BayesHeart: A Probabilistic Approach for Robust, Low-Latency Heart Rate Monitoring on Camera Phones

Xiangmin Fan[1,2] , Jingtao Wang[1,2]
[1]Department of Computer Science
[2]Learning Research and Development Center (LRDC)
University of Pittsburgh, Pittsburgh, PA, USA
{xiangmin, jingtaow}@cs.pitt.edu

ABSTRACT

Recent technological advances have demonstrated the *feasibility* of measuring people's heart rates through commodity cameras by capturing users' skin transparency changes, color changes, or involuntary motion. However, such raw image data collected during everyday interactions (e.g. gaming, learning, and fitness training) is often noisy and intermittent, especially in mobile contexts. Such interference causes increased error rates, latency, and even detection failures for most existing algorithms. In this paper, we present BayesHeart, a probabilistic algorithm that extracts both heart rates and distinct phases of the cardiac cycle directly from raw fingertip transparency signals captured by camera phones. BayesHeart is based on an adaptive hidden Markov model, requires minimal training data and is user-independent. Through a comparative study of twelve state-of-the-art algorithms covering the design space of noise reduction and pulse counting, we found that BayesHeart outperforms existing algorithms in both accuracy and speed for noisy, intermittent signals.

Author Keywords

Heart rate; mobile phone; camera phone; hidden Markov model; healthcare

ACM Classification Keywords

H5.2. Information interfaces and presentation (e.g., HCI): User Interfaces

General Terms

Design, Experimentation, Human Factors

INTRODUCTION

Recent technological advances have demonstrated the *feasibility* of extracting people's heart rates through commodity cameras by capturing users' skin transparency changes [4, 25, 31], thermal changes [11], facial color changes [27], or involuntary motion [3]. As one crucial physiological signal, resting heart rate (RHR) is a key indicator of health condition [8, 10], fitness [9, 17], and expected life span [18, 32]. Heart rate and variations of heart rate have also been used to predict human emotion [26], cognitive workload [30], stress [6, 21], and attention. As a result, low cost, ubiquitous heart rate monitoring has enormous opportunities in healthcare, human-computer interaction [30], balancing physical exertion [23], and computer assisted tutoring [20, 36].
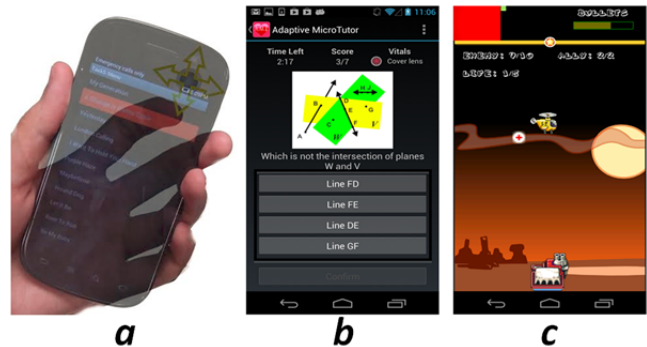

Figure 1: BayesHeart can detect heart rate *implicitly* from *intermittent* mobile interactions. (*a*) LensGesture [37] enhanced UI navigation; (*b*) mobile intelligent tutoring system; (*c*) LivePulse Games (LPG) [14].

Despite compelling advantages in cost, availability, portability, and ease-of-use [14, 28], commodity camera based heart rate monitoring techniques, when running on smartphones, also face three major challenges when compared with traditional approaches such as Electrocardiogram (ECG) and pulse oximetry. First, the indirectly captured Blood Volume Pulse (BVP) signals by cameras are noisier due to interference in background illumination, motion, and contact pressure. Second, the region of interests (ROI), be it human face, head, fingertip, or body, may not be always in the viewport of the camera. Brief but *intermittent* appearance of ROI (e.g. 2 – 20 seconds) will be the norm in many interaction scenarios. Such interference causes increased error rates, latency, and even detection failures when using certain existing algorithms. Last, the *sampling rate* of commodity camera based heart rate monitoring is bounded by the camera's frame rate (30 Hz), which is much lower when compared

with dedicated ECG (125-1500 Hz) devices and pulse oximeters (125-250 Hz). Such low sample rate prevents the adoption of state-of-the-art noise cancelation techniques in heart rate monitoring.

To address these challenges, we present BayesHeart, a probabilistic algorithm that extracts both heart rates and distinct phases of the cardiac cycle directly from noisy, intermittent ROI signals (e.g. fingertip transparency changes, facial color images) captured by camera phones. BayesHeart can extract heart rate from everyday mobile interactions *implicitly*. We use the term "*implicit*" to differentiate our envisioned scenarios with those that require users to mount sensing equipment, "explicitly" launch monitoring apps, and spend an uninterrupted amount of time in data collection. For example, Figure 1.a illustrates LensGesture [37] enhanced UI navigation. A user's heart rate can be inferred when scrolling menu item via back-of-device finger gestures on the camera lens; Figure 1.b shows a mobile intelligent tutoring system which adapts question difficulty based on both learners' stress inferred from heart rates and answer history; Figure 1.c are examples of LivePulse Games (LPG) [14]. LPG integrate users' camera lens covering actions as a control channel for game play and derive users' heart rates from lens-covering actions. In these scenarios, BayesHeart can infer users' heart rates as a *side effect* during everyday tasks.

BayesHeart uses an adaptive hidden Markov model, requiring no user-specific training. BayesHeart has four major advantages when compared with approaches in the existing literature: 1) lower latency and bootstrap time; 2) higher accuracy under noisy and incomplete data; 3) Easier integration with application scenarios that only capture ROI implicitly or intermittently [14, 37]; 4) Joint extraction of both heart rate and distinct phases of the cardiac cycles.
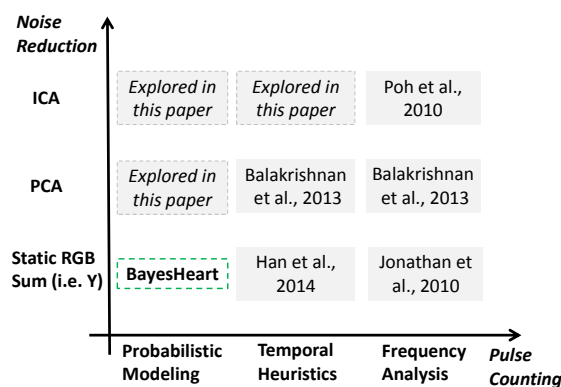
RELATED WORK

Figure 2. The design space of commodity camera based heart rate detection techniques.

Electrocardiography (ECG or EKG) and pulse oximetry (i.e. blood oxygen saturation or blood oxygenation) are the most widely adopted approaches for heart rate detection. ECG and pulse oximetry devices usually come in the form of chest bands, wrist bands, finger clips, or watches. Although the recent release of wearable computing devices such as Android Wear [1] and Apple Watch [2] opens new opportunities for continual heart rate sensing integrated with smartphones, the price and battery life may prevent their wide adoption in the near future.

Several commodity camera based heart rate detection techniques [3, 4, 5, 11. 14, 16, 19, 25, 27, 31] have arisen in recent years. These approaches are capable of extracting heart rate from consumer-grade webcams or smartphones, hence enabling enormous opportunities in health care, fitness training, and affective computing [26].

Along this line, researchers have shown the feasibility of extracting heart rate from finger transparency changes, i.e. photoplethysmography (PPG), captured by a smartphone camera [4, 19, 25]. Jonathan et al. [19] proposed to analyze fingertip video via Fast Fourier transform (FFT), however, the authors didn't conduct a formal evaluation of the accuracy. Poh et al. [27] successfully inferred heart rates by analyzing facial color changes captured by a webcam. Poh's algorithm first used Independent Component Analysis (ICA) to construct a less noisy signal channel from three R/B/G channels, then used FFT and thresholding in the frequency domain for pulse counting. Similarly, Balakrishnan [3] used PCA for noise reduction, frequency domain power analysis for channel selection, and a moving window in temporal domain for peak detection when analyzing involuntary head motions in video. It's also possible to measure heart rate by analyzing facial thermal changes [11]. With the popularization of camera phones, such camera based approaches have already become wildly popular when compared with solutions that rely on dedicated hardware. For example, Instant Heart Rate [16], a commercial camera based PPG app, attracted over 25 million users within two years.

Despite variations in underlying sensing mechanisms, most of today's algorithms adopt a two-step workflow, i.e. 1) noise reduction and 2) heart beat counting. The noise reduction step intends to diminish noise from digitizers, ambient light, body tissue, and motion. Commonly used noise reduction techniques include independent component analysis (ICA) [27], principle component analysis (PCA) [3], smoothing filters [4, 33], and heuristics [14]. The heart beat counting step leverages either temporal domain techniques (peak thresholding [3, 4, 33], heuristic based peak counting [14]) or frequency domain techniques (e.g., Fast Fourier Transform [3, 27]). Figure 2 shows the design space of commodity camera based cardiac pulse detection and the relationship of BayesHeart with existing techniques.

Unfortunately, although the two-step workflow works well on continual and relatively clean signals, it may break when dealing with *implicit* and *intermittent* mobile interaction scenarios (Figure 1). Figure 3 illustrates representative signals (i.e. (a) high quality signals, (b) noisy signals, and (c) intermittent signals) captured from such scenarios.
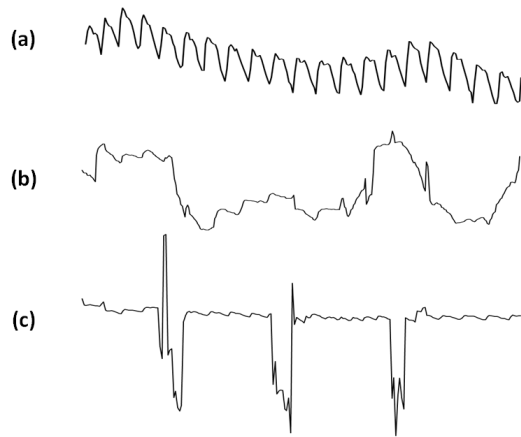
Figure 3. Sample PPG signals captured from a mobile camera (a: high quality signals; b: noisy signals; c: intermittent signals).

In the noise reduction step, component analysis techniques (ICA and PCA) require constructing and updating a linear transformation matrix from historical data. A 30-second window [27] will cause at least the same amount of bootstrap time and increased latency. The calculation of PCA or ICA transformation matrix becomes more challenging when dealing with intermittent signals that only last 5 - 20 seconds in each session.

In the heart beat counting step, an FFT based approach [27] requires *continual* signals meaning that it will break when handling intermittent signals. Meanwhile, temporal domain counting techniques only leverage the *amplitude* properties of pulse peaks and ignore the *temporal regularity* of pulse wave forms. As a result, both peak thresholding and peak counting techniques are sensitive to motion-induced noises, which are hard to eliminate during the noise reduction step.

BayesHeart uses probabilistic modeling to address challenges of existing algorithms in contexts of *noisy*, *implicit* and *intermittent* PPG signals captured by commodity cameras in everyday settings. Unique contributions of BayesHeart include: 1) The usage of an Adaptive Hidden Markov Model to extract both heart rates and distinct phases of the cardiac cycle directly from raw signals; 2) The usage of *discrete local trend* features to achieve both simplified model training and improved robustness; 3) Designing an effective 2/4-state model selection paradigm to exploit both the *temporal regularity* and the *intra-person diversity* in signals. We also advance the state-of-the-art by identifying the design space of commodity-camera based heart rate detection and presenting a comparative study of BayesHeart, existing algorithms, and their variants.

Hidden Markov models have been used in analyzing cardiac pulse signals (e.g., ECG and PPG) before, either for analyzing signal properties [7, 15], or for diagnosing disease [24]. BayesHeart is unique in that existing techniques focused on *continual clinical* signals captured

from *medical-grade* devices collected by researchers or health-care providers. Thanks to the sampling rate (150–1500 Hz), form factors (wrist bands, chest bands, or finger clips), and usage scenarios (in-door, continual measurements in stationary postures), extrinsic noise and interruptions were not a major concerns in existing research. By comparison, BayesHeart explores the opportunities of *implicit* and *intermittent* heart rate monitoring for normal people in everyday environments. Existing techniques for analyzing medical grade ECG/PPG signals do not generalize directly to low frequency (below 30 Hz) and noisy signals captured by commodity cameras. We invented two novel techniques in BayesHeart to overcome new challenges encountered: 1) We defined four novel "local trend" features as the input to our learning algorithm. The absolute signal magnitude feature used by HMM models in ECG and PPG research do not work well on noisy and intermittent signals from commodity cameras; 2) We designed a simple but effective 2/4-state adaptive model to capture signal variances in the unique context of commodity camera based mobile interactions.

THE BAYESHEART ALGORITHM

Background

The underlying theory behind photoplethysmographic (PPG) imaging is as follows: the heart pumps fresh blood to the capillary vessels of a human body during systole in each cardiac cycle. Such blood volume changes lead to changes in fingertip transparency, which can be detected by the built-in camera of the mobile phone when the user covers the lens of the camera with her fingertip [12, 13, 19]. Therefore the changes of finger transparency can be viewed as a *generative process*, in which there are natural correlations between different regions of the PPG waveform and dedicated cardiac phase.
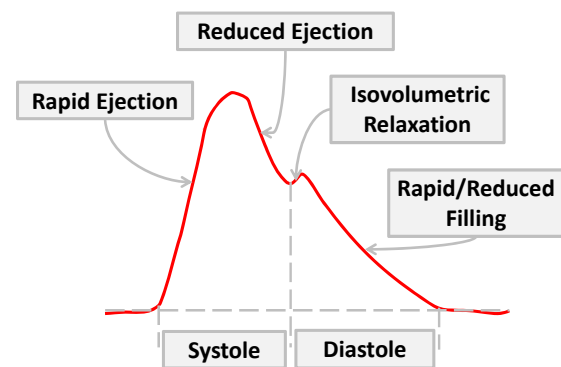

Figure 4. One-cycle waveform associated with the physical activities in one cardiac cycle.

Pulse Modeling

BayesHeart relies on a hidden Markov model to capture the temporal regularity of the different stages in cardiac cycles (hidden) and the finger transparency changes (observable). After training the model, given new observations, we can segment the observations into states by calculating an optimal alignment via Viterbi decoding [29]. Then heart

rate can be estimated by extracting the duration of each cardiac cycle from the derived cardiac alignment.
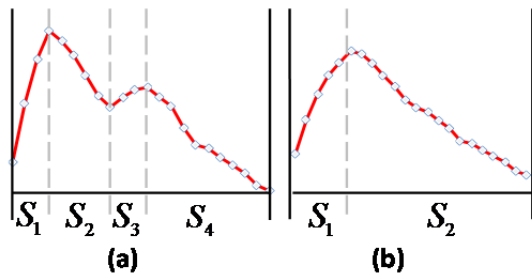
Hidden States

Figure 5. States selection based on the waveform shapes.

According to PPG imaging [22, 33], a typical cardiac cycle includes four distinct stages (Figure 4): 1) rapid ejection (i.e. systolic upstroke); 2) reduced ejection (i.e. systolic downstroke), 3) isovolumetric relaxation (i.e. a small upstroke caused by dicrotic notch); and 4) rapid/reduced filling (i.e. diastolic downstroke). Therefore it is a natural choice to use a 4-state hidden Markov model with each state corresponding to one cardiac stage (Figure 5.a). However, due to extrinsic noise and variations in tissue/skin reflections, the third cardiac stage (isovalumetric relaxation) can be hard to identify in the waveforms captured by commodity cameras. In such situations, the waveform of one cardiac cycle only shows two distinct phases (Figure 5.b): rapid ejection (i.e., systolic upstroke) and reduced ejection plus the whole diastole phase (i.e., a long downstroke). Hence we propose an adaptive 2/4 state model to capture both the subtlety and diversity of waveforms.

We do not consider models with more than 4 states for two reasons. 1) Although the model likelihood on training data improves as the number of states increases, it usually comes at the cost of data overfitting (i.e., more parameters) and more training samples. 2) Our main purpose is to estimate the *duration* of cardiac cycles rather than to analyze subtle changes within one cycle. For this purpose, fine grained segmentations, at the cost of more training data and increased algorithm complexity, won't bring us additional insight.

Observations
Unlike existing research on clinical ECG/PPG analysis [7, 15], we choose not to use the absolute observations (i.e. brightness of finger transparency) in our model because such absolute scales are sensitive to both the environmental illumination changes and motion-induced noise. Instead, we choose the "*local trend*" of each sample point as a more *robust* feature as our model observations. Interestingly, this feature is more *expressive* than the absolute scale in our context. For instance, the wave-form generates much more *increasing* observations in the rapid ejection stage than the filling stage. Such regularities encoded in the "*local trend*" feature are easier to capture by BayesHeart. We further

define four types of discrete observations (o1 – o4) from the "*local trend*" feature (Figure 6, o1 represents increasing observations, o2 represents local minimum observations, o3 represents maximum observations, and o4 represents decreasing observations).
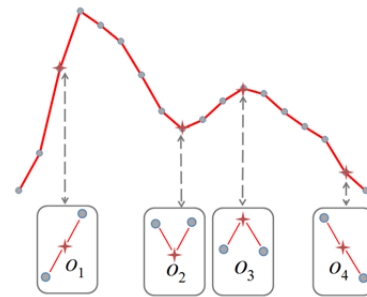

Figure 6. Four types of observations.

Mathematical formulation
BayesHeart is a discrete left-right HMM [29] defined as follows:
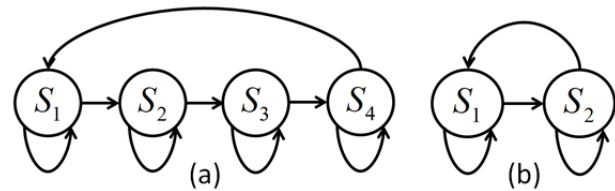

Figure 7. 4-state model (a) and 2-state model (b)

1) *N*, the number of states in the model

We denote the individual states as $S = S_1, \ldots, S_N$ and the state at time t as s_t. As mentioned above, there are two models (i.e., 4-state model and 2-state model, Figure 7) in our approach.

2) *M*, the number of distinct observation symbols per state,

In our case, *M*=4. We denote the individual symbols as $O = \{O_1, O_2, O_3, O_4\}$.

3) The initial state distribution $\pi = \{\pi_i\}$ where

$$\pi_i = P\{s_1 = S_i\}, 1 \le i \le N$$

4) The state transition probability distribution $A = \{a_{ij}\}$ where

$$a_{ij} = P\{s_{t+1} = S_j | s_t = S_i\}, \qquad 1 \le i, j \le N$$

In BayesHeart, we add order constraints within each cardiac cycle by setting $a_{ij} = 0$ for the (i, j) pairs in which $i > j$. The only exception is $a_{N1} > 0$ which enables the model to start new cycles.

5) The observation symbol probability distribution in state j, $B = \{b_j(k)\}$, where

$$b_j(k) = P\{o_k \text{ at } t \mid s_t = S_k\}, 1 \le j \le N,$$

$$1 \le k \le M$$

The hidden Markov model in BayesHeart can be characterized in terms of three probability measures, A, B and π. For convenience, we denote the 4-state model as $\lambda_1 = (A_1, B_1, \pi_1)$ and the 2-state model as $\lambda_2 = (A_2, B_2, \pi_2)$ in a compact way.

Parameter estimation (model training)
To use the Baum-Welch method to estimate model parameters $\lambda = (A, B, \pi)$, we estimate initial distribution of π as the temporal duration of each cardiac state in the training data. For the state transition probability distribution A, we assign a high probability (i.e., 0.8) of remaining in the same state and low probabilities (i.e., 0.2) to transitioning between states. We set $a_{ij} = 0$ (except for a_{N1}) when $i > j$ because the cardiac stages appear sequentially and cannot be reversed.

We train the 4-state model λ_1 and the 2-state model λ_2 separately and use a model selection process detailed in the next section at runtime.

Heart Rate Estimation
After deriving the underlying model via offline training, the BayesHeart runtime includes four phases: 1) model selection; 2) state sequence generation; 3) cardiac pulse interval calculation; and 4) post-processing.

Model selection
BayesHeart uses the first 5 seconds[1] of observations for model selection. We leverage the Bayesian information criterion (BIC) [35] to find the better model λ^* from $\{\lambda_1, \lambda_2\}$ at the same time to prevent the 4-state model from overfitting the observations.

$$BIC_\lambda = -2 \cdot \ln \Pr(o^*|\lambda) + k \cdot (\ln(n) - \ln(2\pi))$$

$$\lambda^* = \underset{\lambda \in \{\lambda_1, \lambda_2\}}{\operatorname{argmin}} BIC_\lambda$$

The model selection step does not introduce extra latency because it can be run in parallel with the state sequences generation step discussed later.

State sequences generation
In this step, we leverage the Viterbi algorithm to infer the optimal state sequence that is most likely to generate the observations by maximizing $Pr(o, s|\lambda)$.

Cardiac cycle/distinct phases extraction
We define the transition from the last state to the first state (i.e. $S_N \rightarrow S_1$) as the start of a new cardiac cycle and mark all of such transitions in the derived state sequence. Therefore, the duration d between two adjacent marks is the

---

[1] We assume that owner change is rare for mobile devices and context (location, environment, etc.) change happens at the scale of minutes or hours rather than seconds. The model selection process can run more frequently when necessary.

duration of one cycle (Figure 8). The instant heart rate estimate is:

$$Instant\ Heart\ Rate\ (bpm) = \frac{60000(ms)}{d\ (ms)}$$

Here bpm means beats per minutes. It is worth noticing that BayesHeart extracts distinct phases in each cardiac cycle in parallel with the heart rate estimation processing (Figure 8).
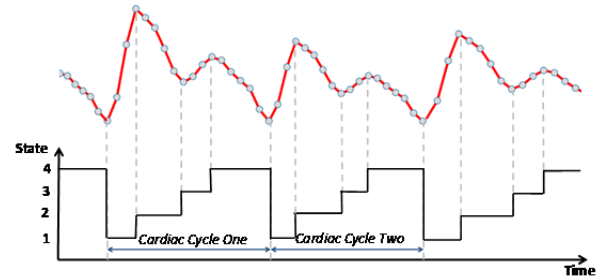

Figure 8. Cardiac cycle/ distinct phase extraction from the underlying state sequence.

Post-processing
Despite the robustness of our BayesHeart algorithm, we still impose two simple heuristics to reduce outliers in extreme situations. If either of the heuristics is violated, BayesHeart rejects the current estimation and outputs the first valid estimation in history.

Heuristic 1: Valid heart rates are within the range of [30, 300] bpm.

Heuristic 2: The maximal change between two adjacent bpm estimates should not be more than (k=5) bpm.

Intermittent signals
As highlighted in the introduction section, the intermittent appearance of ROI (e.g. 2 – 30 seconds) is the norm in many interaction scenarios (e.g., mobile gaming, mobile tutoring, etc.). Therefore we investigate how to extract heart rate via intermittent covering actions. There are three problems when dealing with intermittent covering: 1) how to detect users' covering actions (i.e., when users are covering the lens); 2) how to deal with the noise introduced by intermittent covering actions; 3) once users' covering actions can be detected, how to estimate heart rate based on several separate pieces of signals.

For the first problem, we leverage a fast and reliable linear classification model proposed in [37] to detect the lens covering gesture. The model uses the global mean and standard deviation of all the pixels in an image frame to infer whether the user is covering the lens or not with high accuracy (i.e., 97.9%). After this step, we get a set of data sequences (i.e., observation sequences), with each sequence corresponding to one covering action.

For the second problem, we find that most of the noise is generated by finger movements and pressure changes (e.g., at the beginning of each covering action). Therefore we

apply two techniques to reduce the extreme noise that appears in intermittent signals: 1) Discard the observation sequences corresponding to the covering actions which last less than 2 seconds; 2) For the data sequences that are longer than 2 seconds, discard the first 1 second of data for each covering action (Figure 9, a).
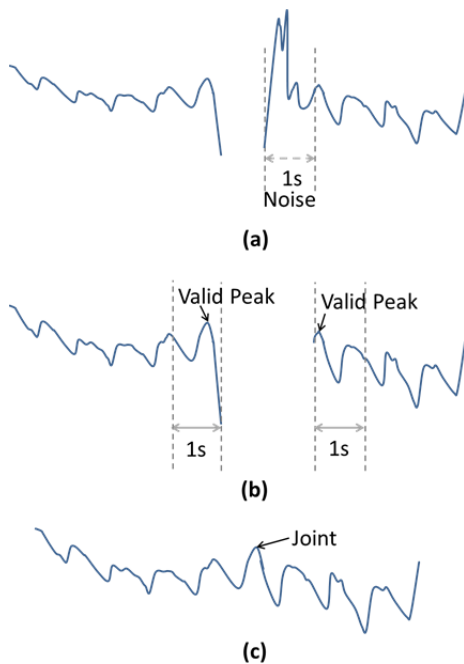

(a)

(b)

(c)

Figure 9. Additional steps with intermittent covering. a) Remove noise at the beginning of each covering action. b) Find two valid peaks. c) Connect the two valid peaks.

For the third problem, we have two strategies to extract heart rate through separate segments of signals: 1) Treat each sequence independently and apply HMM on each of them; 2) Concatenate the sequences to one whole sequence and then apply HMM on it. The first one may lead to lower utilization of the signal because we cannot make use of the front/back ends of each separate piece if the ends do not appear in a complete cycle in that piece. Therefore we try to concatenate them together to get a continuous and complete waveform.

One question emerges regarding how to concatenate separate pieces of signals. One brute force method is to directly connect the last sample of the previous piece with the first sample of the next piece. However, this cannot ensure that the concatenation of the two pieces can form one cycle and may introduce inaccurate estimations. Therefore we use a heuristic to concatenate two pieces trying to make use of the data at the very end/beginning of the signal pieces. The heuristic attempts to find the last valid peak of the previous piece and the first valid peak of the next piece and link them by joining the two peaks together. Since the normal resting heart rate for adults ranges from 60-100 bpm, valid peaks are close to 600ms-1000ms apart; therefore we assume that a valid peak can be

found within a window with the size of 1000ms. So we apply a window at the end of the previous piece as well as the beginning of the next piece and localize the valid peak by choosing the local-maximum point (i.e., o3) with the largest amplitude within the windows (1000ms). (Figure 9, b). Then these two valid peaks are connected together to concatenate the two signals (Figure 9, c).

EVALUATION

We carried out two studies to evaluate the efficacy of BayesHeart. In the first study we investigated the impact of different design choices on BayesHeart, such as the inclusion of an adaptive model and the inclusion of the post processing step on both accuracy and latency. The second study was a comparison of the state-of-the-art algorithms in the design space (Figure 2) with BayesHeart. We report results on both accuracy and latency for both normal signals and intermittent signals.

Data Collection

We collected data from 20 subjects (7 female) from department mailing lists of a local university. The participants were between 23 and 45 years old (mean = 27.8, $\sigma = 4.9$). We collected two types of PPG data via the built-in camera of a smartphone: 1) 10 minutes static covering and 2) 10 minutes intermittent covering. For intermittent covering, participants were asked to cover the lens for 5-10 seconds and then move their finger away for 1-3 seconds, and to repeat this process for a total of 10 minutes. Participants used one hand to operate the mobile phone and we attached a pulse oximeter on their other hand to collect the ground truth heart rate data. Each participant was paid $5 for their time.

We used a Google Nexus smartphone running Android 4.1 for data collection. The Google Nexus has a 4.65 inch, 720*1280 pixels display, 1.2 GHz dual core ARM Cortex-A9 processor. It has a 5 mega-pixel back camera and an LED flash light. We set the built-in camera in preview mode, capturing color images of 144x176 pixel at 30 fps (frames per second). We sample 800 pixels evenly distributed in each frame and use the RGB/YUV sum of these 800 pixels to estimate the brightness of the frame. In this way we derive a set of time-stamped ROI signal vectors. We resample the data by linear interpolation to 30Hz to compensate for the jitter effect of the video stream.

The pulse oximeter in the experiment was a CMS 50D with USB port. CMS 50D is an FDA-approved, medical grade device. The accuracy of CMS 50D for pulse ratio was +/- 2 bpm.

Benchmarking Metrics

In the first part of the evaluation, we use 50% of the data (i.e., for each subject, we use 5-minute static covering data and 5-minute intermittent covering data) for training and the other 50% of the data for testing. For the second part, we use leave-one-out cross-validation (LOOCV) to test

user-independent performance of BayesHeart compared with previous work.

We use ***mean error rate (MER)*** to measure algorithm accuracy. To derive the MER of a given algorithm/configuration, we compare the estimated heart rate with the gold standard *every second* and report the average. ***Estimation latency*** is defined as the time an algorithm needed to generate the first accurate heart rate estimation (+/- 5% when compared with the gold standard). We also define *utilization rate* to model the robustness of an algorithm. It is defined as the ratio between the quantity of samples from which BayesHeart derived valid cardiac cycles (i.e., cycles that are not eliminated in the post-processing step and are used in heart rate calculation) and the total quantity of samples in use.
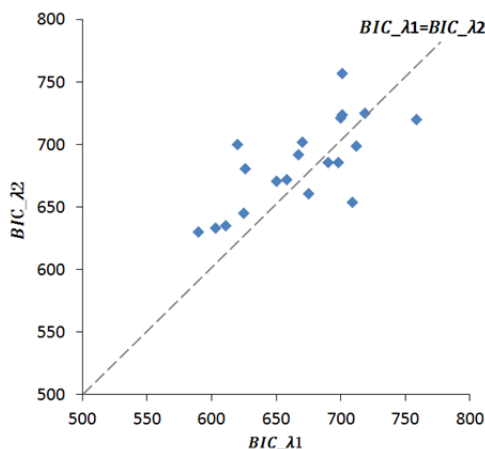
Figure 10. Plots of *BIC* values of 20 subjects. BayesHeart chooses a 4-state model for subjects above the decision boundary and a 2-state model for all others.

Understanding BayesHeart

Model Selection
Figure 10 shows the scatter plot of the Bayesian information criterion (BIC) for the two models. BayesHeart chose 6 subjects for the 2-state model and 14 subjects for the 4-state model.
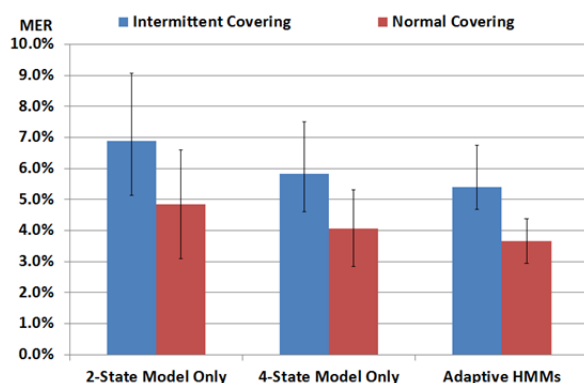
Figure 11. Mean error rates of BayesHeart with different types of models.

Figure 11 illustrates the impact of model selection on the accuracy for both the normal covering condition and intermittent covering condition. The three conditions shown are BayesHeart with 2-state model only, with 4-state model only and with adaptive HMMs.

Two-way within-subjects ANOVA shows that both signal types ($F_{1,19}$=75.36, p<0.01) and the model selection techniques ($F_{2,18}$=16.25, p<0.01) have a significant main effect on MER. Pairwise comparisons also show significant differences in MER between adaptive HMMs and 2-state model only. In summary, the model selection technique in BayesHeart leads to the lowest mean error rates in both normal (3.66% vs. 4.08% vs. 4.84%) and intermittent covering conditions (5.39% vs. 5.84% vs. 6.90%).

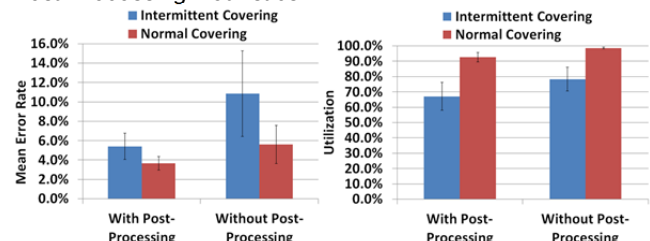*Post-Processing Heuristics*

Figure 12: (left) Mean error rate (right) and utilization rate of BayesHeart with/without post-processing and with normal/intermittent covering data.

Figure 12 shows the impact of BayesHeart post-processing on MER and utilization rates. The post-processing step eliminates outliers of cardiac intervals. By eliminating unreasonable estimates, it leads to the reduction of utilization rate (Figure 12.b, 11.2% and 5.96% reductions in intermittent covering condition and normal covering condition, respectively). Accordingly, it improves the accuracy (Figure 12.a), especially for intermittent signals (5.39% vs. 10.86%). Two-way within-subject ANOVA shows that the MER can be significantly affected by the inclusion of post-processing ($F_{1,19}$=98.96, p<0.01).
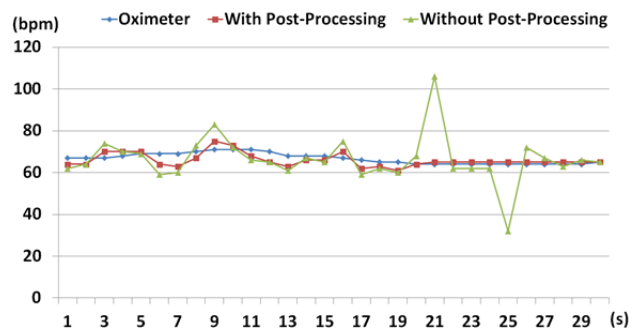
Figure 13. Heart rate estimates generated by the pulse oximeter and our algorithm during 30 seconds.

Figure 13 is a 30-second sample sequence, illustrating heart rate estimates from BayesHeart (both with and without post-processing) against the gold standard.
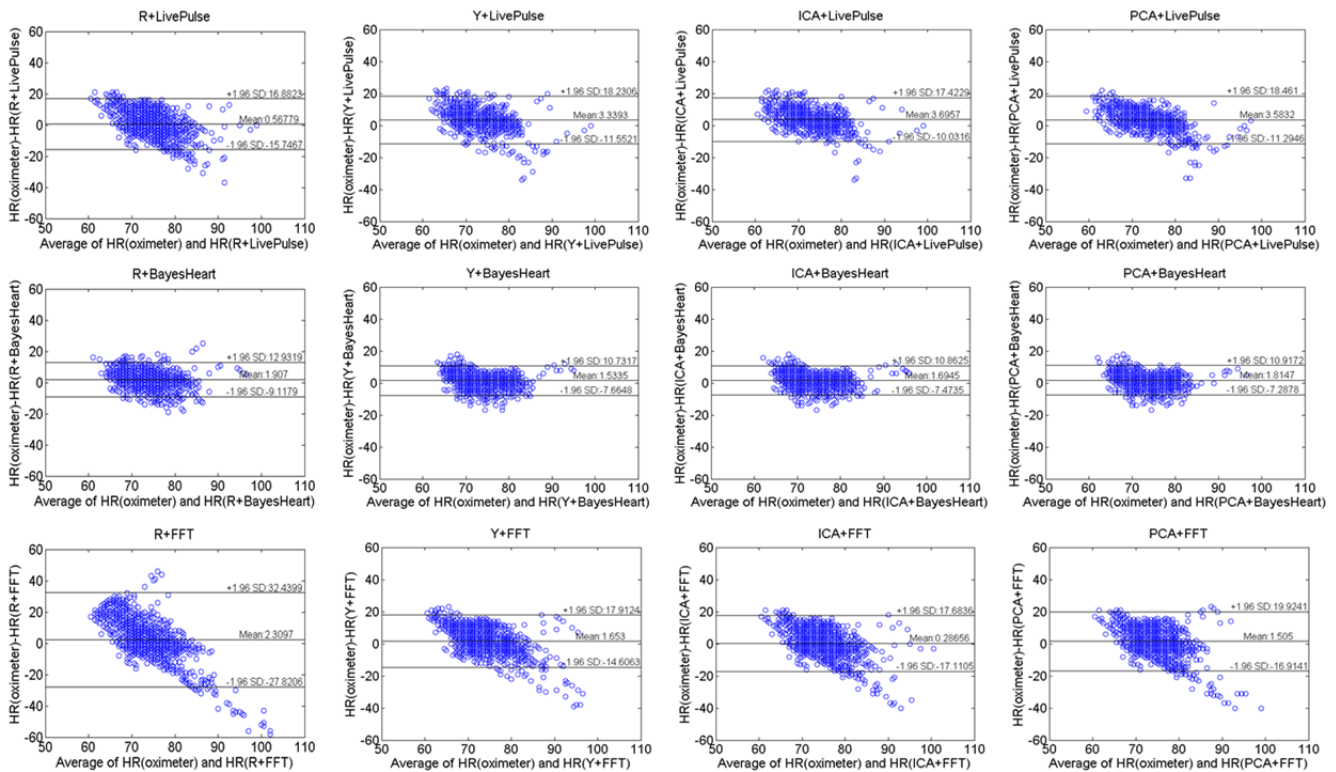
Figure 14. Bland-Altman plots demonstrating the agreement between heart rate measurements obtained from the nine state-of-the-art algorithms (with intermittent covering data) and the pulse oximeter. The lines represent the mean and 95% limits of agreement.

Design Space Exploration

We conducted a comparative study of twelve state-of-the-art algorithms (Figure 2) in the design space of extracting cardiac pulses from commodity cameras. Such a study is important because existing literature [3, 4, 14, 19, 27] focused primarily on the *feasibility* and *non-comparative evaluation* of proposed scenarios. With the popularization of wearable devices and affect/emotion-aware intelligent mobile apps, it is imperative for researchers to gain a deeper understanding of the state of the art. To the best of our knowledge, this is the first systematic study and exploration in commodity camera based heart rate monitoring.

We explore a space of 4 noise reduction techniques by 3 pulse counting techniques for a total of 12 algorithms.

Noise reduction methods we investigated include: **1**) Using the Red channel only (baseline condition [33]); **2**) Using the Y (brightness) channel (static-weighted sum of the R, G, B channels); **3**) Using the ICA technique (dynamic-weighted sum of the R, G, B channels by maximizing channel independence); **4**) Using the PCA technique [2] (dynamic-weighted sum of the R, G, B channels by maximizing channel variance).

Pulse counting techniques include: **A**) LivePulse (temporal domain, heuristic based counting)[3] [14]; **B**) FFT (frequency domain counting, window size = 6 sec[4]); **C**) BayesHeart (temporal domain, probabilistic model based alignment).

Ignoring subtle variations in signal preprocessing and post-processing, existing algorithms can be represented as combining one noise reduction technique and one pulse counting technique. For example, the facial color based method by Poh et al [27, 28] can be represented as **3B** (their baseline condition was **1B**). **The** facial motion based method by Balakrishnan et al [3] is **4A**[5]**.** LivePulse Games [14] used **2A** and the default BayesHeart algorithm is **2C.** Such a comparison will answer questions such as: Will the PCA/ICA based noise reduction technique be effective for detecting pulse amidst motion induced noise? Will adding a PCA/ICA based noise reduction technique improve the performance of BayesHeart even more? Quantitatively,

---

[2] We chose the most periodic channel by a method discussed in [3]. The periodicity of a signal is defined as the percentage of total spectral power accounted for by the frequency with maximal power.

[3] The LivePulse algorithm [14] is a heuristic based outlier removal and local peaks/valleys counting algorithm. LivePulse can be treated as a manually optimized, temporal domain adaptive thresholding algorithm.

[4] The choice of window size involves tradeoffs between frequency resolution, time accuracy, and latency.

[5] The original signal in [3] was head motion, but the same algorithm can be used to process PPG signals and vice-versa.

what's the impact of an algorithm chosen in each step on the overall performance?

We tested the 12 combinations of algorithms with both normal and intermittent covering signals and analyzed both the accuracy and latency for each method. Considering the importance of post-processing heuristics shown in previous sections, in order to minimize confounding factors in the study, we applied the same post-processing heuristics used in the default BayesHeart algorithm in all 12 algorithms.
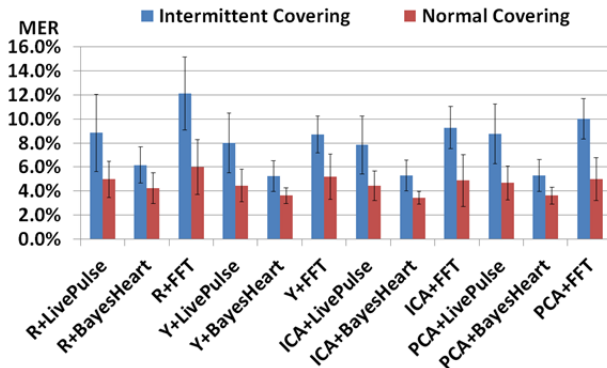

Figure 15. Mean error rates (MER) of algorithms.

Accuracy

Figure 14 shows the Bland-Altman plots demonstrating the agreement between heart rate estimates generated by the 12 algorithms and gold standard with intermittent covering data. The lines represent the mean and 95% limits of agreement. By comparing between different columns (i.e., different noise reduction techniques) we can see that using Y(**2**)/ICA(**3**)/PCA(**4**) could reduce the error compared to using R(**1**) directly. For example, when using BayesHeart on R channel, the mean bias was 1.91 bpm with 95% limits of agreement -9.12 to 12.93 bpm. The mean bias was reduced to 1.53 bpm with 95% limits of agreement -7.66 to 10.73 bpm when using Y(**2**). At the same time, by comparing different rows (i.e., different pulse counting methods) we can find that BayesHeart(**C**) can lower the errors compared to LivePulse(**A**)/FFT(**B**). For example, when applied on Y(**2**), LivePulse(**A**)'s mean bias was 3.34 bpm with 95% limits of agreement -11.55 to 18.23 bpm and FFT(**B**)'s mean bias was 1.61 bpm with 95% limits of agreement -14.61 to 17.91 bpm. In comparison, BayesHeart(**C**) reduced the mean bias to 1.53 bpm with 95% agreement -7.66 to 10.73 bpm.

Figure 15 shows the corresponding MERs. For intermittent covering, Y+BayesHeart(**2C**) has the lowest MER (5.23%), followed by PCA+BayesHeart(**4C**) (5.30%) and ICA+BayesHeart(**3C**) (5.31%). R+FFT(**1B**) has the highest MER (12.14%). For normal covering, ICA+BayesHeart(**3C**) has the lowest MER (3.44%), followed by PCA+BayesHeart(**4C**) (3.63%) and Y+BayesHeart(**2C**) (3.66%). A three-way (signal type vs. noise reduction method vs. pulse counting method) ANOVA shows that signal type has a significant main effect on MERs

($F_{1,19}$=620.94, p<0.01). Noise reduction techniques ($F_{3,17}$=38.29, p=0.02) and pulse counting methods ($F_{2,18}$=257.55, p<0.01) also have a significant main effect on MERs. Not surprisingly, all the 12 algorithms are significantly more accurate when dealing with normal signals.

Among noise reduction techniques, the mean error rates corresponding to Y(**2**), ICA(**3**), PCA(**4**) and R(**1**) (averaged in both signal types) are 5.87%, 5.86%, 6.23% and 7.06%, respectively. Pairwise comparisons show that both the static weighted sum approach in Y(**2**) and the dynamic weight sum approach in ICA(**3**)/PCA(**4**) are significantly better than R(**1**) despite signal quality (p<0.01). Such improvements may be caused by the increased equivalent pixel area in Y(**2**)/ ICA(**3**)/PCA(**4**). Although ICA(**3**) has a lower MER than Y(**2**), the difference is not significant (p=0.92). We attribute that to the non-linear nature of skin/tissue reflection and the latency involved in calculating the transformation matrix, which was in turn used to capture the dynamic nature of extrinsic noises.
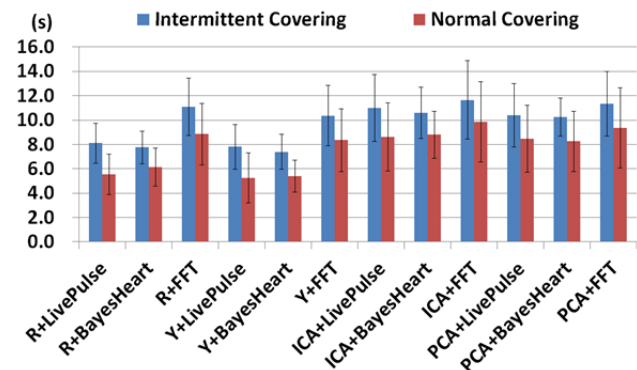

Figure 16. Average latency of algorithms.

Among pulse counting techniques, the mean error rates corresponding to LivePulse(**A**), FFT(**B**), and BayesHeart(**C**) (averaged in both signal types) are 6.50%, 7.65% and 4.62%, respectively. Pairwise comparisons show that BayesHeart can lower MERs significantly compared with LivePulse (p=0.01). And LivePulse can also significantly lower MERs compared with FFT (p<0.01). The reasons include: 1) the heuristic based method (e.g., LivePulse), although simple, could not capture both the diversity and regularity of signal with dealing with increased amount of signals; 2) FFT is the most sensitive technology to noise; the low-sampling rate (30HZ) could be one reason that led to the bad performance of FFT. Besides, the fixed-size window in FFT may also introduce increased noise when dealing with brief, highly intermittent signals; 3) BayesHeart exploits additional information in trellis structure, the state transition cost, and temporal regularity in signals through a simple yet robust probabilistic model; such increased "*signal/noise ratio*" becomes critical when dealing with increased extrinsic noise.

Estimation Latency

Figure 16 shows the average estimation latency of the 12 methods. For intermittent covering, Y+BayesHeart(**2C**) leads to the lowest latency (7.59s), followed by R+BayesHeart(**1C**) (8.02s) and Y+LivePulse(**2A**) (8.23s). ICA+FFT(**3B**) has the highest latency (12.21s). For normal covering, Y+LivePulse(**2A**) has the lowest latency (5.72s), followed by Y+BayesHeart(**2C**) (5.79s) and R+LivePulse(**1A**) (6.04s). PCA+FFT(**4B**) has the highest latency (9.88s).
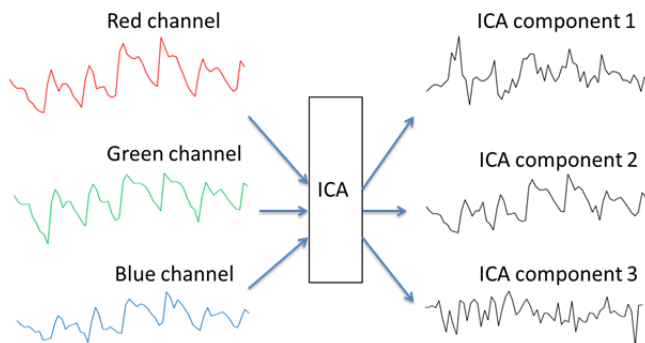
.

Figure 17. The components generated from ICA based noise reduction algorithm on 5-second R/G/B signal.

Three-way ANOVA shows that signal quality ($F_{1,19}=411.44$, $p<0.01$), noise reduction techniques ($F_{3,17}=73.43$, $p<0.01$) and pulse counting methods ($F_{2,18}=55.51$, $p<0.01$) all have a significant main effect on estimation latency. Pair-wise comparisons show that R(**1**) (8.42s) and Y(**2**) (7.84s) are significantly faster than ICA(**3**) (10.51s) and PCA(**4**) (10.12s) ($p < 0.01$). And there is no significant difference in latency between R(**1**) and Y(**2**), between ICA(**3**) and PCA(**4**). This is because ICA(**3**)/PCA(**4**) require additional time to construct and update the transformation matrix in use. In our experiments, ICA(**3**)/PCA(**4**) lead to 3-10 seconds of additional latency for normal signals or 5-12 seconds of additional latency for intermittent signals. Figure 17 shows an example of the components generated from ICA based on a 5-second normal covering R/G/B signal. We can see that when the transformation matrix in ICA(**3**) is outdated or out-of-sync, ICA could lead to *increased* noise

For pulse-counting methods, pairwise comparisons show that LivePulse and BayesHeart can significantly lower latencies when compared with FFT ($p<0.01$). This is because the relatively small number of signal samples and low sampling rate have a negative impact on the frequency domain resolution of FFT. Such low frequency domain resolution leads to less accurate estimates. Therefore the corresponding algorithms require more samples in order to derive accurate estimates.

CONCLUSION AND FUTURE WORK

We present BayesHeart, an accurate, low-latency probabilistic approach for heart rate monitoring via commodity cameras. Major contributions of the paper include: 1) we demonstrated both the feasibility and the quantitative performance of BayesHeart to measure heart rate via camera phone. 2) We reported and discussed the tradeoffs of BayesHeart in detail, especially *how* and *why* BayesHeart is robust to noisy and intermittent signals. 3) By decoupling existing camera based heart rate monitoring techniques into two steps, i.e. noisy reduction and cardiac pulse counting, we compared existing technologies side-by-side highlighting both their relationships and new opportunities. 4) In a 20-subject experiment, we systematically evaluated the state-of-the-art algorithms covering the design space regarding accuracy and latency performance.

BayesHeart is an initial step towards capturing, analyzing, and using physiological signals implicitly from everyday interactions. We believe there are tremendous opportunities if computer user interfaces could be aware of our physiological conditions [34], affect, and emotions. For example, when integrated with a Massive Open Online Courses (MOOC) mobile client, BayesHeart can be used infer the stress levels and cognitive workload of learners. Taking advantage of such information will be beneficial to both learners and instructors. Error correction algorithms in today's software touch keyboards (STK) can adapt to users' stress levels and cognitive workload in addition to holding postures [38] for better accuracy and less frustration. We hope the algorithmic design space we identified could inspire new ideas for the research community in the future.

In addition to accurately quantifying the average duration of each cardiac cycle, BayesHeart also provides opportunities to analyze different sub-phases within each cycle. The static prosperities, relative distributions, and temporal variations of sub-phases may be used as physiological makers of users' emotional states, health conditions, and even personal identities.

To encourage follow-up research and more creative usage of physiological signals in everyday interactions, we have released both the source code (written in Matlab and Java) of BayesHeart and data used in our comparative study to the public under a BSD license at URL http://mips.lrdc.pitt.edu/BayesHeart .

REFERENCES
1. Android Wear Smart Watch OS, http://developer.android.com/wear

2. Apple Watch https://www.apple.com/watch/

3. Balakrishnan, G., Durand, F., & Guttag, J. Detecting Pulse from Head Motions in Video. In *Proc. CVPR 2013*, 3430-3437.

4. Banitsas, K., Pelegris, P., Orbach, T., Cavouras, D., Sidiropoulos, K., & Kostopoulos, S. A simple algorithm to monitor hr for real time treatment applications. In *Proc. ITAB 2009,* 1-5.

5. Cardiograph for iOS, https://itunes.apple.com/us/app/cardiograph/id441079429?ls=1&mt=8

6. Choi, J., & Gutierrez-Osuna, R. Using heart rate monitors to detect mental stress. In *BSN 2009. Sixth International Workshop,* 219-223.

7. Coast, D. A., Stern, R. M., Cano, G. G., & Briller, S. A. An approach to cardiac arrhythmia analysis using hidden Markov models. *Biomedical Engineering, IEEE Transactions on* 37.9 (1990): 826-836.

8. Cook, S., Togni, M., et al, High heart rate: a cardiovascular risk factor?, European heart journal, vol. 27(20), 2006.

9. Cooper, K., Pollock, M., et al, Physical Fitness Levels vs Selected Coronary Risk Factors A Cross-Sectional Study, The Journal of the American Medical Association (JAMA), vol. 236, No. 2, July, 1976.

10. Fox, K., Borer, J. S., Camm, A. J., Danchin, N., Ferrari, R., Sendon, J. L. L., ... & Tendera, M. (2007). Resting heart rate in cardiovascular disease. *Journal of the American College of Cardiology*, *50*(9), 823-830.

11. Garbey, M., Sun, N., Merla, A., & Pavlidis, I. Contact-free measurement of cardiac pulse based on the analysis of thermal imagery. Biomedical Engineering, IEEE Transactions on 54.8 (2007): 1418-1426.

12. Gregoski, M. J., Mueller, M., Vertegel, A., Shaporev, A., Jackson, B. B., Frenzel, R. M., ... & Treiber, F. A. Development and validation of a smartphone heart rate acquisition application for health promotion and wellness telehealth applications. *International journal of telemedicine and applications*,*2012*, 1.

13. Grimaldi, D., Kurylyak, Y., Lamonaca, F., & Nastro, A. Photoplethysmography detection by smartphone's videocamera. In *Proc. IDAACS 2011,* 488-491.

14. Han, T., Xiao, X., Shi, L., Canny, J., & Wang, J. Balancing Accuracy and Fun: Designing Engaging Camera Based Mobile Games for Implicit Heart Rate Monitoring. In *Proc. of CHI 2015*.

15. Hughes, N. P., Tarassenko L., and Roberts S.J. Markov Models for Automated ECG Interval Analysis. In *NIPS*. 2003.

16. Instant Heart Rate for iOS, https://itunes.apple.com/app/instant-heart-rate-measure/id395042892?mt=8

17. Janz, K., Dawson, J., Mahoney, L., Tracking physical fitness and physical activity from childhood to adolescence: the Muscatine study, Medicine and Science in Sports and Exercise, vol. 32(7), 2000.

18. Jensen, M. T., Suadicani, P., Hein, H. O., & Gyntelberg, F. (2013). Elevated resting heart rate, physical fitness and all-cause mortality: a 16-year follow-up in the Copenhagen Male Study. *Heart*, *99*(12), 882-887.

19. Jonathan, E., & Leahy, M. (2010). Investigating a smartphone imaging unit for photoplethysmography. *Physiological measurement*, *31*(11), N79.

20. Jraidi, I., Chaouachi, M., Frasson, C., A Dynamic Multimodal Approach for Assessing Learners' Interaction Experience, In Proc. ACM ICMI 2013.

21. Mark, G., et al, Stress and Multitasking in Everyday College Life: An Empirical Study of Online Activity, In Proc. CHI 2014.

22. McGhee, B. H., & Bridges, E. J. (2002). Monitoring arterial blood pressure: what you may not know. *Critical Care Nurse*, *22*(2), 60-79.

23. Mueller, F., Vetere, F., et al, Jogging over a Distance between Europe and Australia, In Proc. UIST 2010.

24. Pan, S. T., Hong, T. P., & Chen, H. C. Ecg signal analysis by using hidden markov model. In *Proc. iFUZZY 2012,* 288-293.

25. Pelegris, P., Banitsas, K., Orbach, T., & Marias, K. A novel method to detect heart beat rate using a mobile phone. In *Proc. EMBC 2010,* 5488-5491.

26. Picard, R.W., Vyzas, E., Healey, J.: Toward Machine Emotional Intelligence: Analysis of Affective Physiological State. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(10), 2001.

27. Poh, M. Z., McDuff, D. J., & Picard, R. W. (2010). Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics Express*, *18*(10), 10762-10774.

28. Poh, M. Z., McDuff, D., & Picard, R. A medical mirror for non-contact health monitoring. In *ACM SIGGRAPH 2011 Emerging Technologies*, 2.

29. Rabiner, Lawrence R. A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE 77.2 (1989): 257-286.

30. Rowe, D.R., Sibert, J., and Irwin, D. Heart Rate Variability: Indicator of User State as an Aid to Human-Computer Interaction. In Proc. CHI 1998.

31. Scully, C., Lee, J., Meyer, J., Gorbach, A. M., Granquist-Fraser, D., Mendelson, Y., & Chon, K. H. (2012). Physiological parameter monitoring from optical recordings with a mobile phone. *Biomedical Engineering, IEEE Transactions on*, *59*(2), 303-306.

32. Seccareccia, F., Pannozzo, F., Dima, F., Minoprio, A., Menditto, A., Lo Noce, C., & Giampaoli, S. (2001). Heart rate as a predictor of mortality: the MATISS project. *American Journal of Public Health*, *91*(8), 1258-1263.

33. Singh, R. B., Nakade, P. K., & Bhosale, P. R. Heart Rate Measurement Through Photoplethysmography, In *Proc. BEATS 2010*, 170-174.

34. Sun, D., Canny, J., Paredes, P., MouStress: Detecting Stress from Mouse Motion, In Proc. CHI 2014.

35. Wit, Ernst, Edwin van den Heuvel, and Jan‐Willem Romeijn. 'All models are wrong...': an introduction to model uncertainty. *Statistica Neerlandica* 66.3 (2012): 217-236.

36. Woolf, B., Burleson, W., Arroyo, I., Dragon, T., Cooper, D., and Picard, R., Affect-aware Tutors: Recognising and Responding to Student Affect. Intl J Learning Technology 4, 3 (2009).

37. Xiao, X., Han, T., & Wang, J. LensGesture: augmenting mobile interactions with back-of-device finger gestures. In *Proc. ICMI 2013*, 287-294.

38. Yin, Y., Ouyang, T., Partridge, K., Zhai, S., Making touchscreen keyboards adaptive to keys, hand postures, and individuals: a hierarchical spatial backoff model approach, In *Proc. CHI 2013*, 2775-2784.