An Image-based method for phase estimation, gating and temporal super-resolution of cardiac ultrasound

*

*

Abstract. In this paper, we present a novel purely image-based method for the estimation of instantaneous cardiac and respiratory phases from cardiac ultrasound videos and demonstrate its utility in gating and temporal super-resolution. We first use a novel strategy based on inter-frame similarity to transform the complex high-dimensional image sequence into a 1D time series with the same periodicity characteristics. Next, we use a trend extraction technique to decompose this time series into a trend component encoding respiratory motion alone and a residual component encoding the beating heart motion alone. Next, we use the Hilbert transform to estimate the instantaneous cardiac and respiratory phases of each frame from these components. Next, we present a robust non-parametric regression technique that uses these phase estimates to gate out video frames predominantly influenced by respiratory motion. Lastly, we use a kernel regression with a novel kernel to learn an image manifold parameterized by instantaneous cardiac phase that can be sampled to generate a single-cycle video at a higher temporal-resolution. We demonstrate and validate our methods using cardiac ultrasound videos and associated ECG recordings of 6 mice.

1 Introduction

Ultrasound is emerging as an increasingly effective tool for rapid non-invasive visual assessment of cardiac structure and function.

2 Method

In this section, we present the theory underlying the proposed methods for instantaneous phase estimation, gating and temporal super-resolution of cardiac ultrasound videos. Specifically, we begin by describing our method for the estimation of instantaneous cardiac and respiratory phases in Section 2.1. We then present a robust method that uses these phase estimates to gate out video frames predominantly influenced by respiratory motion in Section 2.2. Finally, in Section 2.3, we present a kernel regression method that uses the phase-tagged image sequence to reconstruct a single-period video at higher-temporal resolution.

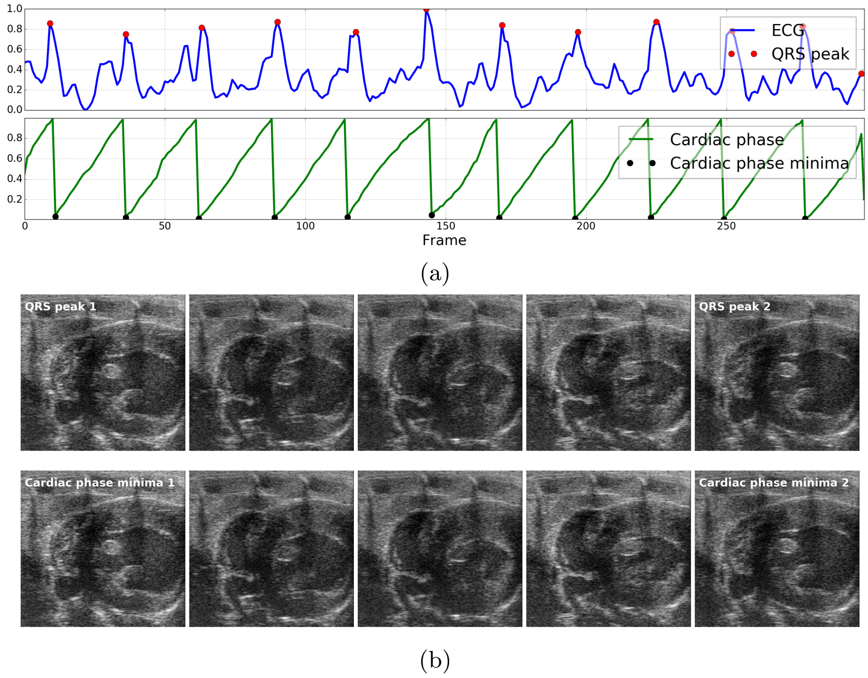


Fig. 1. Illustration of correspondence between ECG and the instantaneous cardiac phase estimated using the proposed method: (a) ECG signal (blue) and the cardiac phase estimated using our method overlaid with peaks of the QRS complex (red circle) and the corresponding cardiac phase minima (black circle), (b) Five video frames evenly spaced in time between the first and second QRS peaks of the ECG signal (top-row) and the corresponding cardiac phase minima (bottom-row) constituting one cardiac cycle.

2.1 Estimation of instantaneous cardiac and respiratory phases

While there have been numerous efforts for the estimation of instantaneous phase and/or frequency in periodic univariate time series data (wherein a single-variable is measured at each time point)[3, 6, 7], there are not many methods that tackle this problem in a multivariate setting which is the case with periodic cardiac ultrasound videos wherein several thousands of variables (pixel intensities) could be measured at each time point. Our strategy was to find a way to simplify this complex multi-variate problem by transforming it into a univariate one and take advantage of existing univariate methods to solve the problem.

We first compute the similarity between all pairs of frames in the given periodic image sequence containing N images/frames to create a matrix $A \in R^{N \times N}$ where in the element $A(i, j)$ is equal to the similarity between the i^{th} and j^{th} frames of the given sequence. Here in, we use normalization correlation also known as the Pearson correlation coefficient to quantify inter-frame similarity but, in principle, any of the similarity metrics used in image registration al-

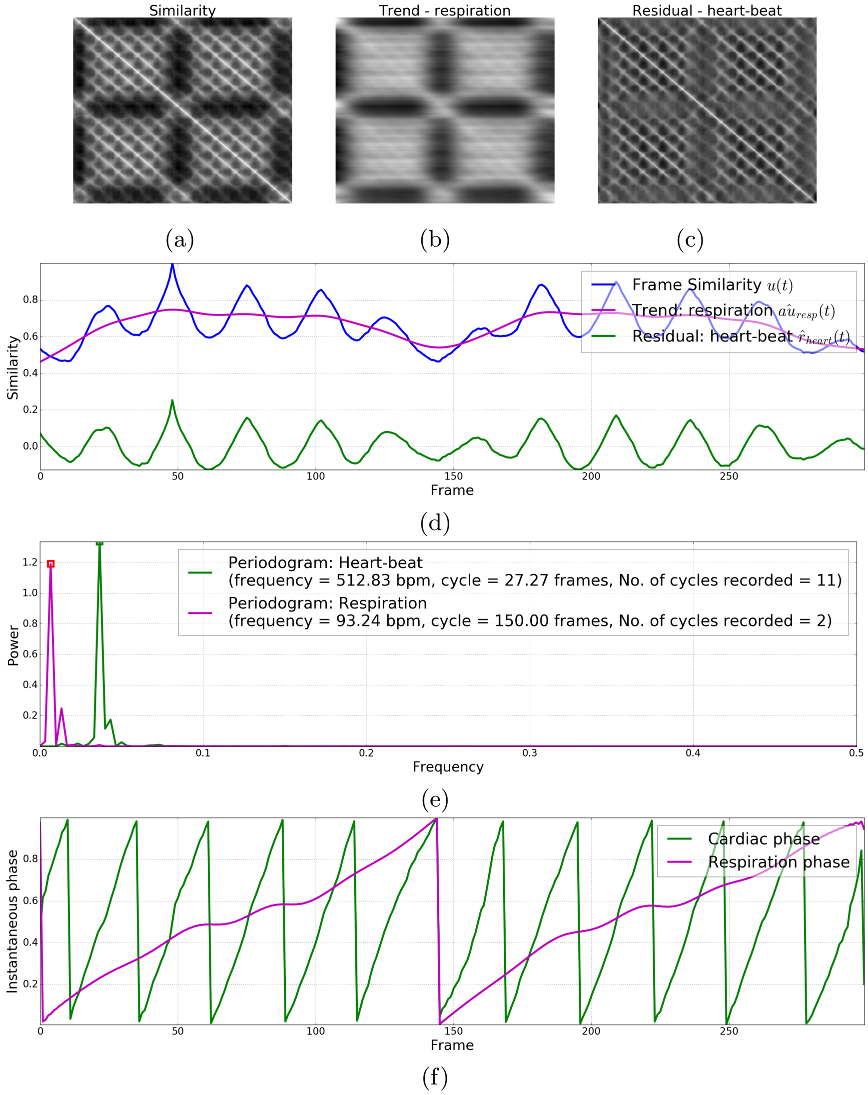


Fig. 2. Illustration of our phase estimation method: (a) Inter-frame similarity matrix, (b) Trend component matrix corresponding to respiratory motion, (c) Residual component matrix corresponding to beating heart motion, (d) Similarity profile chosen for phase estimation with associated trend/respiration and residual/heart-beat components, (e) Periodogram of heart-beat and respiration components along with periodicity characteristics (e.g. frequency, cycle duration) calculated for the peak frequency, and (f) Instantaneous cardiac and respiratory phases estimated using the Hilbert transform.

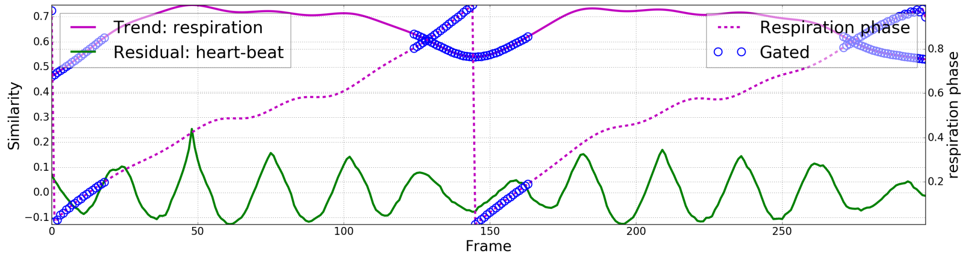
gorithms [4] can potentially be chosen. Each row in the inter-frame similarity matrix A can now be seen as a univariate time series that measures the similarity of the corresponding frame in the image sequence with all other frames. If the similarity metric is chosen with care and if the corresponding frame is not significantly corrupted we can expect this time series to encompass the predominant periodicity characteristics of the original image sequence. Two kinds of periodic motion are present in cardiac ultrasound videos, one corresponding to the beating heart motion and the other corresponding to respiratory motion. Figure 2(a) shows the inter-frame similarity matrix of one of our cardiac ultrasound videos computed using the normalized correlation metric wherein these two sources of periodicity can be clearly observed in many rows/columns.

Next, we use a trend extraction technique [1] called the Hodrick-Prescott (HP) filter [5], widely used in econometrics, to decompose the frame similarity signal $u(t)$ corresponding to each row of A into a sum of two components: (i) Trend component $\tau_{resp}(t)$ encompassing the periodicity characteristic of only the lower-frequency respiratory motion, and (ii) Residual component $r_{heart}(t)$ encompassing the periodicity characteristic of only the higher-frequency beating heart motion. The HP filter performs the decomposition of $u(t) = \tau_{resp}(t) + r_{heart}(t)$ by solving the following optimization problem:

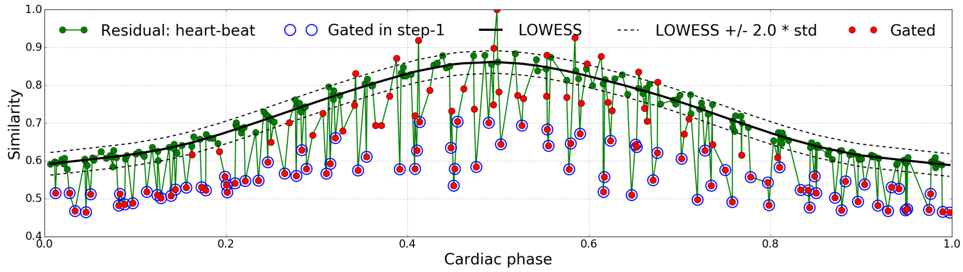
$$\arg\min_{\tau(t)} \left[\sum_{t=1}^{N} \left(u(t) - \tau_{resp}(t)\right)^2 + \lambda \sum_{t=1}^{N-1} \left(\nabla^2 \tau_{resp}(t)\right)^2 \right] \tag{1}$$

where $\nabla^2 \tau_{resp}(t) = \tau_{resp}(t+1) - 2\tau_{resp}(t) + \tau_{resp}(t-1)$ is the second-order difference or derivative that penalizes curvature of the trend signal. We use a ridge-regression implementation of the HP filter [8] with $\lambda = 6400$. Let A_{resp} and A_{heart} be the matrices whose rows contain the trend/respiratory and residual/heart-beat components, respectively, of the frame similarity signals in the corresponding rows of the similarity matrix A. Figures 2(b,c) show the trend $A_{res[}$ and residual A_{heart} matrices for one of our cardiac ultrasound videos. To minimize the effect of any noise on phase estimation, we pick frame similarity signal corresponding to the row of A_{heart} whose periodogram or power-frequency distribution has minimum entropy. Let $\hat{u}(t)$ be the chosen frame similarity signal and let $\hat{\tau}_{resp}(t)$ and $\hat{r}_{heart}(t)$ be its associated trend/respiration and residual/heart-beat component signals, respectively. Figure 2(d) shows the selected frame similarity signal (blue) along with the associated trend/respiration (pink) and residual/heart-beat (green) components. Figure 2(e) shows the periodogram of these two components that are nicely segregated in frequency domain. The periodicity characteristics (e.g. frequency in beats-per-minute (bpm), cycle duration in frames, number of cycles) of heart-beat and respiration can be calculated based on the maximum-power frequency of their periodograms.

Next, considering the narrow-band nature (see Figure 2) of the derived trend/respiratory and residual/heart-beat signals, we use the Hilbert transform [6] to estimate the instantaneous phase of each frame. Specifically, we compute the instantaneous intra-period phase $\phi(t) \in [-\pi, \pi)$ of the periodic time series $x(t)$ using its Hilbert transform $H_x(t)$ as follows: $\phi(t) = arctan\left(\frac{H_x(t)}{x(t)}\right)$ and map $\phi(t)$ from the range $[-\pi, \pi)$ to $[0, 1)$. Let $\hat{\phi}_{heart}(t)$ and $\hat{\phi}_{resp}(t)$ denote the in-



(a)



(b)

Fig. 3. Illustration of our two-step method for gating out respiratory frames: (a) Frames F_{cutoff} discarded (blue circles) in the first step overlaid with the trend/respiration, residual/heart-beat, and instantaneous respiratory phase signals (b) residual/heart-beat signal $\hat{r}_{heart}(t)$ vs cardiac phase $\hat{\phi}_{heart}(t)$ overlaid with frames F_{cutoff} discarded in step-1 (blue circle), LOWESS fit $L(\phi)$ (black solid), upper and lower bounds or 95% confidence interval $(L(\phi) \pm 2.0 * \hat{\sigma}_L)$ of valid non-respiratory frames (black dotted), and frames F_{resp} (red circles) gated at the end of step-2.

stantaneous cardiac and respiratory phases computed from the trend/respiration $\hat{\tau}_{resp}(t)$ and residual/heart-beat $\hat{r}_{heart}(t)$ signals, respectively. Figure 2(f) shows the instantaneous cardiac and respiratory phases computed from the heart-beat and respiration signals in Figure 2(d).

2.2 Gating out respiratory frames

Once the cardiac and respiratory phases of each frame has been estimated, it can be used to guide the extraction of quantitative measurements to a desired part/point of the periodic cycle, a process commonly referred to as gating. In this section, we present a robust two-step method that uses these phase estimates to gate out video frames predominantly influenced by respiratory motion.

In the first step, based on the observation that the predominant influence of respiratory motion occurs around the troughs/minima of the trend/respiration signal $\hat{\tau}_{resp}(t)$ (see pink curve in Fig. 2(d)) corresponding to the respiratory phase $\hat{\phi}_{resp}(t) = 0$, we perform a rough initial gating by discarding the frames $F_{cutoff} = \left\{ t \mid \hat{\phi}_{resp}(t) < c \parallel \hat{\phi}_{resp}(t) > (1 - c) \right\}$ whose respiratory phase distance from $\hat{\phi}_{resp}(t) = 0$ is below a specified cutoff value c that we set to 0.2. Figure 3(a) shows the discarded frames F_{cutoff} overlaid (blue circles) on the trend/respiration $\hat{r}_{resp}(t)$, residual/heart-beat $\hat{r}_{heart}(t)$, and respiratory phase $\hat{\phi}_{resp}(t)$ signals.

In the second step, we learn a robust mapping $L(\phi) : [0, 1) \rightarrow R$ from cardiac phase to the residual/heart-beat component of the frame similarity signal by fitting a robust non-parametric regression model called Locally weighted regression (LOWESS) to the dataset $\left\{ \left(\hat{\phi}_{heart}(t), \hat{r}_{heart}(t) \right) \mid t \notin F_{cutoff} \right\}$ containing the pair of the cardiac phase $\hat{\phi}_{heart}(t)$ and residual/heart-beat $\hat{r}_{heart}(t)$ signal values of all frames that do not belong to the set of frames F_{cutoff} discarded in the first step above. Next, we compute a robust estimate of the standard deviation/dispersion $\hat{\sigma}_L$ of non-respiratory frames around the LOWESS fit based on the median absolute value of the difference between the LOWESS fit $L(\hat{\phi}_{heart}(t))$ and residual/heart-beat signal $\hat{r}_{heart}(t)$ for all frames. Lastly, we gate out all frames $F_{resp} = \left\{ t \mid |L(\hat{\phi}_{heart}(t)) - \hat{r}_{heart}(t)| > k \times \hat{\sigma}_L \right\}$ for which the absolute difference between the LOWESS fit and heart-beat signal value is greater than $k = 0.2$ times the standard deviation $\hat{\sigma}_L$.

2.3 Model to generate images by cardiac phase for super-resolution

In this section, we present a kernel regression model to reconstruct the image at any cardiac phase. This model can then be used to generate a single-cycle video representative of the subject's heart-beat at a desired higher temporal resolution to achieve temporal super-resolution.

Given a cardiac ultrasound video $I(t) : \{1, ..., N\} \rightarrow R^m$ of N images with m pixels each, we extract the residual/heart-beat component $\hat{r}_{heart}(t)$ of the frame similarity signal, estimate the instantaneous cardiac phase $\hat{\phi}_{heart}(t)$ of each frame, and compute the robust LOWESS fit $L(\phi) : [0, 1) \rightarrow R$ that maps cardiac phase to the heart-beat signal as described in Sections 2.1 and 2.2. We then use Nadarya-Watson kernel regression[2] to learn a phase-parametrized image manifold in the form of a function $M(\phi) : [0, 1) \rightarrow R^m$ that maps any cardiac phase ϕ to an image using kernel-weighted local average as follows:

$$M(\phi) = \frac{\sum_{t=1}^{N} K\left(\phi, \hat{\phi}_{heart}(t)\right) I(t)}{\sum_{t=1}^{n} K\left(\phi, \hat{\phi}_{heart}(t)\right)} \tag{2}$$

where kernel $K\left(\phi, \hat{\phi}_{heart}(t)\right) = exp\left\{-\frac{|\phi - \hat{\phi}_{heart}(t)|^2}{2\sigma_{\phi}^2}\right\} \times exp\left\{-\frac{|L(\phi) - \hat{r}_{heart}(t)|^2}{2\sigma_L^2}\right\}$ is defined as the product of two radial-basis function (RBF) kernels. The first RBF kernel weighs images inversely proportional to their distance (accounting for periodicity) in the cardiac phase space. Its bandwidth σ_{ϕ} is set equal to a constant $k_{\phi} = 0.4$ times the median difference in cardiac phase between consecutive frames of the given image sequence. The second RBF kernel weighs images inversely proportional to their deviation from the LOWESS prediction $L(\phi)$ in the residual/heart-beat signal space. Its bandwidth σ_L is set equal to a constant $k_L = 2.0$ times the robust estimate of standard deviation/dispersion $\hat{\sigma}_L$ of non-respiratory frames around the LOWESS fit as described in Section 2.2.

A representative single-cycle video of the subject's heart-beat can now be reconstructed by sampling the manifold $M(\phi)$ at any desired resolution from the range $[0, 1)$ of the cardiac phase. Note that this approach will facilitate temporal super-resolution only if the set of images observed in each distinct period of the given image sequence differ in cardiac phase by some non-zero amount. Also, the amount of temporal super-resolution possible will depend both on the number of cycles captured and the amount by which one can expect the corresponding images to be perturbed in phase across cycles.

3 Results

VID	R^2	Mean ± Std	Median ± IQR
1	0.9995	1.55 ± 0.89	1.00 ± 1.00
2	0.9995	1.36 ± 1.15	1.00 ± 2.00
3	0.9996	1.45 ± 0.89	2.00 ± 1.00
4	0.9998	0.73 ± 0.62	1.00 ± 1.00
5	0.9990	2.09 ± 1.78	1.00 ± 2.00
6	0.9992	1.40 ± 1.74	1.00 ± 0.75

Table 1. My caption

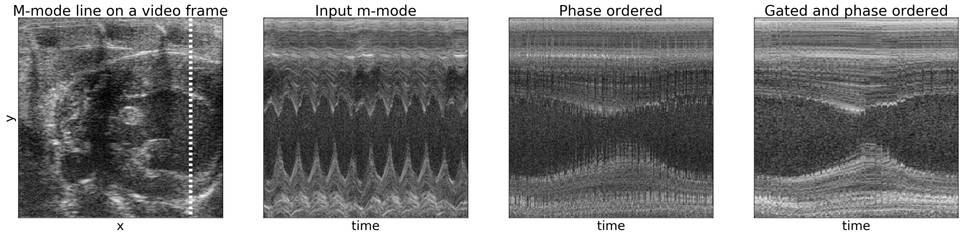
VID	LOOCV (50) NCORR Mean ± Std	QRS Peak Frame NCORR Mean ± Std
1	0.8562 ± 0.0366	0.7981 ± 0.0532
2	0.8558 ± 0.0525	0.7877 ± 0.0503
3	0.8592 ± 0.0348	0.7849 ± 0.0457
4	0.8713 ± 0.0348	0.8395 ± 0.0486
5	0.8655 ± 0.0312	0.8196 ± 0.0476
6	0.8736 ± 0.0278	0.8270 ± 0.0778

Table 2. My caption

4 Discussion and future work

References

1. Alexandrov, T., Bianconcini, S., Dagum, E.B., Maass, P., McElroy, T.S.: A review of some modern approaches to the problem of trend extraction. Econometric Reviews 31(6), 593–624 (2012)



(a)

Fig. 4. Visualization of m-mode frames ordered by estimated cardiac phase

2. Bishop, C.M.: Pattern Recognition and Machine Learning, vol. 4 (2006)
3. Boashash, B.: Estimating and interpreting the instantaneous frequency of a signal. ii. algorithms and applications. Proceedings of the IEEE 80(4), 540–568 (Apr 1992)
4. Goshtasby, A.: Similarity and Dissimilarity Measures. In: Image Registration SE - 2, pp. 7–66. Advances in Computer Vision and Pattern Recognition, Springer London (2012)
5. Hodrick, R., Prescott, E.: Postwar u.s. business cycles: An empirical investigation. Journal of Money, Credit and Banking 29(1), 1–16 (1997), `http://EconPapers.repec.org/RePEc:mcb:jmoncb:v:29:y:1997:i:1:p:1-16`
6. Lu, W.k., Zhang, C.K.: Robust estimation of instantaneous phase using a time-frequency adaptive filter. GEOPHYSICS 78(1), O1–O7 (2013)
7. Luo, Y., Al-Dossary, S., Marhoon, M., Alfaraj, M.: Generalized hilbert transform and its applications in geophysics. The Leading Edge 22(3), 198–202 (2003)
8. Yamada, H.: Ridge regression representations of the generalized hodrick-prescott filter. JOURNAL OF THE JAPAN STATISTICAL SOCIETY 45(2), 121–128 (2015)