Locally adapted spatio-temporal deformation model for dense motion estimation in periodic cardiac image sequences

Bertrand Delhay[1], Patrick Clarysse[1], Isabelle E. Magnin[1]

Creatis, CNRS UMR #5220, INSERM U630, INSA Bât. Blaise Pascal
F-69621 Villeurbanne Cedex, France,
delhay@creatis.insa-lyon.fr

Abstract. We recently introduced a continuous state space parametric model of spatio-temporal transformations and an algorithm, based on Kalman filtering, to represent motion in an image sequence describing a periodic phenomena. One advantage of this method is to simultaneously take into account all the sequence frames to robustly estimate the parameters of a unique spatial and periodic-temporal model. However, in 3D+time, a large number of parameters is required. In this paper, we propose a criterion based on motion energy to locally adapt the trajectory model and thus the temporal complexity of the model. The influence of the model order is illustrated on true 2D+time Magnetic Resonance Images (MRI) of the heart in order to motivate the proposed adaptative criteria. Quantitative results of the proposed adapted spatio-temporal motion model are given on synthetic 2D+time MRI sequences. Preliminary experiments show a significant impact notably regarding the parameter saving while preserving the accuracy of the motion estimates.

1 Introduction

Cardiac motion estimation and modeling is particularly helpful for myocardial function analysis. The last two decades are marked by an important progress of image acquisition devices making possible to better explore the dynamics of moving organs. Magnetic Resonance Imaging (MRI) and Multiple Detector Computed Tomography (MDCT) provide meaningful information about the 3D anatomy and contractile function of the heart. 3D+time segmentation and motion estimation from cardiac images are recognized as a difficult pre-requisite tasks for quantitative analysis of cardiac function. Most of the published works does not explicitly take into account the time dimension and proceeds iteratively from one time point to the next. This results in inconsistent material point trajectories. However, it is clear, especially for cardiac motion recovery purpose, that introducing some realistic temporal constraints will greatly improve the estimated motion pattern [1]. In cardiac motion analysis, some recent works are extension of 2D/3D image registration methods to 2D+time/3D+time [2, 3]. Time axis is here considered as a supplementary axis which is not qualitatively

different from the other spatial axes. In a similar way, in [4], the tracking of the left ventricle is performed with a 4D B-Spline model whose knots fit extracted features from Tagged MRI. In a previous paper [5], we introduced a framework based on state-space formulation and Kalman filtering by imposing a temporal consistency of the estimated motion over the whole image sequence and the whole image space. The periodicity and continuity of the motion is therefore insured but one of the main drawbacks of the method might be the number of involved parameters for motion description. This particular point makes difficult to process a full 3D+time sequence with standard computers. In this paper, we propose to locally adapt the complexity of the trajectories according to the magnitude of the underlying motion in order to focus on the description of the motion in regions which contain the meaningful information. The paper is organized as follows: in the method section, we briefly recall the model and algorithm proposed in [5] and introduce a criteria and a new method to locally adapt the complexity of the motion description. The next section illustrates the influence of the harmonic decomposition order with real 2D+time MRI and give quantitative results about parameters profit and accuracy of the motion estimation with synthetic 2D+time MRI sequences.

2 Method

Our method relies on the temporal modeling of free form deformations with a harmonic decomposition of the control point trajectories. After briefly recalling the principle, we focus on the proposal of this paper which consists in adapting the decomposition order of the model according to relevant information present in the image sequence.

2.1 Spatio-Temporal Model

Let consider a periodized image sequence $\mathcal{I} = \{\mathbf{I}_j, j = 0..J\}$ (Fig. 1) where $j \in \mathbb{Z}^+$ refers to the discrete time axis and where each image $\mathbf{I}_j$ belongs to $\mathbb{R}^d$ and corresponds to the discrete time point t_j. This sequence results from the observation of cardiac motion with tomographic imaging modalities like cardiac cine-MRI or MDCT.

A Lagragian formulation of the motion consists in describing the path of each material point $\mathcal{P}$ during the motion. The time-dependent coordinates $\boldsymbol{x}(t)$ of $\mathcal{P}$ is expressed according to the reference coordinates $\boldsymbol{x}_0$ at reference time t_0. The spatio-temporal transformation $\boldsymbol{\varphi}$ is a one to one continuous mapping defined by: $\mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}^d$; $\boldsymbol{x}(t) = \boldsymbol{\varphi}(\boldsymbol{x}_0, t)$. In our work, the non-rigid continuous spatial transformation φ is modeled by Free Form Deformations (FFD) [6, 7]. It warps an image by moving an underlying set of control points (CP) distributed over a regular grid. Instead of considering the displacement of each CP, we are looking for the whole trajectories they have to follow to fully describe the periodical spatio-temporal cardiac motion under study:

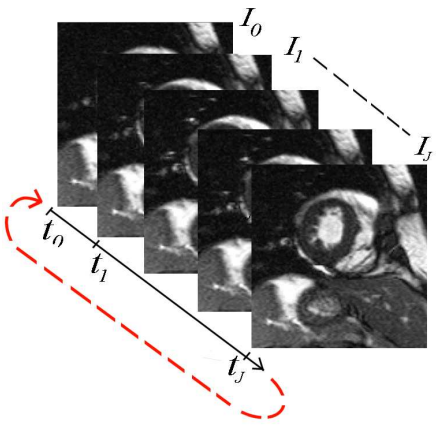


Fig. 1. Periodization of the image sequence to provide a pseudo infinite set of observations

$$\begin{aligned} \boldsymbol{\varphi}(\boldsymbol{x}_0, t) &= \boldsymbol{x}_0 + \boldsymbol{u}(\boldsymbol{x}_0, t) \\ \boldsymbol{\varphi}(\boldsymbol{x}_0, t) &= \boldsymbol{x}_0 + \sum_{k \in K_{inf}(\boldsymbol{x}_0)} \boldsymbol{\xi}_k(t) \beta^k(\boldsymbol{x}_0). \end{aligned} \tag{1}$$

The displacement $\boldsymbol{u}(\boldsymbol{x}_0, t)$ of point $\mathcal{P}$ at time t is modeled by a linear combination of a tensor product of interpolating functions β^k (cubic B-Spline functions in our case) and the position $\boldsymbol{\xi}_k(t)$ at time t of the CPs. k stands for CP index and K_{inf} is the subset of the CPs which influence the motion of $\mathcal{P}$. Thus, the transformation is continuous, smooth and semi-local.

$\boldsymbol{\xi}_k(t)$ is a time dependent function representing the d-dimensional path of CP with index k. In [5], we proposed to express it as a finite sum of periodic harmonic functions (or truncated Fourier series):

$$S_{\boldsymbol{\xi}_k}(t) = \boldsymbol{a}_k^0 + \sum_{n=1}^{N} \left[\boldsymbol{a}_k^n \cos(2\pi n f t) + \boldsymbol{b}_k^n \sin(2\pi n f t) \right], \tag{2}$$

where f is the frequency of the motion to be estimated and N the decomposition order. The truncation to the first N coefficients of the decomposition is equivalent to apply a low-pass filter to the trajectory signal. Small values of N result in very smooth CP trajectories while higher values increase the complexity of the trajectories. The periodicity property of such a decomposition is very interesting for the beating heart analysis. Furthermore, velocity and acceleration can be directly derived and exploited.

2.2 Parameters Estimation

Spatio-temporal motion estimation for the whole image sequence is performed by estimating the coefficients of the Fourier series for all the CPs (size = K). All the parameters are considered as stochastic signals and stored into the state vector $\mathbf{X}$ (eq.(3)) of a state-space formulation (eq.(4)):

$$\mathbf{X}^t = \left[\boldsymbol{a}_0^0\ \boldsymbol{a}_0^1\ ...\ \boldsymbol{a}_0^N\ \boldsymbol{b}_0^1\ ...\ \boldsymbol{b}_0^N\ |\ ...\ |\ \boldsymbol{a}_{K-1}^0\ \boldsymbol{a}_{K-1}^1\ ...\ \boldsymbol{a}_{K-1}^N\ \boldsymbol{b}_{K-1}^1\ ...\ \boldsymbol{b}_{K-1}^N\right]. \quad (3)$$

$$\begin{aligned} \mathbf{X}_{j+1} &= \mathbf{A}_j\mathbf{X}_j + \Gamma\mathbf{v}_j \\ \mathbf{Z}_j &= \mathbf{C}_j\mathbf{X}_j + \mathbf{w}_j \end{aligned} \quad (4)$$

The state vector is not directly measurable but it can be recursively estimated according to a set of successive measurements $\mathbf{Z}_j$. Kalman filter [8] allows to take into account the whole sequence in a recursive way so that we only need results of a previous estimation to perform the next Kalman filter iteration. The image sequence $\mathcal{I}$ is periodized ($\mathcal{I}|\mathcal{I}|\mathcal{I}|\mathcal{I}$..., Fig.(1)) to give a pseudo infinite set of observations. At each innovation step, the new measurement $\mathbf{Z}_j$ is computed from non rigid FFD based registration between image $\boldsymbol{I}_j$ and reference image $\boldsymbol{I}_0$. The Kalman filter provides a useful prediction of the state to initialize the registration at each time point.
The registration similarity criterion must be chosen according to the expected relation between the images to be registered. In the monomodal case, we can assume that the photometric level of material points remains almost constant during motion. The sum of squared differences (SSD) measure has therefore been chosen:

$$SSD_j(\boldsymbol{I}_0, \boldsymbol{I}_j, \boldsymbol{\xi}(t_j)) = \sum_{\boldsymbol{x}\in\Omega} (\boldsymbol{I}_0(\boldsymbol{x}_0) - \boldsymbol{I}_j(\boldsymbol{\varphi}(\boldsymbol{x}_0, t_j)))^2 \quad (5)$$

with Ω the image overlapping domain. Optimization of the criterion is performed through a gradient descent algorithm. The obtained value for $\boldsymbol{\xi}(t_j)$ stands for the new measurement $\mathbf{Z}_j$ to be introduced in the Kalman filter.

The registration algorithm relies on two multi-level pyramidal representations. A first multiresolution pyramid $\mathcal{P}_1$ decomposes the successive observations within the image sequence $\mathcal{I}$, applying a low-pass Gaussian filter to each image I_k independently and then, decimating the number of pixels (or voxels). The second pyramid $\mathcal{P}_2$ allows for the multiscale decomposition of the spatio-temporal FFD transformation [9, 10, 3] (see [5] for more details).

2.3 Harmonic order adaptation

The total number of parameters increases with the number of CPs and with the order of decomposition of their trajectories. Let consider a $\mathbb{R}^d$ space, a warping

grid of size M and a harmonic decomposition order N. The transformation is thus defined by $(2N+1) \times d \times M^d$ parameters.
We observed that for meaningful areas in the image sequence, a decomposition order of $N = 4$ (*i.e.* 9 parameters to describe a CP trajectory path for $d = 2$) is largely sufficient to recover the spatio-temporal motion. The model described in section 2.1 implies that the number of parameters is the same for each CP whatever the region it influences. Indeed, the CPs outside of the heart generally experience a very short trajectory. An adaptation of the harmonic decomposition would avoid spending too much time for meaningless regions of the image and, at the same time, reduce the parameter number of the spatio-temporal model.
We propose to estimate the energy associated to the trajectory of CP with index k through the Parseval formula:

$$E_k = \left|\boldsymbol{a}_k^0\right|^2 + \frac{1}{2}\sum_{n=1}^{N}(|\boldsymbol{a}_k^n|^2 + |\boldsymbol{b}_k^n|^2) \tag{6}$$

The motion estimation starts with $N = 0$ so that $\boldsymbol{a}_k^0$ is the only parameter to be estimated for each CP. This value corresponds to the mean position of the k^{th} CP all over the spatio-temporal motion. The overall algorithm (algorithm 1) consists in alternatively increasing the transformation scale and the image resolution. The algorithm starts with the lower scale of the transformation pyramid and the lower image resolution. After convergence of the Kalman filter at this step, a higher level of the transformation is considered. The previous state vector, which contains the estimated parameters, is used to initialize the state vector corresponding to the current transformation scale. The projection operation is performed according to the following scheme:

$$\mathbf{X}^{l+1} = \left(\uparrow_2 \mathbf{X}^l\right) * \mathbf{H}, \tag{7}$$

where $\uparrow_2$ stands for the upsampling operation, l the pyramid level and where,

$$\mathbf{H} = \frac{1}{\sqrt{2}}\left[\frac{1}{8} \quad \frac{1}{2} \quad \frac{3}{4} \quad \frac{1}{2} \quad \frac{1}{8}\right], \tag{8}$$

is called the mirror filter [11] whose coefficients depend on the interpolating function β.
The harmonic adaptation step occurs just before the transformation level change. To this aim, the motion energy of each CP is computed according to Eq(6). If $E_k > \mu$ where μ is a fixed threshold, the harmonic order N_k is incremented. Until now, μ is experimentaly fixed, but some decision criteria are being investigated

Data: $\mathcal{I}$, configuration file and input parameters
Initializations ($\mathcal{P}_1$ computation, Image Gradient, $N_k = 0$);
while *CurrentLevel* < *TotalCurrentLevel* **do**
  Initialize Kalman Filter;
  while *KalmanIteration* < *MaxKalmanIteration* **do**
    Prediction of the State at time k+1;
    Initialize registration parameters according to the prediction;
    Perform Gradient Descent Search;
    Filtering step;
    KalmanIteration = *KalmanIteration* + 1;
  end
  if *Increase transformation level* **then**
    Locally adapt the decomposition order N_k;
    Project parameters onto next transformation level;
    Increase $\mathcal{P}_2$ level;
    CurrentLevel = *CurrentLevel* + 1;
  else
    Increase $\mathcal{P}_1$ level;
    CurrentLevel = *CurrentLevel* + 1;
  end
end
Result: Spatio-temporal model parameters;

Algorithm 1: Flowchart of the estimation/prediction algorithm with local adaptation of the trajectory complexity

3 Results

For all the tests, the level number of both the image pyramid and the transformation pyramid are set to 4. The number of cycles used for Kalman convergence is fixed to 7 for each level of the pyramids (In general, 4 or 5 cycles are sufficient).

3.1 Harmonic decomposition order influence with true 2D+time MRI sequences

The ability to capture coarse motion from a reduced spatio-temporal model order will be studied on a true patient sequence. This sequence has been acquired using a cine MRI acquisition (1.5T Siemens Magnetom Vision scanner, Helsinki Medical imaging Center) and is composed of 28 time points covering the cardiac cycle (Fig 2, first row). A short axis slice was selected in the middle part of the heart, between the base and the apex. Image dimensions were 160×160, spatial resolution was $1mm \times 1mm$ and temporal resolution was $30ms$.
Figure 2 shows the values of E_k (Eq.(6)) at different transformation scales (FFD grid size 5×5, 7×7, 11×11, 19×19, respectively). Only the first term $\boldsymbol{a}_k^0$ of the Fourier series (*i.e.* model order $N = 0$) is considered for all CPs. This

corresponds to the mean CP position over the full cycle. It is clear on this figure that interesting regions, where motion occurs, are well detected even at order $N = 0$.

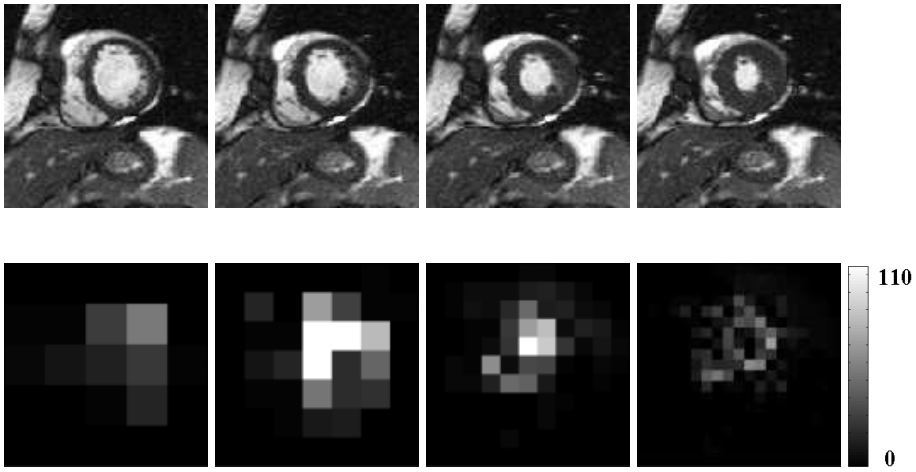
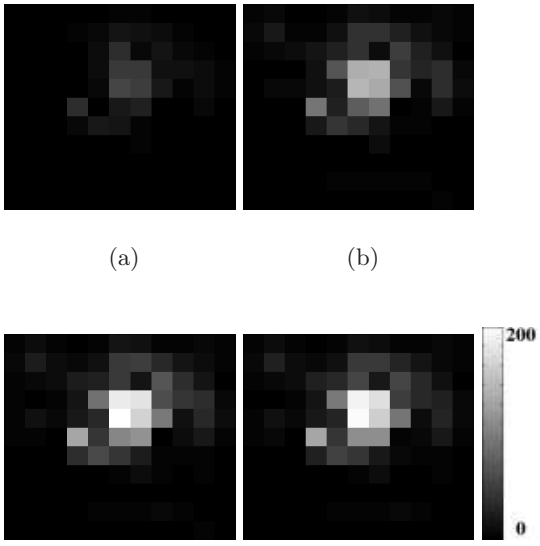


Fig. 2. The first row shows four images from the true MRI sequence during systole. The second rows illustrates the energy E_k at decomposition order $N = 0$ (mean term) for all the CPs. The intensity value in images a-b-c-d) respectively corresponds to the energy value at different scales of the transformation (one pixel per CP). The FFD grid size is 5×5, 7×7, 11×11, 19×19, respectively. The colorbar is expressed in mm^2.

Figure 3 illustrates E_k maps for a fixed transformation scale according to the decomposition order. We experimentally observed that order $N = 4$ is enough to describe the most complex trajectories in the sequence. This value corresponds to 9 parameters to describe trajectories in each direction and for each CP. It is a good trade-off between the smoothness of the estimated trajectories and the number of images in the processed sequences.

3.2 Quantitative results with synthetic 2D+time MRI sequences

Algorithm 1 was applied on a synthetic 2D sequence of a realistic beating heart. This sequence was generated from an actual 2D MRI short axis slice. The synthetic motion has been computed from a spatio-temporal analytical model described in [12]. Dimensions of the images are 160×160 pixels and spatial resolutions are $1mm \times 1mm$. Gaussian noise (O mean, $\sigma = 5$ square intensity unit) has been independently added to each image of the generated sequence. The theoretical spatio-temporal motion field is used as a reference to assess the performance



(a) (b)

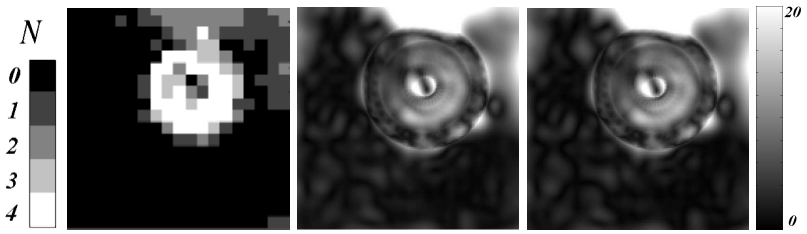
(c) (d)

Fig. 3. Illustration of the energy E_k with model order $N = 0$ (a), $N = 1$ (b), $N = 2$ (c) and $N = 4$ (d) for all the CPs and a transformation scale 11×11.

of the proposed algorithm in terms of mean quadratic (MQE) and mean angular errors (MAE) over the whole sequence. Because of the high number of involved parameters, standard computer architectures (with 1Gb of RAM) fail to converge when N is high and constant for all the CPs. In that case, there is a trade-off to find between the final FFD grid resolution (spatial resolution) and the decomposition order (temporal resolution). This problem is overcome with the proposed method. The Table 1 quantitatively compares our method, the case where $N = 4$ for all CPs of a 19×19 FFD grid (computed on a SGI, 64GB RAM) and an acceptable configuration for a standard computer architecture (Grid Size $= 19 \times 19$, $N = 2$). Results show that the parameter number is significantly reduced with the proposed method. This important properties allow to reach the final transformation scale while considering a decomposition order $N = 4$ for some of the crucial points for the final spatio-temporal motion model. A cartography for N is given in Fig 4.a). The accuracy is even increased (Tab. 1), especially for meaningful areas of the processed sequence as illustrated in Fig 4.b) c). This is because parameter reduction results in a low-pass filtering of the background noise.

Table 1. MQE (in mm) and MAE (in degree) between the reference and the estimated motions fields. All the 22 time-points have been considered to compute the errors. The last column gives the number of parameters required to obtain the results.

	Quadratic error	Angular Error	Number Of Parameters
Adaptative decomposition	0.24 ± 0.28	3.7 ± 8.0	1 830
Grid Size: 19×19; $N = 4$	0.25 ± 0.29	3.8 ± 8.7	6 498
Grid Size: 19×19; $N = 2$	0.29 ± 0.29	4.1 ± 8.6	3 610



(a) (b) (c)

Fig. 4. Cartography of the harmonic order used for the final motion model a). Dark pixels correspond to small N values and bright pixels correspond to high values. b) and c) display the MQE map (mm) for reference and estimated motion fields, respectively. b) corresponds to the configuration where the FFD grid size is 19×19 and $N = 4$. c) corresponds to the map associated to the proposed algorithm.

4 Conclusion

We presented a method to adapt the complexity of a spatio-temporal motion model to the content of the image sequence. The parameters of a continuous state space parametric model are estimated by an algorithm, based on Kalman filtering from an image sequence describing a periodic phenomena. All the sequence frames are taken into account to infer the spatio-temporal model. Reduction of the parameter number using the proposed approach allows to obtain a better accuracy. Complementary tests have to be conducted to evaluate the influence of other method's parameters. As applied to cardiac imaging, such an approach allows for the direct determination of motion parameters that can be exploited for clinical interpretation and diagnosis, helping for instance in the detection of contraction abnormalities. The parameters profit with the proposed scheme makes now possible to conduct tests with 3D image sequences.

5 Acknowledgments

We truly thank J. Lötjönen and the Helsinki Medical imaging Center for providing us cine-MR data of the heart, and the Research group GDR STIC-Santé of the CNRS. This

work has been supported by the French research project ACI-AGIR (http://www.aci-agir.org/) and the Région Rhônes Alpes through the PP3-I3M project of cluster ISLE. It is also part of the French ANR (http://www.agence-nationale-recherche.fr/) project GWENDIA.

References

1. J. Montagnat and H. Delingette, "4D deformable models with temporal constraints : application to 4D cardiac image segmentation," *Medical Image Analysis*, vol. 9, no. 1, pp. 87–100, February 2005.
2. D. Perperidis, R. Mohiaddin, and Rueckert D., "Spatio-temporal free-form registration of cardiac MR image sequences," *Medical Image Analysis*, vol. 9, no. 5, pp. 441–456, 2005.
3. M. J. Ledesma-Carbayo, J. Kybic, M. Desco, A. Santos, M. Sühling, P. Hunziker, and M. Unser, "Spatio-temporal nonrigid registration for ultrasound cardiac motion estimation," *IEEE Transactions on Medical Imaging*, vol. 24, no. 9, pp. 1113–1126, September 2005.
4. J. Huang, D. Abendschein, V. Davila-Roman, and A. Amini, "Spatio-temporal tracking of myocardial deformations with a 4D B-spline model from tagged MRI," *IEEE Transactions on Medical Imaging*, vol. 18, no. 10, pp. 957–972, 1999.
5. B. Delhay, P. Clarysse, C. Pera, and I.E. Magnin, "A spatio-temporal deformation model for dense motion estimation in periodic cardiac image sequences," in *From Statistical Atlases to Personalized Models: Understanding Complex Diseases in Populations and Individuals, Satellite Workshop MICCAI 2006*, Copenhagen, Denmark, October 2006.
6. T. W. Sederberg and S. R. Parry, "Free-form deformation of solid geometric models," in *SIGGRAPH '86: Proceedings of the 13th Annual Conference on Computer Graphics and Interactive Techniques*, New York, USA, 1986, pp. 151–160, ACM Press.
7. D. Rueckert, L. I. Sonoda, C. Hayes, and D. L. G. Hill, "Nonrigid registration using free form deformation: application to breast MR images," *IEEE Transactions on Medical Imaging*, vol. 11, pp. 712–721, 1999.
8. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.
9. V. Noblet, C. Heinrich, F. Heitz, and J.-P. Armspach, "3D deformable image registration: a topology preservation scheme based on hierarchical deformation models and interval analysis optimization," *IEEE Transactions on Image Processing*, vol. 14, no. 5, pp. 553–565, 2005.
10. T. Rohlfing, C. R. Maurer, D. A. Bluemke, and M. A. Jacobs, "Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint," *IEEE Transactions on Medical Imaging*, vol. 22, pp. 730–741, 2003.
11. S. Mallat, *A wavelet tour of signal processing, 2nd edition*, Academic Press, San Diego, 1998.
12. P. Clarysse, C. Basset, L. Khouas, P. Croisille, D. Friboulet, C. Odet, and I. E. Magnin, "2D spatial and temporal displacement field fitting from cardiac MR tagging," *Medical Image Analysis*, vol. 3, pp. 253–268, 2000.