Emotion Recognition using Wireless Signals

by

Mingmin Zhao

B.S. in Computer Science
Peking University, 2015

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2017

Signature redacted

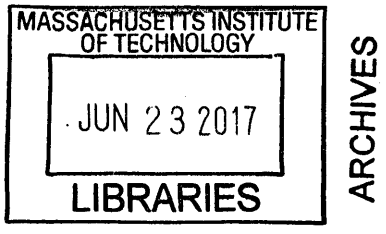
Author .....
Department of Electrical Engineering and Computer Science
May 19, 2017

Signature redacted

Certified by .....
Dina Katabi
Andrew and Erna Viterbi Professor of Computer Science
Thesis Supervisor

Signature redacted

Accepted by .....
Leslie A. Kolodziejski
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students



Emotion Recognition using Wireless Signals

by

Mingmin Zhao

Submitted to the Department of Electrical Engineering and Computer Science
on May 19, 2017, in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering and Computer Science

Abstract

This thesis demonstrates a new technology that can infer a person's emotions from RF signals reflected off his body. EQ-Radio transmits an RF signal and analyzes its reflections off a person's body to recognize his emotional state (happy, sad, etc.). The key enabler underlying EQ-Radio is a new algorithm for extracting the individual heartbeats from the wireless signal at an accuracy comparable to on-body ECG monitors. The resulting beats are then used to compute emotion-dependent features which feed a machine-learning emotion classifier. We describe the design and implementation of EQ-Radio, and demonstrate through a user study that its emotion recognition accuracy is on par with state-of-the-art emotion recognition systems that require a person to be hooked to an ECG monitor.

Thesis Supervisor: Dina Katabi
Title: Andrew and Erna Viterbi Professor of Computer Science

acknowledgments

I would like to acknowledge the contributions of Prof. Dina Katabi and Prof. Fadel Adib, without whom this work would not have been possible.

I am grateful to the members of the NETMIT for their insightful discussions and to all the human subjects for their participation in our experiments.

I also would like to acknowledge my family members and my girl friend, Yining Wang for their love and support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Emotion recognition is an emerging field that has attracted much interest from both the industry and the research community [1, 2, 3, 4, 5]. It is motivated by a simple vision: Can we build machines that sense our emotions? If we can, such machines would enable smart homes that react to our moods and adjust the lighting or music accordingly. Movie makers would have better tools to evaluate user experience. Advertisers would learn customer reaction immediately. Computers would automatically detect symptoms of depression, anxiety, and bipolar disorder, allowing early response to such conditions. More broadly, machines would no longer be limited to explicit commands, and could interact with people in a manner more similar to how we interact with each other.

Existing approaches for inferring a person's emotions either rely on audiovisual cues, such as images and audio clips [6, 3, 7], or require the person to wear physiological sensors like an ECG monitor [8, 9, 10, 11]. Both approaches have their limitations. Audiovisual techniques leverage the outward expression of emotions, but cannot measure inner feelings [12, 9, 13]. For example, a person may be happy even if she is not smiling, or smiling even if she is not happy. Also, people differ widely in how expressive they are in showing their inner emotions, which further complicates this problem [14]. The second approach recognizes emotions by monitoring the physiological signals that change with our emotional state, e.g., our heartbeats. It uses on-body sensors – e.g., ECG monitors – to measure these signals and correlate their changes with joy, anger, etc. This approach is more correlated with the person's inner feelings since it taps into the interaction between the autonomic nervous system and the heart rhythm [15, 16]. However, the use of body sensors is cumbersome and can interfere with user activity and emotions, making this approach unsuitable for regular usage.

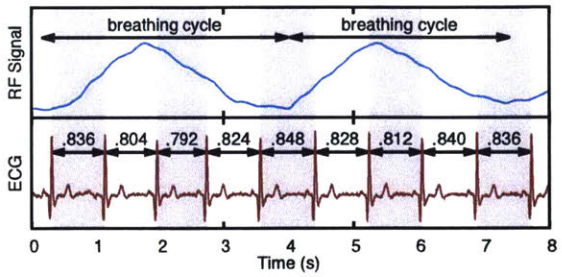


Figure 1-1: **Comparison of RF signal with ECG signal.** The top graph plots the RF signal reflected off a person's body. The envelop of the RF signal follows the inhale-exhale motion. The small dents in the signal are due to heartbeats. The bottom graph plots the ECG of the subject measured concurrently with the RF signal. Individual beats are marked by grey and white shades. The numbers report the beat-length in seconds. Note the small variations in consecutive beat lengths.

In this thesis, we introduce a new method for emotion recognition that achieves the best of both worlds –i.e., it directly measures the interaction of emotions and physiological signals, but does not require the user to carry sensors on his body.

Our design uses RF signals to sense emotions. Specifically, RF signals reflect off the human body and get modulated with bodily movements. Recent research has shown that such RF reflections can be used to measure a person's breathing and average heart rate without body contact [17, 18, 19, 20, 21]. However, the periodicity of the heart signal (i.e., its running *average*) is of little relevance to emotion recognition. Specifically, to recognize emotions, we need to measure the minute variations in *each individual beat length* [15, 22, 12].

Yet, extracting individual heartbeats from RF signals incurs multiple challenges, which can be seen in Fig. 1-1. First, RF signals reflected off a person's body are modulated by both breathing and heartbeats. The impact of breathing is typically orders of magnitude larger than that of heartbeats, and tends to mask the individual beats (see the top graph in Fig. 1-1); to separate breathing from heart rate, past systems operate over multiple seconds (e.g., 30 seconds in [17]) in the frequency domain, forgoing the ability to measure the beat-to-beat variability. Second, heartbeats in the RF signal lack the sharp peaks which characterize the ECG signal, making it harder to accurately identify beat boundaries. Third, the difference in inter-beat-intervals (IBI) is only a few tens of milliseconds. Thus, individual beats have to be segmented to within a few milliseconds. Obtaining such accuracy is particularly difficult in the absence of sharp features that identify the beginning or end of a heartbeat. Our goal is to address these challenges to enable RF-based emotion recognition.

We present EQ-Radio, a wireless system that performs emotion recognition using RF reflections off a person's body. EQ-Radio's key enabler is a new algorithm for extracting

individual heartbeats and their differences from RF signals. Our algorithm first mitigates the impact of breathing. The intuition underlying our mitigation mechanism is as follows: while chest displacement due to the inhale-exhale process is orders of magnitude larger than minute vibrations due to heartbeats, the acceleration of breathing is smaller than that of heartbeats. This is because breathing is usually slow and steady while a heartbeat involves rapid contraction of the muscles (which happen at localized instances in time). Hence, EQ-Radio operates on the acceleration of RF signals to dampen the breathing signal and emphasize the heartbeats.

Next, EQ-Radio needs to segment the RF reflection into individual heartbeats. In contrast to the ECG signal which has a known expected shape (see the bottom graph in Fig. 1-1), the shape of a heartbeat in RF reflections is unknown and varies depending on the person's body and exact posture with respect to the device. Thus, we cannot simply look for a known shape as we segment the signal; we need to learn the beat shape as we perform the segmentation. We formulate the problem as a joint optimization, where we iterate between two sub-problems: the first sub-problem learns a template of the heartbeat given a particular segmentation, while the second finds the segmentation that maximizes resemblance to the learned template. We keep iterating between the two sub-problems until we converge to the best beat template and the optimal segmentation that maximizes resemblance to the template. Finally, we note that our segmentation takes into account that beats can shrink and expand and hence vary in beat length. Thus, the algorithm finds the beat segmentation that maximizes the similarity in the morphology of a heartbeat signal across consecutive beats while allowing for flexible warping (shrinking or expansion) of the beat signal.

We have built EQ-Radio into a full-fledged emotion recognition system. EQ-Radio's system architecture has three components: The first component is an FMCW radio that transmits RF signals and receives their reflections. The radio leverages the approach in [17] to zoom in on human reflections and ignore reflections from other objects in the scene. Next, the resulting RF signal is passed to the beat extraction algorithm described above. The algorithm returns a series of signal segments that correspond to the individual heartbeats. Finally, the heartbeats – along with the captured breathing patterns from RF reflections – are passed to an emotion classification sub-system as if they were extracted from an ECG monitor. The emotion classification sub-system computes heartbeat-based and respiration-based features recommended in the literature [10, 12, 9] and uses an SVM classifier to differentiate among various emotional states.

We evaluate EQ-Radio by conducting user experiments with 30 subjects. We design our experiments in accordance with the literature in the field [10, 12, 9]. Specifically, the subject is asked to evoke a particular emotion by recalling a corresponding memory (e.g., sad or happy memories). She/he may use music or photos to help evoking the appropriate memory. In each experiment, the subject reports the emotion she/he felt, and the period during which she/he felt that emotion. During the experiment, the subject is monitored using both EQ-Radio and a commercial ECG monitor. Further, a video is taken of the subject then passed to the Microsoft image-based emotion recognition system [23].

Our experiments show that EQ-Radio's emotion recognition is on par with state-of-the-art ECG-based systems, which require on-body sensors [8]. Specifically, if the system is trained on each subject separately, the accuracy of emotion classification is 87% in EQ-Radio and 88.2% in the ECG-based system. If one classifier is used for all subjects, the accuracy is 72.3% in EQ-Radio and 73.2% in the ECG-based system.[1] For the same experiments, the accuracy of the image-based system is 39.5%; this is because the image-based system performed poorly when the emotion was not visible on the subject's face.

Our results also show that EQ-Radio's performance is due to its ability to accurately extract heartbeats from RF signals. Specifically, even errors of 40-50 milliseconds in estimating heartbeat intervals would reduce the emotion recognition accuracy to 44% (as we show in Fig. 7-8 in §7.3). In contrast, our algorithm achieves an average error in inter-beat-intervals (IBI) of 3.2 milliseconds, which is less than 0.4% of the average beat length.

1.1 Contributions

This thesis makes three contributions:

- To our knowledge, this is the first work that demonstrates the feasibility of emotion recognition using RF reflections off one's body. As such, this work both expands the scope of wireless systems and advances the field of emotion recognition.
- This work introduces a new algorithm for extracting individual heartbeats from RF reflections off the human body. The algorithm presents a new mathematical formulation of the problem, and is shown to perform well in practice.
- This work also presents a user study of the accuracy of emotion recognition using RF reflections, and an empirical comparison with both ECG-based and image-based emotion recognition systems.

[1] The ECG-based system and EQ-Radio use exactly the same classification features but differ in how they obtain the heartbeat series. In all experiments, training and testing are done on different data.

Chapter 2

Background & Related Work

Emotion Recognition: Recent years have witnessed a growing interest in systems capable of inferring user emotions and reacting to them [24, 25]. Such systems can be used for designing and testing games, movies, advertisement, online content, and human-computer interfaces [26, 27]. These systems operate in two stages: first, they extract emotion-related signals (e.g., audio-visual cues or physiological signals); second, they feed these signals into a classifier in order to recognize emotions. Below, we describe prior art for each of these stages.

Existing approaches for extracting emotion-related signals fall under two categories: audiovisual techniques and physiological techniques. Audiovisual techniques rely on facial expressions, speech, and gestures [6, 28]. The advantage of these approaches is that they do not require users to wear any sensors on their bodies. However, because they rely on outwardly expressed states, they tend to miss subtle emotions and can be easily controlled or suppressed [10]. Moreover, vision-based techniques require the user to face a camera in order for them to operate correctly. On the other hand, physiological measurements, such as ECG and EEG signals, are more robust because they are controlled by involuntary activations of the autonomic nervous system (ANS) [29]. However, existing sensors that can extract these signals require physical contact with a person's body, and hence interfere with the user experience and affect her emotional state. In contrast, EQ-Radio can capture physiological signals without requiring the user to wear any sensors by relying purely on wireless signals reflected off her/his body.

The second stage in emotion recognition systems involves extracting emotion-related features from the measured signals and feeding these features into a classifier to identify a user's emotional state. There is a large literature on extracting such features from both au-

diovisual and physiological measurements [30, 31, 32]. Beyond feature extraction, existing classification approaches fall under two categories. The first approach gives each emotion a discrete label: e.g., pleasure, sadness, or anger. The second approach uses a multidimensional model that expresses emotions in a 2D-plane spanned by *valence* (i.e., positive vs. negative feeling) and *arousal* (i.e., calm vs. charged up) axes [33, 10]. For example, anger and sadness are both negative feelings, but anger involves more arousal. Similarly, joy and pleasure are both positive feelings, but the former is associated with excitement whereas the latter refers to a state of contentment. EQ-Radio adopts the valence-arousal model and builds on past foundations to enable emotion recognition using RF signals.

Finally, another class of emotion recognition techniques relies on smartphone usage patterns (calling, application usage, etc.) to infer user daily mood or personality [34, 35]; however, these techniques operate at much large time scales (days or months) than EQ-Radio, which recognizes emotions at minute-scale intervals.

RF-based Sensing: RF signals reflect off the human body and are modulated by body motion. Past work leverages this phenomenon to sense human motion: it transmits an RF signal and analyzes its reflections to track user locations [36], gestures [37, 38, 39, 40, 41, 42], activities [43, 44], and vital signs [17, 18, 45]. Past proposals also differ in the transmitted RF signals: Doppler radar [18, 45], FMCW [36, 17], and WiFi [37, 38]. Among these techniques, FMCW has the advantage of separating different sources of motion in the environment. Thus, FMCW is more robust for extracting vital signs and enables monitoring multiple users simultaneously; hence, EQ-Radio uses FMCW signals for capturing human reflections.

Our work is closest to prior art that uses RF signals to extract a person's breathing rate and average heart rate [18, 45, 19, 20, 21, 18, 45, 19, 20, 21, 46, 47, 17]. In contrast to this past work, which recovers the average period of a heartbeat (which is of the order of a second), emotion recognition requires extracting the individual heartbeats and measuring small variations in the beat-to-beat intervals with millisecond-scale accuracy. Unfortunately, prior research that aims to segment RF reflections into individual beats either cannot achieve sufficient accuracy for emotion recognition [48, 49, 50] or requires the monitored subjects to hold their breath [51]. In particular, past work that does not require users to hold their breath has an average error of 30-50 milliseconds [50, 48, 49], which is of the same order or larger than the variations in the beat-to-beat intervals themselves, precluding emotion recognition (as we show empirically in §7.3). EQ-Radio's heartbeat segmentation algorithm builds on this past literature yet recovers heartbeats with a mean accuracy of 3.2 milliseconds, hence achieving an order of magnitude reduction in errors in comparison

to past techniques. This high accuracy is what enables us to deliver the first emotion recognition system that relies purely on wireless signals.

Chapter 3

EQ-Radio Overview

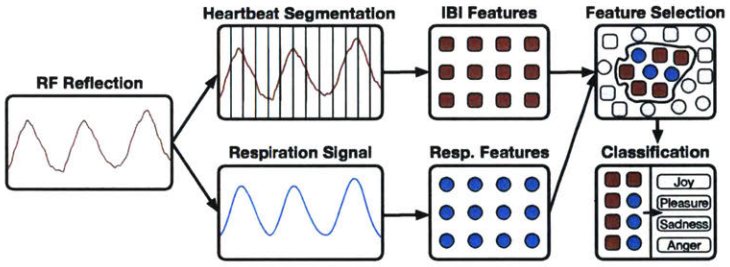


Figure 3-1: **EQ-Radio Architecture.** EQ-Radio has three components: a radio for capturing RF reflections (§4), a heartbeat extraction algorithm (§5), and a classification subsystem that maps the learned physiological signals to emotional states (§6).

EQ-Radio is an emotion recognition system that relies purely on wireless signals. It operates by transmitting an RF signal and capturing its reflections off a person's body. It then analyzes these reflections to infer the person's emotional state. It classifies the person's emotional state according to the known arousal-valence model into one of four basic emotions [33, 10]: anger, sadness, joy, and pleasure (i.e., contentment).

EQ-Radio's system architecture consists of three components that operate in a pipelined manner, as shown in Fig. 3-1:

- *An FMCW radio,* which transmits RF signals and captures their reflections off a person's body.
- *A beat extraction algorithm,* which takes the captured reflections as input and returns a series of signal segments that correspond to the person's individual heartbeats.
- *An emotion-classification subsystem,* which computes emotion-relevant features from the captured physiological signals – i.e., the person's breathing pattern and heartbeats – and uses these features to recognize the person's

emotional state.

In the following sections, we describe each of these components in detail.

Chapter 4

Capturing the RF Signal

EQ-Radio operates on RF reflections off the human body. To capture such reflections, EQ-Radio uses a radar technique called Frequency Modulated Carrier Waves (FMCW) [36]. There is a significant literature on FMCW radios and their use for obtaining an RF signal that is modulated by breathing and heartbeats [17, 52, 53]. We refer the reader to [17] for a detailed description of such methods, and summarize below the basic information relevant to this thesis.

The radio transmits a low power signal and measures its reflection time. It separates RF reflections from different objects/bodies into buckets based on their reflection time. It then eliminates reflections from static objects which do not change across time and zooms in on human reflections. It focuses on time periods when the person is quasi-static. It then looks at the phase of the RF wave which is related to the traveled distance as follows [54]:

$$\phi(t) = 2\pi \frac{d(t)}{\lambda},$$

where $\phi(t)$ is the phase of the signal, λ is the wavelength, $d(t)$ is the traveled distance, and t is the time variable. The variations in the phase correspond to the compound displacement caused by chest expansion and contraction due to breathing, and body vibration due to heartbeats.[1]

The phase of the RF signal is illustrated in the top graph in Fig. 1-1. The envelop shows the chest displacements as the inhale-exhale process. The small dents are due to minute skin vibrations associated with blood pulsing. EQ-Radio operates on this phase signal.

[1] When blood is ejected from the heart, it exercises a force on the rest of the body causing small jitters in the head and skin, which are picked up by the RF signal [17].

Chapter 5

Beat Extraction Algorithm

Recall that a person's emotions are correlated with small variations in her/his heartbeat intervals; hence, to recognize emotions, EQ-Radio needs to extract these intervals from the RF phase signal described above.

The main challenge in extracting heartbeat intervals is that the morphology of heartbeats in the reflected RF signals is unknown. Said differently, EQ-Radio does not know how these beats look like in the reflected RF signals. Specifically, these beats result in distance variations in the reflected signals, but the measured displacement depends on numerous factors including the person's body and her exact posture with respect to EQ-Radio's antennas. This is in contrast to ECG signals where the morphology of heartbeats has a known expected shape, and simple peak detection algorithms can extract the beat-to-beat intervals. However, because we do not know the morphology of these heartbeats in RF a priori, we cannot determine when a heartbeat starts and when it ends, and hence we cannot obtain the intervals of each beat. In essence, this becomes a chicken-and-egg problem: if we know the morphology of the heartbeat, that would help us in segmenting the signal; on the other hand, if we have a segmentation of the reflected signal, we can use it to recover the morphology of the human heartbeat.

This problem is exacerbated by two additional factors. First, the reflected signal is noisy; second, the chest displacement due to breathing is orders of magnitude higher than the heartbeat displacements. In other words, we are operating in a low SINR (signal-to-interference-and-noise) regime, where "interference" results from the chest displacement due to breathing.

To address these challenges, EQ-Radio first processes the RF signal to mitigate the interference from breathing. It then formulates and solves an optimization problem to

recover the beat-to-beat intervals. The optimization formulation neither assumes nor relies on perfect separation of the respiration effect. In what follows, we describe both of these steps.

5.1 Mitigating the Impact of Breathing

The goal of the preprocessing step is to dampen the breathing signal and improve the signal-to-interference-and-noise ratio (SINR) of the heartbeat signal. Recall that the phase of the RF signal is proportional to the composite displacement due to the inhale-exhale process and the pulsing effect. Since displacements due to the inhale-exhale process are orders of magnitude larger than minute vibrations due to heartbeats, the RF phase signal is dominated by breathing. However, the acceleration of breathing is smaller than that of heartbeats. This is because breathing is usually slow and steady while a heartbeat involves rapid contraction of the muscles. Thus, we can dampen breathing and emphasize the heartbeats by operating on a signal proportional to acceleration as opposed to displacement.

By definition, acceleration is the second derivative of displacement. Thus, we can simply operate on the second derivative of the RF phase signal. Since we do not have an analytic expression of the RF signal, we have to use a numerical method to compute the second derivative. There are multiple such numerical methods which differ in their properties. We use the following second order differentiator because it is robust to noise [55]:

$$f_0'' = \frac{4f_0 + (f_1 + f_{-1}) - 2(f_2 + f_{-2}) - (f_3 + f_{-3})}{16h^2}, \tag{5.1}$$

where f_0'' refers to the second derivative at a particular sample, f_i refers to the value of the time series i samples away, and h is the time interval between consecutive samples.

In Fig. 5-1, we show an example RF phase signal with the corresponding acceleration signal. The figure shows that in the RF phase, breathing is more pronounced than heartbeats. In contrast, in the acceleration signal, there is a periodic pattern corresponding to each heartbeat cycle, and the breathing effect is negligible.

5.2 Heartbeat Segmentation

Next, EQ-Radio needs to segment the acceleration signal into individual heartbeats. Recall that the key challenge is that we do not know the morphology of the heartbeat to bootstrap this segmentation process. To address this challenge, we formulate an optimization problem

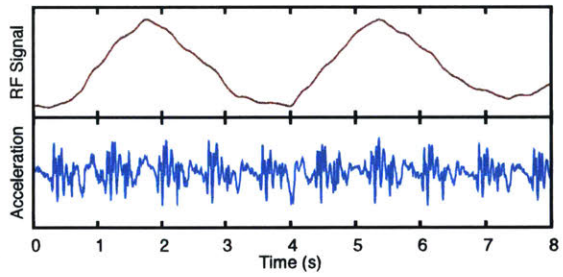


Figure 5-1: **RF Signal and Estimated Acceleration.** The figure shows the RF signal (top) and the acceleration of that signal (bottom). In the RF acceleration signal, the breathing motion is dampened and the heartbeat motion is emphasized. Note that while we can observe the periodicity of the heartbeat signal in the acceleration, delineating beat boundaries remains difficult because the signal is noisy and lacks sharp features.

that jointly recovers the morphology of the heartbeats and the segmentation.

The intuition underlying this optimization is that successive human heartbeats should have the same morphology; hence, while they may stretch or compress due to different beat lengths, they should have the same overall shape. This means that we need to find a segmentation that minimizes the differences in shape between the resulting beats, while accounting for the fact that we do not know a priori the shape of a beat and that the beats may compress or stretch. Further, rather than seeking locally optimal choices using a greedy algorithm, our formulation is an optimization problem over all possible segmentations, as described below.

Let $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ denote the sequence of length n. A segmentation $\mathcal{S} = \{s_1, s_2, ...\}$ of $\boldsymbol{x}$ is a partition of it into non-overlapping contiguous subsequences (segments), where each segment s_i consists of $|s_i|$ points.

In order to identify each heartbeat cycle, our idea is to find a segmentation with segments most similar to each other –i.e., to minimize the variation across segments. Since statistical variance is only defined for scalars or vectors with the same dimension, we extend the definition for vectors with different lengths as follows.

Definition 5.2.1. *Variance of segments* $\mathcal{S} = \{s_1, s_2, ...\}$ *is*

$$Var(\mathcal{S}) = \min_{\boldsymbol{\mu}} \sum_{s_i \in \mathcal{S}} \|s_i - \omega(\boldsymbol{\mu}, |s_i|)\|^2, \tag{5.2}$$

where $\omega(\boldsymbol{\mu}, |s_i|)$ *is linear warping*[1] *of* $\boldsymbol{\mu}$ *into length* $|s_i|$.

Note that the above definition is exactly the same as statistical variance when all the

[1]Linear warping is realized through a cubic spline interpolation [56].

segments have the same length. $\boldsymbol{\mu}$ in the definition above represents the central tendency of all the segments –i.e., a template for the beat shape (or morphology).

The goal of our algorithm is to find the optimal segmentation $\mathcal{S}^*$ that minimizes the variance of segments, which can be formally stated as follows:

$$\mathcal{S}^* = \arg\min_{\mathcal{S}} \mathrm{Var}(\mathcal{S}). \tag{5.3}$$

We can rewrite it as the following optimization problem:

$$\begin{aligned} \underset{\mathcal{S},\boldsymbol{\mu}}{\text{minimize}} \quad & \sum_{s_i \in \mathcal{S}} \|s_i - \omega(\boldsymbol{\mu}, |s_i|)\|^2, \\ \text{subject to} \quad & b_{\min} \leq |s_i| \leq b_{\max},\ s_i \in \mathcal{S}, \end{aligned} \tag{5.4}$$

where b_{min} and b_{max} are constraints on the length of each heartbeat cycle.[2] It is trying to find the optimal segmentation $\mathcal{S}$ and template (i.e., morphology) $\boldsymbol{\mu}$ that minimize the sum of the square differences between segments and template. This optimization problem is difficult as it involves both combinatorial optimization over $\mathcal{S}$ and numerical optimization over $\boldsymbol{\mu}$. Exhaustively searching all possible segmentations has exponential complexity.

5.3 Algorithm

Instead of estimating the segmentation $\mathcal{S}$ and the template $\boldsymbol{\mu}$ simultaneously, our algorithm alternates between updating the segmentation and template, while fixing the other. During each iteration, our algorithm updates the segmentation given the current template, then updates the template given the new segmentation. For each of these two sub-problems, our algorithms can obtain the global optimal with linear time complexity.

Update segmentation $\mathcal{S}$. In the l-th iteration, segmentation $\mathcal{S}^{l+1}$ is updated given template $\boldsymbol{\mu}^l$ as follows:

$$\mathcal{S}^{l+1} = \arg\min_{\mathcal{S}} \sum_{s_i \in \mathcal{S}} \|s_i - \omega(\boldsymbol{\mu}^l, |s_i|)\|^2. \tag{5.5}$$

Though the number of possible segmentations grows exponentially with the length of $\boldsymbol{x}$, the above optimization problem can be solved efficiently using dynamic programming. The recursive relationship for the dynamic program is as follows: if D_t denotes the minimal cost

[2] b_{min} and b_{max} capture the fact that human heartbeats cannot be indefinitely short or long. The default setting of b_{min} and b_{max} is 0.5s and 1.2s respectively.

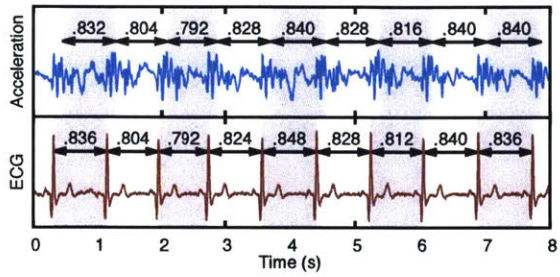


Figure 5-2: **Segmentation Result Compared to ECG.** The figure shows that the length of our segmented beats in RF (top) is very similar to the length of the segmented beats in ECG (bottom). There is a small delay since the ECG measures the electric signal of the heart, whereas the RF signal captures the heart's mechanical motion as it reacts to the electric signal.

of segmenting sequence $\boldsymbol{x}_{1:t}$, then:

$$D_t = \min_{\tau \in \tau_{t,\mathbf{B}}} \{D_\tau + \|\boldsymbol{x}_{\tau+1:t} - \omega(\boldsymbol{\mu}, t-\tau)\|^2\}, \tag{5.6}$$

where $\tau_{t,\mathbf{B}}$ specifies possible choices of τ based on segment length constraints. The time complexity of the dynamic program based on Eqn. 5.6 is $O(n)$ and the global optimum is guaranteed.

Update template $\boldsymbol{\mu}$. In the l-th iteration, template $\boldsymbol{\mu}^{l+1}$ is updated given segmentation $\mathcal{S}^{l+1}$ as follows:

$$\begin{aligned} \boldsymbol{\mu}^{l+1} &= \arg\min_{\boldsymbol{\mu}} \sum_{s_i \in \mathcal{S}^{l+1}} \|s_i - \omega(\boldsymbol{\mu}, |s_i|)\|^2 \\ &= \arg\min_{\boldsymbol{\mu}} \sum_{s_i \in \mathcal{S}^{l+1}} |s_i| \cdot \|\boldsymbol{\mu} - \omega(s_i, m)\|^2 \end{aligned} \tag{5.7}$$

where m is the required length of template. The above optimization problem is a weighted least squares with the following closed-form solution:

$$\boldsymbol{\mu}^{l+1} = \frac{\sum_{s_i \in \mathcal{S}^{l+1}} |s_i| \omega(s_i, m)}{\sum_{s_i \in \mathcal{S}^{l+1}} |s_i|} = \frac{1}{n} \sum_{s_i \in \mathcal{S}^{l+1}} |s_i| \omega(s_i, m) \tag{5.8}$$

Fig. 5-2 shows the final beat segmentation for the data in Fig. 5-1. The figure also shows the ECG data of the subject. The segmented beat length matches the ECG of the subject to within a few milliseconds. There is a small delay since the ECG measures the electric signal of the heart, whereas the RF signal captures the heart's mechanical motion as it reacts to the electric signal [57].

Initialization. Initialization is typically important for optimization algorithms; however, we found that our algorithm does not require sophisticated initialization. Our algorithm

Algorithm 1 Heartbeat Segmentation Algorithm

Input: Sequence $\boldsymbol{x}$ of n points, heart rate range $\mathbf{B}$.
Output: Segments $\mathcal{S}$, template $\boldsymbol{\mu}$ of length m.

Initialize $\boldsymbol{\mu}^0$ as zero vector
$l \leftarrow 0$ ▷ number of iterations
repeat
  $\mathcal{S}^{l+1} \leftarrow$ UPDATESEGMENTATION$(\boldsymbol{x}, \boldsymbol{\mu}^l)$
  $\boldsymbol{\mu}^{l+1} \leftarrow$ UPDATETEMPLATE$(\boldsymbol{x}, \mathcal{S}^{l+1})$
  $l \leftarrow l + 1$
until convergence
return $\mathcal{S}^l$ and $\boldsymbol{\mu}^l$

procedure UPDATESEGMENTATION$(\boldsymbol{x}, \boldsymbol{\mu})$
  $\mathcal{S}_0 \leftarrow \emptyset$
  $D_0 \leftarrow 0$
  for $t \leftarrow 1$ to n **do**
    $\tau^* \leftarrow \arg\min_{\tau \in \tau_{t,\mathbf{B}}} \{D_\tau + \|\boldsymbol{x}_{\tau+1:t} - \omega(\boldsymbol{\mu}, t-\tau)\|^2\}$
    $D_t \leftarrow D_{\tau^*} + \|\boldsymbol{x}_{\tau^*+1:t} - \omega(\boldsymbol{\mu}, t-\tau)\|^2$
    $\mathcal{S}_t \leftarrow \mathcal{S}_{\tau^*} \cup \{\boldsymbol{x}_{\tau^*+1:t}\}$
  return $\mathcal{S}_n$

procedure UPDATETEMPLATE$(\boldsymbol{x}, \mathcal{S})$
  $\boldsymbol{\mu} \leftarrow \frac{1}{n} \sum_{s_i \in \mathcal{S}} |s_i| \omega(s_i, m)$
  return $\boldsymbol{\mu}$

can converge quickly with both random initialization and zero initialization. We choose to initialize the template $\boldsymbol{\mu}^0$ as the zero vector.

Running time analysis. The pseudocode of our algorithm is presented in 1. The complexity of this algorithm is $O(kn)$, where k is the number of iterations the algorithm takes before it converges. The algorithm is guaranteed to converge because the number of possible segmentations is finite and the cost function monotonically decreases with each iteration before it converges. In practice, this algorithm converges very quickly: for the evaluation experiments reported in §7, the number of iteration k is on average 8 and at most 16.

Finally, we note that the overall algorithm is not guaranteed to achieve a global optimum, but each of the subproblems achieves its global optimum. In particular, as detailed above, the first subproblem has a closed form optimal solution, and the second subproblem can be solved optimally with a dynamic program. As a result, the algorithm converges to a local optimum that works very well in practice as we show in §7.2.

Chapter 6

Emotion Classification

After EQ-Radio recovers individual heartbeats from RF reflections, it uses the heartbeat sequence along with the breathing signal to recognize the person's emotions. Below, we describe the emotion model which EQ-Radio adopts, and we elaborate on its approach for feature extraction and classification.

(a) 2D Emotion Model: EQ-Radio adopts a 2D emotion model whose axes are *valence* and *arousal*; this model serves as the most common approach for categorizing human emotions in past literature [33, 10]. The model classifies between four basic emotional states: Sadness (negative valence and negative arousal), Anger (negative valence and positive arousal), Pleasure (positive valence and negative arousal), and Joy (positive valence and positive arousal).

(b) Feature Extraction: EQ-Radio extracts features from both the heartbeat sequence and the respiration signal. There is a large literature on extracting emotion-dependent features from human heartbeats [10, 9, 58], where past techniques use on-body sensors. These features can be divided into time-domain analysis, frequency-domain analysis, time-frequency analysis, Poincaré plot [59], Sample Entropy [60], and Detrend Fluctuation Analysis [61]. EQ-Radio extracts 27 features from IBI sequences as listed in Table 6.1. These particular features were chosen in accordance with the results in [10]. We refer the reader to [10, 58] for a detailed explanation of these features.

EQ-Radio also employs respiration features. To extract the irregularity of breathing, EQ-Radio first identifies each breathing cycle by peak detection after low pass filtering. Since past work that studies breathing features recommends time-domain features [9], EQ-Radio extracts the time-domain features in the first row of Table 6.1.

(c) Handling Dependence: Physiological features differ from one subject to another for the same emotional state. Further, those features could be different for the same subject on different days. This is caused by multiple factors, including caffeine intake, sleep, and baseline mood of the day.

In order to extract better features that are user-independent and day-independent, EQ-Radio incorporates a baseline emotional state: neutral. The idea is to leverage changes of physiological features instead of absolute values. Thus, EQ-Radio calibrates the computed features by subtracting for each feature its corresponding values calculated at the neutral state for a given person on a given day.

(d) Feature Selection and Classification: As mentioned earlier, the literature has many features that relate IBI to emotions. Using all of those features with a limited amount of training data can lead to over-fitting. Selecting a set of features that is most relevant to emotions not only reduces the amount of data needed for training but also improves the classification accuracy on the test data.

Previous work on feature selection [9, 10] uses wrapper methods which treat the feature selection problem as a search problem. However, since the number of choices is exponentially large, wrapper methods have to use heuristics to search among all possible subsets of relevant features. Instead, EQ-Radio uses another class of feature selection mechanisms, namely embedded methods [62]; this approach allows us to learn which features best contribute to the accuracy of the model while training the model. To do this, EQ-Radio uses l_1-SVM [63] which selects a subset of relevant features while training an SVM classifier. Table 6.1 shows the selected IBI and respiration features in bold and italic respectively. The performance of the resulting classifier is evaluated in §7.3.

Domain	Name
Time	Mean, Median, SDNN,**pNN50**, RMSSD, SDNNi, meanRate, *sdRate*, HRVTi, *TINN*.
Frequency	Welch PSD: **LF/HF**, peakLF, peakHF. Burg PSD: **LF/HF**, peakLF, peakHF. Lomb-Scargle PSD: **LF/HF**, peakLF, peakHF.
Poincaré	SD_1, $\mathbf{SD_2}$, $\mathbf{SD_2/SD_1}$.
Nonlinear	$\mathbf{SampEn_1}$, $\mathbf{SampEn_2}$, $\mathbf{DFA_{all}}$, DFA_1, DFA_2.

selected IBI features in **bold**;
selected respiration features in *italic*.

Table 6.1: **Features used in EQ-Radio.**

Chapter 7

Evaluation

In this section, we describe our implementation of EQ-Radio and its empirical performance with respect to extracting individual heartbeats and recognizing human emotional states. All experiments were approved by our IRB.

7.1 Implementation

We reproduced a state-of-the-art FMCW radio designed by past work on wireless vital sign monitoring [17]. The device generates a signal that sweeps from 5.46 GHz to 7.25 GHz every 4 milliseconds, transmitting sub-mW power. The parameters were chosen as in [17] such that the transmission system is compliant with FCC regulations for consumer electronics. The FMCW radio connects to a computer over Ethernet. The received signal is sampled (digitized) and transmitted over the Ethernet to the computer. EQ-Radio's algorithms are implemented on an Ubuntu 14.04 computer with an i7 processor and 32 GB of RAM.

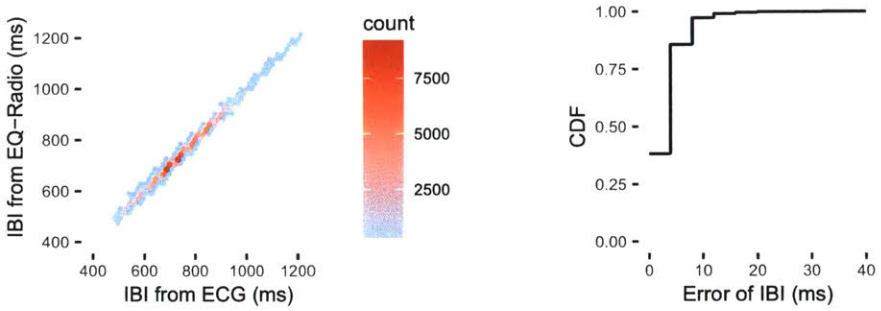
7.2 Evaluation of Heartbeat Extraction

First, we would like to assess the accuracy of EQ-Radio's segmentation algorithm in extracting heartbeats from RF signals reflected off a subject's body.

Experimental Setup

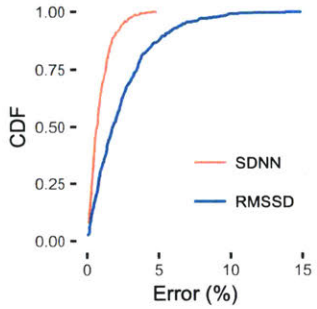
Participants: We recruited 30 participants (10 females). Our subjects are between 19~77 years old. During the experiments, the subjects wore their daily attire with different fabrics.

Experimental Environment: We perform our experiments in 5 different rooms in a standard office building. The evaluation environment contains office furniture including desks, chairs,



(a) Scatterplot of IBI estimates for EQ-Radio vs. ECG

(b) CDF of error in IBI



(c) Error in emotion-related features

Figure 7-1: **Comparison of IBI Estimates Using EQ-Radio and a Commercial ECG Monitor.** The figure shows various metrics for evaluating EQ-Radio's heartbeat segmentation accuracy in comparison with an FDA-approved ECG monitor. Note that the CDF in (b) jumps at 4 ms intervals because the RF signal was sampled every 4 ms.

couches, and computers. The experiments are performed while other users are present in the room. The change in the experimental environment and the presence of other users had a negligible impact on the results because the FMCW radio described in §4 eliminates reflections from static objects (e.g., furniture) and isolates reflections from different humans [17].

Metrics: To evaluate EQ-Radio's heartbeat extraction algorithm, we use metrics that are common in emotion recognition:

- *Inter-Beat-Interval (IBI):* The IBI measures the accuracy in identifying the boundaries of each individual beat.
- *Root Mean Square of Successive Differences (RMSSD):* This metric focuses on differences between successive beats. It is computed as $RMSSD = \sqrt{1/n \sum (IBI_{i+1} - IBI_i)^2}$, where n is the number of beats in the sum and i is a beat index. RMSSD is typically used as a measure of the parasympathetic nervous activity that controls the heart [64].

We calculate RMSSD for IBI sequences in a window of 2 minutes.

- *Standard Deviation of NN Intervals (SDNN):* The term NN-interval refers to the inter-beat-interval (IBI). Thus, SDNN measures the standard deviation of the beat length over a window of time. We use a window of 2 minutes.

Baseline: We obtain the ground truth for the above metrics using a commercial ECG monitor. We use the AD8232 evaluation board with a 3-lead ECG monitor to get the ECG signal. The synchronization between the FMCW signal and the ECG signal is accomplished by connecting both devices to a shared clock.

Accuracy in comparison to ECG

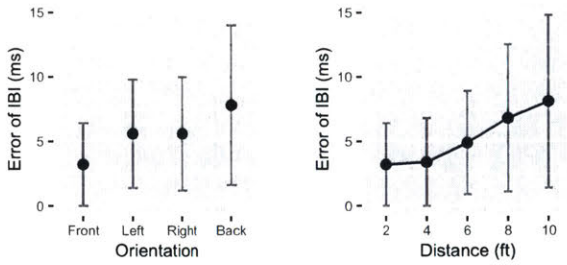
We run experiments with 30 participants, collecting over 130,000 heart beats. Each subject is simultaneously monitored with EQ-Radio and the ECG device. We process the data to extract the above three metrics.

We first compare the IBIs estimated by EQ-Radio to the IBIs obtained from the ECG monitor. Fig. 7-1(a) shows a scatter plot where the x and y coordinates are the IBIs derived from EQ-Radio and the ECG respectively. The color indicates the density of points in a specific region. Points on the diagonal have identical IBIs in EQ-Radio and ECG, while the distance to the diagonal is proportional to the error. It can be visually observed that all points are clustered around the diagonal, and hence EQ-Radio can estimate IBIs accurately irrespective of the their lengths.

We quantitatively evaluate the errors in Fig. 7-1(b), which shows a cumulative distribution function (CDF) of the difference between EQ-Radio's IBI estimate and the ECG-based IBI estimate for each beat. The CDF has jumps at 4ms intervals because the RF signal was sampled every 4ms.[1] The CDF shows that the 97^{th} percentile error is 8ms. Our results further show that EQ-Radio's mean IBI estimation error is 3.2 ms. Since the average IBI in our experiments is 740 ms, on average, EQ-Radio estimates a beat length to within 0.43% of its correct value.

In Fig. 7-1(c), we report results for beat variation metrics that are typically used in emotion recognition. The figure shows the CDF of errors in recovering the SDNN and RMSSD from RF reflections in comparison to contact-based ECG sensors. The plots show that the median error for each of these metrics is less than 2% and that even the 90^{th} percentile error is less than 8%. The high accuracy of these emotion-related metrics suggests that EQ-Radio's emotion recognition accuracy will be on par with contact-based techniques,

[1]The actual sampling rate of our receiver is 1MHz. However, because each FMCW sweep takes 4ms, we obtain one phase measurement every 4ms. For a detailed explanation, please refer to [17].



(a) Error in IBI vs. orientation (b) Error in IBI vs. distance

Figure 7-2: **Error in IBI with Different Orientations and Distances.** (a) plots the error in IBI as a function of the user's orientation with respect to the device. (b) plots the error in IBI as a function of the distance between the user and the device.

as we indeed show in §7.3.

Accuracy for different orientations & distances

In the above experiments, the subject sat relatively close to EQ-Radio, at a distance of 3 to 4 feet, and was facing the device. It is desirable, however, to allow emotion recognition even when the subject is further away or is not facing the device.

Thus, we evaluate EQ-Radio's beat segmentation accuracy as a function of orientation and distance. First, we fix the distance to 3 feet and repeat the above experiments for four different orientations: subject faces the device, subject has his back to the device, and the subject is facing left or right (perpendicular) to the device. We plot the median and standard deviation of EQ-Radio's IBI estimate for these four orientations in Fig. 7-2(a). The figure shows that, across all orientations, the median error remains below 8ms (i.e., 1% of the beat length). As expected, however, the accuracy is highest when the user directly faces the device.

Next, we test EQ-Radio's beat segmentation accuracy as a function of its distance to the subject. We run experiments where the subject sits on a chair at different distances from the device. Fig. 7-2(b) shows the median and standard deviation error in IBI estimate as a function of distance. Even at 10 feet, the median error is less than 8 ms (i.e., 1% of the beat length).

7.3 Evaluation of Emotion Recognition

In this section, we investigate whether EQ-Radio can accurately classify a person's emotions based on RF reflections off her/his body. We also compare EQ-Radio's performance with

more traditional emotion classification methods that rely on ECG signals or images.

Experimental Setup

Participants: We recruited 12 participants (6 females). Among them, 6 participants (3 females) have acting experience of 3∼7 years. People with acting experience are more skilled in emotion management, which helps in gathering high-quality emotion data and providing a reference group [9]. All subjects were compensated for their participation, and all experiments were approved by our IRB.

Experiment design: Obtaining high-quality data for emotion analysis is difficult, especially in terms of identifying the ground truth emotion [9]. Thus, it is crucial to design experiments carefully. We designed our experiments in accordance with previous work on emotion recognition using physiological signals [10, 9]. Specifically, before the experiment, the subjects individually prepare stimuli (e.g., personal memories, music, photos, and videos); during the experiment, the subject sits alone in one out of the 5 conference rooms and elicits a certain emotional state using the prepared stimuli. Some of these emotions are associated with small movements like laughing, crying, smiling, etc.[2] After the experiment, the subject reports the period during which she/he felt that type of emotion. Data collected during the corresponding period are labeled with the subject's reported emotion.

Throughout these experiments, each subject is monitored using three systems: 1) EQ-Radio, 2) the AD8232 ECG monitor, and 3) a video camera focused on the subject's face.

Ground Truth: As described above, subjects are instructed to evoke a particular emotion and report the period during which they felt that emotion. The subject's reported emotion is used to label the data from the corresponding period. These labels provide the ground truth for classification.

Baselines: We compare EQ-Radio's emotion classification to more traditional emotion recognition approaches based on ECG signals and image analysis. We describe the details of these systems in the corresponding sub-sections.

Metrics & Visualization: When tested on a particular data point, the classifier outputs a score for each of the considered emotional states. The data point is assigned the emotion that corresponds to the highest score. We measure *classification accuracy* as the percent of test data that is assigned the correct emotion.

[2]We note that the differentiation filter described in §5.1 mitigates such small movements. However, it cannot deal with larger body movements like walking. Though the FMCW radio we used can isolate signals from different users, as we show in §7.2, for better elicitation of emotional state, there is no other user in the room during this experiment.

We visualize the output of the classification as follows: Recall that the four emotions in our system can be represented in a 2D plane whose axes are *valence* and *arousal*. Each emotion occupies one of the four quadrants: Sadness (negative valence and negative arousal), Anger (negative valence and positive arousal), Pleasure (positive valence and negative arousal), and Joy (positive valence and positive arousal). Thus, we can visualize the classification result for a particular test data by showing it in the 2D valence-arousal space. If the point is classified correctly, it would fall in the correct quadrant.

For any data point, we can calculate the valence and arousal scores as: $S_{\text{valence}} = \max(S_{\text{joy}}, S_{\text{pleasure}}) -$

$\max(S_{\text{sadness}}, S_{\text{anger}})$ and $S_{\text{arousal}} = \max(S_{\text{joy}}, S_{\text{anger}}) - \max(S_{\text{pleasure}}, S_{\text{sadness}})$, where S_{joy}, S_{pleasure}, S_{sadness}, and S_{anger} are the classification score output by the classifier for the four emotions. For example, consider a data point with the following scores $S_{\text{joy}} = 1$, $S_{\text{pleasure}} = 0$, $S_{\text{sadness}} = 0$, and $S_{\text{anger}} = 0$ –i.e., this data point is one unit of pure joy. Such data point falls on the diagonal in the upper right quadrant. A data point that has a high joy score but small scores for other emotions would still fall in the joy quadrant, but not exactly on the diagonal. (Check Fig. 7-4 for an example.)

EQ-Radio's emotion recognition accuracy

To evaluate EQ-Radio's emotion classification accuracy, we collect 400 two-minute signal sequences from 12 subjects, 100 sequences for each emotion. We train two types of emotion classifiers: a person-dependent classifier, and a person-independent classifier. Each person-dependent classifier is trained and tested on data from a particular subject. Training and testing are done on mutually-exclusive data points using leave-one-out cross validation [65]. As for the person-independent classifier, it is trained on 11 subjects and tested on the remaining subject, and the process is repeated for different test subjects.

We first report the person-dependent classification results. Using the valence and arousal scores as coordinates, we visualize the results of person-dependent classification in Fig. 7-3. Different types of points indicate the label of the data. We observe that emotions are well clustered and segregated, suggesting that these emotions are distinctly encoded in valence and arousal, and can be decoded from features captured by EQ-Radio. We also observe that the points tend to cluster along the diagonal and anti-diagonal, showing that our classifiers have high confidence in the predictions. Finally, the accuracy of person-dependent classification for each subject is also shown in the figure with an overall average accuracy of 87.0%.

The results of person-independent emotion classification are visualized in Fig. 7-4.

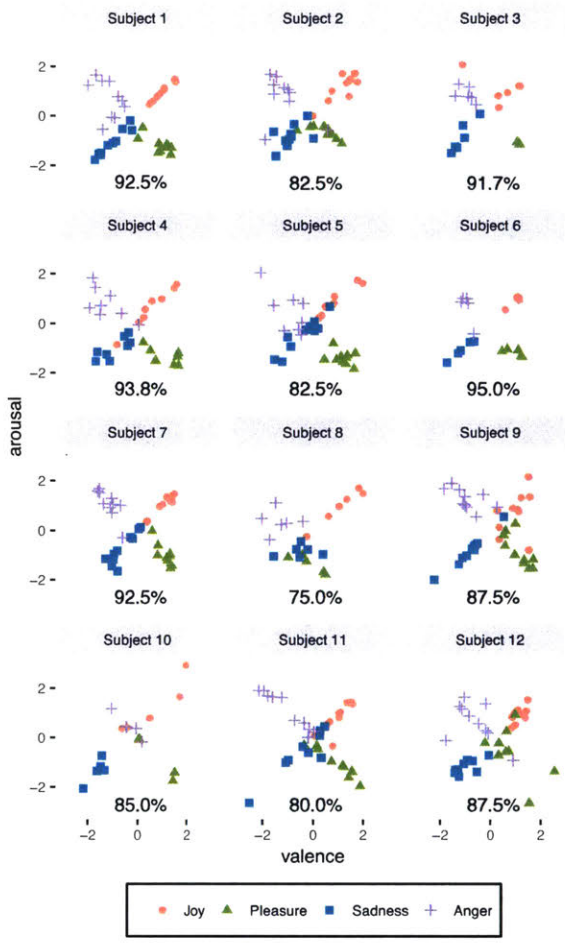


Figure 7-3: **Visualization of EQ-Radio's Person-dependent Classification Results.** The figure shows the person-dependent emotion-classification results for each of our 12 subjects. The x-axis in each of the scatter plots corresponds to the valence, and the y-axis corresponds to the arousal. For each data point, the label is our ground truth, and the coordinate is the classification result. At the bottom of each sub-figure, we show the classification accuracy for the corresponding subject.

EQ-Radio is capable of recognizing a subject's emotion with an average accuracy of 72.3% purely based on data from other subjects, meaning that EQ-Radio succeeds in learning person-independent features for emotion recognition.

As expected, the accuracy of person-independent classification is lower than the accuracy of person-dependent classification. This is because person-independent emotion recognition is intrinsically more challenging since an emotional state is a rather subjective conscious experience that could be very different among different subjects. We note, however, that our accuracy results are consistent with the literature both for the case of person-dependent and person-independent emotion classifications [8]. Further, our results present the first

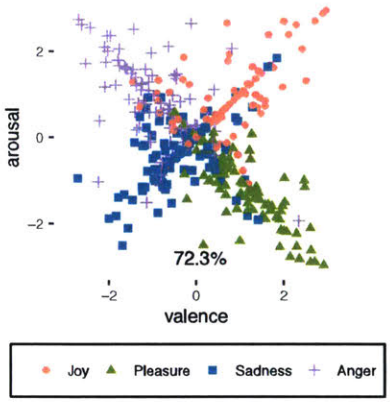
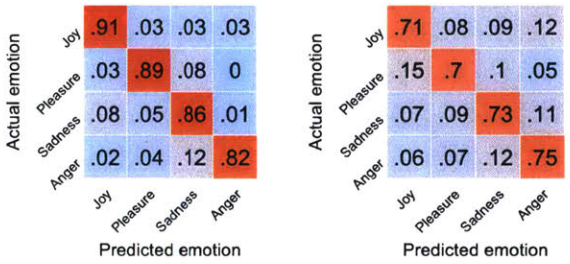


Figure 7-4: **Visualization of EQ-Radio's Person-independent Classification Results.** The figure shows the results of person-independent emotion-classification. The x-axis corresponds to valence, and the y-axis corresponds to arousal.



(a) Person-dependent (b) Person-independent

Figure 7-5: **Confusion Matrix of Person-dependent and Person-independent Classification Results.** The diagonal of each of these matrices shows the classification accuracy and the off-diagonal grid points show the confusion error.

demonstration of RF-based emotion classification.

To better understand the classification errors, we show the confusion matrix of both person-dependent and person-independent classification results in Fig. 7-5. We find that EQ-Radio achieves comparable accuracy in recognizing the four types of emotions. We also observe that EQ-Radio typically makes fewer errors between emotion pairs that are different in both valence and arousal (i.e., joy vs. sadness and pleasure vs. anger).

Emotion recognition accuracy versus data source

It is widely known that gathering data that genuinely corresponds to a particular emotional state is crucial to recognizing emotions and that people with acting experience are better at emotion management. We would like to test whether there is a difference in the performance of EQ-Radio's algorithms in classifying the emotions of actors vs. non-actors, as well as in classifying the emotions of males vs. females. We evaluate the performance of

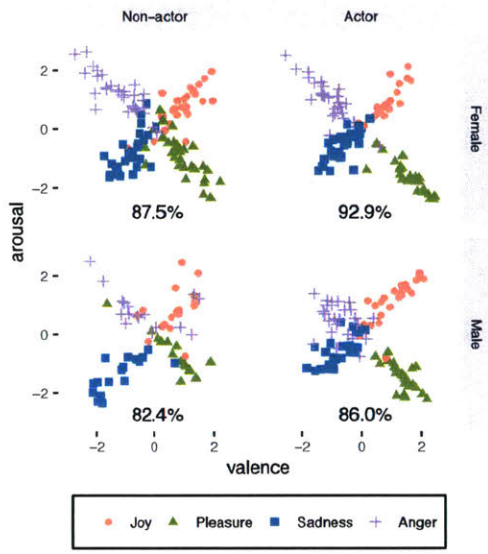


Figure 7-6: **Visualization of EQ-Radio's Group-dependent Classification Results.** The figure shows the results of EQ-Radio's classification within 4 different groups, defined by gender and acting experience. The x-axis corresponds to valence and the y-axis corresponds to arousal.

a specific group of subjects in terms of mutual predictability/consistency, i.e., we predict the emotion label of a data point by training on data obtained from within the same group only. Fig. 7-6 shows our results. These results show that our emotion recognition algorithm works for both actors and non-actors, and for both genders. However, the accuracy of this algorithm is higher for actors than non-actors and for females than males. This could suggest that actors/females have better emotion management skills or that they are indeed more emotional.

EQ-Radio versus ECG-based emotion recognition

In this section, we compare EQ-Radio's emotion classification accuracy with that of an ECG-based classifier. Note that both classifiers use the same set of features and decision making process. However, the ECG-based classifier uses heartbeat information directly extracted from the ECG monitor. In addition, we allow the ECG monitor to access the breathing signal from EQ-Radio and use EQ-Radio's breathing features. This mirrors today's emotion monitors which also use breathing data but require the subject to wear a chest band in order to extract that signal.

The results in Table 7.1 show that EQ-Radio achieves comparable accuracy to emotion recognition systems that use on-body sensors. Thus, by using EQ-Radio, one can eliminate body sensors without jeopardizing the accuracy of emotion recognition based on physiological signals.

Method	Person-dependent	Person-independent
EQ-Radio	87%	72.3%
ECG-based	88.2%	73.2%

Table 7.1: **Comparison with the ECG-based Method.** The table compares the accuracy of EQ-Radio's person-dependent and person-independent emotion classification accuracy with the emotion classification accuracy achieved using the ECG signals (combined with the extracted respiration features).

EQ-Radio versus vision-based emotion recognition

In order to compare the accuracy of EQ-Radio with vision-based emotion recognition systems, we use the Microsoft Project Oxford Emotion API to process the images of the subjects collected during the experiments, and analyze their emotions based on facial expressions. Since the Microsoft Emotion API and EQ-Radio use different emotion models, we use the following four emotions that both systems share for our comparison: joy/pleasure, sadness, anger, and neutral. For each data point, the Microsoft Emotion API outputs scores for eight emotions. We consider their scores for the above four shared emotions and use the label with highest score as their output.

Fig. 7-7 compares the accuracy of EQ-Radio (both person-dependent and person-independent) with the Microsoft Emotion API. The figure shows that that the Microsoft Emotion API does not achieve high accuracy for the first three categories of emotions, but achieves very high accuracy for neutral state. This is because vision-based methods can recognize an emotion only when the person explicitly expresses it on her face, and fail to recognize the innermost emotions and hence they report such emotions as neutral. We also note that the Microsoft Emotion API has higher accuracy for positive emotions than negative ones. This is because positive emotions typically have more visible features (e.g., smiling), while negative emotions are visually closer to a neutral state.

Emotion recognition versus accurate beat segmentation

Finally, we would like to understand how tolerant emotion recognition is to errors in beat segmentation. We take the ground truth beats derived from the ECG monitor and add to them different levels of Gaussian noise. The Gaussian distribution has zero mean and its standard deviation varies between 0 and 60 milliseconds. We re-run the person-dependent emotion recognition classifier using these noisy beats. Fig. 7-8 shows that small errors in estimating the beat lengths can lead to a large degradation in classification accuracy. In particular, an error of 30 milliseconds in inter-beat-interval can reduce the accuracy of emotion recognition by over 35%. This result emphasizes the importance of extracting the

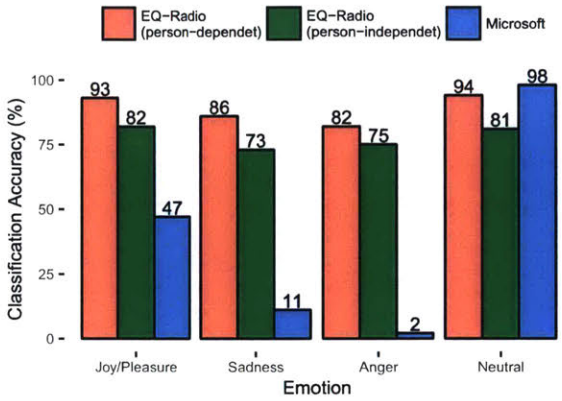


Figure 7-7: **Comparison of EQ-Radio with Image-based Emotion Recognition.** The figure shows the accuracies (on the y-axis) of EQ-Radio and Microsoft's Emotion API in differentiating among the four emotions (on the x-axis).

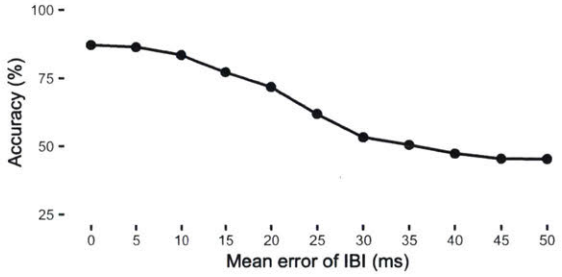


Figure 7-8: **Impact of Millisecond Errors in IBI on Emotion Recognition.** The figure shows that adding small errors to the IBI values (x-axis) significantly reduces the classification accuracy (y-axis). Given that we have four classes, a random guess would have 25% accuracy.

individual beats and delineating their boundaries at an accuracy of a few milliseconds.[3]

[3]Note that given that we have four classes, a random guess would have 25% accuracy. Adding small errors to the IBI values significantly reduces the classification accuracy. The accuracy converges to about 40% instead of 25% because the respiration features are left intact.

Chapter 8

Conclusion

This thesis presents a technology capable of recognizing a person's emotions by relying on wireless signals reflected off her/his body. We believe this marks an important step in the nascent field of emotion recognition. It also builds on a growing interest in the wireless systems' community in using RF signals for sensing, and as such, the work expands the scope of RF sensing to the domain of emotion recognition. Further, while this work has laid foundations for wireless emotion recognition, we envision that the accuracy of such systems will improve as wireless sensing technologies evolve and as the community incorporates more advanced machine learning mechanisms in the sensing process.

We also believe that the implications of this work extend beyond emotion recognition. Specifically, while we used the heartbeat extraction algorithm for determining the beat-to-beat intervals and exploited these intervals for emotion recognition, our algorithm recovers the entire human heartbeat from RF, and the heartbeat displays a very rich morphology. We envision that this result paves way for exciting research on understanding the morphology of the heartbeat both in the context of emotion-recognition as well as in the context of non-invasive health monitoring and diagnosis.

Bibliography

[1] Deba Pratim Saha, Thomas L Martin, and R Benjamin Knapp. Towards incorporating affective feedback into context-aware intelligent environments. In *Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on*, pages 49–55. IEEE, 2015.

[2] Roddy Cowie, Ellen Douglas-Cowie, Nicolas Tsapatsoulis, George Votsis, Stefanos Kollias, Winfried Fellenz, and John G Taylor. Emotion recognition in human-computer interaction. *Signal Processing Magazine, IEEE*, 18(1):32–80, 2001.

[3] Samira Ebrahimi Kahou, Xavier Bouthillier, Pascal Lamblin, Caglar Gulcehre, Vincent Michalski, Kishore Konda, Sébastien Jean, Pierre Froumenty, Yann Dauphin, Nicolas Boulanger-Lewandowski, et al. Emonets: Multimodal deep learning approaches for emotion recognition in video. *Journal on Multimodal User Interfaces*, pages 1–13, 2015.

[4] Rosalind W Picard. Affective computing: from laughter to ieee. *Affective Computing, IEEE Transactions on*, 1(1):11–17, 2010.

[5] Hillary Anger Elfenbein and Nalini Ambady. On the universality and cultural specificity of emotion recognition: a meta-analysis. *Psychological bulletin*, 128(2):203, 2002.

[6] Zhihong Zeng, Maja Pantic, Glenn I Roisman, and Thomas S Huang. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(1):39–58, 2009.

[7] Klaus R Scherer. Vocal communication of emotion: A review of research paradigms. *Speech communication*, 40(1):227–256, 2003.

[8] S Jerritta, M Murugappan, R Nagarajan, and Khairunizam Wan. Physiological signals based human emotion recognition: a review. In *Signal Processing and its Applications (CSPA), 2011 IEEE 7th International Colloquium on*, pages 410–415. IEEE, 2011.

[9] Rosalind W Picard, Elias Vyzas, and Jennifer Healey. Toward machine emotional intelligence: Analysis of affective physiological state. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(10):1175–1191, 2001.

[10] Jonghwa Kim and Elisabeth André. Emotion recognition based on physiological changes in music listening. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(12):2067–2083, 2008.

[11] Foteini Agrafioti, Dimitrios Hatzinakos, and Adam K Anderson. Ecg pattern analysis for emotion detection. *Affective Computing, IEEE Transactions on*, 3(1):102–115, 2012.

[12] Rafael A Calvo and Sidney D'Mello. Affect detection: An interdisciplinary review of models, methods, and their applications. *Affective Computing, IEEE Transactions on*, 1(1):18–37, 2010.

[13] Paul Ekman, Wallace V Friesen, and Phoebe Ellsworth. *Emotion in the human face: Guidelines for research and an integration of findings.* Elsevier, 2013.

[14] Todd B Kashdan, Anjali Mishra, William E Breen, and Jeffrey J Froh. Gender differences in gratitude: Examining appraisals, narratives, the willingness to express emotions, and changes in psychological needs. *Journal of personality*, 77(3):691–730, 2009.

[15] Daniel S Quintana, Adam J Guastella, Tim Outhred, Ian B Hickie, and Andrew H Kemp. Heart rate variability is associated with emotion recognition: direct evidence for a relationship between the autonomic nervous system and social cognition. *International Journal of Psychophysiology*, 86(2):168–172, 2012.

[16] Sylvia D Kreibig. Autonomic nervous system activity in emotion: A review. *Biological psychology*, 84(3):394–421, 2010.

[17] Fadel Adib, Hongzi Mao, Zachary Kabelac, Dina Katabi, and Robert C Miller. Smart homes that monitor breathing and heart rate. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 837–846. ACM, 2015.

[18] Amy D Droitcour, Olga Boric-Lubecke, and Gregory TA Kovacs. Signal-to-noise ratio in doppler radar system for heart and respiratory rate measurements. *Microwave Theory and Techniques, IEEE Transactions on*, 57(10):2498–2507, 2009.

[19] Rich Fletcher and Jing Han. Low-cost differential front-end for doppler radar vital sign monitoring. In *Microwave Symposium Digest, 2009. MTT'09. IEEE MTT-S International*, pages 1325–1328. IEEE, 2009.

[20] Neal Patwari, Lara Brewer, Quinn Tate, Ossi Kaltiokallio, and Maurizio Bocca. Breathfinding: A wireless network that monitors and locates breathing in a home. *Selected Topics in Signal Processing, IEEE Journal of*, 8(1):30–42, 2014.

[21] Ossi Kaltiokallio, Huseyin Yigitler, Riku Jantti, and Neal Patwari. Non-invasive respiration rate monitoring using a single cots tx-rx pair. In *Information Processing in Sensor Networks, IPSN-14 Proceedings of the 13th International Symposium on*, pages 59–69. IEEE, 2014.

[22] Richard D Lane, Kateri McRae, Eric M Reiman, Kewei Chen, Geoffrey L Ahern, and Julian F Thayer. Neural correlates of heart rate variability during emotion. *Neuroimage*, 44(1):213–222, 2009.

[23] Microsoft project oxford emotion api. `https://www.projectoxford.ai/emotion`.

[24] Genevieve Alelis, Ania Bobrowicz, and Chee Siang Ang. Exhibiting emotion: Capturing visitors' emotional responses to museum artefacts. In *Design, User Experience, and Usability. User Experience in Novel Technological Environments*, pages 429–438. Springer, 2013.

[25] Antonio Fernández-Caballero, José Miguel Latorre, José Manuel Pastor, and Alicia Fernández-Sotos. Improvement of the elderly quality of life and care through smart emotion regulation. In *Ambient Assisted Living and Daily Activities*, pages 348–355. Springer, 2014.

[26] Daniel McDuff, Rana El Kaliouby, Jeffrey F Cohn, and Rosalind W Picard. Predicting ad liking and purchase intent: Large-scale analysis of facial responses to ads. *Affective Computing, IEEE Transactions on*, 6(3):223–235, 2015.

[27] Mohammad Soleymani, Maja Pantic, and Thierry Pun. Multimodal emotion recognition in response to videos. *Affective Computing, IEEE Transactions on*, 3(2):211–223, 2012.

[28] Moataz El Ayadi, Mohamed S Kamel, and Fakhri Karray. Survey on speech emotion recognition: Features, classification schemes, and databases. *Pattern Recognition*, 44(3):572–587, 2011.

[29] Bradley M Appelhans and Linda J Luecken. Heart rate variability as an index of regulated emotional responding. *Review of general psychology*, 10(3):229, 2006.

[30] Martin Oswaldo Mendez, Matteo Matteucci, Vincenza Castronovo, Luigi Ferini-Strambi, Sergio Cerutti, and Anna Bianchi. Sleep staging from heart rate variability: time-varying spectral features and hidden markov models. *International Journal of Biomedical Engineering and Technology*, 3(3-4):246–263, 2010.

[31] Paolo Melillo, Marcello Bracale, Leandro Pecchia, et al. Nonlinear heart rate variability features for real-life stress detection. case study: students under stress due to university examination. *Biomed Eng Online*, 10(1):96, 2011.

[32] Alan Jovic and Nikola Bogunovic. Electrocardiogram analysis using a combination of statistical, geometric, and nonlinear heart rate variability features. *Artificial intelligence in medicine*, 51(3):175–186, 2011.

[33] Peter J Lang. The emotion probe: studies of motivation and attention. *American psychologist*, 50(5):372, 1995.

[34] Robert LiKamWa, Yunxin Liu, Nicholas D Lane, and Lin Zhong. Moodscope: building a mood sensor from smartphone usage patterns. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, pages 389–402. ACM, 2013.

[35] Gokul Chittaranjan, Jan Blom, and Daniel Gatica-Perez. Who's who with big-five: Analyzing and classifying personality traits with smartphones. In *2011 15th Annual International Symposium on Wearable Computers*, pages 29–36. IEEE, 2011.

[36] Fadel Adib, Zachary Kabelac, Dina Katabi, and Robert C Miller. 3d tracking via body radio reflections. In *11th USENIX Symposium on Networked Systems Design and Implementation*, 2014.

[37] Fadel Adib and Dina Katabi. See through walls with wifi! In *Proceedings of the ACM SIGCOMM 2013*, pages 75–86, New York, NY, USA, 2013. ACM.

[38] Qifan Pu, Sidhant Gupta, Shyamnath Gollakota, and Shwetak Patel. Whole-home gesture recognition using wireless signals. In *Proceedings of the 19th annual international conference on Mobile computing & networking*, pages 27–38. ACM, 2013.

[39] Li Sun, Souvik Sen, Dimitrios Koutsonikolas, and Kyu-Han Kim. Widraw: Enabling hands-free drawing in the air on commodity wifi devices. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, pages 77–89. ACM, 2015.

[40] Kamran Ali, Alex Xiao Liu, Wei Wang, and Muhammad Shahzad. Keystroke recognition using wifi signals. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, pages 90–102. ACM, 2015.

[41] Teng Wei and Xinyu Zhang. mtrack: High-precision passive tracking using millimeter wave radios. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, pages 117–129. ACM, 2015.

[42] Heba Abdelnasser, Moustafa Youssef, and Khaled A Harras. Wigest: A ubiquitous wifi-based gesture recognition system. In *Computer Communications (INFOCOM), 2015 IEEE Conference on*, pages 1472–1480. IEEE, 2015.

[43] Wei Wang, Alex X Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. Understanding and modeling of wifi signal based human activity recognition. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, pages 65–76. ACM, 2015.

[44] Yan Wang, Jian Liu, Yingying Chen, Marco Gruteser, Jie Yang, and Hongbo Liu. E-eyes: device-free location-oriented activity identification using fine-grained wifi signatures. In *Proceedings of the 20th annual international conference on Mobile computing and networking*, pages 617–628. ACM, 2014.

[45] Amy D Droitcour, Olga Boric-Lubecke, Victor M Lubecke, Jenshan Lin, and Gregory TA Kovacs. Range correlation and i/q performance benefits in single-chip silicon doppler radars for noncontact cardiopulmonary monitoring. *Microwave Theory and Techniques, IEEE Transactions on*, 52(3):838–848, 2004.

[46] Alberto Zaffaroni, Philip De Chazal, Conor Heneghan, Patricia Boyle, Patricia Ronayne Mppm, and Walter T McNicholas. Sleepminder: an innovative contact-free device for the estimation of the apnoea-hypopnoea index. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pages 7091–9094. IEEE, 2009.

[47] Philip De Chazal, Emer O Hare, Niall Fox, and Conor Heneghan. Assessment of sleep/wake patterns using a non-contact biomotion sensor. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 514–517. IEEE, 2008.

[48] Wansuree Massagram, Victor M Lubecke, Anders Host-Madsen, and Olga Boric-Lubecke. Assessment of heart rate variability and respiratory sinus arrhythmia via doppler radar. *Microwave Theory and Techniques, IEEE Transactions on*, 57(10):2542–2549, 2009.

[49] Wei Hu, Zhangyan Zhao, Yunfeng Wang, Haiying Zhang, and Fujiang Lin. Noncontact accurate measurement of cardiopulmonary activity using a compact quadrature doppler radar sensor. *Biomedical Engineering, IEEE Transactions on*, 61(3):725–735, 2014.

[50] Olga Boric-Lubecke, Wansuree Massagram, Victor M Lubecke, Anders Host-Madsen, and Branka Jokanovic. Heart rate variability assessment using doppler radar with linear demodulation. In *Microwave Conference, 2008. EuMC 2008. 38th European*, pages 420–423. IEEE, 2008.

[51] T. Sakamoto, R. Imasaka, H. Taki, T. Sato, M. Yoshioka, K. Inoue, T. Fukuda, and H. Sakai. Feature-based correlation and topological similarity for interbeat interval estimation using ultra-wideband radar. *Biomedical Engineering, IEEE Transactions on*, 2015.

[52] Laura Anitori, Ardjan de Jong, and Frans Nennie. Fmcw radar for life-sign detection. In *Radar Conference, 2009 IEEE*, pages 1–6. IEEE, 2009.

[53] Octavian Postolache, Pedro Silva Girão, Gabriela Postolache, and Joaquim Gabriel. Cardio-respiratory and daily activity monitor based on fmcw doppler radar embedded in a wheelchair. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 1917–1920. IEEE, 2011.

[54] David Tse and Pramod Viswanath. *Fundamentals of wireless communication.* Cambridge university press, 2005.

[55] Noise robust differentiators for second derivative estimation. `http://www.holoborodko.com/pavel/downloads/NoiseRobustSecondDerivative`.

[56] Sky McKinley and Megan Levine. Cubic spline interpolation. *College of the Redwoods*, 45(1):1049–1060, 1998.

[57] Andrew D Wiens, Mozziyar Etemadi, Shuvo Roy, Liviu Klein, and Omer T Inan. Toward continuous, noninvasive assessment of ventricular function and hemodynamics: Wearable ballistocardiography. *Biomedical and Health Informatics, IEEE Journal of*, 19(4):1435–1442, 2015.

[58] U Rajendra Acharya, K Paul Joseph, Natarajan Kannathal, Choo Min Lim, and Jasjit S Suri. Heart rate variability: a review. *Medical and biological engineering and computing*, 44(12):1031–1051, 2006.

[59] Peter Walter Kamen, Henry Krum, and Andrew Maxwell Tonkin. Poincare plot of heart rate variability allows quantitative display of parasympathetic nervous activity in humans. *Clinical science*, 91(2):201–208, 1996.

[60] Douglas E Lake, Joshua S Richman, M Pamela Griffin, and J Randall Moorman. Sample entropy analysis of neonatal heart rate variability. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 283(3):R789–R797, 2002.

[61] Thomas Penzel, Jan W Kantelhardt, Ludger Grote, Jörg-Hermann Peter, and Armin Bunde. Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea. *Biomedical Engineering, IEEE Transactions on*, 50(10):1143–1151, 2003.

[62] Isabelle Guyon and André Elisseeff. An introduction to variable and feature selection. *The Journal of Machine Learning Research*, 3:1157–1182, 2003.

[63] Ji Zhu, Saharon Rosset, Trevor Hastie, and Rob Tibshirani. 1-norm support vector machines. *Advances in neural information processing systems*, 16(1):49–56, 2004.

[64] Juan Sztajzel et al. Heart rate variability: a noninvasive electrocardiographic method to measure the autonomic nervous system. *Swiss medical weekly*, 134:514–522, 2004.

[65] Pierre A Devijver and Josef Kittler. *Pattern recognition: A statistical approach*, volume 761. Prentice-Hall London, 1982.