

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

\*

\*

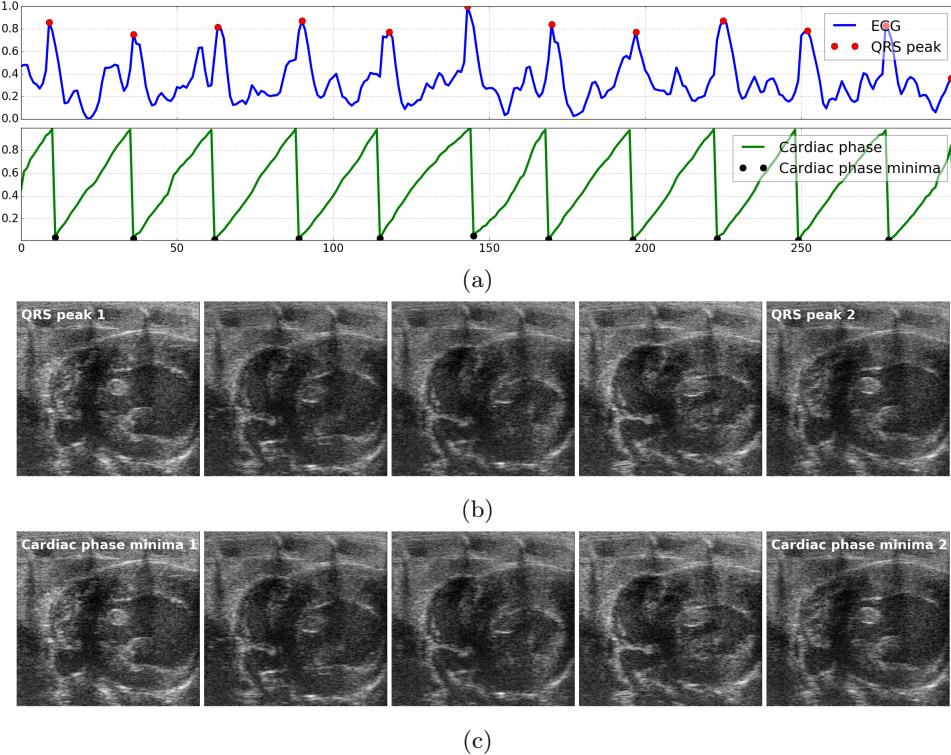
**Abstract.** In this paper, we present a novel image-based method for the estimation of instantaneous cardiac and respiratory phases from cardiac ultrasound videos and demonstrate its utility in gating and temporal super-resolution. We first use a novel strategy based on inter-frame similarity to transform the complex high-dimensional periodic cardiac ultrasound 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 along. Next, we use the Hilbert transform to estimate the instantaneous cardiac and respiratory phases of each frame from these signals. 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

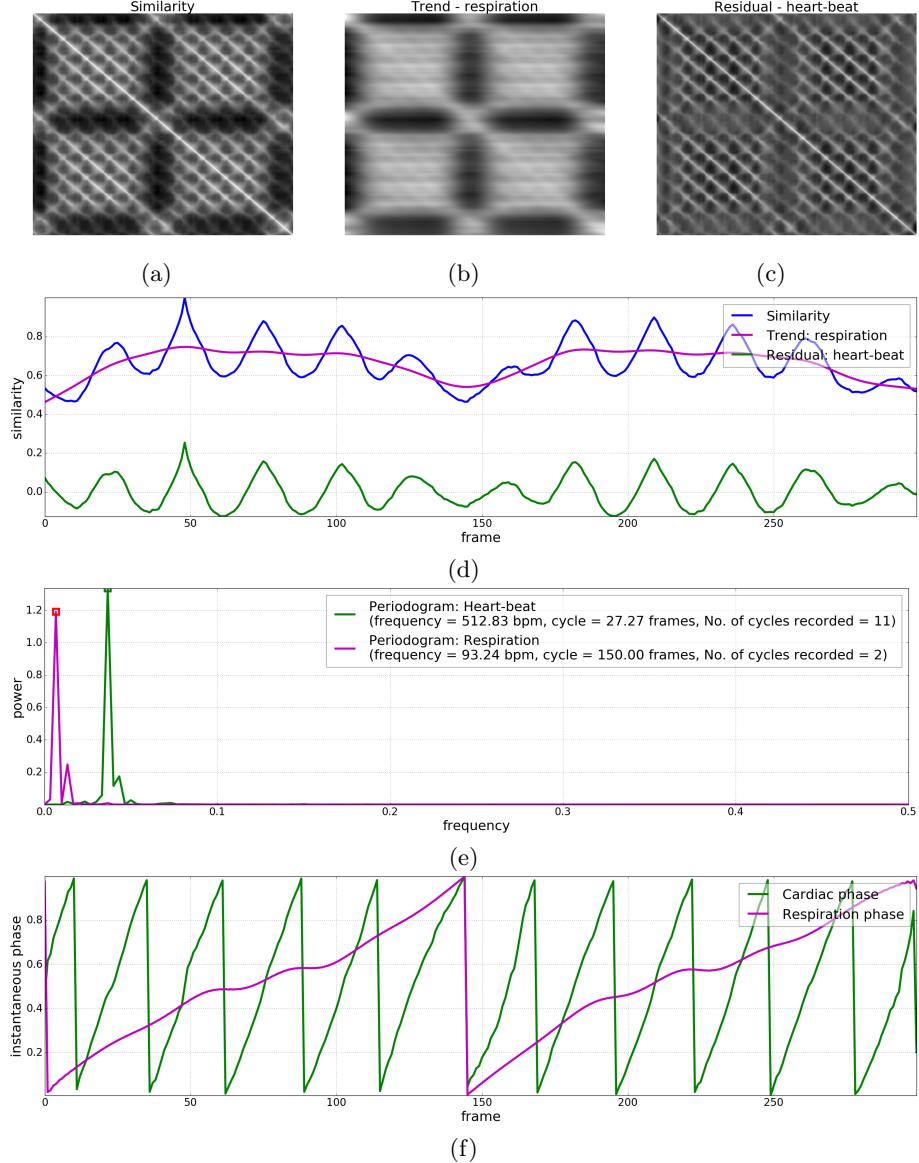
In this section, we present the theory underlying the proposed methods for instantaneous phase estimation, gating and temporal super-resolution of cardiac ultrasound videos. Specifically, we begin by describing our method for the estimation of instantaneous cardiac and respiratory phases in Section 2.1. 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 using only image data. Also shown are the peaks of the QRS complex (pink circle) and the minima of the estimated cardiac phase (black circle) that correspond to the QRS peaks, (b) Five video frames evenly spaced in time between the first and second QRS peaks of the ECG signal constituting one cardiac cycle, and (c) Five video frames evenly spaced in time between first and second minima of the estimated cardiac phase 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 (intensities of pixels) could be measured at each time point. Our strategy was to find a way to simplify this seemingly complex multi-variate problem by transforming it into a univariate one and take advantage of existing univariate methods to solve the problem at hand.



**Fig. 2.** Illustration of the proposed phase estimation method: (a) Inter-frame similarity matrix, (b) Trend component corresponding to the respiratory motion, (c) Seasonal component corresponding to the beating heart motion, (d) The similarity profile chosen for phase estimation along with the associated trend/respiration and seasonal/heart-beat components, (e) Periodogram of the heart-beat and respiration components along with the periodicity characteristics such as frequency and cycle duration, and (f) Instantaneous phase of the heart-beat and respiration components.

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 algorithms [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 the frame similarity signal of 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:

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

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  $A_{resp}$  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. Through Figures 2(b-e), notice how well the two periodic sources have been decoupled from the raw frame similarity signal. Also, in Figure 2, notice

that most of the energy in the periodogram or power-frequency distribution of the respiration and heart-beat components is concentrated within a narrow frequency band. Furthermore, notice how these 1D signals capture the periodicity characteristics (e.g. frequency in beats-per-minute (bpm), period or cycle duration in frames, number of cycles) of the respective sources of motion in the input high-dimensional periodic image sequence.

Next, considering the narrow-band nature 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:

$$\phi(t) = \arctan \left( \frac{H_x(t)}{x(t)} \right) \quad (2)$$

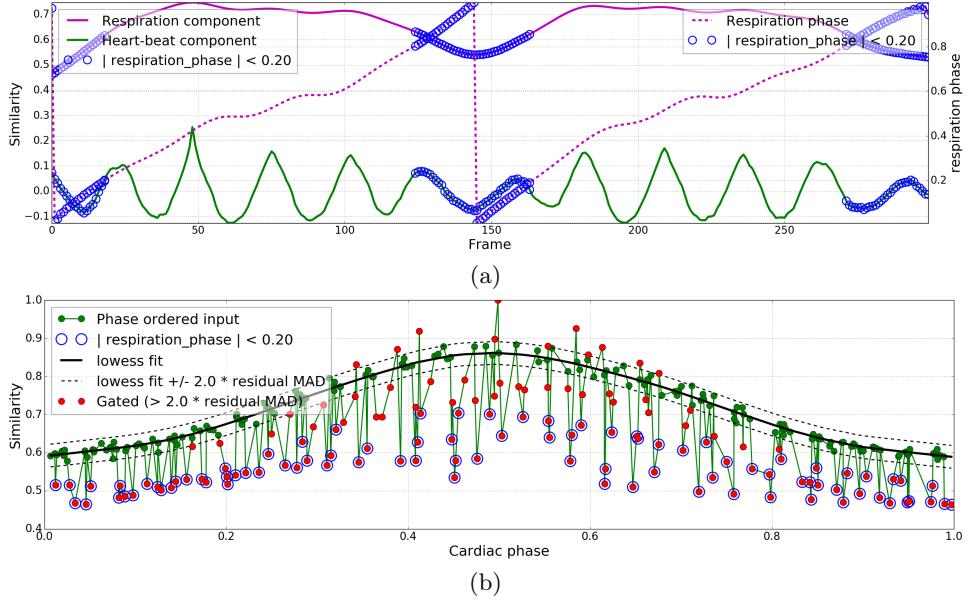
and map  $\phi(t)$  from the range  $[-\pi, \pi]$  to  $[0, 1]$ . Let  $\hat{\phi}_{heart}(t)$  and  $\hat{\phi}_{resp}(t)$  denote the instantaneous cardiac and respiratory phases computed from the trend/respiration  $\hat{r}_{resp}(t)$  and residual/heart-beat  $\hat{r}_{heart}(t)$  signals, respectively, using the aforementioned approach. Figure 2(f) shows the instantaneous cardiac and respiratory phases computed from the heart-beat and respiration signals in in Figure 2(d).

## 2.2 Gating out respiratory frames

Once the instantaneous cardiac and respiratory phases of each frame has been estimated, it can be used to facilitate the extraction of quantitative measurements from 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  $\hat{r}_{resp}(t)$  (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  $F_{cutoff} = \{t \mid \hat{\phi}_{resp}(t) < c \text{ } || \text{ } \hat{\phi}_{resp}(t) > (1 - c)\}$  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  $S(\phi) : [0, 1] \rightarrow R$  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 (LOESS) to the dataset  $\{(\hat{\phi}_{heart}(t), \hat{r}_{heart}(t)) \mid t \notin F_{cutoff}\}$  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 the median absolute value



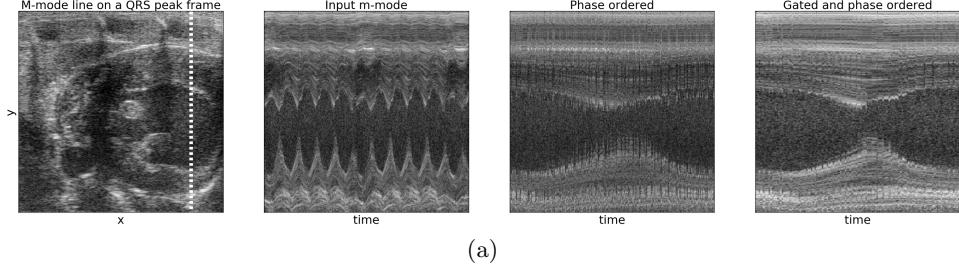
**Fig. 3.** Illustration of our method for gating out respiratory frames: (a) Frames  $F_{cutoff}$  discarded (blue circles) in the first step overlaid with the trend/respiration, residual/heart-beat, and instantaneous respiratory phase signals (b)

$\gamma_{MAD}$  of the difference between the LOESS fit  $S(\hat{\phi}_{heart}(t))$  and residual/heart-beat signal  $\hat{r}_{heart}(t)$  for all frames as a robust estimate of the dispersion of non-respiratory frames around the LOESS fit. Lastly, we gate out all frames  $F_{resp} = \{t \mid |S(\hat{\phi}_{heart}(t)) - \hat{r}_{heart}(t)| > k \times \gamma_{mad}\}$  for which the difference between the LOESS fit and heart-beat signal value is greater than  $k \times \gamma_{mad}$  (we set  $k = 2.0$ ).

### 2.3 Learning to reconstruct image at any cardiac phase

Once the instantaneous cardiac phase of each frame has been estimated, it can be used to enable spatio-temporal super-resolution by considering the set of images within each period of a periodic image sequence as distinct observations of the same data. In this section, we present a method that uses a phase-tagged periodic image sequence to reconstruct a single-period video at a higher temporal resolution.

Given a de-noised periodic sequence  $\{L_1, \dots, L_n\}$  of  $n$  images (with  $m$  pixels each) along with their estimated intra-period phases  $\{\phi_1, \dots, \phi_n\}$ , we first use Nadarya-Watson kernel regression[2] to construct a phase-parameterized image manifold in the form of a function  $M(\phi) : [0, 1] \rightarrow R^m$  that maps any intra-



(a)

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

period phase  $\phi$  to an image as follows:

$$M(\phi) = \frac{\sum_{i=1}^n K(\phi, \phi_i) L_i}{\sum_{i=1}^n K(\phi, \phi_i)} \quad (3)$$

where we define the kernel  $K$  using a gaussian with standard-deviation  $\sigma$  equal to a multiple of the average distance between consecutive phases after sorting them. Note that, we account for the periodic nature of the intra-period phase whenever we measure distance between two phase values. A single-period video can now be reconstructed by sampling the manifold  $M(\phi)$  at any desired resolution from the range  $[0, 1]$  of intra-period phase. Note that the aforementioned process will facilitate temporal super-resolution only if the set of images observed in each distinct period of the given image sequence differ in intra-period phase by some non-zero amount less than its temporal sampling interval. Also, the amount of temporal super-resolution possible will depend both on the number of periods observed and the maximum amount by which one can expect the images to be perturbed in phase across observations.

### 3 Results

### 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)
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)