

Monogenic Signal for Cardiac Motion Analysis from Tagged Magnetic Resonance Image Sequences

Martino Alessandrini¹, Hervé Liebgott¹, Adrian Basarab², Patrick Clarysse¹, Olivier Bernard¹

¹CREATIS, CNRS UMR 5220, INSERM U1044, INSA-Lyon, Université Lyon 1, France

²IRIT, CNRS UMR 5505, Toulouse, France

Abstract

This paper presents a novel algorithm for the analysis of heart motion from tagged magnetic resonance images. The displacement is estimated from the monogenic phase and is therefore robust to possible variations of the local image energy. A local affine model accounts for the typical contraction, torsion and shear of myocardial tissue. An effective B-spline multiresolution strategy automatically selects the scale returning the most consistent velocity estimate. The multiresolution strategy together with a least-squares estimate of the monogenic orientation make the algorithm robust under image noise. Results on realistic simulated images show the proposed algorithm to return more accurate velocity estimates than the SinMod algorithm, itself shown more accurate and robust than the state-of-the-art Harp method.

1. Introduction

Tagged MRI (tMRI) is currently the gold-standard technique for quantification of myocardial contractility [1, 2]. With this technique, cardiac tissue is marked with magnetically saturated tagging lines or grids that deform with the underlying tissue during the cardiac cycle, thus providing details on the myocardial motion.

With time elapsing, the grid loses contrast and sharpness. This is why state-of-the-art techniques for motion estimation from tMRI sequences exploit the image phase rather than the less trustworthy pixel intensity. The popular algorithms HARP (harmonic phase) [1] and SinMod (sine-wave modeling) [2] belong to this family of methods. In particular, SinMod was shown to outperform HARP in [1]. Both methods make use of a directional bandpass filter bank to build first-order harmonic images. From the latter, the displacement is estimated with Fourier-based disparity measures. The limit of directional filtering is that it assumes tags direction and spacing to be constant on all the sequence while, in practice, these values may change locally due to the heart deformation. In this context, we pro-

pose here a novel motion estimation algorithm based on a more sophisticated and flexible image processing tool, called *monogenic signal*.

The monogenic signal extends the analytic signal concept to multiple dimensions [3]. Locally, it decomposes the image into the structural features of *phase* and *orientation* and in its local *amplitude*. These features are computed by locally approximating the image as a 1D monochromatic wave propagating in the direction of maximum image energy variation. Interestingly, due to the periodic patterns, the monochromatic wave model provides a realistic representation for the structure of tMRI images. This makes the monogenic signal an excellent framework for developing techniques dedicated to their analysis. To our knowledge this is the first study investigating this possibility.

Our algorithm computes the displacement by tracking variations in the monogenic phase. The computation is made locally on a spatially moving window. A local affine model accounts for the typical contraction, torsion and shear components of myocardial tissue. An effective B-spline multiresolution strategy automatically selects the scale returning the most consistent velocity estimate. The multiscale window choice together with an unconventional least-squares orientation estimate improves the robustness under image noise. Besides that the employment of the phase reduces the sensitivity to brightness variations.

The paper proceeds as follows. In Section 2 the monogenic signal is briefly summarized. Section 3 describes the proposed motion estimation algorithm. In Section 4 implementation details are discussed. Section 5 assesses the performance of the proposed algorithm and compares it with SinMod. Concluding remarks are left to Section 6.

2. Monogenic signal analysis of tagged MRI images

The monogenic signal provides an extension of the standard analytic signal for 2D gray-scale images. The adopted image model is [4]:

$$I(\mathbf{x}) = A(\mathbf{x}) \cos(\varphi(\mathbf{x})), \quad (1)$$

where $\mathbf{x} = [x, y]$ is the spatial coordinate vector, $A(\mathbf{x})$ is the *local amplitude* and $\varphi(\mathbf{x})$ is the *local phase*. Additionally, intrinsic dimensionality one is assumed, *i.e.*, the local variations of I are concentrated along a single direction, defined by the *local orientation* $\theta(\mathbf{x})$. Amplitude, phase and orientation are computed from the responses to three 2D spherical quadrature filters (SQFs) [4] obtained likewise: one *even* rotation invariant bandpass $b_e(\mathbf{x})$ filter and two *odd* bandpass filters $b_{o1}(\mathbf{x})$ and $b_{o2}(\mathbf{x})$. The odd filters are computed from the Riesz transform of the even filter, see [3, 4] for details. Several SQFs families exist in the literature. Here, the difference of Poisson (DoP) kernel is adopted [4].

Monogenic features are obtained as:

$$\theta(\mathbf{x}) = \arctan\left(\frac{q_2(\mathbf{x})}{q_1(\mathbf{x})}\right), \quad (2)$$

$$\varphi(\mathbf{x}) = \arctan\left(\frac{|\mathbf{q}(\mathbf{x})|}{p(\mathbf{x})}\right), \quad (3)$$

and $A(\mathbf{x}) = \sqrt{p^2(\mathbf{x}) + |\mathbf{q}(\mathbf{x})|^2}$, where $p(\mathbf{x}) = (I * b_e)(\mathbf{x})$, $q_1(\mathbf{x}) = (I * b_{o1})(\mathbf{x})$, $q_2(\mathbf{x}) = (I * b_{o2})(\mathbf{x})$, $\mathbf{q}(\mathbf{x}) = [q_1(\mathbf{x}), q_2(\mathbf{x})]^T$ and “*” denotes 2D convolution. Monogenic phase and orientation can be conveniently combined in the *phase vector* $\mathbf{r}(\mathbf{x}) = [r_1(\mathbf{x}), r_2(\mathbf{x})] = \varphi(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x})$, with $\mathbf{n}(\mathbf{x}) = [\cos(\theta(\mathbf{x})), \sin(\theta(\mathbf{x}))]^T$ [4].

The *local frequency*, defined as the derivative of the phase along \mathbf{n} , can also be computed as [4]:

$$f \triangleq (\nabla\varphi)^T \cdot \mathbf{n} = \frac{p\nabla^T\mathbf{q} - \mathbf{q}^T\nabla p}{p^2 + |\mathbf{q}|^2}, \quad (4)$$

where $\nabla = [\partial_x, \partial_y]^T$. Dependency on \mathbf{x} is implied.

In this study, as recommended in [5], we replace the classical point-wise estimate of θ (2), with a robust least-squares estimate. The least-squares orientation estimate is obtained by maximizing the directional Hilbert transform [5] averaged over a local neighborhood v_σ . The maximization problem is solved by the eigenvector associated with the largest eigenvalue of the 2×2 matrix $\mathbf{T}(\mathbf{x})$, with entries:

$$[\mathbf{T}(\mathbf{x})]_{nm} = \int_{\mathbb{R}^2} v_\sigma(\mathbf{x}' - \mathbf{x}) q_n(\mathbf{x}') q_m(\mathbf{x}') d\mathbf{x}', \quad (5)$$

with $n, m = \{1, 2\}$. The matrix \mathbf{T} can be assimilated to a Riesz-transform counterpart of the standard structure tensor. See [5] for details. Here, v_σ corresponds to a Gaussian kernel with variance σ^2 .

In the case of tagged MRI images, the commonly adopted image model is [1, 2]:

$$I(\mathbf{x}) = A(\mathbf{x}) \cos(\omega_0^T \mathbf{x}) \quad (6)$$

where $A(\mathbf{x})$ is the image without the tag pattern and $\omega_0 = \omega_0 \mathbf{n}$, with \mathbf{n} the tags direction and ω_0 is related to the tags spacing l_0 by $\omega_0 = 2\pi/l_0$.

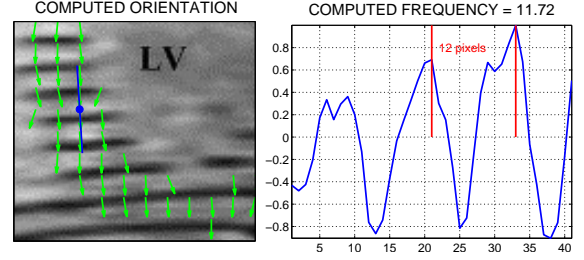


Figure 1. (a) tMRI image with horizontal tags. The computed orientation is represented as green arrows. The blue segment defines a 1D neighborhood selected by the local orientation. (b) The blue line shows the image intensity profile along the blue segment of figure (a). The local tags spacing is of 12 pixels, very well estimated by the monogenic frequency ($l = 1/f = 11.72$ pixels).

The tMRI image model (6) is closely related to the monogenic image model in (1). In particular, it directly satisfies the assumption of local dimensionality one. This allows to establish a direct correspondence between the previously introduced monogenic features and the geometrical properties of tMRI images. Namely, monogenic orientation computes the local tags direction while monogenic frequency estimates the local tags spacing (cf. Fig. 1). These structural features will be exploited in the derivation of a robust motion estimation algorithm in the next section.

3. Proposed motion estimation algorithm

Following [4], we compute displacement $\mathbf{d} = [d_1, d_2]$ between two frames assuming the conservation of the image phase over time, *i.e.* $\varphi(\mathbf{x}, t+1) = \varphi(\mathbf{x} - \mathbf{d}, t)$. Since, according to (6), $\varphi(\mathbf{x}, t) = \omega_0 \mathbf{n}^T \mathbf{x}$, the equation for the displacement becomes $-\varphi_t(\mathbf{x}) = \omega_0 \mathbf{n}^T \mathbf{d}$, where we have called $\varphi_t(\mathbf{x}) = \varphi(\mathbf{x}, t+1) - \varphi(\mathbf{x}, t)$. By pre-multiplying both terms by \mathbf{n} , we obtain the phase conservation expressed in terms of the phase vector $\mathbf{r}(\mathbf{x})$ [4]. Then, assuming all points translate of the same quantity \mathbf{d}_0 within a local window w centered in $\mathbf{x}_0 = [x_0, y_0]$, the following linear system of equations is obtained:

$$\langle \mathbf{J} \rangle_w \mathbf{d}_0 = -\langle \mathbf{r}_t \rangle_w, \quad \mathbf{J} = \omega_0 \mathbf{n} \mathbf{n}^T \quad (7)$$

where $\langle \mathbf{v} \rangle_w$ denotes the weighted average $\int_{\Omega} w(\mathbf{x} - \mathbf{x}_0) \mathbf{v}(\mathbf{x}) d\mathbf{x}$. In (7), $\omega_0 = 2\pi f$, where f is computed as in (4) and \mathbf{n} is defined by the monogenic orientation θ . The term \mathbf{r}_t is computed from the SQFs responses as in [4]:

$$\mathbf{r}_t = \frac{p_1 \mathbf{q}_2 - \mathbf{q}_1 p_2}{|p_1 \mathbf{q}_2 - \mathbf{q}_1 p_2|} \arctan\left(\frac{|p_1 \mathbf{q}_2 - \mathbf{q}_1 p_2|}{p_1 p_2 + \mathbf{q}_1^T \mathbf{q}_2}\right) \quad (8)$$

where subscripts 1 and 2 denote the time instant t and $t+1$ respectively.

We conclude this section by noting that the maximum displacement that our algorithm can compute is equal to one half the tags spacing $l_0/2$. Beyond that value, due to the periodicity of the pattern, the velocity cannot be determined unambiguously.

3.1. Affine model

Clearly, the simple translation model employed by Felsberg [4] is too restrictive in a general context. Also, its validity is heavily dependent on the choice of the size of w . The solution we propose is to replace the constant motion assumption with an affine model. A part of translations, this accounts for rotation, expansion, compression and shear. Considering a window w centered at $(x_0, y_0) = (0, 0)$, the affine model is written:

$$\mathbf{d}(\mathbf{x}) = \mathbf{B}(\mathbf{x})\mathbf{u}, \quad \mathbf{B} = \begin{bmatrix} 1 & 0 & x & y & 0 & 0 \\ 0 & 1 & 0 & 0 & x & y \end{bmatrix}, \quad (9)$$

where $\mathbf{u} = [d_{10}, d_{20}, d_{1x}, d_{1y}, d_{2x}, d_{2y}]^T$ is the new unknown vector: d_{10} and d_{20} correspond to the translation of the window center and $d_{ik} = \partial_k d_i$.

Plugging (9) into (7) leads to an underdetermined system of equations. The solution is then obtained by pre-multiplying both terms by \mathbf{B}^T , i.e. $\langle \mathbf{M} \rangle_w \mathbf{u} = \langle \mathbf{b} \rangle_w$, with $\mathbf{M} = \mathbf{B}^T \mathbf{J} \mathbf{B}$ and $\mathbf{b} = -\mathbf{B}^T \mathbf{r}_t$.

3.2. Multiscale choice of window size

The choice of the window size is a tedious issue connected with local techniques: the assumed motion model (translational or affine) may not hold when the window is too big, otherwise, the adoption of an excessively small window may result in the well known *aperture problem*. To circumvent this issue, in [6] Sühling *et al.* proposed a multiscale strategy for locally choosing the most consistent window size. This is based on the possibility of computing the image moments, i.e., the entries of the system matrix \mathbf{M} and the vector \mathbf{b} , at multiple scales, by using an efficient B-spline *coarse-to-fine* strategy. In particular, they are obtained from window functions w that are progressively scaled and subsampled by a factor 2 in each dimension. More precisely, at scale j , the window $w^j(\mathbf{x} - \mathbf{x}_0) = w((\mathbf{x} - 2^j \mathbf{x}_0)/2^j)$ is employed, where w is written as the separable product of two B-spline functions.

By doing so, at each scale $J_f \leq j \leq J_c$ ($J_f \geq 0$) a solution \mathbf{u}^j can be computed. Among the scales considered, the \mathbf{u}^j producing the smallest residual error $\|\mathbf{M}\mathbf{u}^j - \mathbf{b}\|_{\ell_2}/\|w\|_{\ell_1}$ is retained as the final displacement estimate. Whenever necessary, bi-cubic interpolation is employed to obtain a dense motion field. With this strategy, the scale providing the most consistent motion estimate is selected.

Table 1. Endpoint Error ($\mu \pm \sigma$) in pixels on 8 simulated sequences.

SEQUENCE	ALGORITHM	
	Proposed	SinMod
D30	0.152 ± 0.121	0.215 ± 0.145
D30F20	0.082 ± 0.072	0.128 ± 0.112
D30R10T01P0	0.264 ± 0.149	0.363 ± 0.199
D30R20T01P0	0.462 ± 0.239	0.970 ± 1.129
D30R20T01P0F20	0.209 ± 0.139	0.344 ± 0.224
R20F20	0.244 ± 0.164	0.416 ± 0.264
R10	0.161 ± 0.087	0.220 ± 0.090
R20	0.104 ± 0.072	0.174 ± 0.122

3.3. Iterative displacement refinement

The hypothesis of small displacements employed in differential techniques may be inadequate whenever the displacement is substantial or the image intensity profile is non-linear. A possible way to deal with this limitation is to implement a form of Gauss-Newton optimization: the current estimate is used to undo the motion, and then the estimator is reapplied to the warped images to find the residual displacement [7]. When applied iteratively, this procedure can improve the estimation accuracy considerably. We employed the aforementioned refinement scheme in the algorithm presented.

4. Implementation details

The number of iterations of the refinement scheme of Section 3.3 was 3 in all the experiments. The center frequency of the SQFs was chosen close to the image first harmonic, given by $1/l_0$ ($l_0 \approx 6$ pixels in our experiments). The multiscale window choice was implemented by considering fifth-order B-splines and scales $j = \{2, 3, 4, 5\}$. A value $\sigma = 2$ was used for the robust computation of the monogenic orientation. The proposed algorithm has been implemented in MATLAB (R2011b, The MathWorks, Natick, MA).

5. Results

The algorithm was tested on realistic simulated cardiac tMRI sequences with grid patterns for which the benchmark motion was known and the performance was compared to the one of SinMod [2]. Synthetic sequences were generated with the ASSESS software [8]. The characteristic tag-fading effect, not considered in ASSESS, was also taken into account in this study. The effect was obtained by adjusting the image's histogram limits on each frame so as to match those of a real sequence taken as a template.

The estimation error was assessed with the endpoint error $EE = \|\mathbf{d} - \hat{\mathbf{d}}\|_{\ell_2}$, measured in pixels, where \mathbf{d} denotes

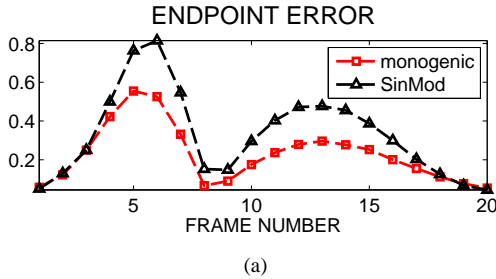


Figure 2. Average endpoint error on R20F20.

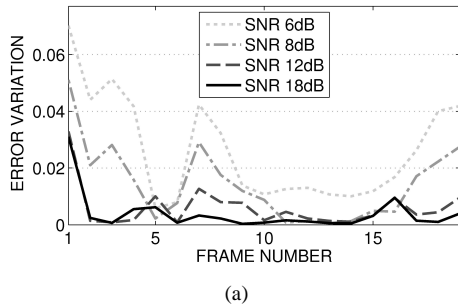


Figure 3. Performance for decreasing values of SNR.

the estimated displacement and \bar{d} the benchmark displacement.

The results obtained on 8 simulated sequences are summarized in Table 1. For each algorithm the parameters were optimized to return the smallest average error on the sequence D30R20T01P0F20. The name of each sequence reflects the values of the parameters used for its generation, namely: contraction/expansion (D), rotation (R), thickening (T), frame-rate (F) and healthy (P0) or pathological (P3) state. Greater detail on their meaning can be found in [8]. These results show that the proposed algorithm systematically returns the estimate with the smallest mean value and variance, which is a proof of precision and reliability.

The algorithm performance on each frame of a simulated sequence is reported in Fig. 2. The figure shows that the obtained estimates are constantly more precise than the ones returned by SinMod. The behavior is similar on all the considered sequences.

The sensitivity to noise was also evaluated. To this end, we contaminated the frames of sequence R20F20 with additive Rician noise [9]. Fig. 3 reports the endpoint error variation for decreasing values of SNR. The results are based on 15 independent noise realizations. These results show that the performance of the proposed algorithm remains virtually unchanged. The robustness against noise stems from two factors: the multiscale window choice and the robust monogenic orientation. The first guarantees that the integration scale is optimized locally so as to minimize

the noise effect on the velocity determination, while the second ensures a more robust computation of the monogenic features.

The computational time for the optimal parameters set was of 0.55 s/image (image size, 256×256 pixels²) for the adopted unoptimized MATLAB implementation.

6. Discussions and conclusion

We have described a novel algorithm for the analysis of heart motion from tMRI images. On considered data, the algorithm has been shown to be more accurate than the recent SinMod algorithm.

We conclude by noting that the model (6) is perfectly suited for line-tags. In the case of the grid-tags, a second wave perpendicular to the first should be included in the image model. This would suggest investigating the use of 2D extensions of the monogenic signal as the one in [10]. Nonetheless, the results presented here show that, even in the most common grid-tag case, the monogenic-phase-based algorithm presented still produces relevant estimates.

References

- [1] Osman N, McVeigh E, Prince J. Imaging heart motion using harmonic phase mri. *IEEE TMI* 2000;19(3):186–202.
- [2] Arts T, Prinzen F, Delhaas T, Milles J, Rossi A, Clarysse P. Mapping displacement and deformation of the heart with local sine-wave modeling. *IEEE TMI* 2010;29(5):1114–1123.
- [3] Felsberg M, Sommer G. The monogenic signal. *IEEE TSP* 2001;49(12):3136–3144.
- [4] Felsberg M. Optical flow estimation from monogenic phase. *Proceedings of the 1st International Conference on Complex motion, IWCM'04*. 2004; 1–13.
- [5] Unser M, Sage D, Van De Ville D. Multiresolution monogenic signal analysis using the riesz laplace wavelet transform. *IEEE TIP* 2009;18(11):2402–2418.
- [6] Sühling M, Arigovindan M, Jansen C, Hunziker P, Unser M. Myocardial motion analysis from b-mode echocardiograms. *IEEE TMI* 2005;14(4):525–536.
- [7] Bergen JR, Anandan P, Hanna KJ, Hingorani R. Hierarchical model-based motion estimation. In *Proceedings ECCV* 1992; 237–252.
- [8] Clarysse P, Tafazzoli J, Delachartre P, Croisille P. Simulation based evaluation of cardiac motion estimation methods in tagged-mr image sequences. *Journal of Cardiovascular Magnetic Resonance* 2011;13:360.
- [9] Smal I, Carranza-Herrezuelo N, Klein S, Niessen W, Meijering E. Quantitative comparison of tracking methods for motion analysis in tagged mri. In *Proc. of ISBI* 2011; 345–348.
- [10] Wietzke L, Sommer G. The signal multi-vector. *Journal of Mathematical Imaging and Vision* 2010;37:132–150.