{1}{p^{I_2^{\mathbf{u}}}}\right) + \frac{1}{\left(p^{I_2^{\mathbf{u}}}\right)^2}f''\left(\frac{1}{p^{I_2^{\mathbf{u}}}}\right)\right]\frac{\partial p^{I_2^{\mathbf{u}}}}{\partial i_2} \tag{24}$$

and in Table 2 for the non-integral case where

$$\frac{\partial L_3^{\mathbf{u}}}{\partial i_2} = \frac{1}{\left(p^{I_2^{\mathbf{u}}}\right)^3}\psi''\left(\frac{1}{p^{I_2^{\mathbf{u}}}}\right)\frac{\partial p^{I_2^{\mathbf{u}}}}{\partial i_2}. \tag{25}$$

One can verify that Eqs. (24) and (25) are dual. The structure of this flow is similar to (12) and (13), the statistical control being here based on the variation of the information content of $I_2^{\mathbf{u}}$, analyzed via the term $L_3^{\mathbf{u}}$. When $f = f_{\mathrm{KL}}$, one obtains the flow for the Shannon marginal entropy $H_{\mathrm{S}}(I_2^{\mathbf{u}})$.

The flow (21) can be finally written as

$$\partial_{\mathbf{u}}\mathscr{S}(\mathbf{u}) = -V_3^{I_1,I_2^{\mathbf{u}}}(\mathbf{x})\nabla I_2^{\mathbf{u}}(\mathbf{x}), \tag{26}$$

$$V_3^{I_1,I_2^{\mathbf{u}}}(\mathbf{x}) := w_0^{I_1,I_2^{\mathbf{u}}}\left[w_1^{I_1,I_2^{\mathbf{u}}}V_1^{I_1,I_2^{\mathbf{u}}}(\mathbf{x}) - w_2^{I_1,I_2^{\mathbf{u}}}V_2^{I_2^{\mathbf{u}}}(\mathbf{x})\right]. \tag{27}$$

When $f = f_{\mathrm{KL}}$, Eqs. (26) and (27) describe the gradient flow for ECC, NMI and exclusive information. It is easily shown that they reduce to (12) and (13) with null expectation and

$$\frac{\partial L_1^{\mathbf{u}}}{\partial i_2} = w_3^{I_1,I_2^{\mathbf{u}}}\left[w_4^{I_1,I_2^{\mathbf{u}}}\frac{\partial \log p^{I_1,I_2^{\mathbf{u}}}}{\partial i_2} - w_5^{I_1,I_2^{\mathbf{u}}}\frac{\partial \log p^{I_2^{\mathbf{u}}}}{\partial i_2}\right], \tag{28}$$

where the weighting coefficients $w_i^{I_1,I_2^{\mathbf{u}}}$ $(i \in \{3,4,5\})$ are defined in Table 4. Studying their asymptotic behavior when $\mathbf{u} \to \mathbf{u}^*$ enlightens interesting properties of ECC, NMI and exclusive information. Assuming perfect registration, i.e., the random variables I_1 and $I_2^{\mathbf{u}^*}$ obey the same distribution, and recalling the property $H_{\mathrm{S}}(X,X) = H_{\mathrm{S}}(X)$, it is straightforward to prove that: $\lim_{\mathbf{u}\to\mathbf{u}^*} w_4^{I_1,I_2^{\mathbf{u}}} = 2$ and $\lim_{\mathbf{u}\to\mathbf{u}^*} w_5^{I_1,I_2^{\mathbf{u}}} = 1$. Comparing this result to the MI case (16), for which one has: $w_4^{I_1,I_2^{\mathbf{u}}} = w_5^{I_1,I_2^{\mathbf{u}}} = 1$, shows that MI, on the one hand, and ECC, NMI and Z, on the other hand, belong to different families of Shannon information measures. We conjecture that the increased robustness of normalized information measures to image overlap (Maes et al., 1997; Studholme et al., 1999) originates from the specific weighting of $\frac{\partial}{\partial i_2}\log p^{I_1,I_2^{\mathbf{u}}}$ and $\frac{\partial}{\partial i_2}\log p^{I_2^{\mathbf{u}}}$. This observation can be extended to f-informations vs. ECC_f, NMI_f and exclusive f-informations (Petitjean, 2003c), taking into account that: $\lim_{\mathbf{u}\to\mathbf{u}^*} w_1^{I_1,I_2^{\mathbf{u}}} = 2$ and $\lim_{\mathbf{u}\to\mathbf{u}^*} w_2^{I_1,I_2^{\mathbf{u}}} = 1$.

3.1.4. Linear parametric motion models

Similar computations can be carried out for deriving the expressions of the 1st variation of generalized information measures over linear parametric motion spaces. The resulting flows can be expressed under the following unified form:

Table 3
Weighting coefficients involved in the variational derivative of generalized normalized information measures

Criterion	$w_1^{I_1,I_2^{\mathbf{u}}}$	$w_2^{I_1,I_2^{\mathbf{u}}}$	$w_0^{I_1,I_2^{\mathbf{u}}}$
ECC_f	2	$\mathrm{ECC}_f(I_1,I_2^{\mathbf{u}})$	$\frac{1}{I_f(I_1,I_1)+I_f(I_2^{\mathbf{u}},I_2^{\mathbf{u}})}$
NMI_f	$\mathrm{NMI}_f(I_1,I_2^{\mathbf{u}})$	$\frac{\mathrm{NMI}_f(I_1,I_2^{\mathbf{u}})\mathrm{ECC}_f(I_1,I_2^{\mathbf{u}})}{2}$	$\frac{\mathrm{NMI}_f(I_1,I_2^{\mathbf{u}})}{I_f(I_1,I_1)+I_f(I_2^{\mathbf{u}},I_2^{\mathbf{u}})}$
Z_f	2	1	1

$$\partial_{\Theta}\mathscr{S}(\mathbf{u}^{\Theta}) = -\int_{\underline{D}} V^{I_1,I_2^{\mathbf{u}}}(\mathbf{x})\left[\left(\nabla I_2^{\mathbf{u}}\right)^t\mathbf{B}\right](\mathbf{x})\,\mathrm{d}\mathbf{x}, \tag{29}$$

where $V^{I_1,I_2^{\mathbf{u}}} := V_1^{I_1,I_2^{\mathbf{u}}}$ for f-informations and $V^{I_1,I_2^{\mathbf{u}}} := V_3^{I_1,I_2^{\mathbf{u}}}$ for normalized generalized information measures. The flow in the parameter space appears to be the sum of local contributions along the projection $[(\nabla I_2^{\mathbf{u}})^t\mathbf{B}]$ of $\nabla I_2^{\mathbf{u}}$ onto the basis $\mathbf{B}$ of $\mathscr{T}$. Notice that, for $f=f_{\mathrm{KL}}$, the resulting expressions differ from Viola's approach (Viola and Wells, 1997). This is because we *exactly* optimize the MI similarity functional instead of using an ergodic approximation as in (Viola and Wells, 1997).

3.2. First-order regularization

Non-parametric and high-order parametric motion models require explicit regularization. Since regularizing functionals involving low-order motion derivatives allow for reduced over-smoothing artifacts and computationally efficient schemes, 1st-order differential stabilizers have been retained. Moreover, in our application context, NRR is targeted towards motion estimation instead of geometric alignment as in most situations. Ensuring the physical consistency of recovered motion is, therefore, a key requirement. This implies using a functional $\mathscr{R}(\mathbf{u})$ which guarantees that smoothing is confined within the boundaries of moving structures. This condition is not met by classical stabilizers (e.g., thin-plate (Bookstein, 1989; Rueckert et al., 1999), linear elasticity (Bajcsy and Kovačič, 1989; Christensen et al., 1996), viscous fluid (Christensen et al., 1996b)) which do not incorporate discontinuity information and prove to be excessively diffusive, resulting into inter-region smoothing even though yielding excellent alignment of anatomical structures. Applied to cardiac tagged MR data, these stabilizers deliver unrealistic vector fields which exhibit strong blurring artifacts in the vicinity of endocardial and epicardial boundaries. These artifacts originate from mixing unrelated motion information between myocardium and blood pools or thoracic structures.

As in Hermosillo (2002), we have used the Nagel–Enkelmann oriented smoothness constraint (Nagel and Enkelmann, 1986), which is a classical 1st-order geometric stabilizer involving the Jacobian $\nabla\mathbf{u}$ of $\mathbf{u}$:

$$\mathscr{R}(\mathbf{u}) := \frac{1}{2}\mathrm{trace}(\nabla\mathbf{u}^t T_{I_1}\nabla\mathbf{u}), \tag{30}$$

$$T_{I_1} := \frac{1}{\|\nabla I_1\|^2 + 2\beta}\left[\left(\beta + \|\nabla I_1\|^2\right)\mathbb{I}\mathrm{d} - \nabla I_1\nabla I_1^t\right]. \tag{31}$$

Here, $\mathbb{I}\mathrm{d}$ denotes the identity tensor and $\beta>0$ is a contrast parameter. The tensor T_{I_1} defines a projector onto the hyperplane orthogonal to ∇I_1, inhibiting regularization along ∇I_1. Consequently, $\mathscr{R}(\mathbf{u})$ promotes intra-region smoothing, yielding motion estimates localized within image objects. The first variation of $\mathscr{R}(\mathbf{u})$ is

$$\partial_{\mathbf{u}}\mathscr{R}(\mathbf{u}) = -\nabla\cdot(T_{I_1}\nabla\mathbf{u}). \tag{32}$$

An alternative choice is half-quadratic regularization (Hellier et al., 2001), which allows for motion discontinuity preservation.

Table 4
Weighting coefficients involved in the variational derivative of normalized information measures

Criterion	$w_4^{I_1,I_2^{\mathbf{u}}}$	$w_5^{I_1,I_2^{\mathbf{u}}}$	$w_3^{I_1,I_2^{\mathbf{u}}}$
ECC	2	$\frac{2}{NMI(I_1,I_2^{\mathbf{u}})}$	$\frac{1}{H_S(I_1)+H_S(I_2^{\mathbf{u}})}$
NMI	$NMI(I_1,I_2^{\mathbf{u}})$	1	$\frac{1}{H_S(I_1,I_2^{\mathbf{u}})}$
Z	2	1	1

3.3. Implementation issues

We have chosen a simple and computationally efficient implementation which has proven to perform robustly. Obviously, more sophisticated choices can be made at every stage, potentially yielding improvements to be assessed. The hyperparameters of the method have been tuned empirically. We have verified experimentally that using a fixed image acquisition protocol allows to guarantee the robustness of these settings.

- *Estimating the probability distributions:* Empirical pdf estimates derived from normalized histograms have been used. The joint histogram of I_1 and $I_2^{\mathbf{u}}$ over the domain of overlap $\underline{D}$ is constructed by binning intensity pairs using 256 bins for both images. Bilinear or partial volume (Maes et al., 1997) interpolation can be used for estimating $I_2^{\mathbf{u}}$ at transformed positions $\mathbf{u}(\mathbf{x})$, yielding similar performances. Marginal pdf are obtained by summing over rows and columns of the (smoothed) normalized joint histogram. A known side effect is that the resulting estimate of p^{I_1} does depend on $\mathbf{u}$.
- *Computing the statistical forces:* Isotropic Parzen Gaussian kernels $\mathscr{K}_1$ and $\mathscr{K}_2$, with respective standard deviations σ_1 and σ_2, have been retained, yielding Gaussian derivatives in the luminance domain. The latter are computed by smoothing the joint histogram using a 4th-order IIR approximation of the Gaussian filter (Deriche, 1993), and estimating the derivatives $\frac{\partial L_1^{\mathbf{u}}}{\partial i_2}$ and $\frac{\partial L_2^{\mathbf{u}|i_1}}{\partial i_2}$ using central finite differences (FD). The expectation $E_{I_1}(\cdot)$ w.r.t. the random variable I_1 is estimated empirically as an arithmetic average w.r.t. the variable i_1. The Parzen kernel widths, which have a critical impact on convergence and accuracy (D'Agostino et al., 2003), have been empirically set to $\sigma_1 = \sigma_2 = 3$. Alternatively, the leave-k-out cross-validation technique proposed in (Hermosillo, 2002) can be applied to marginal histograms to automate their optimal determination.
- *Computing the image derivatives:* Image gradients are robustly estimated using the Canny–Deriche filter with smoothing parameter set to 1.5.
- *Computing the regularization forces:* The Nagel–Enkelmann tensor is computed analytically from Eq. (31) with $\beta = 10000$. Central FD are used for estimating the partial derivatives of $\mathbf{u}$ and discretizing (32). The regularization parameter has been empirically set to $\lambda = 20$.
- *Discretizing the flow:* Gradient descent is performed via a crude Euler explicit scheme

$$\mathbf{u}^t(\mathbf{x}) = \mathbf{u}^{t-1}(\mathbf{x}) - \Delta t \partial_{\mathbf{u}} \mathscr{J}(\mathbf{u}^{t-1}(\mathbf{x}))$$

  with fixed time step $\Delta t = 0.015$.
- *Multiresolution implementation:* In order to deal with large displacements, a standard coarse-to-fine strategy based on Gaussian dyadic pyramids for luminance and displacement has been considered (Lester and Arridge, 1999). We have used 3 resolution levels, including full resolution.
- *Testing for convergence:* Since generalized information measures do not contain spatial information, testing convergence by monitoring their evolution along iterations is not practical. Instead, we monitor the variations of displacement estimates between two iterations, and decide for convergence by comparing against a predefined threshold (consistently dilated according to the resolution level). In practice, we have used the following norm:

$$\|\mathbf{u}^t - \mathbf{u}^{t-1}\|_{\mathscr{F}} := \max_{x \in \underline{D}} \left(\|\mathbf{u}^t(\mathbf{x}) - \mathbf{u}^{t-1}(\mathbf{x})\|_{L^2} \right).$$

Computation time per exam, typically consisting of 6 sequences (3 short-axis (SA) and 3 long-axis (LA) levels) of 18 images of size 80×80, is about 30 min on a PC with a PIII 1.2 GHz processor. Extracting systolic motions at 3 SA levels requires $\approx$60 s.

4. Quantifying myocardial deformations

Performing NRR between successive image pairs of a tagged MR sequence $(I_t)_{0 \leqslant t \leqslant T}$ delivers frame-to-frame displacement estimates $(\mathbf{u}^*_{t,t+1})_{0 \leqslant t < T}$. Integrating them over time starting from ED ($t = 0$) yields absolute displacements $(\mathbf{u}^*_t)_{0 \leqslant t \leqslant T}$. Systolic motions can be derived in this way, or, for static studies, by direct NRR between ED and ES frames.[6] Quantifying myocardial deformations can then be achieved either from frame-to-frame or absolute displacements.

4.1. Computing deformation parameters

The myocardium has a complex structure consisting of locally parallel muscle cells (myocytes), a vascular network and a dense plexus of connective tissue (Glass et al., 1991). Myocytes are tied together by a collagenous network and bundled together into fibers. Fibers are organized into a nested architecture with a continuously varying transmural orientation distribution: fibers in the mid wall are circumferential whereas subendocardial fibers are longitudinally directed (Nielsen et al., 1991). From a mechanical point of view, the myocardium is therefore an inhomogeneous, anisotropic and nearly incompressible material characterized by a complex stress-strain relationship correlated to fiber structure (Glass et al., 1991; Hu et al., 2003). The left ventricle (LV) deformations comprise radial thickening, circumferential shortening, torsion, and longitudinal shortening. For the right ventricle (RV), the longitudinal fiber orientation governs the dominant motion.

[6] Experiments have shown that the two approaches yield similarly accurate results, indicating that error propagation artifacts inherently induced by integration are limited. This can be explained by the conjunction of two factors: (i) as shown in Section 5.2, the proposed NRR technique delivers very accurate frame-to-frame displacements estimates; (ii) tags undergo a stationary demagnetization process and are persistent within the myocardium during systole; this eliminates the tedious problem of devising robust integration schemes w.r.t. missing information.

Local deformation information is derived either from the recovered displacement fields themselves (0th-order attributes), or from their spatial variations (1st-order attributes) encoded by the strain tensor $\mathbf{E}(\mathbf{u})$, defined as:

$$\mathbf{E}(\mathbf{u}) := \tfrac{1}{2}(\nabla\mathbf{u} + \nabla\mathbf{u}^t + \nabla\mathbf{u}^t\nabla\mathbf{u}). \quad (33)$$

Letting $\mathbf{d}$ and $\mathbf{d}'$ denote unit vector fields over D ($\|\mathbf{d}\| = \|\mathbf{d}'\| = 1$, $\mathbf{d} \neq \mathbf{d}'$):

- 0th-order attributes allow for assessing directional displacements $(\mathbf{d}^t\mathbf{u})\mathbf{d}$. For SA views, radial displacement (resp. rotation/torsion along the LA of the heart) is defined by letting $\mathbf{d}$ be the vector field $\mathbf{n}$ (resp. $\mathbf{t}$) normal (resp. tangent) to myocardial boundaries. For LA views, radial (resp. longitudinal) displacement is obtained by letting $\mathbf{d}$ be the vector field $\mathbf{z}^\perp$ (resp. $\mathbf{z}$) normal (resp. parallel) to LA.
- Relevant 1st-order attributes comprise radial ($\mathbf{d} = \mathbf{n}$), circumferential ($\mathbf{d} = \mathbf{t}$) and longitudinal ($\mathbf{d} = \mathbf{z}$) strains $Q_{\mathbf{d}} := \mathbf{d}^t\mathbf{E}(\mathbf{u})\mathbf{d}$ and shears $Q_{\mathbf{d},\mathbf{d}'} := \mathbf{d}^t\mathbf{E}(\mathbf{u})\mathbf{d}'$, and eigenvectors/eigenvalues of $\mathbf{E}(\mathbf{u})$ which define extremal deformation directions/magnitudes.

In practice, the fields $\mathbf{t}$ and $\mathbf{n}$ are derived either from a polar coordinate system located at the center of gravity of the LV (Fig. 3), or from the gradient of a distance function over a binary segmentation map of the myocardium (cf. Section 4.2). The latter solution applies to the whole heart and accounts for its intrinsic geometry whereas the former is restricted to LV analysis and involves an extrinsic geometry. For non-parametric motions, the strain tensor $\mathbf{E}(\mathbf{u})$ is computed numerically using central FD. For parametric motions, a closed-form expression onto the local motion basis can be derived, allowing analytic computations.

Averaging these various attributes over each region of a layered regional model of the myocardium yields segmental descriptions of the deformation. Due to the lack of consensus regarding RV segmentation, which is made difficult by its relative thinness and complex geometry, only LV segmental measurements are reported in this paper. They refer to a 3 SA-levels, 16-segments LV model compliant with the American Heart Association (AHA) recommendations (Cerqueira et al., 2002), shown in Fig. 4.

For the healthy heart, both individualized and population-related measurements have been investigated during the whole cardiac cycle. The latter have been derived by averaging individual segmental parameters over a reference database (cf. Section 5.1).

4.2. Segmenting and tracking the myocardium

Using a stabilizer with intra-region smoothing properties allows to considerably decrease the bias of surrounding tissues onto displacement estimates within the myocardium. Since NRR operates over the whole image domain, additional motion data are recovered within the heart chambers and over the thorax. For myocardial function quantification purpose, the latter are useless and must be discarded to reduce computations and enhance visualization. Myocardium segmentation is thus a necessary step.

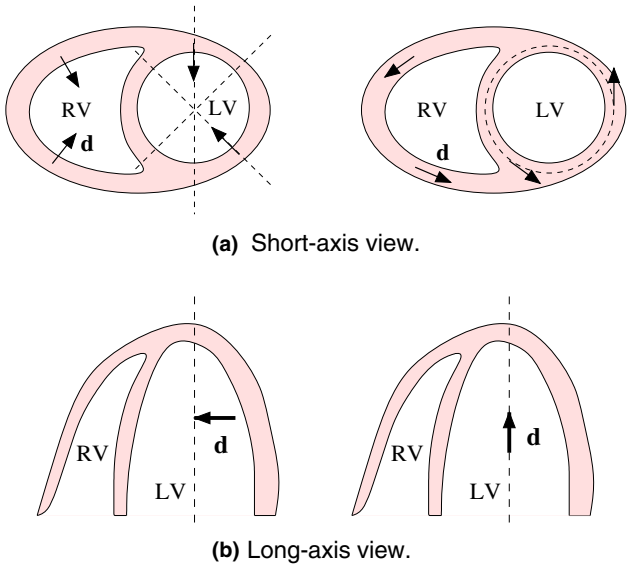

Fig. 3. Relevant orientations for deformation assessment in (a) short-axis and (b) long-axis views. For the LV, a cylindrical coordinate system with origin at the LV gravity center and along the long-axis of the heart is often used.

To this end, rather than processing the tagged MR data in which the tagging pattern interferes with myocardial boundaries, we use the jointly acquired cine MR images which provide a direct access to heart anatomy. Segmentation is then achieved within the same NRR framework as for motion estimation by:

1. segmenting the ED image by any technique of choice (Suri, 2000); in our implementation, we have favored mathematical morphology and used a classical watershed segmentation scheme;
2. performing frame-to-frame NRR over the whole cine MR sequence, starting from the ED image; thanks to the use of a uniform acquisition protocol for tagged and cine MR data (cf. Section 5.1), the information metrics and hyperparameter set used for motion estimation are applicable;
3. integrating the estimated frame-to-frame displacements over time to produce absolute displacements;
4. warping the ED segmentation mask with the absolute displacement fields to generate segmentations in subsequent phases.

For static studies, NRR is performed between the ED and ES frames only, yielding directly absolute systolic displacements.

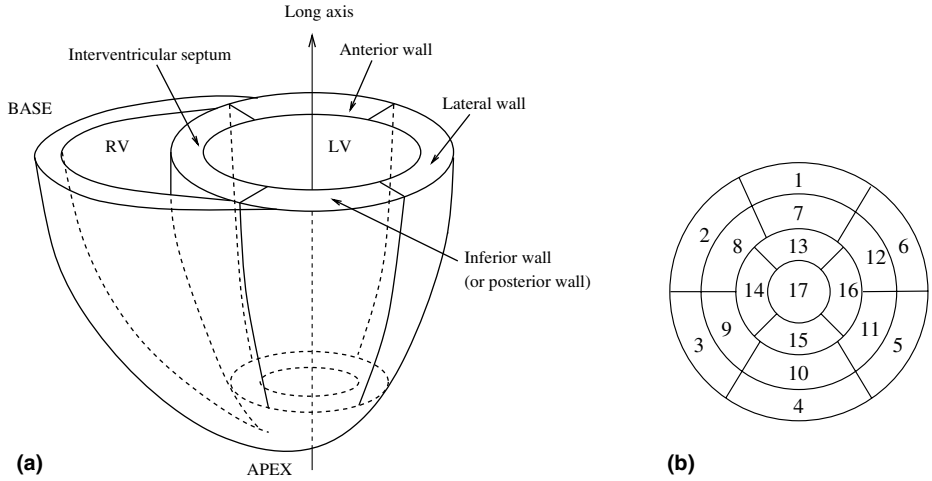

Fig. 4. (a) Myocardial walls and (b) bull-eye representation of the 3-levels, 16-segments LV model recommended by the AHA. Segments #1–6, #7–12 and #13–16 refer to the basal, mid and apical levels, respectively. Segment #17, associated with the apex, is optional and has not been used.

5. Results

5.1. Materials

Experiments have been performed on a cardiac MRI database comprising healthy and pathological data. Healthy exams originate from an homogeneous population of 12 volunteers (9 men/3 women, mean age $\pm$ SD $= 30 \pm 4$ years). Pathological data consist of acute ischemia, dilated cardiomyopathies (DCM) and sarcoïdosis. Tagged MR data were acquired during multiple breath holds on a 1.5T GE scanner using the SPAMM (SPAtial Modulation of Magnetization) (Axel and Dougherty, 1989) tagging technique with ECG triggering.[7,8] For each patient, 10 base-to-apex SA and 8 LA sequences were captured. A sequence comprised 15–21 images depending on the heart rate. The cine MR sequences required for myocardium segmentation were acquired at the same slice levels in identical conditions.[9] 3 SA levels, in compliance with the AHA recommendations (Cerqueira et al., 2002), and 3 LA levels have been effectively processed.

5.2. Evaluation and information metrics selection

An objective evaluation of the proposed approach has been carried out using simulated tagged MR data. The latter are created by warping real images by realistic synthetic motion data and simulating tag fading out. The synthetic motion consists of the mean displacement component $(\mathbf{u}_t^{\text{atlas}})_{0 \leqslant t \leqslant T}$ of a multi-phase statistical cardiac motion atlas (Petitjean, 2003c; Petitjean et al., 2004). Tag fading out is simulated by histogram specification. Formally, starting from a reference tagged MR sequence $(I_t)_{0 \leqslant t \leqslant T}$, a simulated image $\tilde{I}_t$ at phase t $(0 < t \leqslant T)$ is obtained by warping the initial frame I_0 by $\mathbf{u}_t^{\text{atlas}}$, and by imposing to the transformed image $I_0^{\mathbf{u}_t^{\text{atlas}}}$ the histogram of the real image I_t by means of histogram specification. Accuracy is assessed by comparing the displacement $\mathbf{u}_t^*$, estimated by registering I_0 onto $\tilde{I}_t$, to the ground truth $\mathbf{u}_t^{\text{atlas}}$ using the classical L^2 norm

$$\varepsilon(\mathbf{x}) := \|\mathbf{u}_t^{\text{atlas}}(\mathbf{x}) - \mathbf{u}_t^*(\mathbf{x})\|_{L^2}.$$

Experiments have been carried on simulated SPAMM, CSPAMM and DANTE data. Due to the lack of a theoretical basis for automatically selecting a suitable information metrics, systematic testing has been achieved for various integral and non-integral divergences. Specifically, the I_α, χ_α, Matusita and Renyi divergences (Vajda, 1989) have been tested, for various values of their control parameter. The best results have been obtained for the information metrics f_α associated to I_α-divergence with $\alpha = 1.2$, which was found to allow for accurately recovering both frame-to-frame and systolic motions.

[7] The acquisition parameters were as follows: 285 mm field of view, 256×128 acquisition matrix, 1.48×1.48 mm in-plane resolution, 45 ms time interval, 10 mm slice thickness, 10 mm tag spacing.

[8] Additional datasets were also collected using the CSPAMM (Continuous SPAMM) (Fischer et al., 1993) and DANTE (Delays Alternating with Nutations for Tailored Excitations) (Mosher and Smith, 1990) sequences to assess sensitivity w.r.t. the tagging protocol.

[9] Despite using identical levels, acquisitions at different breath holds potentially yield spatial misalignment of tagged and MR data due to respiratory-induced through-plane motion and in-plane deformations. The former cannot be fixed. Compensating for the latter would require performing NRR between tagged and cine ED frames. Upon visual inspection, such artifacts have proved, however, to be very limited for the cohort and have been neglected. Moreover, accurately dealing with them is not a critical issue, since myocardium segmentation is only meant to enhance the visualization of dense parameter maps and does not influence quantification.

Using the f_α metrics, we have then compared the performances of the f-information, exclusive f-information, generalized ECC and generalized NMI similarity criteria. The conclusion is that these various criteria yield similar results, but exclusive f-information converges significantly faster than f-information (Fig. 5) while requiring a lower computational effort than ECC_f and NMI_f. The convergence speed up factor[10] w.r.t. to I_f varies between 3 and 8. The exclusive f-information criterion has, therefore, been finally retained.

A typical example of frame-to-frame motion recovery from simulated SPAMM data is shown in Fig. 6. The results, synthesized in Table 5, demonstrate that the proposed approach allows to estimate reliable frame-to-frame displacements, with sub-pixel average accuracy and maximal error less than 3 pixels. The same conclusion holds for large magnitude displacements such as systolic motions. Moreover, performance for systolic frames does not depend on the MR-tagging technology. In any case, tags are matched accurately, as illustrated in Fig. 7.

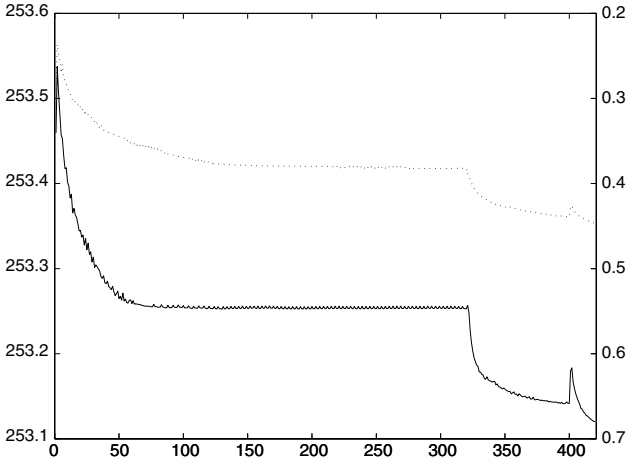

Fig. 5. Evolution of f-information $-I_f$ (dashed line, right scale) vs. exclusive f-information Z_f (plain line, left scale) along iterations for the information metrics f_α ($\alpha = 1.2$) during hierarchical NRR between synthetic SPAMM data. Three resolution levels have been used. In order to highlight convergence speed up, a fixed number of iterations has been imposed at each level. The leaps correspond to transitions towards finer resolutions.

5.3. Non-pathological data

Numerous experiments have been performed on non-pathological data. Fig. 8 shows typical displacement fields recovered from SA sequences at basal, mid and apical levels during systole for a normal heart. When dealing with natural images, assessing the accuracy of the proposed method for motion recovery is difficult due to the lack of reliable ground truth. Instead, we have attempted to objectively assess its performance for quantifying myocardial function. To this end, the literature on myocardial strain measurement has been reviewed, yielding reference values and confidence intervals for the various deformation parameters described in Section 4.1 (Petitjean et al., 2003b).

Quantification has proven to corroborate well-established qualitative results on heart anatomy. Regarding 0th-order attributes, the radial contraction increases on lateral and inferior walls (Fig. 10), reaching 4.5 mm in the apex (Fig. 9). In addition, the septum still contracts while free walls start dilating. This behavior seems to be linked to the RV motion. Rotation sense differs at the apex and the base (Fig. 11). Measurements on LA images (Fig. 12) show that longitudinal contraction mostly takes place at the base and that radial contraction is localized on the LV free walls. An initial translation of the apex towards the base is also observed. This translation has been previously reported in (Moore et al., 2000) with an amplitude of $\approx$2 mm. Whatever the attribute, automated systolic measurements have been found to lie within reference confidence intervals (Petitjean et al., 2003b).

Moreover, quantifying deformations dynamically allows to highlight phenoma which have not been previously described in the literature and require more investigation. Averaging population-specific measurements of radial contraction for each segment over all levels, and for each level over all segments, respectively, clearly underlines the propagation of deformation: contraction undergoes a clockwise motion from anterior to septal wall (Fig. 13); it propagates longitudinally from base to apex in the septum and, in a less obvious manner, from apex to base in each of the remaining segments (Fig. 14). One can notice that electrical activation behaves similarly.

Regarding 1st-order parameters, radial strain is positive, indicating wall thickening, until early diastole for all segments. At ES, radial strain is negative for anterior and lateral walls, and zero or positive for septal and inferior walls, showing that diastolic thinning has started for anterior and lateral walls. As previously noticed in (Moore et al., 2000), radial strain exhibits, however, a high inter-subject variability which may question the clinical usability of this parameter. Circumferential strain increases on apical antero-lateral and basal inferior walls (Fig. 15). Maximal eigenvalues line up with radial strain values, underlining the fact that maximum stretch occurs in the radial direction.

[10] Convergence speed is measured in terms of number of iterations, or, equivalently, in terms of CPU time, given that the additional cost for computing the derivative $\partial_{\mathbf{u}} I_f(I_2^{\mathbf{u}}, I_2^{\mathbf{u}})$ involved in $\partial_{\mathbf{u}} Z_f(I_1, I_2^{\mathbf{u}})$ is marginal compared to the cost for evaluating $\partial_{\mathbf{u}} I_f(I_1, I_2^{\mathbf{u}})$ once Parzen estimates for $I_2^{\mathbf{u}}$ have been computed.

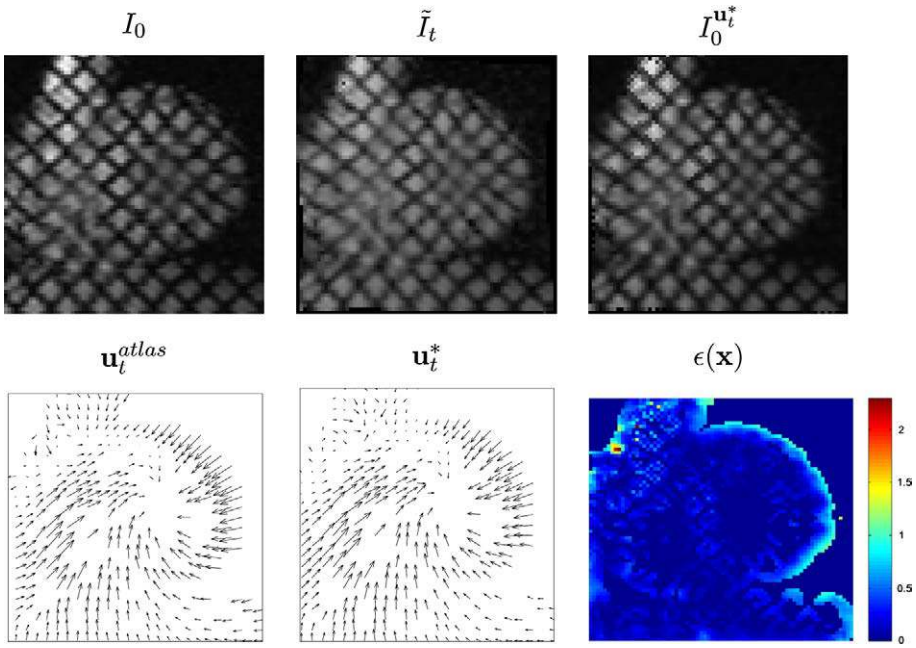

Fig. 6. Evaluation of the accuracy of I_α-information-based NRR ($\alpha = 1.2$) on a pair of consecutive ($t = 1$) simulated SPAMM images (SA view, mid level). The ground truth $\mathbf{u}_t^{\text{atlas}}$ and the recovered motion $\mathbf{u}_t^*$ are shown on the bottom line, together with the error map $\varepsilon(\mathbf{x}) := \|\mathbf{u}_t^{\text{atlas}}(\mathbf{x}) - \mathbf{u}_t^*(\mathbf{x})\|_{L^2}$ (in pixel units).

Table 5
Statistics of the local error $\varepsilon(\mathbf{x}) := \|\mathbf{u}_t^{\text{atlas}}(\mathbf{x}) - \mathbf{u}_t^*(\mathbf{x})\|_{L^2}$ between ground truth $\mathbf{u}_t^{\text{atlas}}$ ($t = 1$) and estimated displacement $\mathbf{u}_t^*$ using I_α-information-based NRR ($\alpha = 1.2$) for various tagging protocols

Tagging sequence	Mean $\pm$ SD	Range
SPAMM	0.12 ± 0.21	0–2.29
CSPAMM	0.16 ± 0.30	0–2.72
DANTE	0.14 ± 0.22	0–2.19

5.4. Pathological cases

The proposed approach has also been applied to pathological cases including acute ischemia and DCM.

For the ischemic case, observing directly the recovered displacement fields shows that motion is reduced to its tangential component in the anterior wall (Fig. 16). This is confirmed by radial contraction measurements, which greatly decrease in this area (Fig. 17). This hypokinetic zone has been diagnosed to be the infarcted region.

Studying the DCM case involves measurements at various levels. Comparing segmental measurements over the cardiac cycle to values for the healthy heart reveals that radial contraction is reduced in the septal and inferior segments, reaching only 1.8 mm, whereas torsion is enhanced and prolonged (Fig. 9). Both strains are lower than normal, on the whole myocardium for radial strain, and on the mid and apical infero-lateral walls for circumferential strain.

6. Discussion

The proposed variational framework for NRR using generalized information measures is generic and readily suited for multimodal image registration. It extends previous results obtained in the particular case of MI, namely the work of Hermosillo et al. (2001) who derived the expression of the MI flow for non-parametric transforms, and the early approach of Viola and Wells (1997) who obtained a similar result for linear parametric transforms using an ergodic approximation of MI. Compared to the latter, our approach yields an exact expression of the MI flow for linear parametric transforms, and highlights its relationship with the non-parametric case.

Within the Shannon framework, we have also derived new results relative to previously proposed information measures such as the ECC, the popular NMI and the exclusive information. Our contributions concern the derivation of their flows over arbitrary transform spaces, and the proof that they belong to a separate class of information measures. These results pave the way for new applicative developments such as assessing the performance of the poorly known ECC criterion or using NMI over non-parametric transform spaces.

Extension outside the Shannon framework provides a complete theoretical framework for optimizing already proposed non-Shannon information measures such as I_α (Pluim, 2000) or Renyi (He et al., 2003) information, and opens an avenue for investigating the suitability of novel information measures to specific application

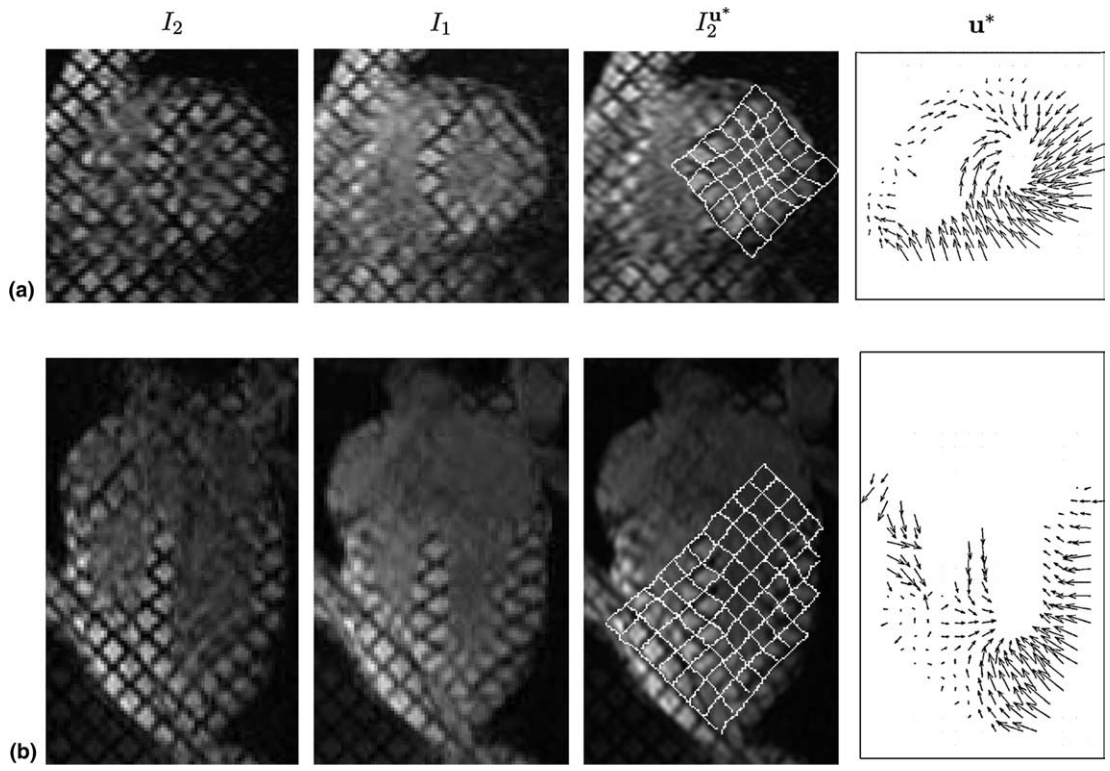

Fig. 7. NRR of (a) SA and (b) LA tagged MR data using exclusive f-information. The f_α information metrics ($\alpha = 1.2$) has been used. A synthetic grid warped by $\mathbf{u}^*$ has been overlaid to enhance tag motion recovery.

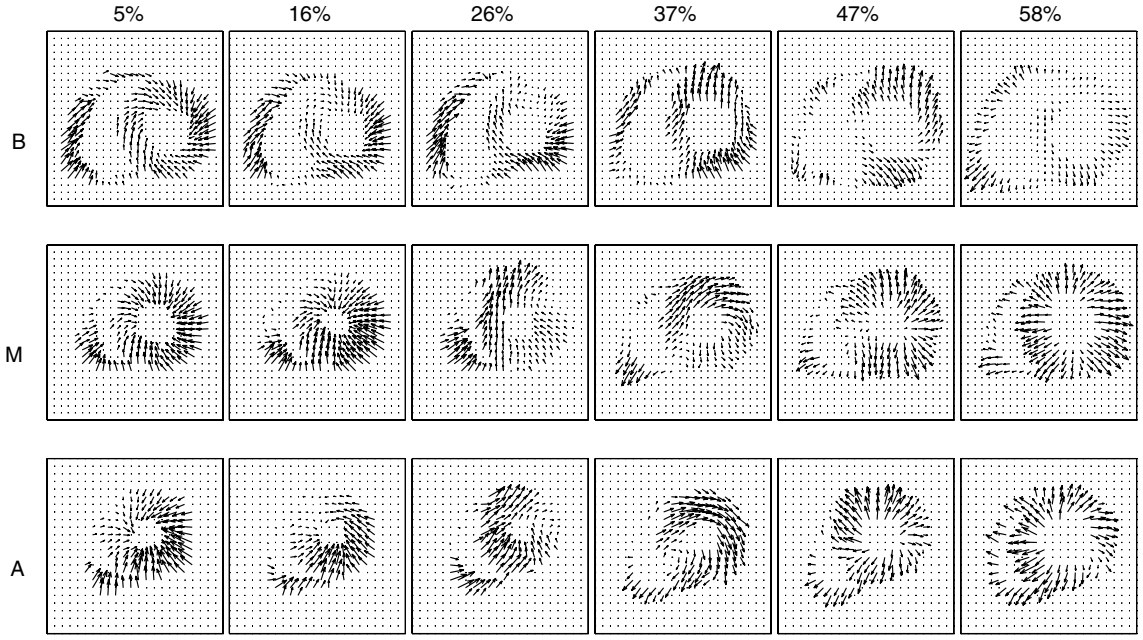

Fig. 8. Myocardial displacements estimated from SA data for a healthy heart at basal (B), mid (M) and apical (A) levels using exclusive f-information-based NRR. The f_α-information metrics ($\alpha = 1.2$) has been used. Time scale is expressed as % of the cardiac cycle.

contexts. Another important contribution concerns the generalization of normalized information measures to non-Shannon metrics $f \neq f_{\mathrm{KL}}$, and the derivation of their flows over arbitrary transform spaces. In particular, we have underlined the relevance of exclusive f-information which allows for faster convergence than f-information, and retains the robustness properties of generalized NMI while yielding a simpler flow.

Similarity criteria based on generalized information measures can obviously be combined with any

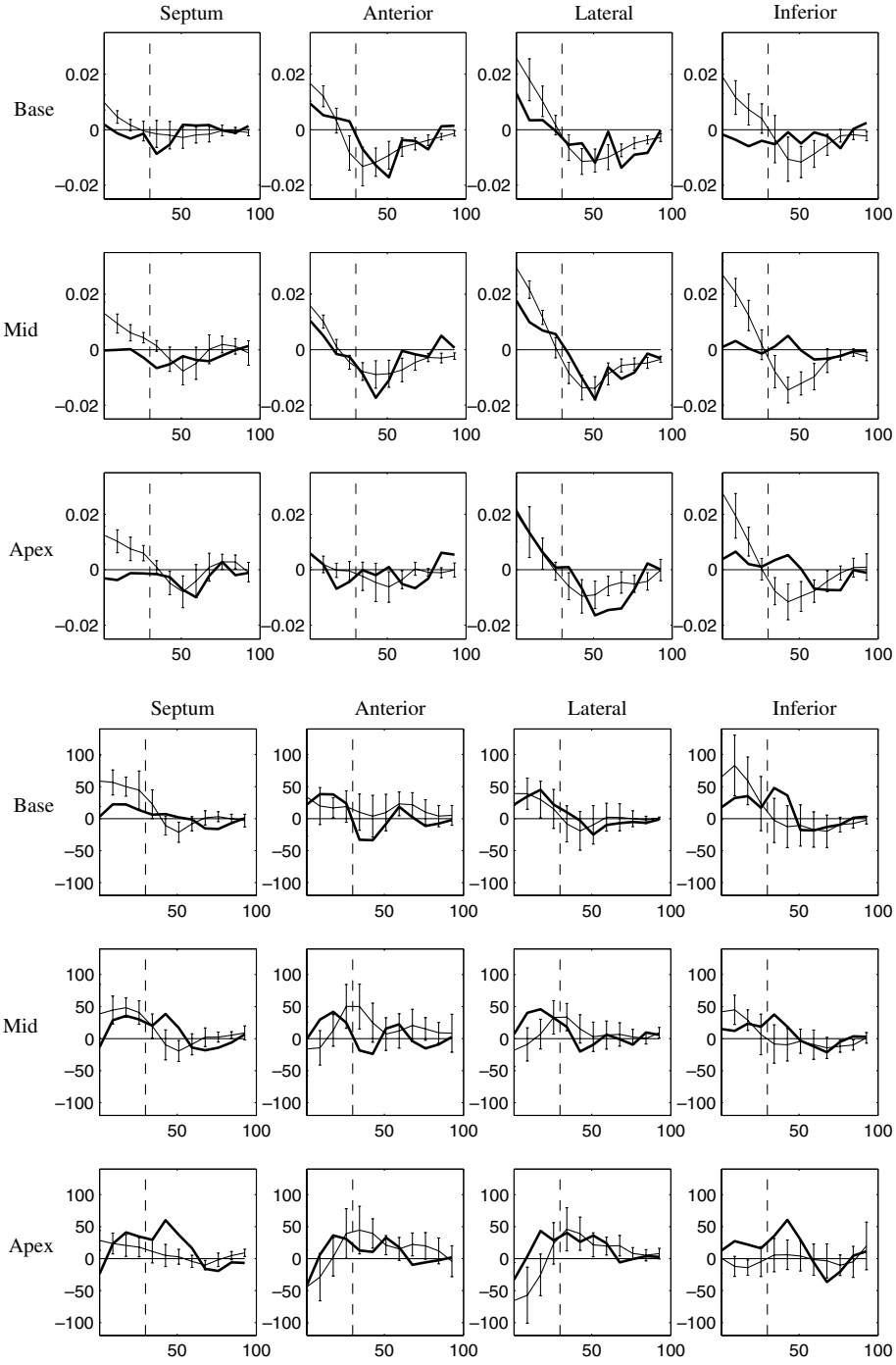

Fig. 9. (Top rows) Radial (in mm/s) and (bottom rows) rotation (in degrees/s) velocities of LV segments during a cardiac cycle for the mean of healthy data (plain line) and for a DCM case (bold line). Standard deviation for healthy data is shown as vertical bars. Time scale is expressed as % of the cardiac cycle. The vertical dashed line indicates end-systole.

regularization model. For instance, using viscous fluid regularization (Christensen et al., 1996b) yields a variational model suited to large displacements which generalizes the approach based on MI proposed by D'Agostino et al. (2003).

Concerning the mathematical properties of the model, the well-posedness of the MI flow has been established by Faugeras and Hermosillo (2001) for non-parametric transforms. A close look on the proof suggests that this result could be easily extended to generalized information measures thanks to the convexity property of integral information metrics and the duality relationship between integral and non-integral f-informations.

In some applications, the use of local statistics allows to deal more efficiently with non-stationary intensity

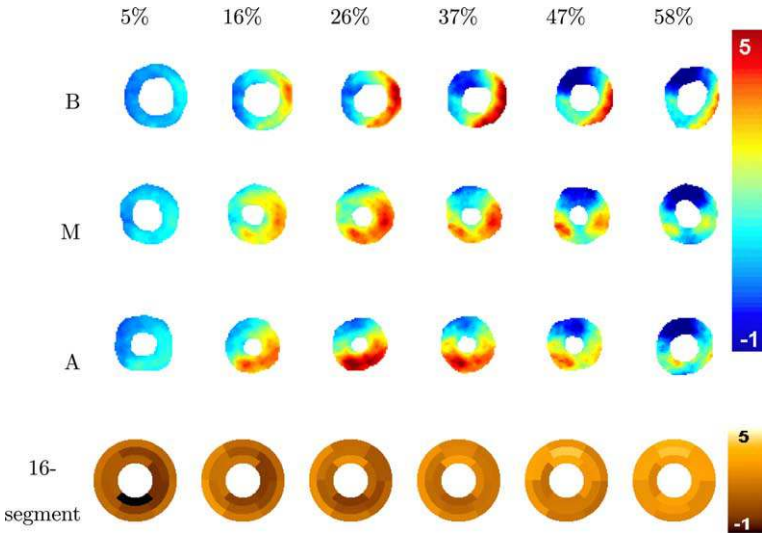

Fig. 10. Local and segmental radial contraction maps from SA data for a healthy heart at basal (B), mid (M) and apical (A) levels. Positive (resp. negative) values indicate inward (resp. outward) motion. Time scale is expressed as % of the cardiac cycle. Motion data, shown in Fig. 8, have been masked outside the LV.

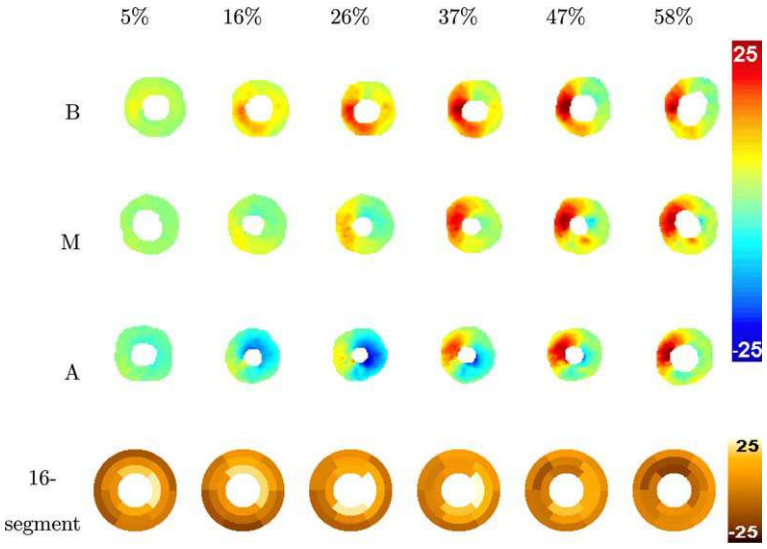

Fig. 11. Local and segmental torsion maps from SA data for a healthy heart at basal (B), mid (M) and apical (A) levels. Positive (resp. negative) values indicate clockwise (resp. counterclockwise) motion. Time scale is expressed as % of the cardiac cycle. Motion data, shown in Fig. 8, have been masked outside the LV.

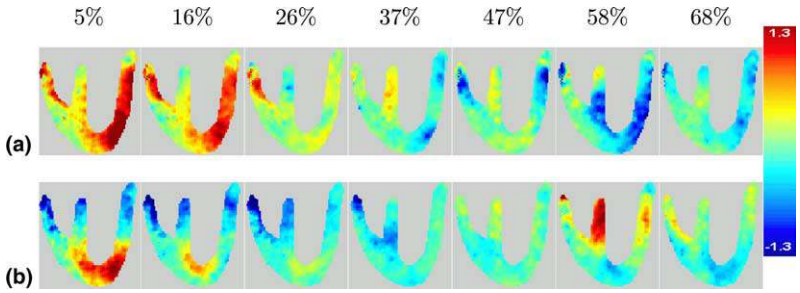

Fig. 12. (a) Radial and (b) longitudinal contraction maps from LA data for a healthy heart. Positive (resp. negative) values indicate (a) inward (resp. outward) motion, and (b) motion towards the apex (resp. base). Time scale is expressed as % of the cardiac cycle.

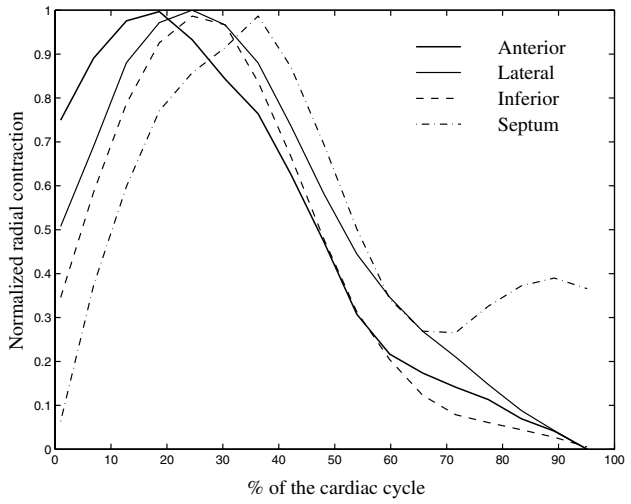

Fig. 13. Circumferential propagation of radial contraction: mean radial contraction averaged over all slice levels for the healthy volunteer group during a cardiac cycle. To underline propagation, measurements for each segment have been normalized to unity. Time scale is expressed as % of the cardiac cycle.

distributions. In particular, non-parametric local Parzen estimates have been proposed by Hermosillo (2002) and inserted into the variational optimization of MI for non-parametric transforms. This approach is easily extended to generalized information measures over arbitrary transform spaces. The interested reader is referred to (Petitjean et al., 2002).

Currently, the main methodological limitation relies in the lack of a theoretical basis for selecting a suitable information metrics. Consequently, one has to resort to systematic empirical testing.

On the applicative side, the proposed approach allows to recover reliable myocardial displacement estimates with pixel average accuracy over the whole cardiac cycle, at any slice level and under arbitrary incidence. Remarkably, measurements are available both on the LV and the RV. The derived deformation parameters are in accordance with the reference values reported in the medical literature (Petitjean et al., 2003b).

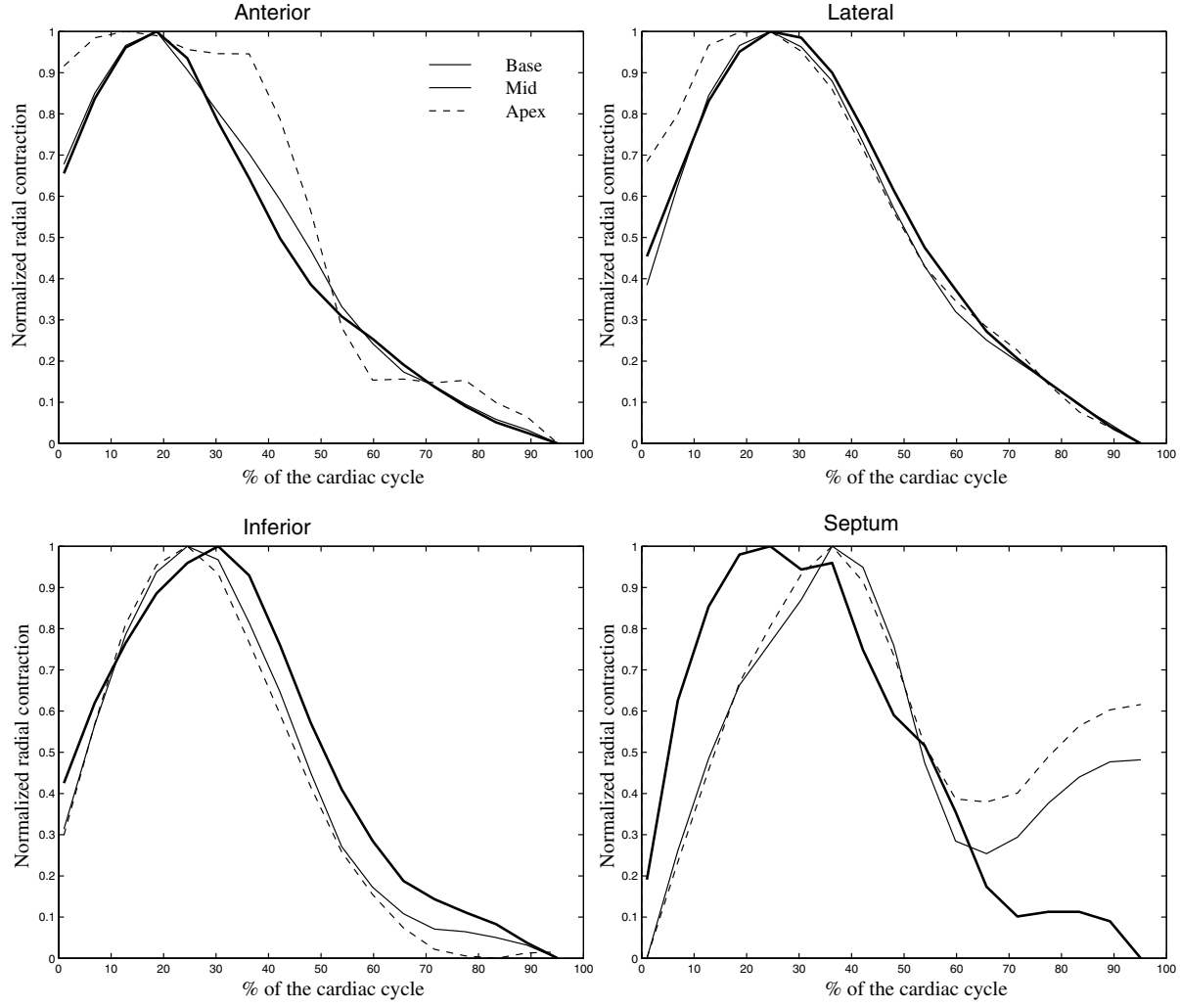

Fig. 14. Longitudinal propagation of radial contraction: mean radial contraction in myocardial segments for the healthy volunteer group during a cardiac cycle. To underline propagation, measurements for each level have been normalized to unity. Time scale is expressed as % of the cardiac cycle.

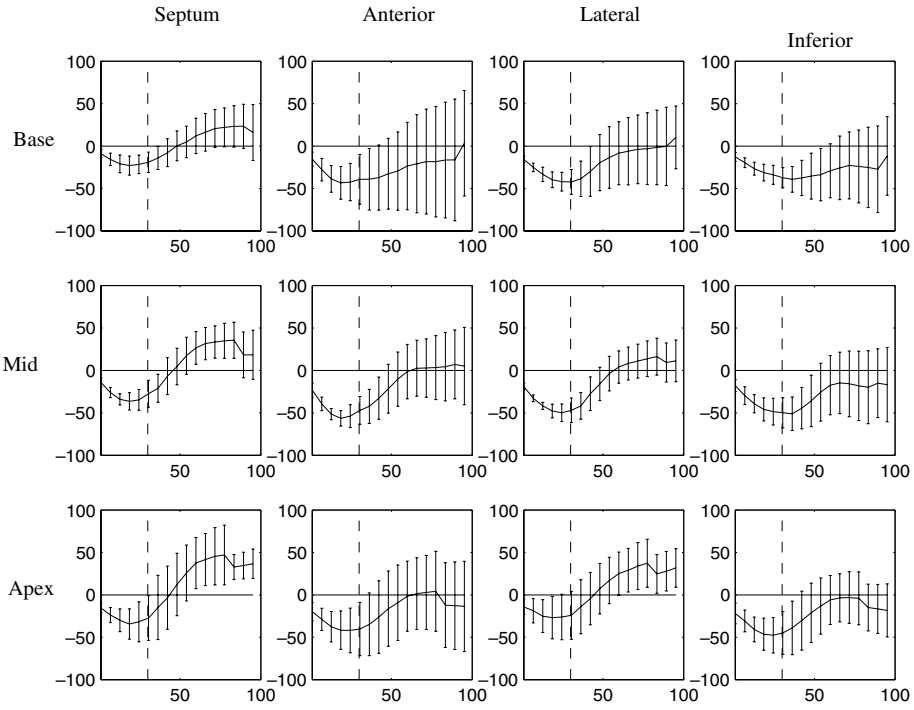

Fig. 15. Circumferential strain (in %) of LV segments at basal, mid and apical levels during cardiac cycle for the mean of healthy data. Time scale is expressed as % of the cardiac cycle.

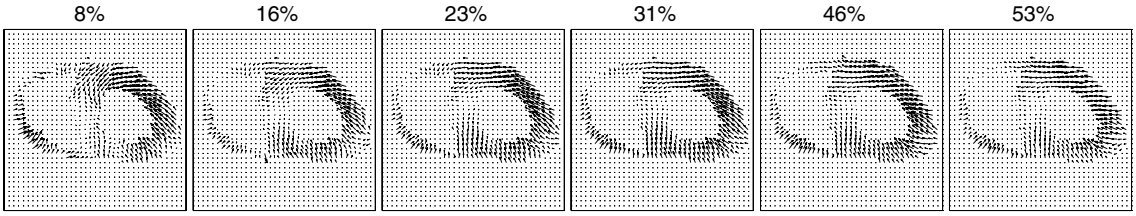

Fig. 16. Myocardial displacement estimated from SA data for an ischemic patient at mid level. Time scale is expressed as % of the cardiac cycle.

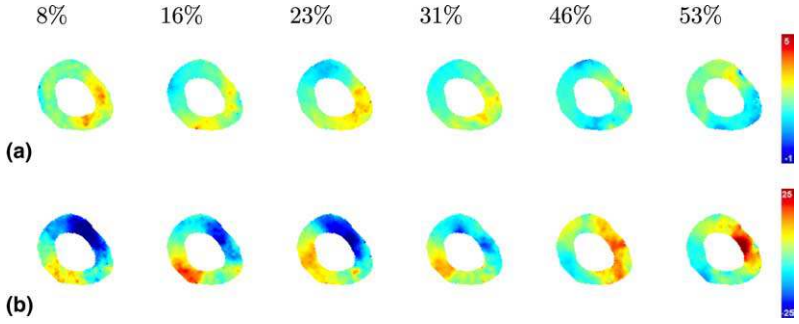

Fig. 17. Local (a) radial contraction and (b) rotation maps from SA data for an ischemic patient at mid level. Time scale is expressed as % of the cardiac cycle. The associated motion data (see Fig. 16) have been masked outside the LV.

The choice of a modeling context for displacement is conditioned by the goal of the motion recovery step. For quantification purpose solely, both non-parametric and parametric motion spaces are admissible, yielding optimization schemes with similar complexities.[11]

Parametric models offer the advantage of being more compact and of allowing an analytical computation of the deformation parameters, with an expected impact on measurement accuracy. As far as we are concerned, myocardial motion modeling, in order to built realistic, very compact, parametric displacements models adapted to the cardiac kinematics, is the ultimate goal (Petitjean et al., 2004). Hence, we have

[11] Except if a stochastic optimization strategy is used in the parametric case as in (Viola and Wells, 1997).

retained a non-parametric displacement model in order to preserve the statistical richness of the motion estimates, instead of filtering them by a parametric representation. Nonetheless, assessing quantification performance using a B-spline representation would be desirable for comparison purpose with the approach of (Chandrashekara et al., 2002).

7. Conclusion

We have presented a novel and efficient method for quantifying myocardial deformations from tagged and cine MR data. This method relies on a novel and generic variational NRR framework using generalized information measures. It is fully unsupervised, thus reproducible, and versatile with respect to the MR tagging technology. It delivers, with pixel average accuracy, measurements over the whole myocardium and the whole cardiac cycle, at any slice level and under arbitrary incidence. Its computational performances are attractive: systolic measurements at 3 levels are obtained in 60 s, and a complete dynamic quantification at 3 SA and 3 LA levels in about 30 min on a PIII platform. Significant improvements are expected from the use of faster gradient descent techniques, code optmization and advent of more powerful processors. For healthy data, measurements have been shown to provide dynamic quantitative insights on the kinematics of heart contraction. For pathological data, preliminary experiments for acute ischemia and DCM have underlined that specific motion parameters deviate from the healthy case. Current developments include: (i) extensive study of the dilated/hypertrophic cardiomyopathies; (ii) integration of SA and LA tagged MR data to derive 3D motion estimates, (iii) joint study of contraction/perfusion and contraction/activation via multimodal NRR, which the current technique is readily suited for.

Appendix A. Computation of the gradient flow for integral *f*-informations

In what follows, we derive the expression of the first variation of an integral f-information functional $\mathscr{S}(\mathbf{u})$ over a non-parametric transform space $\mathscr{T}$ of sufficiently regular mappings over D.

Let $\mathbf{u}+\varepsilon\mathbf{h}$ be a variation of a displacement field $\mathbf{u}\in\mathscr{T}$, for some $\varepsilon>0$ and some mapping $\mathbf{h}\in\mathscr{T}$. From Eq. (9), the Gâteaux derivative of the perturbed functional $\mathscr{S}(\mathbf{u}+\varepsilon\mathbf{h})$ is computed as follows:

$$
\begin{aligned}
-\frac{\partial\mathscr{S}(\mathbf{u}+\varepsilon\mathbf{h})}{\partial\varepsilon} &= \frac{\partial}{\partial\varepsilon}\int_{\Omega^2} p^{I_1}(i_1)p^{I_2^{\mathbf{u}+\varepsilon\mathbf{h}}}(i_2) f\big(\rho^{\mathbf{u}+\varepsilon\mathbf{h}}(i_1,i_2)\big)\,\mathrm{d}i_1\,\mathrm{d}i_2 \\
&= \int_{\Omega^2}\left[\frac{\partial p^{I_2^{\mathbf{u}+\varepsilon\mathbf{h}}}(i_2)}{\partial\varepsilon}p^{I_1}(i_1)f\big(\rho^{\mathbf{u}+\varepsilon\mathbf{h}}(i_1,i_2)\big)+f'\big(\rho^{\mathbf{u}+\varepsilon\mathbf{h}}(i_1,i_2)\big)\right. \\
&\quad\left.\times\ \frac{\partial p^{I_1,I_2^{\mathbf{u}+\varepsilon\mathbf{h}}}(i_1,i_2)}{\partial\varepsilon}-\frac{p^{I_1,I_2^{\mathbf{u}+\varepsilon\mathbf{h}}}(i_1,i_2)}{p^{I_2^{\mathbf{u}+\varepsilon\mathbf{h}}}(i_2)}\frac{\partial p^{I_2^{\mathbf{u}+\varepsilon\mathbf{h}}}(i_2)}{\partial\varepsilon}\right)\Bigg]\mathrm{d}i_1\,\mathrm{d}i_2 \\
&= \underbrace{\int_{\Omega^2} L_1^{\mathbf{u}+\varepsilon\mathbf{h}}(i_1,i_2)\frac{\partial p^{I_1,I_2^{\mathbf{u}+\varepsilon\mathbf{h}}}(i_1,i_2)}{\partial\varepsilon}\mathrm{d}i_1\,\mathrm{d}i_2}_{:=A_1} \\
&\quad+\underbrace{\int_{\Omega^2} L_2^{\mathbf{u}+\varepsilon\mathbf{h}}(i_1,i_2)p^{I_1}(i_1)\frac{\partial p^{I_2^{\mathbf{u}+\varepsilon\mathbf{h}}}(i_2)}{\partial\varepsilon}\mathrm{d}i_1\,\mathrm{d}i_2}_{:=A_2},
\end{aligned}
$$

where $L_1^{\mathbf{u}+\varepsilon\mathbf{h}}$ and $L_2^{\mathbf{u}+\varepsilon\mathbf{h}}$ are defined in Table 1.

Approximating $p^{I_1,I_2^{\mathbf{u}+\varepsilon\mathbf{h}}}$ globally over $\underline{D}$ via a 2-D Parzen estimator with density kernel $\mathscr{K}_2$ (cf. Eq. (10)), we get

$$
\begin{aligned}
\frac{\partial p^{I_1,I_2^{\mathbf{u}+\varepsilon\mathbf{h}}}}{\partial\varepsilon}(i_1,i_2) &= \frac{1}{|\underline{D}|}\int_{\underline{D}}\frac{\partial\mathscr{K}_2}{\partial y}\big(I_1(\mathbf{x})-i_1,I_2^{\mathbf{u}+\varepsilon\mathbf{h}}(\mathbf{x})-i_2\big) \\
&\quad\times\nabla I_2^{\mathbf{u}+\varepsilon\mathbf{h}}(\mathbf{x})\mathbf{h}(\mathbf{x})\,\mathrm{d}\mathbf{x}.
\end{aligned}
$$

Letting $\varepsilon\to 0$, the term A_1 can then be rewritten as a convolution over the intensity space Ω^2

$$
\begin{aligned}
A_1 &= \frac{1}{|\underline{D}|}\int_{\underline{D}}\left[\frac{\partial\mathscr{K}_2}{\partial y}\star L_1^{\mathbf{u}}\right]\big(I_1(\mathbf{x}),I_2^{\mathbf{u}}(\mathbf{x})\big)\nabla I_2^{\mathbf{u}}(\mathbf{x})\mathbf{h}(\mathbf{x})\,\mathrm{d}\mathbf{x} \\
&= \frac{1}{|\underline{D}|}\int_{\underline{D}}\left[\mathscr{K}_2\star\frac{\partial L_1^{\mathbf{u}}}{\partial i_2}\right]\big(I_1(\mathbf{x}),I_2^{\mathbf{u}}(\mathbf{x})\big)\nabla I_2^{\mathbf{u}}(\mathbf{x})\mathbf{h}(\mathbf{x})\,\mathrm{d}\mathbf{x},
\end{aligned}
\tag{A.1}
$$

where $\frac{\partial L_1^{\mathbf{u}}}{\partial i_2}$ is given by (15).

Similarly, approximating $p^{I_2^{\mathbf{u}+\varepsilon\mathbf{h}}}$ globally over $\underline{D}$ via a 1-D Parzen estimator with density kernel $\mathscr{K}_1$ (cf. Eq. (11)), we obtain

$$
\frac{\partial p^{I_2^{\mathbf{u}+\varepsilon\mathbf{h}}}}{\partial\varepsilon}(i_2) = \frac{1}{|\underline{D}|}\int_{\underline{D}}\mathscr{K}'_1(I_2^{\mathbf{u}+\varepsilon\mathbf{h}}(\mathbf{x})-i_2)\nabla I_2^{\mathbf{u}+\varepsilon\mathbf{h}}(\mathbf{x})\mathbf{h}(\mathbf{x})\,\mathrm{d}\mathbf{x}.
$$

Letting $\varepsilon\to 0$, the term A_2 can thus be written as

$$
\begin{aligned}
A_2 &= \frac{1}{|\underline{D}|}\int_{\underline{D}}\left(\int_{\Omega^2}p^{I_1}(i_1)\mathscr{K}'_1\big(I_2^{\mathbf{u}}(\mathbf{x})-i_2\big)L_2^{\mathbf{u}}(i_1,i_2)\,\mathrm{d}i_1\,\mathrm{d}i_2\right) \\
&\quad\times\nabla I_2^{\mathbf{u}}(\mathbf{x})\mathbf{h}(\mathbf{x})\,\mathrm{d}\mathbf{x} \\
&= \frac{1}{|\underline{D}|}\int_{\underline{D}}\left(\int_{\Omega}\left[\mathscr{K}'_1\star L_2^{\mathbf{u}|i_1}\right](I_2^{\mathbf{u}}(\mathbf{x}))p^{I_1}(i_1)\,\mathrm{d}i_1\right) \\
&\quad\times\nabla I_2^{\mathbf{u}}(\mathbf{x})\mathbf{h}(\mathbf{x})\,\mathrm{d}\mathbf{x} \\
&= \frac{1}{|\underline{D}|}\int_{\underline{D}}\left(\int_{\Omega}\left[\mathscr{K}_1\star\frac{\partial L_2^{\mathbf{u}|i_1}}{\partial i_2}\right](I_2^{\mathbf{u}}(\mathbf{x}))p^{I_1}(i_1)\,\mathrm{d}i_1\right) \\
&\quad\times\nabla I_2^{\mathbf{u}}(\mathbf{x})\mathbf{h}(\mathbf{x})\,\mathrm{d}\mathbf{x} \\
&= \frac{1}{|\underline{D}|}\int_{\underline{D}}E_{I_1}\left(\left[\mathscr{K}_1\star\frac{\partial L_2^{\mathbf{u}|i_1}}{\partial i_2}\right](I_2^{\mathbf{u}}(\mathbf{x}))\right)\nabla I_2^{\mathbf{u}}(\mathbf{x})\mathbf{h}(\mathbf{x})\,\mathrm{d}\mathbf{x},
\end{aligned}
\tag{A.2}
$$

where $E_{I_1}(\cdot)$ denotes expectation w.r.t. the random variable I_1, and $L_2^{\mathbf{u}|i_1}(\cdot) := L_2^{\mathbf{u}}(i_1,\cdot)$ for any given intensity value $i_1 \in \Omega$. The symbol $\bigstar$ is reused for denoting 1-D convolution over Ω. One easily shows that $\frac{\partial L_2^{\mathbf{u}|i_1}}{\partial i_2} = -\rho^{\mathbf{u}} f''(\rho^{\mathbf{u}}) \frac{\partial \rho^{\mathbf{u}}}{\partial i_2}$, so that identity (14) holds.

Combining (A.1) and (A.2) yields finally

$$\left.\frac{\partial \mathscr{S}(\mathbf{u}+\varepsilon\mathbf{h})}{\partial \varepsilon}\right|_{\varepsilon=0} = -\int_{\underline{D}} V_1^{I_1,I_2^{\mathbf{u}}}(\mathbf{x}) \nabla I_2^{\mathbf{u}}(\mathbf{x}) \mathbf{h}(\mathbf{x})\, \mathrm{d}\mathbf{x} = \langle \partial_{\mathbf{u}} \mathscr{S}(\mathbf{u}), \mathbf{h} \rangle_{\mathscr{L}^2},$$

where $V_1^{I_1,I_2^{\mathbf{u}}}$ is given by (13). This yields the result (12).

References

Amini, A.A., Chen, Y., Curwen, R., Mani, V., Sun, J., 1998. Coupled B-snake grids and constrained thin-plate splines for analysis of 2-D tissue deformations from tagged MRI. IEEE Transactions on Medical Imaging 17 (3), 344–356.

Amini, A.A., Chen, Y., Elayyadi, M., Radeva, P., 2001. Tag surface reconstruction and tracking of myocardial beads from SPAMM-MRI with parametric B-spline surfaces. IEEE Transactions on Medical Imaging 20 (2), 94–103.

Amini, A.A., Prince, J.L., 2001. Measurement of Cardiac Deformations from MRI: Physical and Mathematical Models. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Axel, L., Dougherty, L., 1989. MR imaging of motion with spatial modulation of magnetization. Radiology 171 (3), 841–845.

Bajcsy, R.K., Kovačič, S., 1989. Multiresolution elastic matching. Computer Vision, Graphics and Image Processing 46, 1–21.

Bookstein, F., 1989. Principal warps: thin-plate splines and the decomposition of deformations. IEEE Transactions on Pattern Analysis and Machine Intelligence 11 (6), 567–585.

Cerqueira, M.D., Weissman, N.J., Dilsizian, C., Jacobs, A.K., Kaul, S., Laskey, W.K., et al., 2002. Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart. Circulation 105, 539–542.

Christensen, G.E., Miller, M.I., Vannier, M.W., 1996. Individualizing neuro-anatomical atlases using a massively parallel computer. IEEE Computer 29 (1), 32–38.

Christensen, G.E., Rabbit, R., Miller, M.I., 1996b. Deformable templates using large deformation kinematics. IEEE Transactions on Image Processing 5 (10), 1435–1447.

Collignon, A., Vandermeulen, D., Suetens, P., Marchal, G., 1995. 3D multi-modality medical image registration using feature space clustering. In: Proceedings Computer Vision, Virtual Reality and Robotics in Medicine (CVRMed'95) (Nice, France) – LNCS 905. Springer-Verlag, Heidelberg, Germany, pp. 195–204.

Chandrashekara, R., Mohiaddin, R.H., Rueckert, D., 2002. Analysis of myocardial motion in tagged MR images using nonrigid image registration. In: Proceedings SPIE Conference on Medical Imaging 2002: Image Processing, San Diego, CA, vol. 4684, pp. 1168–1179.

Clarysse, P., Basset, C., Khouas, L., Croisille, P., Friboulet, D., Odet, C., Magnin, I., 2000. Two-dimensional spatial and temporal displacement and deformation field fitting from cardiac magnetic resonance imaging. Medical Image Analysis 4, 253–268.

D'Agostino, E., Maes, F., Vandermeulen, D., Suetens, P., 2003. A viscous fluid model for multimodal non-rigid image registration using mutual information. Medical Image Analysis 7, 565–575.

Declerck, J., 1997. Etude de la dynamique cardiaque par analyse d'images tridimensionnelle. Ph.D. Dissertation Université de Nice-Sophia-Antipolis, France.

Denney, T.S., 1994. Stochastic estimation of deformable motion from magnetic resonance tagged cardiac images. Ph.D. Dissertation, John Hopkins University, Baltimore, MD.

Denney, T.S., Prince, J.L., 1994. Optimal brightness functions for optical flow estimation of deformable motion. IEEE Transactions on Medical Imaging 3 (2), 178–191.

Denney, T.S., Prince, J.L., 1995. Reconstruction of 3D left ventricular motion from planar tagged cardiac MR images: an estimation theoretic approach. IEEE Transactions on Medical Imaging 14 (4), 625–635.

Denney, T.S., 1999. Estimation and detection of myocardial tags in MR images without prior knowledge of myocardial contours. IEEE Transactions on Medical Imaging 18 (4), 330–344.

Deriche, R., 1993. Recursively implementing the Gaussian and its derivatives. Research Report 1893, INRIA, Sophia-Antipolis, France.

Duncan, J., Ayache, N., 2000. Medical image analysis: progress over two decades and challenges ahead. IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (1), 85–106.

Dougherty, L., Asmuth, J.C., Blom, A.S, Axel, L., Kumar, R., 1999. Validation of an optical flow method for tag displacement estimation. IEEE Transactions on Medical Imaging 18 (4), 359–363.

Faugeras, O., Hermosillo, G., 2001. Well-posedness of eight problems of multi-modal statistical image matching. Research Report 4235 INRIA, Sophia-Antipolis, France.

Fischer, S., McKinnon, G., Maier, S., Boesiger, P., 1993. Improved myocardial tagging contrast. Magnetic Resonance in Medicine 30, 191–200.

Frangi, A., Niessen, W.J., Viergever, M.A., 2001. Three-dimensional modeling for functional analysis of cardiac images: a review. IEEE Transactions on Medical Imaging 20 (1), 2–25.

Frangi, A., Rueckert, D., Duncan, J., 2002. Three-dimensional cardiovascular image analysis. IEEE Transactions on Medical Imaging 21 (9), 1005–1010.

Glass, L., Hunter, P., McCulloch, A.D., 1991. Theory of Heart: Biomechanics Biophysics and Nonlinear Dynamics of Cardiac Function. Springer-Verlag, New York, NY.

Gupta, S.N., Prince, J.L., 1995. On variable brightness optical flow for tagged MRI. In: Proceedings Information Processing in Medical Imaging (IPMI'95), pp. 323–334.

Gupta, S.N., Prince, J.L., Androutsellis-Theotokis, S., 1997. Bandpass optical flow for tagged MRI. In: Proceedings IEEE ICIP'97, San Diego, CA, vol. 3, pp. 364–367.

Guttman, M., Zerhouni, E.A., McVeigh, E.R., 1997. Analysis and visualization of cardiac function from MR images. IEEE Computer Graphics and Applications 17 (1), 55–63.

Guttman, M., Prince, J.L., McVeigh, E.R., 1994. Tag and contour detection in tagged MR images of the left ventricle. IEEE Transactions on Medical Imaging 13 (1), 74–88.

Haber, I., Metaxas, D.N., Axel, L., 2000. Three-dimensional motion reconstruction and analysis of the right ventricle using tagged MRI. Medical Image Analysis 4 (4), 335–355.

Hajnal, J.V., Hill, D.L.G., Hawkes, D.J (Eds.), 2001. Medical Image Registration. CRC Press, Boca Raton, FL.

He, Y., Hamza, A.B., Krim, H., 2003. A generalized divergence measure for robust image registration. IEEE Transactions on Signal Processing 51 (5), 1211–1220.

Hellier, P., Barillot, C., Memin, E., Perez, P., 2001. Hierarchical estimation of a dense deformation field for 3-D robust registration. IEEE Transactions on Medical Imaging 20 (5), 388–402.

Hermosillo, G., Chefd'Hotel, C., Faugeras, O., 2001. A variational approach to multi-modal image matching. Research Report 4117 INRIA, Sophia-Antipolis, France.

Hermosillo, G., 2002. Variational methods for multimodal image matching. Ph.D. Dissertation, Université de Nice-Sophia-Antipolis, France.

Hu, Z., Metaxas, D., Axel, L., 2003. In vivo strain and stress estimation of the heart left and right ventricles from MRI images. Medical Image Analysis 7 (4), 435–444.

Huang, J., Abendschein, D., Dávila-Román, V., Amini, A.A., 1999. Spatio-temporal tracking of myocardial deformations with a 4-D B-spline model from tagged MRI. IEEE Transactions on Medical Imaging 18 (10), 957–972.

Kerwin, W., Prince, J.L., 1999. Tracking MR tag surfaces using a spatiotemporal filter and interpolator. International Journal of Image System and Technology 10 (2), 128–142.

Kerwin, W., Osman, N.F., Prince, J.L., 2001. Image processing and analysis in tagged cardiac MRI. In: Handbook of Medical Imaging. Academic Press, San Diego, CA, pp. 375–391 (Chapter 24).

Kraitchman, D.L., Young, A., Chang, C.N., Axel, L., 1995. Semi-automatic tracking of myocardial motion in MR tagged images. IEEE Transactions on Medical Imaging 14 (3), 422–433.

Kumar, S., Goldgof, D., 1994. Automatic tracking of SPAMM grid and the estimation of deformation parameters from cardiac MR images. IEEE Transactions on Medical Imaging 13 (1), 122–132.

Lester, H., Arridge, S.R., 1999. A survey of hierarchical non-linear medical image registration. Pattern Recognition 32 (1), 129–149.

Leventon, M.E., Crimson, W.E.L., 1998. Multi-modal volume registration using joint intensity distributions. In: Proceedings of the Medical Image Computing and Computer-assisted Intervention (MICCAI'98) (Cambridge, MA) – LNCS 1496. Springer-Verlag, Heidelberg, Germany, pp. 1057–1066.

Lipton, M.J., Bogaert, J., Boxt, L.M., Reba, R.C., 2002. Imaging of ischemic heart disease. European Radiology 12, 1061–1080.

McVeigh, E.R., Öztürk, C., 2001. Imaging myocardial strain. IEEE Signal Processing Magazine 18 (6), 44–56.

Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P., 1997. Multimodality image registration by maximization of mutual information. IEEE Transactions on Medical Imaging 16 (2), 187–198.

Maintz, J.B.A., Viergever, M.A., 1998. A survey of medical image registration. Medical Image Analysis 2 (1), 1–36.

Mäkelä, T., Clarysse, P., Sipilä, O., Pauna, N., Pham, Q.C., Katila, T., Magnin, I.E., 2002. A review of cardiac image registration methods. IEEE Transactions on Medical Imaging 21 (9), 1011–1021.

Moore, C.C., Lugo-Olivieri, C.H., McVeigh, E.R., Zerhouni, E.A., 2000. Three-dimensional systolic strain patterns in the normal left ventricle: characterization with tagged MR imaging. Radiology 214 (2), 453–466.

Mosher, T., Smith, M., 1990. A DANTE tagging sequence for the evaluation of translational sample motion. Magnetic Resonance in Medicine 15, 334–339.

Nagel, H.H., Enkelmann, W., 1986. An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. IEEE Transactions on Pattern Analysis and Machine Intelligence 8 (5), 565–593.

Nielsen, P.M., LeGrice, I.J., Smaill, B.H., Hunter, P.J., 1991. Mathematical model of geometry and fibrous structure of the heart. American Journal of Physiology 260 (4), H1365–H1378.

Osman, N.F., Kerwin, W., McVeigh, E.R., Prince, J.L., 1999. Cardiac motion tracking using CINE harmonic phase (HARP) magnetic resonance imaging. Magnetic Resonance in Medicine 42 (6), 1048–1060.

Osman, N.F., Prince, J.L., 2000. Imaging heart motion using harmonic phase MRI. IEEE Transactions on Medical Imaging 19 (3), 186–202.

Öztürk, C., McVeigh, E.R., 2000. Four dimensional B-spline based motion analysis of tagged cardiac MR images. Physics in Medicine and Biology 45 (6), 1683–1702.

Park, J., Metaxas, D., Young, A.A., Axel, L., 1996. Analysis of left ventricular wall motion based on volumetric deformables models and MRI-SPAMM. Medical Image Analysis 1 (1), 53–71.

Perperidis, D., Rao, A., Lorenzo-Valdés, M., Mohiaddin, R., Rueckert, D., 2003. Spatio-temporal alignment of 4D cardiac MR images. In: Proceedings of the Second International Workshop on Functional Imaging and Modeling of the Heart (FIMH'2003) (Lyon, France) – LNCS 2674. Springer-Verlag, Germany, pp. 203–214.

Petitjean, C., Rougon, N., Prêteux, F., 2002. Non-rigid image registration using generalized information measures. Research Report 02001 GET/INT, Evry, France.

Petitjean, C., Rougon, N., Prêteux, F., Cluzel, P., Grenier, P., 2003a. Measuring myocardial deformations from MR data using information-theoretic non-rigid registration. In: Proceedings of the Second International Workshop on Functional Imaging and Modeling of the Heart (FIMH'2003) (Lyon, France) – LNCS 2674. Springer-Verlag, Germany, pp. 162–172.

Petitjean, C., Rougon, N., Cluzel, P., 2003b. Assessment of myocardial function: a review of quantification methods and results using tagged MRI. Research Report 03009 GET/INT, Evry, France.

Petitjean, C., 2003c. Recalage non rigide d'images par approches variationnelles statistiques. Application à l'analyse et à la modélisation de la fonction myocardique en IRM. Ph.D. Dissertation, Université Paris V – René Descartes, France.

Petitjean, C., Rougon, N., Prêteux, F., Cluzel, P., Grenier, P., 2004. Building and using a statistical 3D motion atlas for analyzing myocardial contraction in MRI. In: Proceedings SPIE Conference on Medical Imaging 2004: Image Processing, San Diego, CA, vol. 5370, pp. 253–264.

Pluim, J.P.W., 2000. Mutual information based registration of medical images. Ph.D. Dissertation, University of Utrecht, The Netherlands.

Pluim, J.P.W., Maintz, J.B.A., Viergever, M.A., 2000. Image registration by maximization of combined mutual information and gradient information. IEEE Transactions on Medical Imaging 19 (8), 809–814.

Pluim, J.P.W., Maintz, J.B.A., Viergever, M.A., 2003. Mutual-information-based registration of medical images: a survey. IEEE Transactions on Medical Imaging 22 (7), 986–1004.

Prince, J.L., McVeigh, E.R., 1992. Motion estimation from tagged MR images sequences. IEEE Transactions on Medical Imaging 11 (2), 238–249.

Rao, A., Sanchez-Ortiz, G.I., Chandrashekara, R., Lorenzo-Valdés, M., Mohiaddin, R., Rueckert, D., 2003. Construction of a cardiac motion atlas from MR using non-rigid registration. In: Proceedings of the Second International Workshop on Functional Imaging and Modeling of the Heart (FIMH'2003) (Lyon, France) – LNCS 2674. Springer-Verlag, Germany, pp. 141–150.

Roche, A., Malandain, G., Pennec, X., Ayache, N., 1998. Multimodal image registration by maximization of the correlation ratio. Research Report 3378 INRIA, Sophia-Antipolis, France.

Roche, A., Malandain, G., Ayache, N., Prima, S., 1999. Towards a better comprehension of similarity measures used in medical image registration. In: Proceedings of the Medical Image Computing and Computer-assisted Intervention (MICCAI'99) (Cambridge, UK) – LNCS 1679. Springer-Verlag, Heidelberg, Germany, pp. 555–566.

Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J., 1999. Nonrigid registration using free-form deformations: application to breast MR images. IEEE Transactions on Medical Imaging 18 (8), 712–721.

Studholme, C., Hill, D.L., Hawkes, D., 1999. An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognition 32 (1), 71–86.

Suri, J.S., 2000. Computer vision, pattern recognition and image processing in left ventricle segmentation: the last 50 years. Pattern Analysis and Applications 3 (3), 209–242.

Vajda, I., 1989. Theory of Statistical Inference and Information. Kluwer Academic Publisher, Dordrecht, The Netherlands.

Viola, P., Wells III, W.M., 1997. Alignment by maximization of mutual information. International Journal of Computer Vision 24 (2), 137–154.

Young, A.A., Kraitchman, D.L., Dougherty, L., Axel, L., 1995. Tracking and finite element analysis of stripe deformation in magnetic resonance tagging. IEEE Transactions on Medical Imaging 14 (3), 413–421.

Zerhouni, E., Parish, D., Rogers, W., Yang, A., Shapiro, E., 1988. Human heart: tagging with MR imaging-A method for non invasive assessment of myocardial motion. Radiology 169 (1), 59–63.