Learning Your Heart Actions From Pulse: ECG Waveform Reconstruction From PPG

Qiang Zhu, *Student Member, IEEE,* Xin Tian, *Student Member, IEEE,* Chau-Wai Wong, *Member, IEEE,* and Min Wu, *Fellow, IEEE*

***Abstract*—*Objective:* In this paper, we study the relation between electrocardiogram (ECG) and photoplethysmogram (PPG) and infer the waveform of ECG via the PPG signals. *Methods:* In order to address this inverse problem, a transform is proposed to map the discrete cosine transform (DCT) coefficients of each PPG cycle to those of the corresponding ECG cycle. The resulting DCT coefficients of the ECG cycle are inversely transformed to obtain the reconstructed ECG waveform. *Results:* The proposed method is evaluated with the different morphologies of the PPG and ECG signals on three benchmark datasets with a variety of combinations of age, weight, and health conditions using different training modes. Experimental results show that the proposed method can achieve a high prediction accuracy greater than 0.92 in averaged correlation for each dataset when the model is trained subject-wise. *Conclusion:* With a signal processing and learning system that is designed synergistically, we are able to reconstruct ECG signal by exploiting the relation of these two types of cardiovascular measurement. *Significance:* The reconstruction capability of the proposed method may enable low-cost ECG screening for continuous and long-term monitoring. This work may open up a new research direction to transfer the understanding of clinical ECG knowledge base to build a knowledge base for PPG and data from wearable devices.**

***Index Terms*—ECG, PPG, inverse problem, DCT.**

I. Introduction

CARDIOVASCULAR disease (CVD) has become the leading cause of human death – about 32% of all deaths worldwide in 2017 according to the Global Burden of Disease results [2]. Statistics also reveal that young people, especially athletes, are more prone to sudden cardiac arrests than before [3]. Those life-threatening cardiovascular diseases often happen outside clinics and hospitals, and the patients are recommended by cardiologists to attend a long-term continuous monitoring program [4].

The electrocardiogram (ECG) is a fundamental tool of clinical practice [5], and the most commonly used cardiovascular diagnostic procedure today. Many modern wearable ECG systems have been developed in recent decades. They are simpler in physical configuration, more reliable than before, and many weigh only a fraction of a pound. However, the materials used to provide excellent signal quality with the electrode (such as the gel, adhesive, and metal stud) are prone to skin irritation and discomfort during prolonged use [6], which restricts the long-term use of the devices.

The photoplethysmogram (PPG) is a noninvasive circulatory signal related to the pulsatile volume of blood in tissue [7]. In common PPG modalities, the tissue is irradiated by a light-emitting diode, and the reflected or transmitted light intensity is measured by a photodetector on the same or the other side of the tissue. A pulse of blood modulates the light intensity at the photodetector, and the PPG is varying in an opposite direction with the volumn of blood [7]. Compared with ECG, PPG is easier to set up and more economical, making it nearly ubiquitous in clinics and hospitals in the form of finger/toe clips and oximeters. In consumer electronics, PPG has increasing popularity in the form of wearable devices that offer continuous and long-term monitoring capability with little to no skin irritations.

The PPG and ECG signals are intrinsically related, considering that the variation of the peripheral blood volume is influenced by the left ventricular myocardial activities, and these activities are controlled by the electrical signals originating from the sinoatrial (SA) node. The timing, amplitude, and shape characteristics of the PPG waveform contain information about the interaction between the heart and connective vasculature. These features have been translated to measure heart rate, heart rate variability, respiration rate [8], blood oxygen saturation [9], blood pressure [10], and to assess vascular function [11]. As the prevailing use of wearable device capturing users' daily PPG signal, we are inspired to utilize this cardiovascular relation to reconstruct the ECG waveform from the PPG measurement. This exploration, if successful, can provide low-cost ECG screening for continuous and long-term monitoring and take advantage of both the rich clinical knowledge base of ECG signal and the easy accessibility of the PPG signal.

Regarding related work, the authors in [12] trained several classifiers to infer the quantized levels of RR, PR, QRS, and QT interval parameters, respectively, from selected time domain and frequency domain features of PPG. Even though the system achieved 90% accuracy on a benchmark hospital dataset, the capability confined to only inferring ECG parameters may restrict a broader use of this art. In this paper, we propose to estimate the waveform of the ECG signal using PPG measurement by learning a signal model that relates the two time series. We first preprocess the ECG and PPG signal pairs to obtain temporally aligned and normalized sets of signals. We then segment the signals into pairs of cycles

Q. Zhu, X. Tian, and M. Wu are with the Department of Electrical and Computer Engineering, University of Maryland, Collge Park, MD, 20742 USA e-mail: {zhuqiang, xtian17, minwu}@umd.edu.

C.-W. Wong is with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC, 27695 USA e-mail: chauwai.wong@ncsu.edu.

A preliminary 4-page version reporting early-stage results of this work was presented in the 2019 IEEE EMBS BHI Conference (BHI'19) [1].

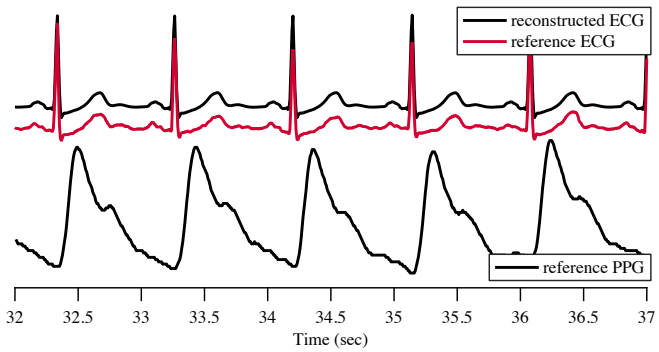


Fig. 1. Upper: a five-second reconstructed ECG signal in test set (black line) vs. the reference ECG signal (red line) using the data from the MIMIC-III database [13]. The two signals were intentionally drawn with an offset in the vertical direction to better reveal the details. Lower: the corresponding PPG signal used to reconstructed the ECG signal.

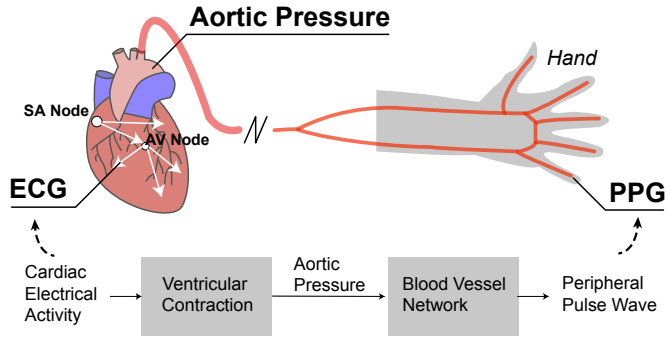


Fig. 2. A visualization of the relationship between the ECG, the aortic pressure, and the PPG.

and train a linear transform that maps the discrete cosine transform (DCT) coefficients of the PPG cycle to those of the corresponding ECG cycle. The ECG waveform is then obtained via the inverse DCT. We evaluate our methodology on two publicly available datasets as well as a self-collected dataset, which in total contains 147 subjects with a wide variety of age, weight, and health conditions using different training modes. As an example, Fig. 1 shows a five-second reconstructed ECG signal in the test set using the proposed method. Note that the reconstructed ECG signal is almost identical with the reference one.

The significance of this work is threefold. First, the statistics of the system performance metrics evaluated on three databases show that our proposed system can reconstruct ECG signals accurately. Second, to the best of our knowledge, this is the first work addressing the problem of reconstructing ECG signals from the PPG signals. It may open up a new direction for cardiac medical practitioners, wearable technologists, and data scientists to leverage a rich body of clinical ECG knowledge and transfer the understanding to build a knowledge base for PPG and data from wearable devices. Third, the technology may enable a more user-friendly, low-cost, continuous, and long-term cardiac monitoring that supports and promotes public health, especially for people with special needs.

The rest of the paper is organized as follows. In Section II, we mathematically model the relationship between the ECG and PPG signals. In Section III, we introduce the proposed system based on the proposed signal model. We test the system and report the experimental results in Section IV, and discuss the possible extension and the limitations of the proposed system in Section V. The conclusion is drawn in Section VI.

II. A Cycle-wise Signal Model for PPG and ECG

In this section, we discuss a physiological model we adopted in this paper to develop the proposed algorithm. As shown in Fig. 2, during each cardiac cycle, the atrioventricular (AV) node receives the electrical signals originated from the SA node. The AV node then transmits this bio-electrical signals through the bundle of His, left bundle branches, and Purkinje fibers to the left ventricular myocardium, causing the depolarization and contraction of the left ventricle. As a result of this process, the pressure of the left ventricle rises and exceeds the aortic pressure, causing the opening of the aortic valves, blood flow from the left ventricle into the aorta, and the corresponding rise of the aortic pressure. Upon closure of the aortic valves, the generated pulse wave transmits the blood to the peripheral parts of our body, such as fingertips or toes, through a network of blood vessels.

A. The ECG Signal and the Aortic Pressure

Consider one specific cardiac cycle. We denote the uniformly sampled cardiac electrical activity as $e(n)$, $n \in [1, L]$, where L is the total number of samples within the cycle. We denote the electrocardiogram measurement recording the potential difference between two electrodes placed on the surface of the skin as $c_y(n)$. Taking into account the human body electrical resistance and the sensor noise, we model the ECG signal $c_y(n)$ as:

$$c_y(n) = \alpha e(n) + v_y(n), \tag{1}$$

where α denotes a subject-specific parameter accounting for the resistance of the electrical path between the heart and the skin surface; $v_y(n)$ denotes the ECG sensor noise, which is modeled as a zero-mean white Gaussian process.

The contraction and relaxation of the heart muscles follow the bio-electrical activities of the heart. These biomechanical activities further modulate the aortic pressure via the opening and closing of the aortic valves. The aortic pressure, denoted as $p_a(n)$, is thus highly correlated with the cardiac electrical activities $e(n)$. To model this correlation, we first map both $e(n)$ and $p_a(n)$ to their frequency domain via type II DCT [14], as DCT has the potential to provide a compact and effective representation of the signals [15]. We then propose to model the relationship of the two signals with a linear transform from the DCT domain of $e(n)$ to that of $p_a(n)$ as:

$$\mathbf{P}_a = \mathbf{H}\mathbf{E}, \tag{2}$$

where $\mathbf{E}$, $\mathbf{P}_a \in \mathbb{R}^{L\times 1}$ are the DCT-II coefficients of $e(n)$ and the aortic pressure $p_a(n)$ respectively. $\mathbf{H} \in \mathbb{R}^{L\times L}$ is the transition matrix.

B. The Pulse Wave and the PPG Signal

When the pulse wave and blood flow travel through our body from the aorta to a peripheral site, it might experience different interactions with the blood vessels, for instance, splitting and pushing. We assume the structure of the blood vessel path of a specific person is time-invariant. Inspired by models for the vocal tract in speech production ([16], Chapter 3), we propose to model this blood vessel channel from the aorta to the peripheral site as a linear time-invariant system. We denote the peripheral pulse signal at a specific body site as $p_p(n)$. We write $p_p(n)$ according to the prior channel assumption as:

$$p_p(n) = b(n) \circledast p_a(n) + v_b(n), \tag{3}$$

where $b(n)$ denotes the impulse response of the channel of blood vessels, and $\circledast$ denotes a symmetric convolution operation [17]. $v_b(n)$ is the zero-mean white Gaussian noise, capturing the variance of this model. The symmetric convolution of $b(n)$ and $p_a(n)$ gives a result that is the same as a linear convolution of the symmetrically left-sided extended version of $b(n)$ and two-sided extended version of $p_a(n)$. The extension of $p_a(n)$ provides smooth boundary values for filtering near its original endpoints. This "folded aliasing" may be preferable in modeling this blood vessel channel effect to the warp-around aliasing of a circular convolution [17].

We assume the PPG sensor attached to the same peripheral site works in the transmissive mode. It means that the photodetector of the PPG sensor is on the other side of the tissue with the light-emitting diode. We assume the light source has a constant intensity of I on the spectral range of the receiver side. We further assume no relative motion between the attached skin and the photodetector, and the contact is tight enough so that the signal is not influenced by the possible environmental illuminations. We write the PPG measurement, denoted as $c_x(n)$, as:

$$c_x(n) = I\left[\tau_0 + \tau_1 p_p(n)\right] + v_x(n), \tag{4}$$

where τ_0 and τ_1 denote the relative transmissive strength of the non-pulsatile components and pulsatile components of tissue, respectively [18] [1]. $v_x(n)$ denotes the PPG sensor noise, which is modeled as a zero-mean white Gaussian process. We can rewrite (4) as:

$$c_x(n) = I_1 p_p(n) + I_0 + v_x(n), \tag{5}$$

where $I_1 = I\tau_1$ and $I_0 = I\tau_0$.

C. The Inverse Model from PPG to ECG

According to the property of the symmetric convolution, a symmetric convolution in time domain can be represented as a pointwise multiplication across the frequency domain of a cosine transform [17]. Combined with the linearity property of the DCT, we may rewrite (3) in frequency domain as:

$$\mathbf{P}_p = \mathbf{B}\mathbf{P}_a + \mathbf{V}_b, \tag{6}$$

[1]Based on the dichromatic model [19], when PPG sensor works in the *reflective mode*, parameters τ_0 and τ_1 in (4) denote the relative reflective strength of the non-pulsatile components and pulsatile components of tissue, respectively [20].

where $\mathbf{P}_p$, $\mathbf{P}_a$, and $\mathbf{V}_b$ are the DCT-II coefficients of $p_p(n)$, $p_a(n)$, and $v_b(n)$ respectively. $\mathbf{B} \triangleq \text{diag}(B_1, B_2, ..., B_L) \in \mathbb{R}^{L\times L}$, where B_k denotes the kth DCT-I coefficient of $b(n)$. We next apply a type II DCT on both sides of (1) and (5) and we arrive at:

$$\mathbf{C}_y = \alpha\mathbf{E} + \mathbf{V}_y \tag{7}$$

$$\mathbf{C}_x = I_1\mathbf{P}_p + \mathbf{I}_0 + \mathbf{V}_x, \tag{8}$$

where $\mathbf{C}_y$, $\mathbf{V_y}$, $\mathbf{C}_x$, $\mathbf{I}_0$ and $\mathbf{V}_x$ denotes the DCT-II coefficients of $c_y(n)$, $v_y(n)$, $c_x(n)$, constant function I_0 and $v_x(n)$ respectively. Assuming the nonsingularity of the matrix $\mathbf{B}$ and $\mathbf{H}$ and according to (2), (6), (7), and (8), we have:

$$\mathbf{C}_y = \mathbf{F}\mathbf{C}_x + \mathbf{C}_0 + \mathbf{V}, \tag{9}$$

where $\mathbf{F} \triangleq \alpha I_1^{-1}\mathbf{H}^{-1}\mathbf{B}^{-1}$, $\mathbf{C}_0 \triangleq -\alpha I_1^{-1}\mathbf{H}^{-1}\mathbf{B}^{-1}\mathbf{I}_0$, and $\mathbf{V} \triangleq \mathbf{V}_y - \alpha\mathbf{H}^{-1}\mathbf{B}^{-1}\left(I_1^{-1}\mathbf{V}_x + \mathbf{V}_b\right)$. When we look individually at each element of $\mathbf{C}_y$, we have:

$$C_y(k) = \mathbf{F}(k)\mathbf{C}_x + C_0(k) + V(k), \quad k \in [1, L], \tag{10}$$

where $\mathbf{F}(k)$ is the kth row of matrix $\mathbf{F}$; $C_0(k)$ and $V(k)$ denote the kth element of $\mathbf{C}_0$ and $\mathbf{V}$, respectively. We know $V(k)$ is a zero-mean Gaussian random variable, as it is a linear combination of zero-mean Gaussian random variables from v_y, v_b, and v_x. According to (10), the relation between the PPG and the ECG signal is captured by a linear model in their frequency domain. We are thus motivated to explore the linear relationships between the DCT coefficients of PPG signal and those of the ECG signals.

III. Methodology

According to the signal model we discussed in the previous section, we propose a system which learns the linear transform $\mathbf{F}$ from pairs of PPG and ECG data. The pipeline of the system is shown in Fig. 3. The pair of PPG and ECG signals are first preprocessed into pairs of synchronized cycles. The cycle pairs are then fed into the training system to learn the transform matrix. We discuss further the details of the system as follows.

A. Preprocessing: Cycle-wise Segmentation

The goal of preprocessing ECG and PPG signals is to obtain temporally aligned and normalized pair of signals, so that the critical temporal features of both waveforms are synchronized to facilitate our investigation. As shown in Fig. 3, the preprocessing phase contains data alignment, signal detrending, cycle-wise segmentation, temporal scaling, and normalization stages that be explained as follows.

a) Data alignment: Considering possible misalignment of the signal pair in each trial, we perform a two-level signal alignment to obtain physically aligned signal pairs. We first estimate the signal delay in the cycle level using the peak features as they are the most distinguishable features within the cycle. We then align the signals to the sample level based on their physical correspondence.

Suppose we have a pair of almost simultaneously recorded PPG and ECG signals, denoted as $\mathbf{x} \in \mathbb{R}^T$ and $\mathbf{y} \in \mathbb{R}^T$ respectively. We name the coordinate of the systolic peak in

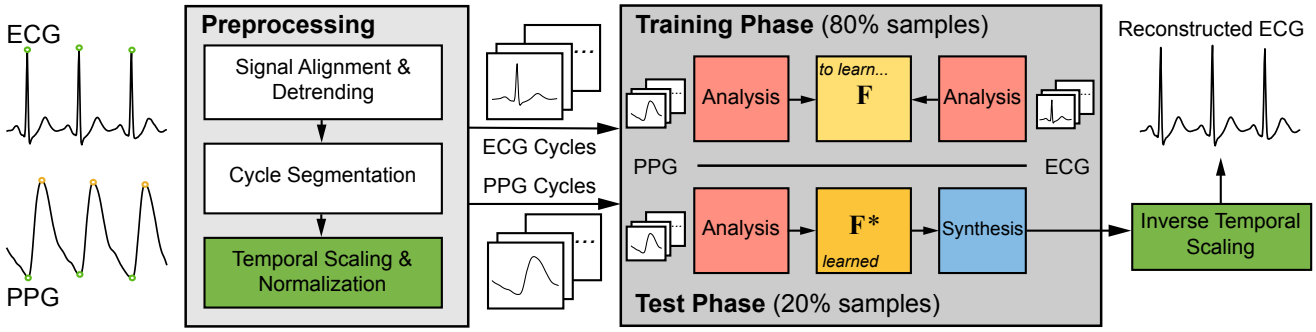


Fig. 3. Flowchart of the proposed system. The ECG and PPG signals are first preprocessed to obtain physically aligned and normalized pairs of cycles. The selected DCT coefficients of 80% pairs of cycles are used for training a linear transform $\mathbf{F}$ which is used in the test phase to reconstruct the ECG signals.

the ith cycle of PPG as $n_{\text{sp}}(i)$ and the R peak of ECG as $n_{\text{rp}}(i)$. The cycle delay m_{delay} is searched for in a discrete interval $\mathbb{D} \triangleq [-k, k]$, where the search radius $k = 5$ as we expect the cycle delay to be small. For each evaluated $m \in \mathbb{D}$, we first preliminarily align the signal with respect to $n_{\text{sp}}(1-m\cdot\mathbb{1}(m<0))$, and $n_{\text{rp}}(1-m\cdot\mathbb{1}(m>0))$. The aligned coordinates of PPG and ECG peaks are $\{n'_{\text{sp}}(m)\}$ and $\{n'_{\text{rp}}(m)\}$. We then estimate the cycle delay $\hat{m}_{\text{delay}}$ by solving the following problem:

$$\hat{m}_{\text{delay}} = \underset{m\in\mathbb{D}}{\operatorname{argmin}} \sum_{i=1}^{i=M-k} \big|n'_{\text{sp}}(i - m\cdot\mathbb{1}(m<0)) - n'_{\text{rp}}(i + m\cdot\mathbb{1}(m>0))\big|, \tag{11}$$

where M denotes the total number of cycles; $\mathbb{1}$ denotes the indicator function. We align the signals by shifting the PPG signal so that the systolic peaks of PPG and the R peaks of ECG are temporally matched.

Next, we align the signal to the sample level according to the R peak of the ECG and the onset point of PPG in the same cycle (the local minimum point before the systolic peak), considering that the R peak corresponds approximately to the opening of the aortic valve, and the onset point of PPG indicates the arrival of the pulse wave [7]. In this way, we eliminate the pulse transit time and align the signals.

b) Detrending: The non-stationary trend in both signals requires additional attention to temporal pattern analysis. A slowly-varying trend can be estimated and then subtracted from the original signals. The trend is assumed to be a smooth, unknown version of $\mathbf{x}$ and $\mathbf{y}$ with a property that its accumulated convexity measured for every point on the signal is as small as possible, namely,

$$\hat{\mathbf{x}}_{\text{trend}} = \underset{\hat{\mathbf{x}}}{\operatorname{argmin}} \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2 + \lambda \|\mathbf{D}_2\hat{\mathbf{x}}\|_2^2, \tag{12}$$

where $\mathbf{x}$ is the original signal, $\hat{\mathbf{x}}_{\text{trend}}$ is the estimated trend in $\mathbf{x}$, λ is a regularization parameter controlling the smoothness of the estimated trend, and $\mathbf{D}_2 \in \mathbb{R}^{T\times T}$ is a Toeplitz matrix that acts as a second-order difference operator. The closed-form solution of (12) is $\hat{\mathbf{x}}_{\text{trend}} = (\mathbf{I} + \lambda\mathbf{D}_2^{\intercal}\mathbf{D}_2)^{-1}\mathbf{x}$, where $\mathbf{I}$ is the identity matrix, Hence, the detrended signal is $\tilde{\mathbf{x}} = \mathbf{x} - \hat{\mathbf{x}}_{\text{trend}}$, and similarly, $\tilde{\mathbf{y}} = \mathbf{y} - \hat{\mathbf{y}}_{\text{trend}}$.

c) Segmentation & Normalization: After the signal alignment and detrending, we segment each cycle of the signal $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ to prepare for the learning phase. In our experiment, we introduce the following two cycle segmentation schemes: SR and R2R.

- *SR*: we segment the signal according to the points which are $1/3$ of the cycle length to the left of the R peaks of the ECG signal. We call this scheme SR as it approximately captures the standard shape of sinus rhythm.
- *R2R*: we segment the signal according to the location of the R peak of the ECG signal to mitigate the reconstruction error in the QRS complex.

After the segmentation, we temporally scale each cycle sample via linear interpolation to make it of length L in order to mitigate the influence of the heart rate variation. We then normalize each cycle by subtracting the sample mean and dividing by the sample standard deviation. We denote the normalized PPG and ECG cycle samples as $\mathbf{c}_x$, $\mathbf{c}_y \in \mathbb{R}^{M\times L}$.

B. Learning a DCT-domain Linear Transform

The right part of Fig. 3 shows our proposed learning framework. In the training phase, we build and train a linear transform to model the relation between the DCT coefficients of PPG and ECG cycles. We then use the trained matrix to reconstruct the ECG waveform in the test phase.

Specifically, we first perform cycle-wise DCT on $\mathbf{c}_x$ and $\mathbf{c}_y$, which yields $\mathbf{C}_x$, $\mathbf{C}_y \in \mathbb{R}^{M\times L}$. Then the first L_x, L_y DCT coefficients of $\mathbf{C}_x$, $\mathbf{C}_y$ are selected to represent the corresponding waveform as the signal energy is concentrated mostly on the lower frequency components per our observation. We denote them as $\tilde{\mathbf{C}}_x \in \mathbb{R}^{M\times L_x}$ and $\tilde{\mathbf{C}}_y \in \mathbb{R}^{M\times L_y}$. We next separate $\tilde{\mathbf{C}}_x$ and $\tilde{\mathbf{C}}_y$ into training and test sets as $\mathbf{C}_{x,\text{train}} \in \mathbb{R}^{M_{\text{train}}\times L_x}$, $\mathbf{C}_{y,\text{train}} \in \mathbb{R}^{M_{\text{train}}\times L_y}$ and $\mathbf{C}_{x,\text{test}} \in \mathbb{R}^{M_{\text{test}}\times L_x}$, $\mathbf{C}_{y,\text{test}} \in \mathbb{R}^{M_{\text{test}}\times L_y}$, where $M_{\text{train}} + M_{\text{test}} = M$.

In the training process, a linear transform matrix $\mathbf{F}^* \in \mathbb{R}^{L_x\times L_y}$ that maps from PPG to ECG DCT coefficients is learned through ridge regression as described below:

$$\mathbf{F}^* = \underset{\mathbf{F}}{\operatorname{argmin}} \|\mathbf{C}_{x,\text{train}}\,\mathbf{F} - \mathbf{C}_{y,\text{train}}\|_{\text{F}}^2 + \gamma\|\mathbf{F}\|_{\text{F}}^2, \tag{13}$$

where $\|\cdot\|_{\text{F}}$ denotes the Frobenius norm of a matrix, and $\gamma > 0$ is a complexity parameter that controls the shrinkage of $\mathbf{F}$ toward zero. The goal of penalizing $\|\mathbf{F}\|_{\text{F}}^2$ is to reduce the variance of the predictions and to avoid overfitting [21].

The analytic solution to (13) is $\mathbf{F}^* = (\mathbf{C}_{x,\text{train}}^\intercal \mathbf{C}_{x,\text{train}} + \gamma \mathbf{I})^{-1} \mathbf{C}_{x,\text{train}}^\intercal \mathbf{C}_{y,\text{train}}$, where $\mathbf{I}$ is the identity matrix.

In the test phase, we apply the optimal linear transform $\mathbf{F}^*$ learned in training stage on $\mathbf{C}_{\text{x,test}}$ and estimate the corresponding DCT coefficients of ECG cycles. We denote the estimate as $\hat{\tilde{\mathbf{C}}}_{y,\text{test}} \triangleq \mathbf{C}_{x,\text{test}}\ \mathbf{F}^*$. To reconstruct ECG, we first augment each row of $\hat{\tilde{\mathbf{C}}}_{y,\text{test}}$ to be in the same dimension as L (by padding zeros). We denote the zero-padded matrix as $\hat{\mathbf{C}}_{y,\text{test}} \in \mathbb{R}^{M_{\text{test}} \times L}$. We then apply inverse DCT to each row of $\hat{\mathbf{C}}_{y,\text{test}}$, interpolate the resulted matrix row by row to its original temporal scale, and concatenate the inversely scaled pieces of cycles to obtain the reconstructed ECG signal $\hat{\mathbf{y}}_{\text{test}}$.

IV. Experiments

A. Experiment I: Capnobase TBME-RR Database

We first used the Capnobase TBME-RR [8] to evaluate the performance of the proposed system. The dataset contains 42 eight-min sessions of simultaneously recorded PPG and ECG measurements from 29 pediatric and 13 adults [2], sampled at 300 Hz. The 42 cases were randomly selected from a larger collection of physiological signals collected during elective surgery and routine anesthesia. Each recorded session corresponds to a unique subject. The PPG signal was acquired on subjects' fingertips via a pulse oximeter. The dataset has a wide variety of patient's age (min: 1, max: 63, median: 14) and weight (min: 9 kg, max: 145 kg, median: 49 kg) and is thus a favorable dataset for testing the performance of our proposed system.

We first pruned the signals according to the human-labeled artifact segments and processed the pairs of ECG and PPG signal using the method introduced in Section III-A to obtain aligned and normalized pairs of the signal cycles. We set $L = 300$ and $L_y = 100$, as most of the diagnostic information of ECG is contained below 100 Hz [5]. We set $\lambda = 500$ and $\gamma = 10$ empirically as they offer the best regularization results in the tasks. In order to test the consistency of the system, we selected the first 80% of each session as the training set and the rest for testing. We used the following two metrics to evaluate the system performance in the test set:

- Relative root mean squared error:

$$\text{rRMSE}(\mathbf{y}_{\text{test}}, \hat{\mathbf{y}}_{\text{test}}) = \frac{\|\mathbf{y}_{\text{test}} - \hat{\mathbf{y}}_{\text{test}}\|_2}{\|\mathbf{y}_{\text{test}}\|_2}, \tag{14}$$

- Pearson's correlation coefficient:

$$\rho(\mathbf{y}_{\text{test}}, \hat{\mathbf{y}}_{\text{test}}) = \frac{(\mathbf{y}_{\text{test}} - \bar{y}_{\text{test}})^\intercal (\hat{\mathbf{y}}_{\text{test}} - \bar{\hat{y}}_{\text{test}})}{\|\mathbf{y}_{\text{test}} - \bar{y}_{\text{test}}\|_2 \, \|\hat{\mathbf{y}}_{\text{test}} - \bar{\hat{y}}_{\text{test}}\|_2}, \tag{15}$$

where $\mathbf{y}_{\text{test}}$, $\bar{\hat{y}}_{\text{test}}$, and $\bar{y}_{\text{test}}$ denote the ECG signal in test set, the average of all coordinates of the vectors $\hat{\mathbf{y}}_{\text{test}}$ and $\mathbf{y}_{\text{test}}$ respectively.

In this study, we evaluate the system in the following two training modes:

[2]Note that the recording in this database is of high signal quality. In cases when the signal is corrupted by noise or subject's motion artifacts, a denoising process is needed to clean the signal before the preprocessing stage.

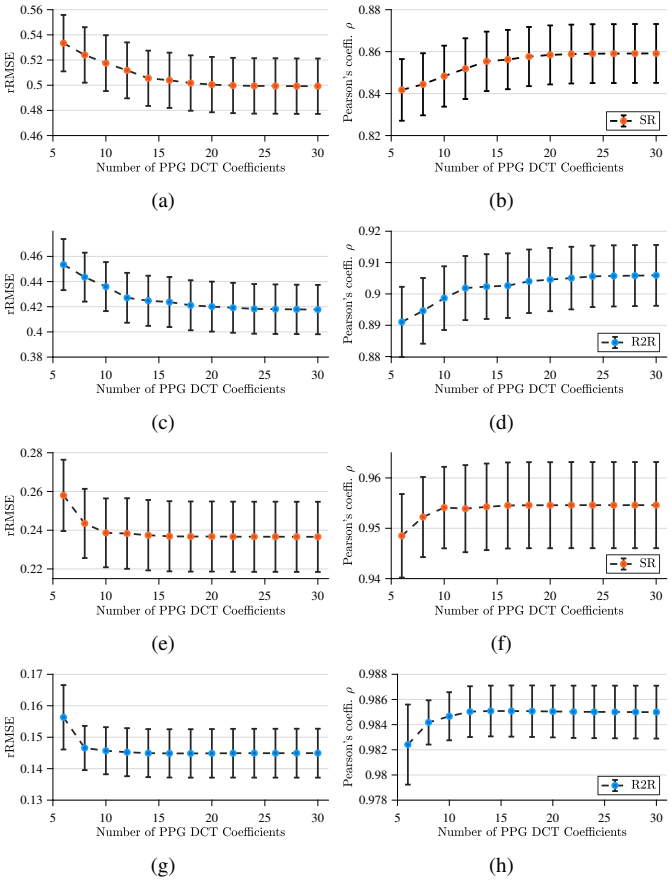


Fig. 4. The line plots give the average of rRMSE in (a), (c), (e), and (g) and ρ in (b), (d), (f), and (h) of all sessions in the test set for different choices of number of PPG DCT coefficient m_1 using SR (a), (b), (e), and (h) and R2R (c), (d), (g), (h) segmentation scheme and SI (a)–(d) and SD (e)-(h) mode respectively. The vertical bars at each data point shows one standard error above and below the sample mean.

- *Subject Independent (SI)* mode: we trained a single linear transform $\mathbf{F}^*$ using all the training data, i.e., the trained model is independent with each subject in the dataset.
- *Subject Dependent (SD)* mode: a linear transform $\mathbf{F}^*$ was trained and tested in each session. In this way, we obtained a subject dependent model for each individual.

We first cross-validated the number of DCT coefficients of the PPG signal L_x used in the learning system. It is clear that the more variables as predictors, i.e., more PPG DCT coefficients are used in the linear system, the better the performance can be achieved in training. However, we can observe from Fig. 4 that the performance of our system in the test set using either SR and R2R becomes saturated as L_x gets more significant from approximate 18 and 12 in the SI and SD mode respectively. The trends of convergence in both modes suggest potential model overfitting. Another observation is that the convergence rate is slower in the SI mode compared with the SD mode. Such observation is expected because the data diversity is much higher in the SI mode than that in the SD mode, and more variables are needed to capture the additional variance in the SI mode. $L_x = 18$ in the SI mode and $L_x = 12$ in the SD mode are thus favorable to us as the system has comparable performance and the model is parsimonious than

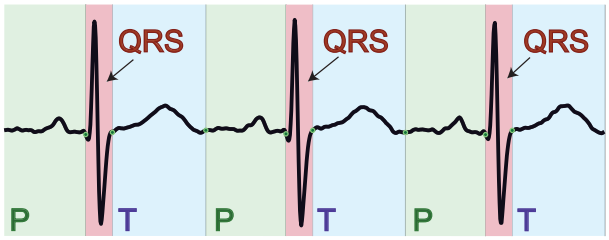


Fig. 5. An example of the ECG segmentation result on three cycles of the signal in the 1st session of TBME-RR database. The green, red, and blue areas in the plot denote the estimated P waves, the QRS waves, and the T waves, respectively. For each cycle, the ratio between the duration of the QRS+T wave is $3/2$ of the duration of the T wave.

those with larger L_x.

The norm of one cycle of ECG signal is usually dominated by that of the QRS complex. This fact of unbalanced signal energy distribution might lead to insufficient evaluation on the P wave and T wave of the ECG signal. To address this problem, we further separated the ECG cycle into and evaluated the system performance on segments of the P wave, QRS wave, and T wave. The evaluation was performed in terms of rRMSE and ρ on each segment as well as using the entire cycle of the signal. Specifically, we adopted the QRS detection algorithm introduced in [22] to locate the onset and endpoint of the QRS complex. We empirically selected the 60% point between the onset points of two adjacent QRS complexes as the separating point for the P and T wave. Fig. 5 shows one example of the ECG segmentation result sampled from the first subject in the database. Note that the onset and endpoint of all waves in each cycle are accurately estimated.

We list the average performance using R2R and SR cycle segmentation schemes in different training modes in Table I and plot the results using the box plots in Fig. 6. In Table I, the performance is characterized by the sample mean and standard deviation of rRMSE and ρ on P, QRS, T, and all waves, where all wave denotes the whole length of the signal including every wave. In addition to the ridge regression learning method we introduced in Section III-B, we also list the performance result using ordinary least squares (OLS) [21] and least absolute shrinkage and selection operator (lasso) [23], respectively. From the statistics, we learn that overall R2R gives better performance than SR, and model trained in the SD mode gives better performance compared with that trained in the SI mode in this dataset as possible subject differences in terms of $\mathbf{H}$ in (2) and $b(t)$ in (3) are expected. The three regression methods, OLS, ridge regression and lasso give comparable performance. In general, R2R outputs comparable results on P and T waves compared with SR, whereas R2R outperforms SR on QRS and all waves. In the SD mode, the average performance in ρ on T wave is about 0.92 using R2R and 0.90 using SR, much higher values than those on the P wave. There are two possible reasons that explain this result. First, compared with the QRS and T waves, the amplitude of the P wave is much smaller. As a result, the P wave becomes more sensitive to the noise compared with the T wave. Second, the shape of the T wave signifies the repolarization of the ventricles, and the ventricular repolarization is correlated with the shape

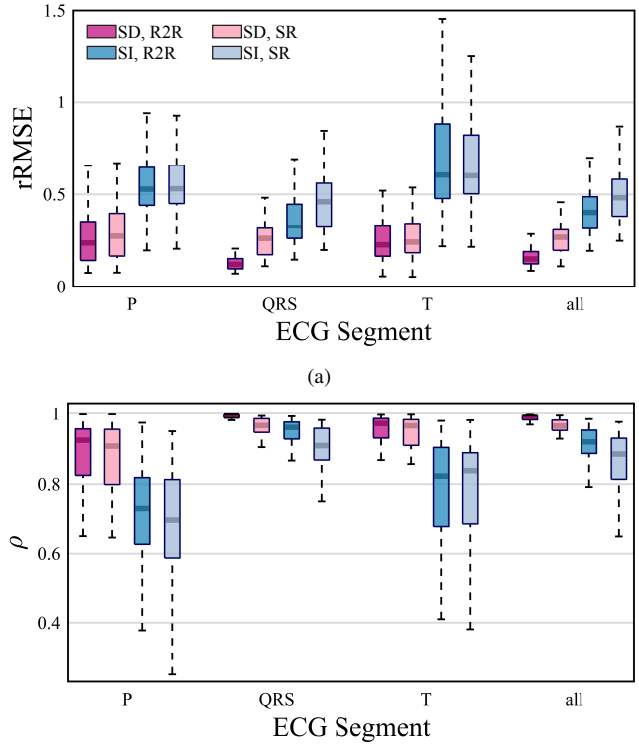


Fig. 6. Comparison of the performance of the proposed method in test set of the TBME-RR database in different combinations of the SR or R2R segmentation schemes and the SD or SI test modes evaluated at P, QRS, T, and all waves. Statistics of the (a) rRMSE and (b) ρ are summarized using the box plots.

of the dicrotic notch in the PPG signal. This is because, during the ventricular repolarization process, the closure of the aortic valve is associated with a small backflow of blood into the ventricle and a characteristic notch in the aortic pressure tracings. This connection between the P wave of ECG and the dicrotic notch of PPG may facilitate the system performance on the P wave.

As an example, we show a five-second segment of the reconstructed ECG waveform in the test set of the first subject in Fig. 7 using the R2R cycle segmentation scheme with $L_x = 18$ in the SI mode and $L_x = 12$ in the SD mode. We choose the first subject to be the example as the system performance evaluated on this subject approximates the average performance over the database. We see from the plot that the system retains most of the shape of the original ECG waveform except for the S peaks in the SI mode and almost perfectly reconstructs the ECG waveform and maintains the location of each PQRST peaks in the SD mode.

In Fig. 8, we plot the rRMSE and ρ of each session concerning subjects' age and weight respectively in two 3-D plots in the SI and SD mode. We then fit a linear model with an interaction term for each combination of training mode and evaluation metric according to the least-squares criterion. An F-test is performed to test whether subjects' profile, i.e., age, weight, and the interaction between age and weight, can significantly affect the performance of the algorithm in each metric and training mode combination. F-test results of small

TABLE I

THE SYSTEM PERFORMANCE IN TEST SET OF THE TBME-RR DATABASE IN TERMS OF MEAN AND STANDARD DEVIATION (IN PARENTHESIS) OF RRMSE AND ρ. R2R SEGMENTATION USING DIFFERENCE COMBINATIONS OF THE TRAINING MODE (SI/SD), THE SEGMENTATION SCHEMES (SR/R2R) AND THE LINEAR REGRESSION METHODS (OLS/RIDGE/LASSO). THE BEST PERFORMED ENTRY IN EACH COLUMN AND TRAINING MODE IS BOLDED FOR BETTER VISUALIZATION. THE ENTRY WITH LOWEST STANDARD DEVIATION WILL BE BOLDED IF THE MEANS OF MULTIPLE ENTRIES ARE IDENTICAL.

		RRMSE				ρ			
		P	QRS	T	all	P	QRS	T	all
TBME-RR	(*SI*)								
	OLS	0.563 (0.197)	0.465 (0.173)	0.736 (0.422)	0.499 (0.142)	0.660 (0.208)	0.879 (0.102)	0.717 (0.262)	0.859 (0.091)
SR	ridge	**0.561** (0.199)	0.465 (0.173)	0.734 (0.423)	0.499 (0.141)	0.659 (0.210)	0.880 (0.101)	0.718 (0.267)	0.859 (0.090)
	lasso	0.565 (0.200)	0.468 (0.173)	0.734 (0.421)	0.502 (0.140)	0.652 (0.210)	0.879 (0.102)	**0.718** (0.266)	0.858 (0.090)
	OLS	0.564 (0.202)	**0.359** (0.139)	0.726 (0.434)	**0.418** (0.124)	0.686 (0.203)	**0.937** (0.059)	0.709 (0.261)	**0.906** (0.061)
R2R	ridge	0.562 (0.204)	0.360 (0.142)	0.722 (0.424)	0.418 (0.126)	**0.687** (0.204)	0.937 (0.060)	0.713 (0.264)	0.906 (0.063)
	lasso	0.564 (0.204)	0.363 (0.142)	**0.721** (0.435)	0.420 (0.125)	0.684 (0.203)	0.937 (0.060)	0.711 (0.267)	0.905 (0.062)
TBME-RR	(*SD*)								
	OLS	0.329 (0.288)	0.275 (0.153)	0.285 (0.213)	0.284 (0.155)	0.825 (0.206)	0.946 (0.071)	0.898 (0.180)	0.947 (0.068)
SR	ridge	0.289 (0.165)	0.273 (0.134)	0.285 (0.173)	0.277 (0.121)	0.836 (0.179)	0.949 (0.061)	0.906 (0.146)	0.951 (0.052)
	lasso	0.314 (0.155)	0.291 (0.139)	0.306 (0.170)	0.294 (0.122)	0.810 (0.188)	0.943 (0.065)	0.899 (0.145)	0.947 (0.054)
	OLS	0.283 (0.180)	0.131 (0.049)	0.285 (0.233)	0.170 (0.080)	0.859 (0.159)	0.990 (0.011)	0.906 (0.173)	0.982 (0.025)
R2R	ridge	**0.273** (0.165)	**0.128** (0.039)	**0.275** (0.178)	**0.165** (0.061)	**0.869** (0.141)	**0.991** (0.006)	**0.917** (0.131)	**0.984** (0.014)
	lasso	0.287 (0.158)	0.135 (0.038)	0.295 (0.173)	0.175 (0.058)	0.856 (0.137)	0.990 (0.006)	0.911 (0.131)	0.983 (0.012)

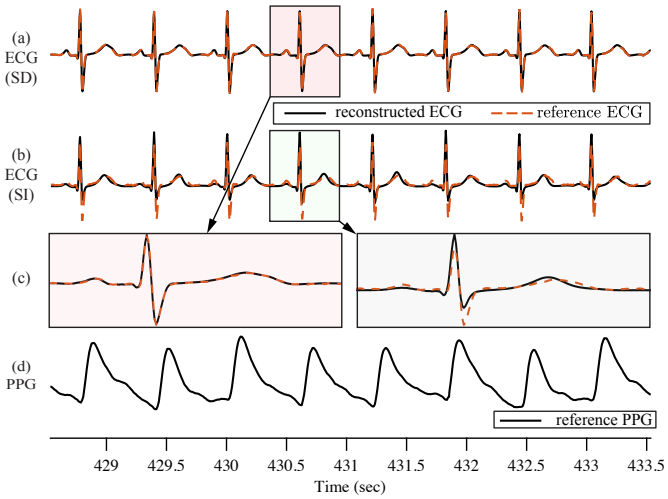


Fig. 7. The reconstructed ECG (black solid line) in (a) the SD and (b) SI and the reference ECG (orange dashed line) waveform of the last 5 seconds of the first session (age: 4 years old, weight: 18 kg) in TBME-RR database. Zoomed-in version of the shaded cycle in each mode is shown in (c). The corresponding PPG waveform is shown in (d).

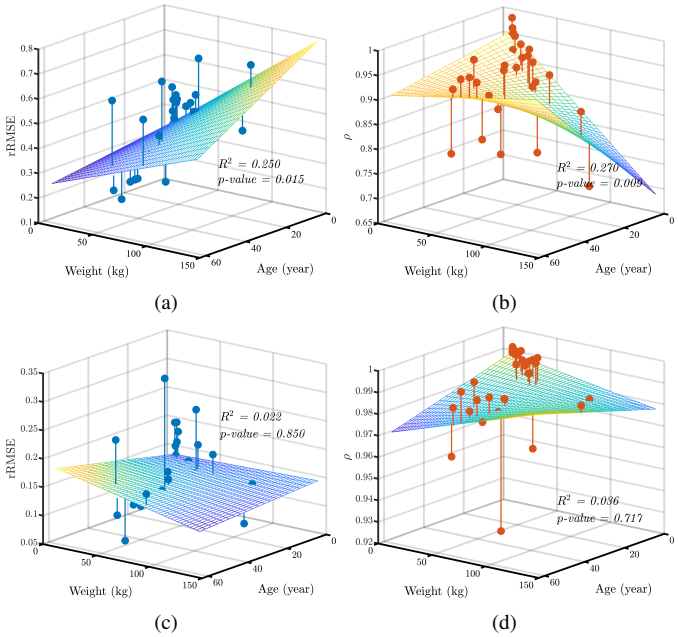


Fig. 8. Scatter plots of (a) RRMSE and (b) ρ vs. subjects' weight and age using R2R scheme. Each sample corresponds to one of 42 sessions. The surface mesh on each plot shows the regressed linear model: RRMSE or $\rho \sim$ intercept + age + weight + age $\times$ weight. The R^2 and the p-value of F-test is shown on each plot.

p-values shown in Figs. 8(a) and 8(b) for the SI mode reveal that the performance of the algorithm is dependent on the combination of subject's age and weight, whereas the large high p-values shown in Figs. 8(c) and 8(d) for the SD mode does not show a strong evidence to reject the hypothesis that the performance of the algorithm is independent of age and weight. Moreover, we notice that the performance tends to be lower as the subject's weight gets larger. This trend of performance degradation might be due to the bias of the training sample that the number of new-borns is much larger than the number of other groups of subjects in the database.

B. Experiment 2: MIMIC-III Database

Medical Information Mart for Intensive Care III (MIMIC-III) [13] is an extensive database comprising vital sign measurements at the bedside documented in MIMIC-III waveform database and part of the patients' profile in the MIMIC-III clinical database. The database is publicly available and encompasses a large population of ICU patients. In this experiment, a subset of the MIMIC-III database was used to

TABLE II

THE SYSTEM PERFORMANCE IN TEST SET OF THE MIMIC-III DATABASE IN TERMS OF SAMPLE MEAN AND STANDARD DEVIATION (IN PARENTHESIS) OF RRMSE AND ρ. R2R SEGMENTATION USING DIFFERENCE COMBINATIONS OF TRAINING MODE (SD/SI) AND LINEAR REGRESSION METHODS (OLS/RIDGE/LASSO). THE BEST PERFORMED ENTRY IN EACH COLUMN AND TRAINING MODE IS BOLDED FOR BETTER VISUALIZATION.

		rRMSE				ρ			
		P	QRS	T	all	P	QRS	T	all
MIMIC-III (*SD*)									
	OLS	0.451 (0.183)	0.320 (0.115)	0.367 (0.175)	0.333 (0.119)	0.807 (0.150)	0.936 (0.045)	0.896 (0.103)	0.935 (0.055)
R2R	ridge	**0.436** (0.175)	0.311 (0.113)	**0.356** (0.169)	0.324 (0.114)	**0.819** (0.141)	0.939 (0.044)	0.903 (0.097)	0.939 (0.053)
	lasso	0.439 (0.171)	**0.310** (0.110)	0.358 (0.162)	**0.324** (0.109)	0.817 (0.139)	**0.940** (0.042)	**0.903** (0.094)	**0.940** (0.049)
MIMIC-III (*SI*)									
	OLS	**0.844** (0.240)	**0.503** (0.166)	**0.773** (0.211)	**0.599** (0.148)	**0.533** (0.252)	0.880 (0.082)	**0.627** (0.318)	**0.790** (0.118)
R2R	ridge	**0.844** (0.240)	**0.503** (0.166)	**0.773** (0.211)	**0.599** (0.148)	0.533 (0.253)	**0.881** (0.082)	**0.627** (0.318)	**0.790** (0.118)
	lasso	**0.844** (0.240)	**0.503** (0.166)	**0.773** (0.211)	**0.599** (0.148)	0.533 (0.253)	**0.881** (0.082)	**0.627** (0.318)	**0.790** (0.118)

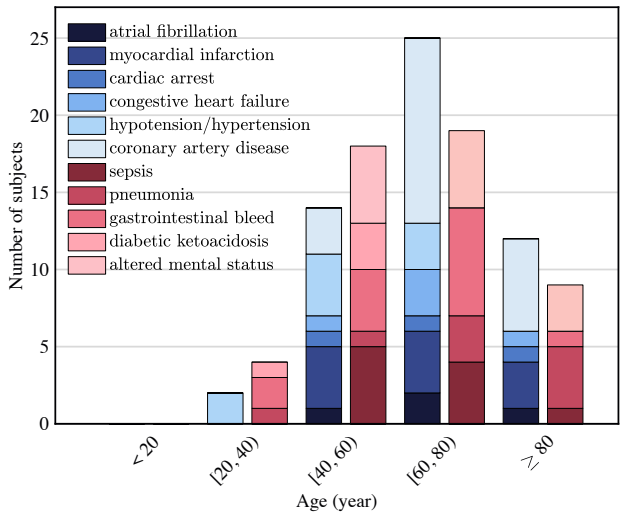


Fig. 9. Distribution of subjects collected from the MIMIC-III database in five age groups and eleven disease types. Within each age group, the cardiac-related diseases are colored as different shades of blue on the left, and the noncardiac-related diseases are colored as different shades of red on the right.

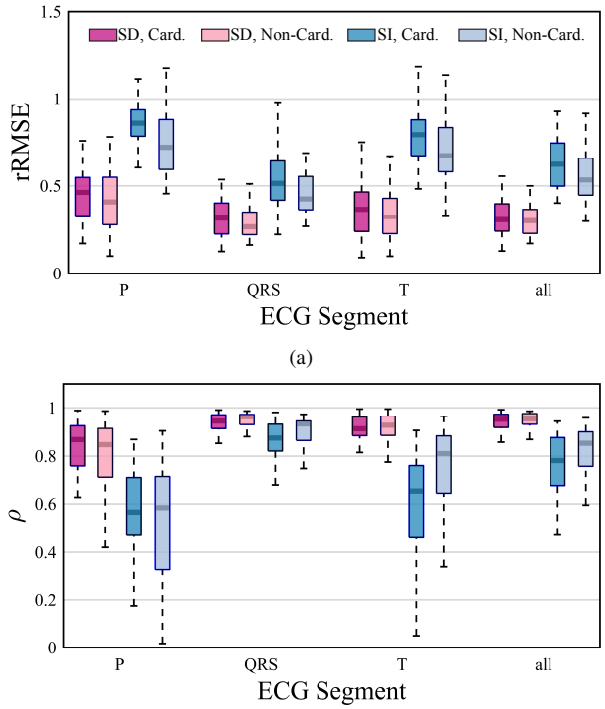


Fig. 10. Comparison of the performance of the proposed method in test set of the MIMIC-III database in different combinations of the disease types and test modes. Statistics of the (a) rRMSE and (b) ρ are summarized using the box plots.

evaluate the system's performance when the subjects were with various cardiac or non-cardiac malfunctions.

Specifically, we selected waveforms that contain both lead II ECG and PPG signals from folder 35 in the MIMIC-III waveform database. Then we linked the selected waveforms with the MIMIC-III clinical database by subject ID to match with the corresponding patient profile. Among the patients, we selected those with specific cardiac/non-cardiac diseases and removed low signal quality PPG/ECG pairs. The resulting collected database consists of 53 patients with six common cardiac diseases and 50 patients with five types of non-cardiac diseases. The distribution of the collected patients is visualized in stacked bar plot based on each one's age group and disease type in Fig. 9. Each patient has three sessions of 5-min ECG and PPG recordings collected within several hours. Cardiac diseases in the resulting database include atrial fibrillation, myocardial infarction, cardiac arrest, congestive heart failure, hypotension, hypertension and coronary artery disease, while non-cardiac diseases are composed of sepsis, pneumonia, gastrointestinal bleed, diabetic ketoacidosis and altered mental status.[3]

In this part of experiment, we evaluate our proposed system in the following two training modes (both under R2R segmentation scheme):

- *Subject Independent (SI)* mode: we trained one linear transform $\mathbf{F}^*$ using training data from patients with

[3]Based on the ICD-9-CM diagnosis codes, we chose those cardiac diseases under the list of "diseases of the circulatory system", which is corresponding to 390-459 in the ICD-9-CM diagnosis codes. For those non-cardiac diseases, we selected them from other categories, including "injury and poisoning", "diseases of the respiratory system", etc.

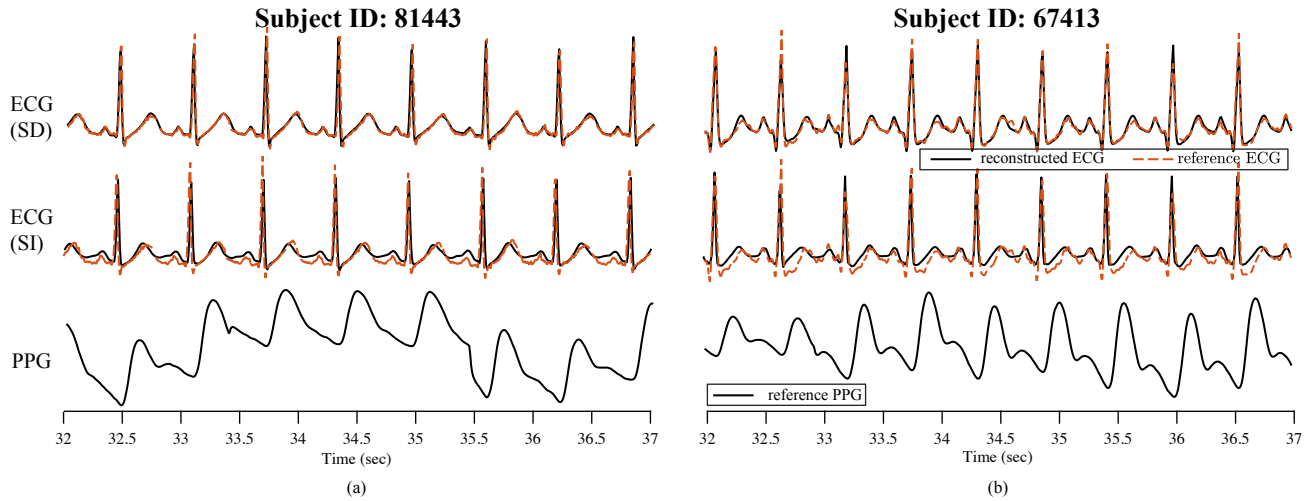


Fig. 11. Two qualitative comparisons between the reconstructed ECG signals tested in the SD (1st row) and SI (2nd row) mode from the MIMIC-III database. (a) The subject is male, 54 years old, and with upper gastrointestinal bleeding. The Pearson's correlation coefficients are 0.969 in the SD mode, and 0.923 in the SI mode. (b) The subject is male, 52 years old, and with congestive heart failure. The correlation coefficients are 0.959 in the SD mode, and 0.881 in the SI mode.

cardiac diseases and another linear transform $\mathbf{F}^*$ from non-cardiac disease patients, i.e., the trained model is independent with each subject in terms of disease type.

- *Subject Dependent (SD)* mode: For each subject, a linear transform $\mathbf{F}^*$ was trained on the first two sessions and tested on the third session. In this way, we obtained a subject dependent model for each individual.

We summarized the average performance in Table II using R2R cycle segmentation scheme in the SD and SI training modes. Results in Table II are characterized by the sample mean and the standard deviation of rRMSE and ρ values on P, QRS, T, and all waves as in the first experiment. The rRMSE and ρ values are also plotted using the box plots in Fig. 10. The statistics reveal that overall non-cardiac cases give better performance than cardiac cases as less variation exists in the morphology of non-cardiac ECG signals. The model trained in the SD mode gives better performance compared with that trained in the SI mode in this dataset, which suggests that $\mathbf{H}$ in (2) and $b(t)$ in (3) may be subject dependent. In general, for the SD mode, the average performance in ρ on T wave is about 0.90 and on QRS wave is about 0.94 using R2R, much higher than those on the P wave, which is in accordance with the first experiment.

In Fig. 11, we show two five-second segments of the reconstructed ECG waveform in the test set from two subjects using the R2R cycle segmentation scheme with $L_x = 18$ in the SI mode and $L_x = 12$ in the SD mode. The first subject is a 54-year-old male with upper gastrointestinal bleeding, and the second subject is a 52-year-old male with congestive heart failure. We see from the plots that the system retains the major shape of the original ECG waveform except for the P waves of the first subject and S waves of the second subject in the SI mode. The system almost perfectly reconstructs the shape of the ECG waveform in the SD mode.

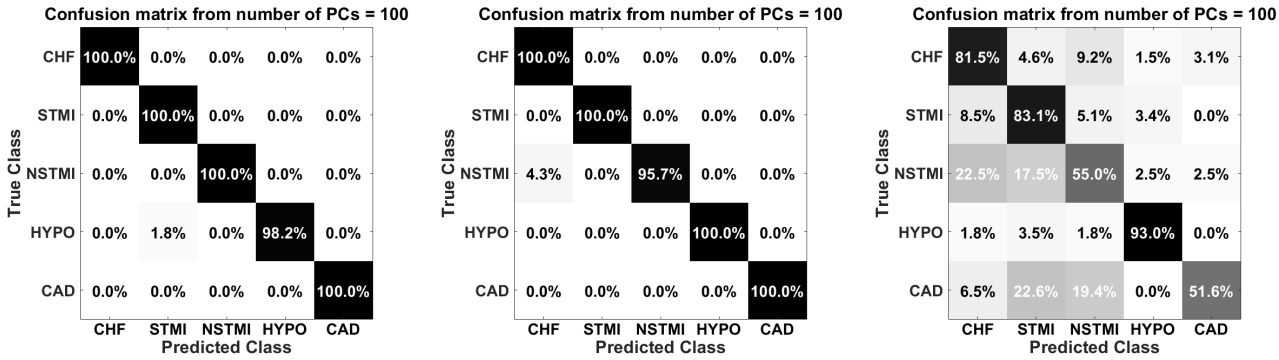
In addition to quantitative analysis of the reconstruction performance by Pearson correlation and rRMSE, we also executed a disease classification experiment on the reconstructed ECG signals to show the potential of our proposed method in applications within biomedical health informatics.

First, from the collected MIMIC-III database, we selected 28 patients with five types of cardiac diseases, including congestive heart failure, ST-segment elevated myocardial infarction, non-ST segment elevated myocardial infarction, hypotension, and coronary artery disease. For each patient, we performed the SD mode ECG reconstruction experiment to obtain the reconstructed ECG signals. To simulate the diagnosis process of cardiologists, we connected the cycle-wise ECG signals into pieces of 30-cycle length for training and classification. The training data is composed of 70 % from the original ECG signals, and the testing data constitutes of the rest 30 % from original ECG signals and all of the reconstructed ECG signals. The detailed distribution of training and testing data concerning disease types are shown in Table III.

We applied PCA for dimensionality reduction and SVM classifier with polynomial kernel from SVM library [24]. The confusion matrices for classification are illustrated in Fig. 12 with the reduced dimension equals to 100. Comparing Figs. 12(a) and 12(b), we conclude that our reconstructed ECG has a comparable classification performance as the original ECG signals. We also include the confusion matrix for original PPG classification in Fig. 12(c) for reference. The superior performance of classification from the reconstructed ECG signals compared to that of the original PPG signal indicates the fidelity of the reconstructed ECG recordings in the presence of cardiac pathologies.

C. Experiment 3: Self-collected UMD Dataset

Next, we test the temporal consistency of the proposed system with the self-collected data using consumer-grade sensors. Two subjects participated in this two-weeks long experiment. One subject is male, 31 years old. The other is female, 23



(a) (b) (c)

Fig. 12. Confusion matrices for classification results using kernel SVM on three types of data: (a) original ECG, (b) inferred ECG, and (c) original PPG.

TABLE III
DISTRIBUTION OF TRAINING AND TESTING DATA FOR DISEASE CLASSIFICATION IN THE MIMIC-III DATASET

Disease	Number of patients	Number of training data	Number of test data (original ECG)	Number of test data (reconstructed ECG)
CHF	7	163 (23.6%)	65 (25.8%)	67 (23.9%)
STMI	7	171 (24.7%)	59 (23.4%)	68 (24.3%)
NSTMI	5	114 (16.5%)	40 (15.9%)	46 (16.4%)
HYPO	5	158 (22.8%)	57 (22.6%)	64 (22.9%)
CAD	4	86 (12.4%)	31 (12.3%)	35 (12.5%)
Total	28	692 (100%)	252 (100%)	280 (100%)

CHF: congestive heart failure
STMI: ST-segment elevated Myocardial infarction
NSTMI: non-ST segment elevated Myocardial infarction
HYPO: hypotension
CAD: coronary artery disease

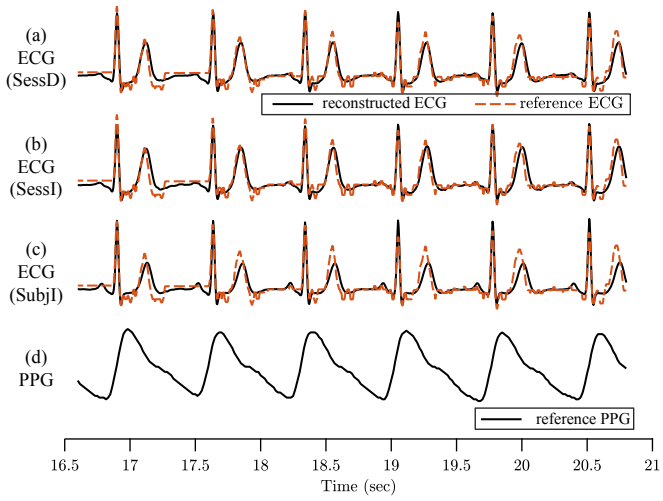


Fig. 13. A qualitative comparison among the reconstructed ECG signals tested in (a) SessD, (b) SessI, and (c) SubjI modes respectively, from the 6th session of the first subject in self-collected database. In (a-c), the black line indicates the reconstructed ECG and the orange dashed line refers to the reference ECG. The Pearson's correlation coefficients for these three cases are 0.937 in SessD, 0.917 in SessI, and 0.869 in SubjI. (d): the corresponding PPG waveform.

years old. Both of them are Asian. According to the most-recent medical examinations received by both subjects, none of them had been diagnosed with any known CVDs or mental illness. We recorded six 5-min sessions for the first subject and seven sessions for the second subject in different times over a two week period. In each session, the subject was asked to wear two devices, namely, EMAY FDA-clear handheld single-lead ECG monitor (Model: EMG-10), and CONTEC pulse oximeter (Model: CMS50E) to record their lead I bipolar ECG signals[4] and finger-tip PPG signals simultaneously. We asked the subject to wear the PPG sensor on his/her index finger of the right hand, and attach the electrodes of the ECG sensor to the palm of the left hand and the back of the right hand. The subjects were asked to sit in front of a table and put their arms on the table as motionless and peacefully as possible to reduce the motion-induced artifacts during the recording time. The sampling rates of the ECG and PPG sensors are 150 and 60 Hz, respectively. We up-sampled both signals to 300 Hz via the bilinear interpolation for consistency consideration and properly aligned the pair of signals.

We evaluate the system performance in the following three training modes:

- *Session Dependent (SessD)* mode: Same as the SD mode we investigated in Section IV-A. $\mathbf{F}^*$ was trained and tested separately in each session.
- *Session Independent (SessI)* mode: The sessions of each subject were first listed chronologically. $\mathbf{F}^*$ was trained on the first 80% of the sessions and was tested on the rest of the sessions in order to maximize the temporal difference of the training and test set.
- *Subject Independent (SubjI)* mode: We combined the subject dependent training sets used in SessI mode and trained a subject independent model to test on the same test set in SessI mode.

In this experiment, we use the R2R segmentation scheme and set $L_x = 12$ in SessD and SessI mode and $L_x = 18$ in SubjI mode. The cycle segmentation process is guided by the peak detection algorithms introduced in [22]. The two algorithms are deployed to detect the R peak of ECG and the

[4]We measure the lead I ECG signal in this experiment considering the easiest accessibility among all leads using the handheld ECG sensor.

TABLE IV
THE SYSTEM PERFORMANCE IN TEST SET OF THE UMD DATABASE IN TERMS OF SAMPLE MEAN AND STANDARD DEVIATION (IN PARENTHESIS) OF RRMSE AND ρ. THE BEST PERFORMED ENTRY IN EACH COLUMN AND TRAINING MODE IS BOLDED FOR BETTER VISUALIZATION.

		rRMSE				ρ			
		P	QRS	T	all	P	QRS	T	all
UMD Data (*SessD*)									
	OLS	0.591 (0.102)	0.230 (0.051)	0.491 (0.103)	0.372 (0.068)	0.620 (0.155)	**0.970** (0.013)	0.863 (0.057)	0.926 (0.028)
R2R	ridge	**0.589** (0.101)	**0.229** (0.051)	**0.490** (0.105)	**0.370** (0.068)	**0.623** (0.147)	**0.970** (0.013)	**0.864** (0.057)	**0.926** (0.027)
	lasso	0.593 (0.102)	0.235 (0.051)	0.494 (0.107)	0.376 (0.071)	0.618 (0.149)	0.968 (0.013)	0.861 (0.058)	0.924 (0.028)
UMD Data (*SessI*)									
	OLS	**0.660** (0.070)	**0.278** (0.021)	0.569 (0.052)	0.427 (0.047)	0.575 (0.125)	**0.966** (0.009)	0.835 (0.039)	0.903 (0.024)
R2R	ridge	0.660 (0.071)	**0.278** (0.021)	**0.567** (0.053)	**0.426** (0.049)	0.575 (0.125)	**0.966** (0.009)	**0.836** (0.039)	**0.904** (0.024)
	lasso	0.664 (0.073)	0.280 (0.022)	0.568 (0.056)	0.428 (0.051)	**0.569** (0.125)	0.965 (0.010)	0.834 (0.041)	0.903 (0.026)
UMS Data (*SubjI*)									
	OLS	**0.724** (0.058)	**0.302** (0.024)	**0.591** (0.111)	**0.447** (0.046)	**0.503** (0.146)	**0.956** (0.013)	**0.830** (0.044)	**0.895** (0.025)
R2R	ridge	0.724 (0.059)	**0.302** (0.024)	**0.591** (0.111)	**0.447** (0.046)	0.503 (0.147)	**0.956** (0.013)	**0.830** (0.044)	**0.895** (0.025)
	lasso	0.725 (0.059)	0.303 (0.025)	0.592 (0.110)	0.448 (0.047)	0.500 (0.146)	0.956 (0.014)	0.829 (0.045)	**0.895** (0.025)

onset point of the PPG signal, respectively. Fig. 13 shows one example of the reconstructed waveforms from the 6th session of the first subject. Note that this session is recorded more than one week after the other sessions. From the qualitative result in 2nd and 3rd rows of Fig. 13, we notice that the reconstructed signals match well with the reference ECG in all waves in the condition of long temporal separation from the training set.

Similar to the previous two experiments, we summarized the average performance in different combinations of training modes and regression methods and evaluate each combination in terms of rRMSE and ρ in P, QRS, T waves respectively. Notice that in general, the system perform best in SessD mode, followed by SessI and SubjI. Again, this difference may suggest possible subject-wise difference of the model parameter $b(t)$, $\mathbf{H}$, or α. Consistent observations in this dataset also include better performance in T wave than P wave, and our conjecture remains with the one claimed in Section IV-A.

V. DISCUSSIONS

A. Cycle Segmentation via PPG

We have evaluated the system in Section IV assuming the availability of the ground truth cardiac cycle information obtained from the ECG signal. We now examine a more practical setting when the cycles are estimated solely from the PPG signal, thereby accounting for the real-world constraint that the reference cycle information is unavailable.

The MIMIC-III database introduced in Section IV-B was adopted in this experiment. We segmented the signal according to the onset points of the PPG signal, considering the onset point represents one of the most distinct features within the PPG cycle. We name this segmentation scheme *O2O*.

To single out the contribution to the reconstruction error due to the discrepancy in the waveform shape rather than the misalignment of the ECG peaks, we evaluate O2O after each reconstructed cycle was post-processed to align with the original ECG signal. This was done by shifting each reconstructed ECG cycle in time so that the original and reconstructed ECG signals were matched according to their R peaks. We list the performance metrics in the SD and SI modes and compare the results with the R2R segmentation in Table V. Note that $\rho = 0.510$ when using O2O segmentation without the peak alignment in the SD mode, and ρ increases to 0.823 once the peak is aligned. The performance statistics reveal that the shape of the waveform is inferred well, and increased error in reconstruction by O2O compared with R2R is mainly due to the misalignment of the signal that has a sample mean and standard deviation of 0.38% and 3.98% in relative cycle length, respectively. This observation is consistent across the SI and SD training modes.

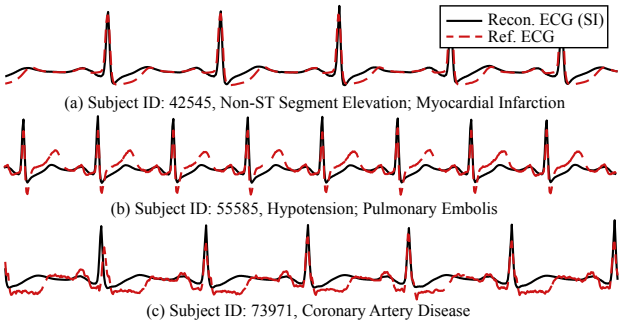
TABLE V
PERFORMANCE COMPARISON USING O2O AND R2R CYCLE SEGMENTATION SCHEMES ON THE MIMIC-III TEST DATASET.

Segmentation	rRMSE (SD)	ρ (SD)	rRMSE (SI)	ρ (SI)
O2O	0.553	0.823	0.689	0.717
R2R	0.324	0.940	0.599	0.790

The disease classification experiment was conducted using the O2O segmentation without the peak alignment. We observed a comparable classification accuracy of the reconstructed ECG signal compared with the result when the model was trained with the R2R segmentation. This observation indicates that the ECG reconstruction deviation does not affect the diagnostic power of the reconstructed ECG signal.

B. Limitations and Extensions of the Proposed Methodology

For some subjects with cardiac complications that influence the morphology of ECG waves, our proposed model in Section II and the corresponding methodology using DCT representations have limitations and may not be able to always faithfully reproduce the ECG signals from PPG, especially when the model is trained in the SI mode. Fig. 14 shows three examples of 5-second long reconstructed ECG signals from



(a) Subject ID: 42545, Non-ST Segment Elevation; Myocardial Infarction

(b) Subject ID: 55585, Hypotension; Pulmonary Embolis

(c) Subject ID: 73971, Coronary Artery Disease

Fig. 14. Three examples of the reconstructed ECG signal of low performance in presence of difference pathologies of the ECG signal. The reconstructed ECGs fail to capture the waveform during the elevation of the T wave (a, b, c), the T wave (b, c), the P wave (c).

the MIMIC-III database using a subject independent model in MIMIC-III database that do not fully capture some detailed characteristics of the original ECG signal. Some other cases that may influence the system performance include motion-induced artifacts and loose contact artifacts in PPG recordings under ambulatory conditions. With a more sophisticated training system and the availability of a larger dataset, we expect such limitations can be addressed.

In order to provide more model flexibility in reconstruction, we foresee that the mapping $\mathbf{F}$ is not limited to a linear transform but can be generalized to nonlinear mappings or transforms (for example, neural networks) and harness more patient data and medical knowledge. Also, the analysis channel of the system is not limited to DCT but can be of other analytical forms, including discrete wavelet transform, discrete Fourier transform, or other parameterized mapping jointly learned with $\mathbf{F}$. With further exploration of datasets with detailed profiles of subjects and larger size of data, a more complex model can be learned based on biomedical, statistical, and physical meanings of the signals to capture the relation of PPG and ECG better. In addition, since ECG is a more adequate and important indicator than PPG for many cardiovascular diseases (CVDs), it has the potential that the developed model, along with the reconstructed ECG, has a significant implication on CVD inference.

VI. Conclusion

This paper has presented a learning-based approach to reconstruct ECG signals from PPG signals based on the synergy of physical model, biomedical knowledge, and data. The algorithm was successfully evaluated in both subject dependent and subject independent fashions on two widely-adopted databases as well as a self-collected database. We have cross-validated the system's hyper-parameters, tested the CVD diagnosis performance using the reconstructed ECG signal, and verified the algorithm's accuracy and consistency at a fine ECG waveform level. As a pilot study, this work demonstrates that with a signal processing and learning system that is designed synergistically, we can precisely reconstruct ECG signal from the more easily obtainable PPG data by exploiting the relation of these two types of cardiovascular related measurement.

References

[1] Q. Zhu, X. Tian, C.-W. Wong, and M. Wu, "ECG reconstruction via PPG: A pilot study," in *IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, Chicago, IL, May 2019.

[2] The Global Burden of Disease: 2017 update. [Online]. Available: http://ghdx.healthdata.org/gbd-data-tool

[3] Sudden death in young people: Heart problems often blamed. [Online]. Available: https://www.mayoclinic.org/diseases-conditions/sudden-cardiac-arrest/in-depth/sudden-death/art-20047571

[4] X. Fan, Q. Yao, Y. Cai, F. Miao, F. Sun, and Y. Li, "Multiscaled fusion of deep convolutional neural networks for screening atrial fibrillation from single lead short ECG recordings," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 6, pp. 1744–1753, nov 2018.

[5] W. B. Fye, "A history of the origin, evolution, and impact of electrocardiography," *American Journal of Cardiology*, vol. 73, no. 13, pp. 937–949, May 1994.

[6] "Maude adverse event report: Conmed corporation invisatrace adult tape wet gel ECG electrodes invisatrace ECG electrodes." [Online]. Available: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfMAUDE/detail.cfm?mdrfoi__id=2170635

[7] A. Reisner, P. A. Shaltis, D. McCombie, and H. H. Asada, "Utility of the photoplethysmogram in circulatory monitoring," *Anesthesiology: The Journal of the American Society of Anesthesiologists*, vol. 108, no. 5, pp. 950–958, May 2008.

[8] W. Karlen, S. Raman, J. M. Ansermino, and G. A. Dumont, "Multiparameter respiratory rate estimation from the photoplethysmogram," *IEEE Trans. on Bio. Eng.*, vol. 60, no. 7, pp. 1946–1953, Jul. 2013.

[9] T. Aoyagi and K. Miyasaka, "Pulse oximetry: its invention, contribution to medicine, and future tasks." *Anesthesia and Analgesia*, vol. 94, no. 1 Suppl, p. S1, 2002.

[10] R. Payne, C. Symeonides, D. Webb, and S. Maxwell, "Pulse transit time measured from the ECG: an unreliable marker of beat-to-beat blood pressure," *Journal of Applied Physiology*, vol. 100, no. 1, pp. 136–141, Jan. 2006.

[11] J. Allen and A. Murray, "Similarity in bilateral photoplethysmographic peripheral pulse wave characteristics at the ears, thumbs and toes," *Physiological Measurement*, vol. 21, no. 3, p. 369, Aug. 2000.

[12] R. Banerjee, A. Sinha, A. D. Choudhury, and A. Visvanathan, "PhotoECG: Photoplethysmography to estimate ECG parameters," in *IEEE International Conf. on Acoustics, Speech and Signal Proc.*, Florence, Italy, May 2014, pp. 4404–4408.

[13] A. E. Johnson, T. J. Pollard, L. Shen, H. L. Li-wei, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. A. Celi, and R. G. Mark, "MIMIC-III, a freely accessible critical care database," *Scientific Data*, vol. 3, p. 160035, 2016.

[14] V. Britanak, P. C. Yip, and K. R. Rao, *Discrete Cosine and Sine Transforms: General Properties, Fast Algorithms and Integer Approximations*. Elsevier, 2010.

[15] H. GholamHosseini, H. Nazeran, and B. Moran, "ECG compression: evaluation of FFT, DCT, and WT performance," *Australas Phys. Eng. Sci. Med.*, vol. 21, no. 4, pp. 186–192, Dec. 1998.

[16] J. R. Deller, J. H. L. Hansen, and J. G. Proakis, *Discrete-Time Processing of Speech Signals*. Wiley-IEEE Press, 2000.

[17] S. A. Martucci, "Symmetric convolution and the discrete sine and cosine transforms," *IEEE Transactions on Signal Processing*, vol. 42, no. 5, pp. 1038–1051, May 1994.

[18] R. R. Anderson and J. A. Parrish, "The optics of human skin," *Journal of Investigative Dermatology*, vol. 77, no. 1, pp. 13–19, Jul. 1981.

[19] S. A. Shafer, "Using color to separate reflection components," *Color Research & Application*, vol. 10, no. 4, pp. 210–218, Dec. 1985.

[20] W. Wang, A. C. den Brinker, S. Stuijk, and G. de Haan, "Algorithmic principles of remote PPG," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 7, pp. 1479–1491, Sep. 2016.

[21] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer, New York, 2009.

[22] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Transaction on Biomedical Engineering*, vol. 32, no. 3, pp. 230–236, Mar. 1985.

[23] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, Jan. 1996.

[24] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, p. 27, 2011.