QUANTITATIVE COMPARISON OF TRACKING METHODS FOR MOTION ANALYSIS IN TAGGED MRI

Ihor Smal, Noemí Carranza-Herrezuelo, Stefan Klein, Wiro Niessen, Erik Meijering

Biomedical Imaging Group Rotterdam,
Departments of Medical Informatics and Radiology
Erasmus MC – University Medical Center Rotterdam, The Netherlands
Email: i.smal@erasmusmc.nl

ABSTRACT

Myocardial tagging in magnetic resonance imaging (MRI) has shown great potential for noninvasive measurement of the motion of a beating heart. A critical issue in exploiting this technology in practice is the availability of robust and accurate methods for tag tracking. In this paper we quantitatively evaluate and compare four motion analysis methods that are frequently used in practice, based on optical flow, harmonic phase MRI, B-snake grids, and non-rigid registration techniques. Experiments on realistic synthetic images and data from three different (pre)clinical experiments show that non-rigid registration methods yield the highest accuracy and robustness among the considered methods.

Index Terms— Optical flow, non-rigid registration, HARP, motion analysis, tracking, tagged MRI.

1. INTRODUCTION

Myocardial tagging using magnetic resonance imaging (MRI) is a well-known noninvasive method for studying regional heart dynamics. It offers great potential for quantitative analysis of a variety of kine(ma)tic parameters that can be used for identification of ischemic and infarcted tissues. Tagging involves the selective magnetic saturation of a plane orthogonal to the final image plane before the standard spin-echo sequence [1]. In the images, tags appear as dark lines (compared to the bright myocardium, see Fig. 1), which move along with the tissue and allow motion quantification of specific areas of the myocardium. The tag pattern amplitude decays with time due to T_1 relaxation. It is known from the imaging protocol that tag lines, which are typically about 1 pixel thick, have an approximately Gaussian intensity profile [2], and the images are corrupted by Rician measurement noise.

The analysis of tagged MR images commonly consists of three tasks [2–6]. First, the endocardial and epicardial contours of the left ventricle (LV) are defined either manually or with semi-automated algorithms [2]. Second, the tag lines are automatically segmented in every image of the sequence. In most cases, the line intersections are used as feature points during the tracking. Third, the tag line positions are used to either fit a parametrized model of cardiac displacement [3], or to reconstruct a dense displacement field [7] from sparse measurements using interpolation techniques based on thin-plate spline or B-spline models [4, 6].

Fig. 1. Examples of images from studies of human hearts, with frames 1, 10, and 20 (out of 20 per heart cycle) clearly showing the tag fading.

The process of identifying and tracking tag intersections manually, to analyze the relative motion of tags and calculate local strain and rotation, is time consuming and laborious. Several computerized methods for motion analysis in tagged MRI (tMRI) have been proposed in the literature [2–6], but a thorough comparison of their relative performance has been lacking. To establish the performance baseline, we compare four existing and conceptually different automated tracking approaches that are frequently used in practice for motion analysis in tMRI. The experiments are based on both realistic synthetic data with known ground truth, as well as real data from three different (pre)clinical studies, and aim to demonstrate the strong and weak points of each method using the same standardized image data.

2. TRACKING METHODS

The four tracking methods compared in this paper, based on optical flow, harmonic phase MRI, B-spline grids, and non-rigid registration methods are briefly described next. All of them do not require indication of LV contours and produce a dense displacement field, which can be used to obtain the

Supported by the European Commission (FP7-grant 201842).

deformation gradient tensor, and to compute any concise description of regional cardiac kinematics.

Optical flow (OF) methods for tag tracking [8, 9] assume that the intensity of image structures is constant under motion, at least for a short duration, which is admissible if the image frames are separated by short time intervals as compared to T_1 recovery (the T_1 signal modulation may also be incorporated in the framework [9] if needed). These methods do not require explicit modeling of tag appearance, but do require the displacements between successive frames to be relatively small, as they are measured by taking local derivatives. Here, we used total variation (TV) regularization of the OF displacement fields, which accounts for anisotropic noise removal and edge preservation, and recently was successfully used for motion estimation in tMRI [8].

Harmonic Phase MRI (HARP) [10] is another frequently used and commercially available (Diagnosoft®HARP™) technique, based on the fact that tagged MR images show distinct spectral peaks in the Fourier domain, each of which contains information about the motion in a certain direction [10]. The inverse Fourier transform of each peak, extracted by a bandpass filter, gives a complex image, whose phase is linearly related to a directional component of the myocardial motion, but "wrapped" in the range $[-\pi, \pi)$. Every point in the image is characterized by two HARP angles, corresponding to two orthogonal displacement components. Tracking proceeds from frame t to $t+1$ by finding a corresponding point that shares the same HARP angles. HARP is fundamentally capable of tracking arbitrary points in the image, but is prone to errors at points near the myocardial boundary, where tag fading and non-myocardial background structures may cause erroneous linking of points that coincidentally have the same pair of HARP angles.

B-snake grids (BSG) is an alternative approach, based on active contour models ("snakes"), that aims at tracking tag lines [3, 6]. The grids are constructed by having horizontal and vertical curves (represented by B-splines) share control points, where the individual movement of control points has only local effect, so that local tissue deformations can be captured without affecting other parts of the curve. The number of control points is usually much smaller than the sampling of the curve on the pixel grid. The grid deforms (by optimizing the control points) to align with the low image intensity points (corresponding to tag lines). The estimated grid in one frame is then taken as the initial configuration for the next frame. Optimization is typically done with the nonlinear conjugate gradient method or quasi-Newton algorithms.

Non-rigid registration (NRR) methods, which also have been applied to tracking and motion analysis [11, 12], aim at finding an optimal transform $\mathbf{T}_t$, which spatially aligns every image I_t with a reference image (in our case the first image of the sequence). To capture the local transformation of the myocardium, free form deformation (FFD) models based on B-splines [12] are used (similar to BSG), which produce a smooth and continuous transform $\mathbf{T}_t$. To measure the degree of alignment between two images, we used mutual information (MI), a well-known and robust similarity measure. The cost function associated with the transformation parameters can be optimized using an iterative optimization method such as the quasi-Newton, conjugate gradient, or stochastic gradient descent methods, usually in a multiresolution setting [13]. Robustness and accuracy are improved by initializing the registration between the reference image and I_t at time t with the transformation obtained at $t-1$. In our implementation the optimization was done using the adaptive stochastic gradient descent method and a 4-level multiresolution strategy.

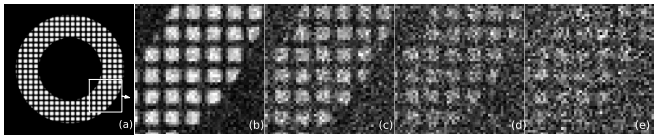

Fig. 2. Examples of synthetic images (Type A) used in our experiments: the initial noise-free image of the tagged annular object (a), and zoomed image regions that demonstrate the modeled tag and myocardium appearance for different SNR levels, which correspond to 18.06 dB (b), 12.04 dB (c), 8.51 dB (d), and 6.02 dB (e).

3. EXPERIMENTAL RESULTS

Evaluation on Synthetic Image Data. The performance of the methods was first evaluated using realistic synthetic image data. Two types of synthetic images, for which the ground truth was available, were used to assess accuracy and robustness as a function of the signal-to-noise ratio (SNR) of the data. The first (Type A) consisted of a sequence of 24 images (256×256 pixels) and modeled motion (radial expansion and contraction) of an annular object with an outer radius $r_{ext} = 80$ pixels, and inner radius $r_c = 50$ pixels, similar to [14, 15]. The intensity levels of the modeled myocardium and tagging, and the level of the Rician noise, were chosen to simulate data with SNRs in the range 6-18 dB (Fig. 2).

As the second type (Type B) of synthetic data we used the previously reported computational phantom [11]. This phantom (Fig. 3) models more complex motion, where in addition to radial expansion and contraction, also rotation is considered to realistically simulate left ventricular motion through the cardiac cycle. For both types of data (A and B), we additionally modeled the fading of the tag pattern during a heart cycle, as typically observed in practice. Exponential decaying of the tag amplitude was used to produce image sequences, where the SNR changed from 18.06 dB (in the initial frame) to 6.02 dB (in the last frame of the sequence).

The parameters for each method were manually optimized for best performance. The accuracy of tag position estimation was measured using the root mean square error (RMSE), averaged for each method over N_R independent runs on the synthetic data with different noise realizations (we used $N_R =$

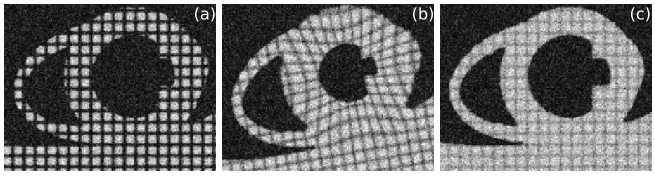

Fig. 3. Example of synthetic data (Type B) with time-varying SNR level that models the tag fading: the highest SNR = 18.06 dB (a) exponentially decays through the sequence to the lowest SNR = 6.02 dB (c).

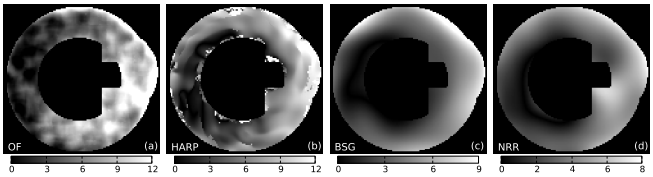

Fig. 4. Examples of the error ("heat") maps produced by each method for Type B data. The intensity values represent the absolute error (in pixel units) between the measured and the ground truth displacement for each point of the myocardium. Results are given for the lowest SNR (6.02 dB) for the frame (Type B data) where the largest deformations occurred.

5). Additionally, the image sequences were manually annotated by 3 independent observers. The results for each method as well as for manual tracking are shown in Tables 1 and 2. Examples of typical errors produced by each method in a single experiment for Type B data are shown in Fig. 4.

The OF-based method demonstrated high sensitivity to SNR changes and produced the largest RMSE (together with BSG). High noise levels complicated the estimation of image derivatives, which the method is based on. Independently of the data type and SNR, the largest errors always occurred at the boundaries of the modeled myocardium (Fig. 4(a)), where the estimation of the intensity derivatives is extremely prone to errors, especially in the case of large boundary displacements (fast expansion or contraction).

The HARP method performed relatively well, even for the lowest SNRs considered in this study. The RMSE was either comparable to or lower than the error made by manual analysis (Table 1). As known from literature, and confirmed by our experiments, HARP is accurate within the myocardial region and generates errors mainly at the boundaries (Fig. 4(b)), where it fails to find tags with corresponding HARP angles in subsequent HARP images. Most of the time, the method produces small and rather isolated "islands" of large errors in the regions close to the myocardial boundaries.

The performance of the BSG method varied considerably depending on SNR. For high SNRs, BSG outperformed HARP, but for low SNRs and more complex data (Type B), it performed worse than OF. On average, the method correctly tracked the tag lines, but in cases of large myocardial displacement (on the order of the distance between tag lines), there was a tendency to incorrectly put some curves in the background, characterized by low image intensities (very similar to the tag lines). Due to the coupling of neighboring

SNR	6.02 dB	8.51 dB	12.04 dB	18.06 dB
Type A (expansion)				
OF	5.70±3.62	3.17±3.01	1.08±0.56	1.02±1.66
HARP	1.06±2.21	0.57±0.40	0.43±0.34	0.36±0.32
BSG	2.90±3.17	2.56±4.46	1.21±2.05	0.61±1.52
NRR	0.65±0.38	0.43±0.24	0.29±0.18	0.18±0.11
Manual	1.37±1.04	0.84±0.48	0.69±0.43	0.55±0.36
Type A (contraction)				
OF	5.62±2.91	3.81±2.03	1.36±0.92	0.45±0.32
HARP	1.99±3.32	1.28±2.09	1.11±2.06	0.88±1.64
BSG	2.17±3.03	1.84±2.97	0.33±0.24	0.26±0.20
NRR	0.66±0.40	0.42±0.25	0.29±0.18	0.19±0.11
Manual	1.36±0.90	1.05±0.77	0.77±0.46	0.62±0.41
Type B (phantom)				
OF	2.61±2.10	1.71±1.26	1.68±0.98	1.12±0.84
HARP	1.99±3.15	1.15±8.56	0.51±0.66	0.39±0.88
BSG	3.33±3.78	1.86±2.52	1.35±2.05	1.12±1.95
NRR	0.67±0.39	0.40±0.23	0.25±0.16	0.13±0.10
Manual	1.24±0.71	0.88±0.55	0.71±0.50	0.61±0.41

Table 1. Results of tracking accuracy assessment (RMSE as a function of SNR) for Type A (expansion/contraction) and Type B data. The numbers represent RMSE ± standard deviation, given in pixel units.

	Type A (expansion)	Type A (contraction)	Type B
OF	1.27±0.85	0.80±0.66	1.46±1.19
HARP	0.58±0.60	1.64±3.08	1.17±1.79
BSG	7.42±9.31	2.60±6.96	9.63±7.99
NRR	0.18±0.11	0.19±0.14	0.34±0.27
Manual	0.70±0.44	0.62±0.39	0.64±0.38

Table 2. Results of tracking accuracy assessment (RMSE ± standard deviation) for Type A (expansion/contraction) and Type B data with time varying SNR level that models the tag fading.

control points of the B-splines and depending on the distance between the control points, such errors can affect large regions within the myocardium.

The tracking method based on NRR, which also uses a B-spline grid but with a cost function (mutual information) that utilizes all available image information, contrary to minimizing the sum of intensity values only along the tag lines (as in the case of BSG), demonstrated the highest accuracy and robustness among the considered methods, for all types of image data and all SNR levels.

Evaluation on Real Image Data. The experiments on real tMRI data were done using the in-house clinical 3T MRI scanner (GE Medical Systems). A SPAMM pulse sequence was used to acquire image data in experiments on healthy and diseased rats and pigs, and diseased human patients. Multiple short-axis view images (of size 256×256 pixels) were collected using the following imaging parameters (human-, rat-, pig-related): repetition time (6.5, 13, 4) msec, echo time (3.1, 4, 1.25) msec, flip angle (12, 7, 11) degrees, slice thickness (8, 1.6, 6) mm, spacing between slices (28, 1.6, 12) mm, pixel size (1.48×1.48, 0.19×0.19, 1.25×1.25) mm^2, frames per heart cycle (20, 24, 20), number of slice positions (3, 7, 4), tag spacing (11.5, 1.5, 6) mm, tag orientation 45° and 135°.

Five tagged MR image sequences of each type (human, rat, pig) were analyzed using HARP and NRR (two methods that demonstrated the highest performance on synthetic data).

	Human	Rat	Pig
HARP	1.65±1.73	1.36±0.93	1.43±1.35
NRR	1.23±1.02	1.33±0.84	1.12±0.84

Table 3. Results of tracking accuracy assessment (RMSE ± standard deviation) using real data from human, pig, and rat studies.

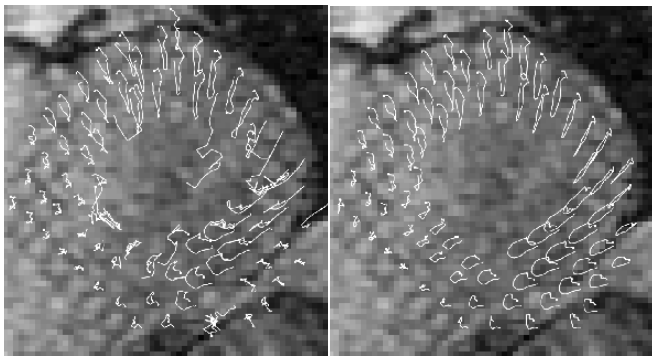

Fig. 5. Tag trajectories obtained using the harmonic phase (HARP) method (left) versus the non-rigid registration (NRR) method (right).

The algorithm parameters were fixed to the same values as in the case of the synthetic data. As no ground truth was available for the real data, we measured the accuracy in comparison with manual tracking produced by experts (10 randomly selected tags per sequence to limit the burden). The results are shown in Table 3. Typical tracking results using NRR and HARP are shown in Fig. 5.

Contrary to NRR, HARP could not cope with the poor image quality, and produced many erroneous tracks: reliable tag tracking was achieved only in the first 3-6 frames of the image sequences (where the tags had not faded yet) or in regions within the myocardium with sufficiently high SNR. The results in Table 3 do not include these extremely erroneous tracks, because of their absence in the manually annotated data (it was extremely difficult and cumbersome even for the human experts to track the myocardial boundary points). In other words, the results in the table are biased towards tags that could be tracked relatively well, and the differences between the methods would likely be much larger if all tags could have been included.

4. CONCLUSIONS

In this paper we have evaluated four conceptually different approaches to tag tracking for motion analysis in MRI, based on optical flow, harmonic phase MRI, B-snake grids, and non-rigid registration. The results of experiments on both synthetic and real data revealed that non-rigid registration yields the best tracking accuracy overall. The HARP method demonstrated good performance only in high-quality image data. The performance of the two other methods was less consistent and highly dependent on image quality. Based on these findings, our future work will focus on the extension of our previously published Bayesian tracker [16], which will use non-rigid registration as a prior model for heart dynamics. The combined method is expected to improve the robustness of the tracking and to further reduce the residual error of non-rigid registration, due to explicit and more accurate modeling of tag appearance and noise.

5. REFERENCES

[1] L. Axel and L. Dougherty, "MR imaging of motion with spatial modulation of magnetization," *Radiology*, vol. 171, pp. 841–845, 1989.

[2] M. A. Guttman, J. L. Prince, and E. R. McVeigh, "Tag and contour detection in tagged MR images of the left ventricle," *IEEE Trans. Med. Imaging*, vol. 13, no. 1, pp. 74–88, 1994.

[3] A. A. Young, D. L. Kraitchman, L. Dougherty, and L. Axel, "Tracking and finite element analysis of stripe deformation in magnetic resonance tagging," *IEEE Trans. Med. Imaging*, vol. 14, no. 3, pp. 413–421, 1995.

[4] S. Kumar and D. Goldgof, "Automatic tracking of SPAMM grid and the estimation of deformation parameters from cardiac MR images," *IEEE Trans. Med. Imaging*, vol. 13, no. 1, pp. 122–132, 1994.

[5] X. Deng and T. S. Denney, Jr., "Three-dimensional myocardial strain reconstruction from tagged MRI using a cylindrical B-spline model," *IEEE Trans. Med. Imaging*, vol. 23, no. 7, pp. 861–867, 2004.

[6] A. A. Amini, Yasheng Chen, R. W. Curwen, V. Mani, and J. Sun, "Coupled B-snake grids and constrained thin-plate splines for analysis of 2-D tissue deformations from tagged MRI," *IEEE Trans. Med. Imaging*, vol. 17, no. 3, pp. 344–356, 1998.

[7] T. S. Denney, Jr. and J. L. Prince, "Reconstruction of 3-D left ventricular motion from planar tagged cardiac MR images: An estimation theoretic approach," *IEEE Trans. Med. Imaging*, vol. 14, no. 4, pp. 625–635, 1995.

[8] N. Carranza-Herrezuelo, A. Bajo, F. Sroubek, C. Santamarta, G. Cristobal, A. Santos, and M. J. Ledesma-Carbayo, "Motion estimation of tagged cardiac magnetic resonance images using variational techniques," *Comput. Med. Imaging Graph.*, vol. 34, pp. 514–522, 2010.

[9] J. L. Prince and E. R. McVeigh, "Motion estimation from tagged MR image sequences," *IEEE Trans. Med. Imaging*, vol. 11, no. 2, pp. 238–249, 1992.

[10] N. F. Osman, W. S. Kerwin, E. R. McVeigh, and J. L. Prince, "Cardiac motion tracking using CINE harmonic phase (HARP) magnetic resonance imaging," *Magn. Reson. Med.*, vol. 42, pp. 1048–1060, 1999.

[11] M. J. Ledesma-Carbayo, J. Kybic, M. Desco, A. Santos, M. Suhling, P. Hunziker, and M. Unser, "Spatio-temporal nonrigid registration for ultrasound cardiac motion estimation," *IEEE Trans. Med. Imaging*, vol. 24, no. 9, pp. 1113–1126, 2005.

[12] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: Application to breast MR images," *IEEE Trans. Med. Imaging*, vol. 18, no. 8, pp. 712–721, 1999.

[13] S. Klein, M. Staring, and J. P. W. Pluim, "Evaluation of optimization methods for nonrigid medical image registration using mutual information and B-splines," *IEEE Trans. Image Process.*, vol. 16, no. 12, pp. 2879–2890, 2007.

[14] T. Chen, X. Wang, S. Chung, D. Metaxas, and L. Axel, "Automated 3D motion tracking using Gabor filter bank, robust point matching and deformable models," *IEEE Trans. Med. Imaging*, vol. 29, no. 1, pp. 1–11, 2010.

[15] W. R. Crum, E. Berry, J. P. Ridgway, U. M. Sivananthan, L. B. Tan, and M. A. Smith, "Simulation of two-dimensional tagged MRI," *J. Magn. Reson. Imaging*, vol. 7, pp. 416–424, 1997.

[16] I. Smal, W. J. Niessen, and E. Meijering, "Particle filtering methods for motion analysis in tagged MRI," in *Proceedings of the IEEE International Symposium on Biomedical Imaging*, 2010, pp. 488–491.