

A Novel Filtering Approach for 3D Harmonic Phase Analysis of Tagged MRI

Xiaokai Wang¹, Maureen L. Stone², Jerry L. Prince^{3,1}, Arnold D. Gomez³

¹ Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, US 21205

² Department of Neural and Pain Sciences, Department of Orthodontics, University of Maryland Dental School, Baltimore, MD, US 21201

³ Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD, US 21218

ABSTRACT

Harmonic phase analysis has been used to perform noninvasive organ motion and strain estimation using tagged magnetic resonance imaging (MRI). The filtering process, which is used to produce harmonic phase images used for tissue tracking, influences the estimation accuracy. In this work, we evaluated different filtering approaches, and propose a novel high-pass filter for volumes tagged in individual directions. Testing was done using an open benchmarking dataset and synthetic images obtained using a mechanical model. We compared estimation results from our filtering approach with results from the traditional filtering approach. Our results indicate that 1) the proposed high-pass filter outperforms the traditional filtering approach reducing error by as much as 50% and 2) the accuracy improvements are especially marked in complex deformations.

Keywords: Imaging, mechanics, brain, tongue

1. INTRODUCTION

Harmonic phase analysis (HARP) is a motion estimation method for tagged MRI that was originally developed to measure 2D cardiac motion.^{1,2} In contrast to traditional registration methods, which depend on image intensity, HARP analysis uses the phase values, which offers increased robustness to noise and fading effects.³ For this reason, phase-based methods have been used to study the motion of organs other than the heart, including the brain,⁴ and the liver.⁵ Motion estimation has also been performed in the tongue to study speech generation.⁹ Much of this previous work has been performed following similar strategies and was used in earlier cardiac work in 2D. However, analysis of 3D tongue motion, as in other non-cardiac applications, requires consideration of fundamental differences in tissue deformation profiles, how they can affect the performance of motion estimation, and what can be done to reduce potential errors.

The fundamental principle behind HARP analysis is to track the *harmonic phase* information defined at each material point in the tissue.^{1,2} Because the harmonic phase is a material property that moves with the tissue, it can be tracked by finding locations of corresponding phase values along a sequence of time frames. The basic steps of HARP analysis are shown in Fig. 1. Before harmonic phase values can be tracked, harmonic phase images are extracted from standard tagged images using filtering in the Fourier domain (k-space). The filter demodulates motion information from the original tagged image; ideally, the filtered harmonic phase image should retain all phase shifts associated with motion, while removing any interference from other sources. For this reason, tracking accuracy in HARP is related directly to the underlying motion, and the characteristics of the filter used to obtain the harmonic phase images.

By design, motion induces changes in the patterns of a tagged image. If motion is rigid, the periodic tagged pattern shifts in the direction of motion along with its harmonic phase values. In contrast, deformation results in stretched or compressed tag patterns, which changes the frequency. The amount of deformation is measured using a tensorial quantity called strain, which corresponds to the spatial gradient of the displacement field. Strain is often the end point of analysis of motion because it is related to changes in shape.

Send correspondence to Xiaokai Wang (xwang173@jhu.edu) or Arnold D. Gomez (adgomez@jhu.edu)

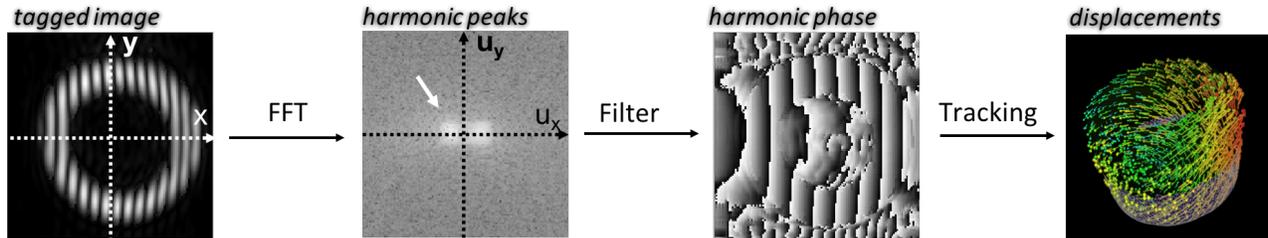


Figure 1. Basics of HARP analysis. The frequency spectrum of tagged images exhibits *harmonic peaks* (white arrow), which are isolated to obtain *harmonic phase images* for tissue tracking. Motion in 3D can be extracted from tagging in three orthogonal directions.

Traditional tags called spatial modulation of magnetization (SPAMM), have been applied as a grid.⁶ While this strategy reduces acquisition time, it produces k -space information with harmonic peaks in each grid direction, increasing the likelihood of interference between them. The conventional HARP filter is designed to reduce possible interference by applying a circular filter centered at harmonic peaks in Fourier space.¹ (Harmonic peaks arise at the tag frequency ω_{tag} due to the periodic nature of the tags). Depending on the type of images used to apply the tags, the radius of this filter can be $0.5 \omega_{\text{tag}}$ (for regular SPAMM images), or ω_{tag} (for complementary SPAMM images).^{1,2} However, this comes at the cost of reducing sensitivity to motion-induced frequency shifts beyond the filter cutoff frequency.

Alternatively, SPAMM can be applied in different volumes with a single tagging direction each. Although this strategy increases acquisition time, it eliminates the influence of other harmonic peaks from additional tagging directions and the possibility of interference from them. Thus, we hypothesize that a high-pass filter is better suited to preserve the motion information, particularly in motion fields that induce large frequency shifts.

2. CONTRIBUTIONS

Previous studies have proposed different bandpass filters for optimizing cardiac HARP in 2D.^{7,8} However, there is a lack of systematic analysis of the effects of harmonic phase image extraction on 3D displacement results using newer approaches to acquire and track tagged images.

We propose a novel approach to process tagged SPAMM images for the analysis of 3D motion. As a large portion of the literature has focused on bandpass filters, the proposed high-pass filter has been previously overlooked. However, our experiments show that a high-pass filter outperforms the traditional HARP filter in terms of tracking accuracy, especially in cases with complex and flexible deformations (e.g., tongue motion). Because the proposed filtering approach exploits the intrinsic characteristics of images independently tagged in individual directions (instead of a single image with a grid), these results also have ramifications in the design of future motion estimation experiments for measuring large 3D deformations.

3. FILTER DESIGN

This study compares the tracking performance of the three filtering approaches shown in Fig. 2. In all the experiments, the filters were applied to MRI slices (SPAMM) interpolated into a common volume to generate a volumetric time sequence using cubic spline interpolation. The first filtering approach is an extension of the traditional HARP filters^{1,2} from circles in 2D to spheres in 3D. For example, if the image is tagged in the x -direction, the corresponding filter is

$$f_{\text{SPx}}(u_x, u_y, u_z) = \begin{cases} 1, & \text{if } (u_x - \omega_{\text{tag}})^2 + u_y^2 + u_z^2 \leq \left(\frac{\omega_{\text{tag}}}{2}\right)^2; \\ 0, & \text{otherwise,} \end{cases} \quad (1)$$

where u_x , u_y , and u_z represent frequency coordinates in the k -space. The second approach comprises extended bandpass filters or ‘slab’ filters, which preserve high frequencies orthogonal to the tagging direction. For instance, in the x -direction, this would result in

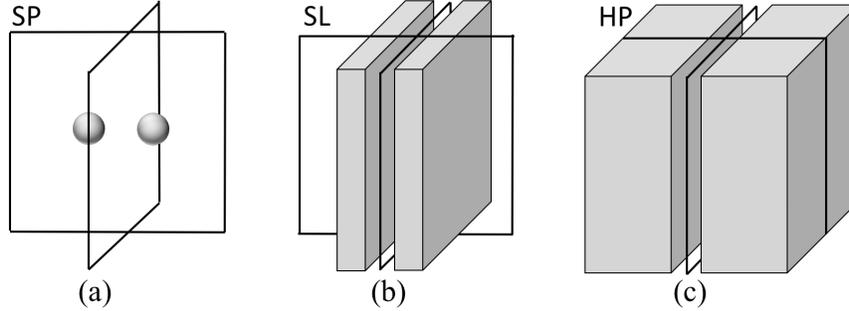


Figure 2. HARP filters. Prior to tracking, the tagged volumes were filtered using spherical (a), ‘slab’ (b), and high-pass filters (c) to extract harmonic phase images. SP: spherical filters; SL: ‘slab’ filters; HP: high-pass filters.

$$f_{\text{SL}_x}(u_x, u_y, u_z) = \begin{cases} 1, & \text{if } (u_x - \omega_{\text{tag}})^2 \leq \left(\frac{\omega_{\text{tag}}}{2}\right)^2; \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

The last approach consists of filtering the images with a high-pass filter. For the x -direction, the filter would be defined as

$$f_{\text{HP}_x}(u_x, u_y, u_z) = \begin{cases} 1, & \text{if } |u_x| \geq \frac{\omega_{\text{tag}}}{2}; \\ 0, & \text{otherwise,} \end{cases} \quad (3)$$

which would preserve frequency content starting at 50% of the first harmonic frequency and above in each orthogonal direction. Note that the above descriptions apply to one tagging direction, and additional filters (e.g., f_{HP_y} and f_{HP_z}) would be necessary to obtain harmonic phase images in the remaining directions (y and z). To reduce edge effects, the filters were smoothed using a Gaussian function of $\sigma = 0.02\omega_{\text{tag}}$ prior to application. Due to the symmetric nature of harmonic peaks, dual filters were used for the positive and negative harmonic peaks, which were combined to reduce phase errors using a method described in the literature.¹⁰

4. EXPERIMENTS

The main hypothesis was tested by performing two experiments. In both cases, tracking was performed using the phase vector incompressible registration algorithm (PVIRA),³ which is a 3D motion estimation method using a phase formulation of the iLogDemons algorithm.¹¹

4.1 Comparison Against Open Benchmarking Database

The goal of the first experiment was to evaluate the performance of the filters in terms of displacement error using experimental MRIs. The experiment used images from a phantom of cardiac motion, which is the conventional application of MR-based motion estimation.

4.1.1 Imaging Data

Tagged MRI slice sets encoded in orthogonal directions were obtained from an open database introduced in a 2011 MICCAI workshop for validation of myocardial tracking algorithms.¹² Sparse slices from a deformable, MRI-compatible phantom were interpolated into three homogeneous volumes, and filtered using the approaches described in Section 3. The 3D cardiac phantom, along with sample tagged slices is shown in Fig. 3. The dataset included eight manually tracked landmarks obtained from two expert observers. Values from the two observers were averaged to obtain a single set of tracked landmarks and used as a ground truth. The median inter-observer variability was reported to be 0.77 mm.¹²

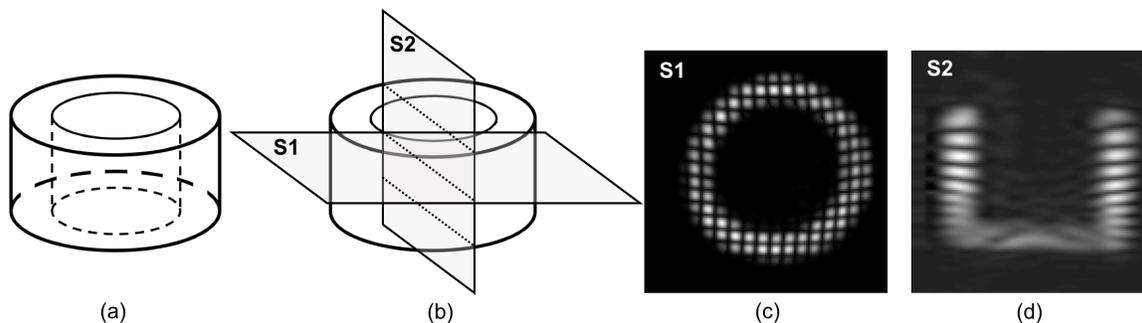


Figure 3. Open access 3D cardiac phantom: (a) is 3D view for the experimental cardiac phantom. S1 (c) and S2 (d) are two orthogonal slices from the phantom (b). For visualization purpose, the vertical tags and the horizontal tags are combined in a single image (c).

4.1.2 Characterization of Tracking Error

Tracking accuracy was measured by comparing the Euclidean distance between the ground truth, and the landmarks evaluated using the displacement fields obtained with each of the filters. Descriptive statistics (including the median error, and error range) were calculated across all eight landmarks and 20 time points. The database was designed with a focus on displacements, and no strain computation was performed given the sparsity of the landmarks.

4.2 Analysis of Simulated Tongue Motion

This experiment focused on analysis of tongue motion, which produces fundamentally different deformation fields than that of the heart.¹² The goal of the experiment was to evaluate the different filters in terms of displacement *and* strain error. To this end, synthetic tagged images of different motion fields were generated using a finite-element model, which has the benefit of producing dense displacement and strain values that can be used as a ground truth. This benchmarking approach has been previously described in the literature.¹³

4.2.1 Finite-Element Simulations

The finite-element mechanical model of the tongue appears in Fig. 4. The model included the mandible, the hyoid bone, and the tongue, which consisted of muscular compartments representing 13 muscles. It was constructed using 256 quadratic hexahedral elements for the deformable tongue, and 3000 linear quadratic elements to represent rigid bones. Material parameters were extracted from the literature.^{13,14} Simulations were generated using FEBio Software,¹⁵ which was set to ramp up contractions producing 2–11 time frames per simulation depending on the amount of deformation.

Motion was produced by assigning active contraction of the muscular compartments according to previous numerical studies,^{13,16} along with manual tuning of the model. The activation intensity was expressed as a percentage of the maximum sarcomeric activation assumed to be 35kPa. A total of five simulations were produced to address speech generation and to approximate rigid motion. The simulations (shown in Fig. 5) included:

1. Semi-rigid tongue motion to simulate minimal tongue deformation (mandible rotation by -3.44°);
2. /s/ as the sound of the letter ‘s’ (1.8% activation of GG, 3.5% activation of SL, 9% activation of T, 9% activation of V, 1.8% activation of GH, mandible Rotation by -0.40°);
3. /k/ as the sound of the letter ‘k’ (30% activation of IL, 30% activation of HG, 80% activation of SG);
4. /a/ as in cat (3% activation of GG, 54% activation of HG, 3% activation of SG, mandible Rotation by -1.38°);
5. /e/ as in tea (7.2% activation of SL, 0.6% activation of T, 60% activation of V, 3% activation of HG, 30% activation of SG);

4.2.2 Generation of Synthetic Data

Nodal displacements from the mechanical model were interpolated onto an imaging grid with $0.7813 \text{ mm} \times 0.6185 \text{ mm} \times 0.7813 \text{ mm}$ resolution. Both Lagrangian and Eulerian displacements were obtained using the first time frame as the reference configuration. The Eulerian displacements were used to deform an atlas T1 image of the human tongue.¹⁷ Prior to deformation, synthetic SPAMM was applied using a tag frequency of 15 mm, which is similar to existing in vivo studies.³ The Lagrangian displacement was used to generate ground truth measurements for displacement and strain. The Green-Lagrange strain tensor¹⁸ was calculated in the imaging grid using finite difference.

4.2.3 Characterization of Strain Error

The images were filtered and tracked using the methods described above. After obtaining displacement results, the Euclidean distance to the ground truth was calculated in each field. Displacement error was defined as the median Euclidean distance within the tongue at a given time frame. Strains were calculated from the results obtained using each of the filters, and compared to the ground truth. Comparisons were based on the difference between the median of the shearing strain γ_{med} from the ground truth and each of the test fields. The scalar quantity γ_{med} was defined as the median of the difference between the first and the third eigenvalue of the strain tensor across each material location.¹⁸ Error was quantified as the spatial median of the difference at a given time frame. To better understand the relationship between deformation and motion estimation performance, error values were compared to γ_{med} via scatter plots.

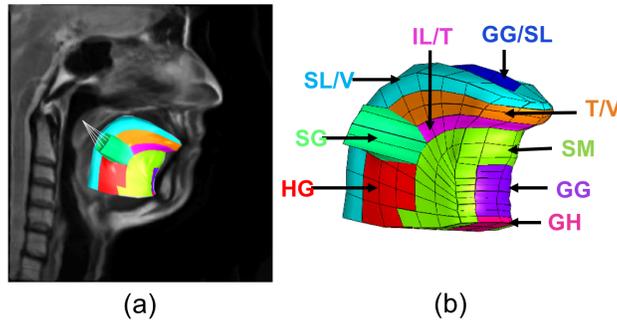


Figure 4. Synthetic FE model of tongue: The synthetic model is used to simulate the muscle activation of the human tongue. According to the tongue anatomy, essential muscles are marked in (b): SL = superior longitudinal; V = verticalis; SG = styloglossus; HG = hyoglossus; GG = genioglossus; GH = geniohyoid; IL = inferior longitudinal; T = transverse; SM = surrounding tissues.

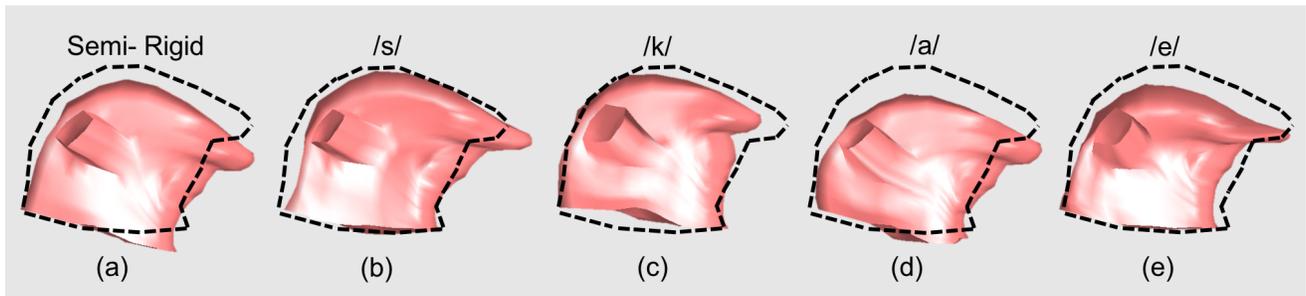


Figure 5. Simulated configurations with increasing complexity by activating different muscles to have semi-rigid motion (a), to pronounce /s/ (b), /k/ (c), /a/ (d), /e/ (e). The black dotted outline represents the original configuration when there is no deformation.

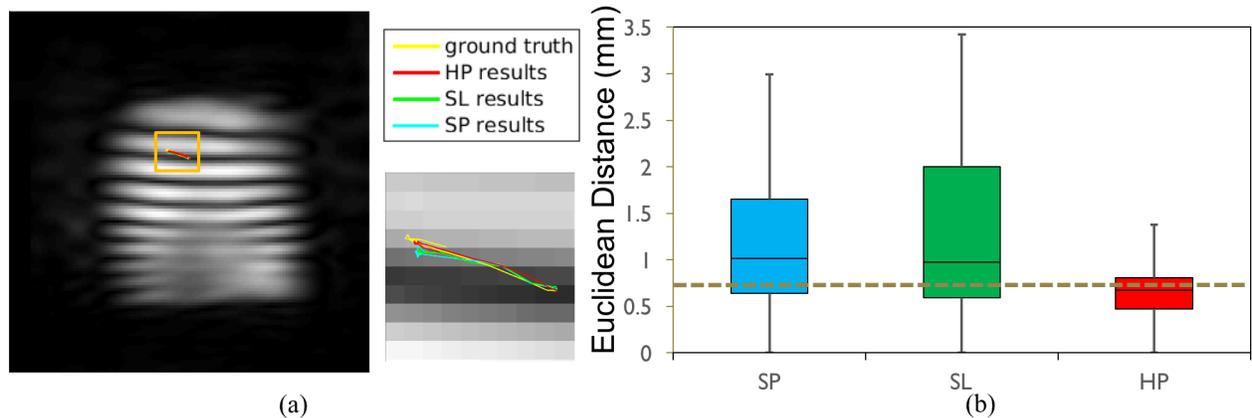


Figure 6. Results for STACOM phantom datasets: (a) 3D Tracking results for one landmark, (b) whole range box-plots of tracking errors. The gray dashed line represents the inter-observer variability.

5. RESULTS AND DISCUSSION

5.1 Comparison Against Open Benchmarking Database

Tracking of one of the eight landmarks appears in Fig. 6(a). The median Euclidean distance (error) across the first 20 time frames was 1.01 ± 0.72 mm when using the traditional HARP filter, 0.97 ± 0.88 mm using the ‘slab’ filter, and 0.67 ± 0.28 mm using the high-pass filter—an improvement over the traditional approach. The tracking trajectory (Fig. 6(a)), shows that the high-pass filter tracks more closely to the ground truth across all time frames. The error distribution also shows that the high-pass approach results in less scatter in error values, as shown in Fig. 6(b).

5.2 Analysis of Simulated Tongue Motion

The displacement error across five simulated cases at the last time frame using the different filters appears in Table 1. The general trend is that the high-pass filter outperforms the others across the different motions. The high-pass filter yielded the largest error (median) in */e/* at 0.162 mm, and the smallest in semirigid motion, at 0.090 mm. In contrast to the general trend, the spherical filter performed less accurately with errors as large as 0.362 mm in the */e/* simulation, and as low as 0.143 mm in the semirigid motion simulation. The slab filter’s performance appeared to be between the extremes. The differences between the filters varied depending on the simulations, the lowest difference was observed in the semirigid simulation, and the largest on the */e/* simulation. Magnitude color maps of the ground truth and the approximated displacements using different filters appears in Fig. 7, and shows the relative better performance of the high-pass filter (Fig. 7(c)), which has the highest resemblance to the ground truth (Fig. 7(a)).

Strain error at the last time frame for each of the simulations appears in Fig. 8(a). Just as in the analysis of displacements, the trend is similar; the high-pass filtering approach yields lower error and less spread. Unlike the analysis of displacements, strain calculation offers an explanation for the difference between the filters across different simulations: Figure 8(b) shows a positive correlation between strain difference and the magnitude of strain (per γ_{med}). As may be expected, higher strains, which are the spatial gradient of deformation, result in larger frequency alterations in the tagging pattern. Given the nature of frequency modulation,⁹ large frequency

Table 1. Displacements error (mm) for the 5 simulated cases

	Semi-Rigid Motion	<i>/s/</i>	<i>/k/</i>	<i>/a/</i>	<i>/e/</i>
SP	0.143 ± 0.123	0.307 ± 0.211	0.330 ± 0.240	0.334 ± 0.230	0.362 ± 0.500
SL	0.108 ± 0.125	0.190 ± 0.163	0.184 ± 0.175	0.203 ± 0.195	0.224 ± 0.368
HP	0.090 ± 0.105	0.113 ± 0.141	0.137 ± 0.144	0.148 ± 0.164	0.162 ± 0.255

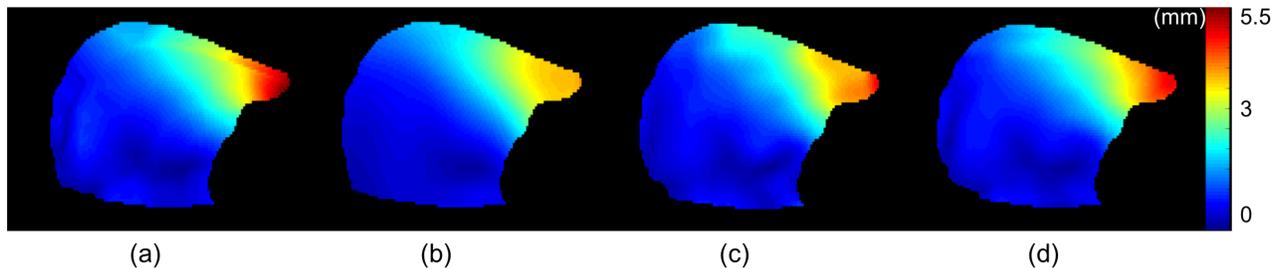


Figure 7. Displacements estimation results at the last time frame: (a) Total Lagrange Displacements map of ground truth; Estimated Lagrange Displacements map using spherical filter in (b), using ‘slab’ filter in (c) and using high-pass filter in (d).

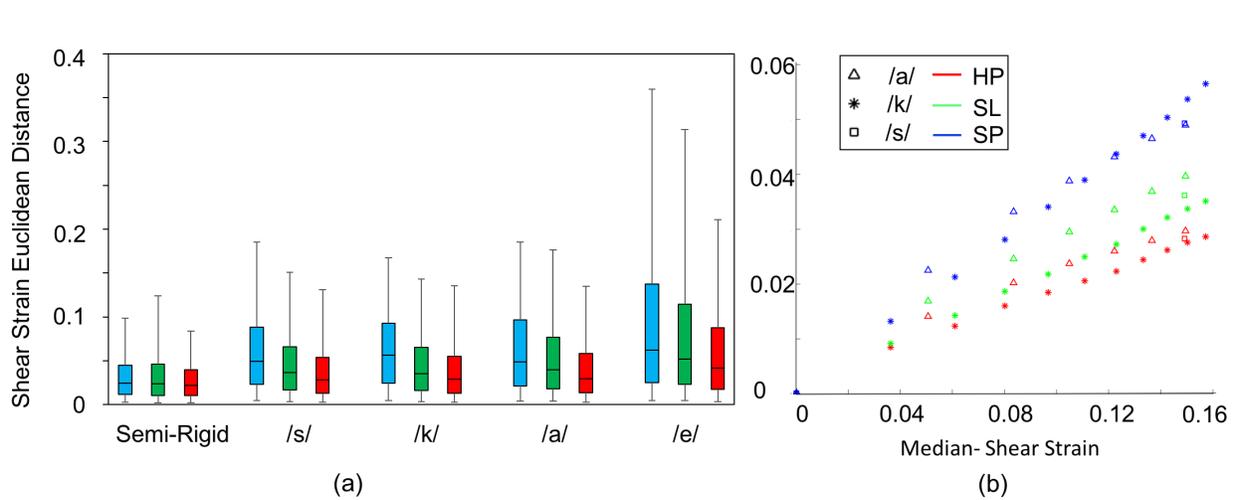


Figure 8. Strain estimation results: (a) is the strain estimation errors for five simulations at the last time frame shown in 5% to 95% boxplot. (b) describes the relationship between strain estimation error and median of shear strain for /a/, /k/, and /s/.

alterations in tagged pattern will result in broader spectral spread from the harmonic peak. Thus, information from regions with larger spread are more likely to fall outside the filter’s cut off range. As hypothesized, the high-pass filter preserves this information, resulting in better tracking performance.

Our results show that displacements and strains within the feasible range in speech generation can result in frequency components that can be outside standard HARP filtering. One way to improve performance, if the scanning time allows it, is to acquire separate images with independent tagging directions. This practice would enable the use of the presented filter. One possible downside may be an increased sensitivity to noise, which will be the subject of future research.

6. CONCLUSION

This work presented a method to extract harmonic phase images that improves motion estimation accuracy. The high-pass method yielded error improvement of as much as 50% and the improvement appeared to be more apparent in complex motion patterns (Fig. 7). When the displacement field is relatively simple, the modulated frequency components accumulate around the first harmonic frequency peak near the band-pass region. However, more complex displacements field leads to more modulations and spreads of the frequency components, where some higher frequency components are discarded by the band-pass filters but maintained by the high-pass approach. Potential weaknesses of our filter are the sensitivity to noise in the acquired images and the requirement for longer image acquisition time to get tagged images in individual directions.

7. ACKNOWLEDGMENTS

The authors would like to thank Aaron Carass for manuscript preparation assistance. Funding is provided by grants R01DC014717 and R01NS055951 from the National Institutes of Health in the United States.

REFERENCES

1. Osman NF, Kerwin WS, McVeigh ER, Prince JL Cardiac motion tracking using CINE harmonic phase (HARP) magnetic resonance imaging. (1999) *Magnetic Resonance in Medicine* 42(6):1048-1060.
2. Osman NF, McVeigh ER, Prince JL Imaging heart motion using harmonic phase MRI. (2000) *IEEE Trans Med Imaging* 19(3):186-202.
3. Xing F, Woo J, Gomez AD, Pham DL, Bayly PV, Stone M, Prince JL Phase vector incompressible registration algorithm (PVIRA) for motion estimation from tagged magnetic resonance images. (2017) *IEEE Trans Med Imaging*:1-13.
4. Feng Y, Abney TM, Okamoto RJ, Pless RB, Genin GM, Bayly PV Relative brain displacement and deformation during constrained mild frontal head impact. (2010) *J Roy Soc Interface* 7(53):1677-1688.
5. Mannelli L, Wilson GJ, Dubinsky TJ, Potter CA, Bhargava P, Cuevas C, Linnau KF, Kolokythas O, Gunn ML, Maki JH Assessment of the liver strain among cirrhotic and normal livers using tagged MRI. (2012) *J of Magnetic Resonance Imaging* 36(6):1522-2586.
6. Moore CC, McVeigh ER, Zerhouni EA Quantitative Tagged Magnetic Resonance Imaging of the Normal Human Left Ventricle. (2000) *Topics in magnetic resonance imaging* 11(6):359-371.
7. Davis CA, Li J, Denney TS Jr Analysis of spectral changes and filter design in tagged cardiac MRI. (2006) *IEEE ISBI: From Nano to Macro*:137-140.
8. Marinelli M, Positano V, Osman NF, Recchia FA, Lombardi M, Landini L Automatic filter design in HARP analysis of tagged magnetic resonance images. (2008) *IEEE ISBI: From Nano to Macro*:1429-1432.
9. Parthasarathy V, Prince JL, Stone M, Murano ZE, NessAiver M Measuring tongue motion from tagged cine-MRI using harmonic phase (HARP) processing. (2007) *J of the Acoustical Society of America* 121(1):491-504.
10. Ryf S, Tsao J, Schwitter J, Stuessi A, Boesiger P Peak-combination HARP: a method to correct for phase errors in HARP. (2004) *J of Magnetic Resonance Imaging* 20(5):874-880.
11. Mansi T, Pennec X, Sermesant M, Delingette H, Ayache N iLogdemons: a demons-based registration algorithm for tracking incompressible elastic biological tissues. (2011) *Int J Comput Vis* 92(1):92-111.
12. Tobon-Gomez C, De Craene M, McLeod K, Tautz L, Shi W, Hennemuth A, Prakosa A, Wang H, Carr-White G, Kapetanakis S, Lutz A, Rasche V, Schaeffter T, Butakoff C, Friman O, Mansi T, Sermesant M, Zhuang X, Ourselin S, Peitgen HO, Pennec X, Razavi R, Rueckert D, Frangi AF, Rhode KS Benchmarking framework for myocardial tracking and deformation algorithms: an open access database. (2013) *Med Image Anal* 17(6):632-648.
13. Ramsey J, Prince JL, Gomez AD Test suite for image-based motion estimation of the brain and tongue. (2017) *Proc. SPIE* 10137:1-8.
14. Stavness I, Lloyd JE, Fels S Automatic prediction of tongue muscle activations using a finite element model. (2012) *J of Biomechanics* 45(16):2841-2848.
15. Maas SA, Ellis BJ, Ateshian GA, Weiss JA FEBio: Finite Elements for Biomechanics. (2012) *J of Biomechanical Engineering* 134(1):1-10.
16. Perrier P, Payan Y, Zandipour M, Perkell J Influences of tongue biomechanics on speech movements during the production of velar stop consonants: A modeling study. (2003) *J of the Acoustical Society of America* 114(3):1582-1599.
17. Woo J, Xing F, Lee J, Stone M, Prince JL A spatio-temporal atlas and statistical model of the tongue during speech from cine-MRI. (2016) *Computer Methods in Biomechanics and Biomedical Engineering : Imaging & Visualization*: 1-12.
18. Spencer AJM Continuum mechanics. (1980) *Longman Group UK Limited, Essex, England*