Emotion Recognition Using Wireless Signals

By Mingmin Zhao, Fadel Adib, and Dina Katabi

Abstract
This paper 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.

1. INTRODUCTION
Emotion recognition is an emerging field that has attracted much interest from both the industry and the research community.13, 18, 22, 35, 40 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,22, 42, 48 or require the person to wear physiological sensors like an Electrocardiogram (ECG) monitor.7, 13, 26, 36 Both approaches have their limitations. Audiovisual techniques leverage the outward expression of emotions, but cannot measure inner feelings.12, 16, 36 For example, a person may be happy even if she is not smiling. Also, people differ widely in how expressive they are in showing their inner emotions, which further complicates this problem.25 The second approach recognizes emotions by monitoring the physiological signals that change with our emotional state. Intuitively, a person’s heart rate increases with anger or excitement; there are also more complex changes that appear as variability in the duration of a heart beat.12, 39 This approach uses on-body sensors – For example, 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.27, 30 However, the use of body sensors is cumbersome and can interfere with user activity and emotions, making this approach unsuitable for regular usage.

In this paper, we introduce a new method for emotion recognition that achieves the best of both worlds – that is, 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 Radio Frequency (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.6, 15, 19, 23, 33 However, the periodicity of the heart signal (i.e., its running average) is not sufficient for emotion recognition. To recognize emotions, we need to measure the minute variations in each individual beat length.12, 29, 39

Yet, extracting individual heartbeats from RF signals incurs multiple challenges, which can be seen in Figure 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 Figure 1); to separate breathing from heart rate, past systems operate over multiple seconds (e.g., 30sec in Ref. Adib et al.5) in the frequency domain, forgoing the ability to measure the beat-to-beat variability. Second, heartbeats in the

Figure 1. Comparison of RF signal with ECG signal. The top graph plots the RF signal reflected off a person’s body. The envelope 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.
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 infers people’s emotions from the radio signals that bounce off their bodies. 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. 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 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 Figure 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 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 Frequency Modulated Carrier Waves (FMCW) radio that transmits RF signals and receives their reflections. The radio leverages the approach in Ref. Adib et al. 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 heartbeat 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 and uses an Support Vector Machine (SVM) classifier to differentiate 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. 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 evoke 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.

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. 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.

2. BACKGROUND AND RELATED WORK

2.1. Emotion recognition

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. 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. 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). 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.

2.2. 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, gestures, and body posture.
activities, and vital signs. Our work is closest to prior art that uses RF signals to extract a person’s breathing rate and average heart rate. 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 or requires the monitored subjects to hold their breath. In particular, past work that does not require users to hold or requires the monitored subjects to hold their breath. We refer the reader to Ref. Adib et al. for a detailed description of such methods, and summarize below the basic information relevant to this paper.

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: \( \phi(t) = \frac{2\pi d(t)}{\lambda}, \) where \( \phi(t) \) is the phase of the signal, \( \lambda \) 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.

The phase of the RF signal is illustrated in the top graph in Figure 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.

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 FMCW. There is a significant literature on FMCW radios and their use for obtaining an RF signal that is modulated by breathing and heartbeats. We refer the reader to Ref. Adib et al. for a detailed description of such methods, and summarize below the basic information relevant to this paper.

Figure 2. EQ-Radio Architecture. EQ-Radio has three components: a radio for capturing RF reflections (Section 4), a heartbeat extraction algorithm (Section 5), and a classification subsystem that maps the learned physiological signals to emotional states (Section 6).

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: \( \phi(t) = \frac{2\pi d(t)}{\lambda}, \) where \( \phi(t) \) is the phase of the signal, \( \lambda \) 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.

5. BEAT EXTRACTION ALGORITHM

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 Signal-to-Interference-and-Noise Ratio (SINR) regime, where “interference” results from the chest displacement due to breathing.

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

$$f'' = \frac{4f_0 + (f_1 + f_3) - 2(f_2 + f_4) - (f_5 + f_7)}{16h^2},$$

where $f''_i$ 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 Figure 3, 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 the corresponding acceleration signal. The figure shows $h$ is the time interval between consecutive samples.

**Figure 3. 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 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.

5.2. Heartbeat segmentation

Next, EQ-Radio segments 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 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. Below we formalize this intuition.

Let $x = (x_1, x_2, \ldots, x_n)$ denote the sequence of length $n$. A segmentation $S = \{s_1, s_2, \ldots\}$ of $x$ is a partition of it into non-overlapping contiguous subsequences (segments), where each segment $s_i$ consists of $|s_i|$ points. The goal of our algorithm is to find the optimal segmentation $S^*$ that minimizes the variance of segments, which can be formally stated as follows:

$$S^* = \arg\min_S \text{Var}(S).$$

We can rewrite it as the following optimization problem:

$$\min_{S, \mu} \sum_{s \in S} \left\| s - \omega(\mu, |s|) \right\|^2,$$

subject to $b_{\min} \leq |s| \leq b_{\max}$, $s \in S$.

The term $\mu$ in the definition above represents the template for the beat shape (i.e., its morphology), and $\omega(\mu, |s|)$ is linear warping of $\mu$ into length $|s|$. The terms $b_{\min}$ and $b_{\max}$ are constraints on the length of each heartbeat cycle. The optimization aims to find the optimal segmentation $S$ and template $\mu$ that minimize the sum of the square differences between segments and template. This optimization problem is difficult as it involves both combinatorial optimization over $S$ and numerical optimization over $\mu$.

Solving the optimization problem. Instead of estimating the segmentation $S$ and the template $\mu$ 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 $S$. In the $l$-th iteration, segmentation $S^{(l)}$ is updated given template $\mu$ as follows:

$$S^{(l+1)} = \arg\min_S \sum_{s \in S} \left\| s - \omega(\mu', |s|) \right\|^2,$$

5. Linear warping is realized through a cubic spline interpolation. $b_{\min}$ 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.

---

b 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.\(^1\)
Though the number of possible segmentations grows exponentially with the length of \( x \), 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 of segmenting sequence \( x_{t:t} \), then:

\[
D_t = \min_{\tau_{t+1,b}} D_{t+1} + \| x_{t:t} - \omega(\mu_t, \tau_{t+1:b}) \|^2,
\]

(5)

where \( \tau_{t+1,b} \) specifies possible choices of \( \tau \) based on segment length constraints. The time complexity of the dynamic program based on Equation 5 is \( O(n) \) and the global optimum is guaranteed.

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

\[
\mu^{l+1} = \arg \min_{\mu} \sum_{s_i \in S^{l+1}} | s_i - \omega(\mu, s_i) |^2
\]

(6)

\[
= \arg \min_{\mu} \sum_{s_i \in S^{l+1}} | s_i | \| \mu - \omega(s_i, m) \|^2
\]

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

\[
\mu^{l+1} = \frac{\sum_{s_i \in S^{l+1}} | s_i | \omega(s_i, m)}{\sum_{s_i \in S^{l+1}} | s_i |} = \frac{1}{n} \sum_{s_i \in S^{l+1}} | s_i | \omega(s_i, m)
\]

(7)

Figure 4 shows the final beat segmentation for the data in Figure 3. 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.47

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.

| Domain         | Name                | Features used in EQ-Radio. |
|----------------|---------------------|-----------------------------|
| 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, SD_1, SD_2/SD_0, SampEn, SampEn_1, SampEn_2, DFA, DFA_. |

Selected IBI features in **bold**; selected respiration features in italics.
limited amount of training data can lead to over-fitting. Thus, EQ-Radio uses $l_1$-SVM$^{50}$ which selects a subset of relevant features while training an SVM classifier. Table 1 shows the selected IBI and respiration features in bold and italic respectively. The performance of the resulting classifier is evaluated in Section 7.2.

7. EVALUATION
All experiments in this section were approved by our IRB.

7.1. Evaluation of heartbeat extraction
First, we assess the accuracy of EQ-Radio’s segmentation algorithm in extracting heartbeats from RF signals.

Experimental setup. Participants: We recruited 30 participants (10 females). The subjects are between 19 and 77 year old. The subjects had no restrictions on their clothing.

Experimental Environment: We perform our experiments in five different rooms in a standard office building. The evaluation environment contains office furniture including desks, chairs, 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 Section 4 eliminates reflections from static objects (e.g., furniture) and isolates reflections from different humans.$^6$

Metrics: To evaluate EQ-Radio’s heartbeat extraction, 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. RMSSD is typically used as a measure of the parasympathetic nervous activity that controls the heart.$^{44}$
- Standard Deviation of NN Intervals (SDNN): The term NN-interval refers to the IBI. Thus, SDNN measures the standard deviation of the beat length over a window of time. We use a window of 2min.

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 ECG device and the FMCW radio are connected to a shared clock to keep them synchronized.

Accuracy in comparison to ECG. We run experiments with 30 participants, collecting over 130,000 heartbeats. 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. Figure 5(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 their lengths.

We quantitatively evaluate the errors in Figure 5(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 each FMCW sweep takes 4ms. The CDF shows that the 97th percentile error is 8ms. Our results further show that EQ-Radio’s mean IBI estimation error is 3.2ms. Since the average IBI in our experiments is 740ms, on average, EQ-Radio estimates a beat length to within 0.43% of its correct value.

In Figure 5(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 90th 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, as we indeed show in Section 7.2.

7.2. Evaluation of emotion recognition

We evaluate EQ-Radio’s ability to recognize emotions.

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. 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. Thus, it is crucial to design experiments carefully. We designed our experiments in accordance with previous work on emotion recognition using physiological signals. 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 five 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. 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) 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.

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

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 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 \( S_{\text{joy}}, S_{\text{pleasure}}, S_{\text{sadness}}, \) and \( S_{\text{anger}} \) 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 \( S_{\text{anger}} = 0 \) — that is, 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 on the diagonal.

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. 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 person-dependent classification in Figure 6. Different types of points indicate different emotions. We observe that emotions are well clustered and segregated, suggesting that they 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 average accuracy of 87.0%.

The results of person-independent emotion classification are shown in Figure 7. EQ-Radio can recognize 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 that 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. Further, our results present the first 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 Figure 8. 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).
Comparison to ECG and vision-based systems. Table 2 shows the accuracy of EQ-Radio in comparison to an ECG-based emotion classifier. 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. The results in Table 2 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.

Next, we compare EQ-Radio with vision-based emotion recognition. 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 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.

Figure 8 shows the accuracy of EQ-Radio (both person-dependent and person-independent) with the Microsoft Emotion API. The results indicate that EQ-Radio achieves comparable accuracy to the ECG-based method. However, EQ-Radio is more convenient to use since it does not require body sensors.

Table 2. Comparison with the ECG-based Method.

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

4 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).
Emotion API. The figure shows 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.

8. CONCLUSION

This paper presents a technology for inferring a person’s emotions from the wireless signals reflected off her/his body. We believe this marks an important step in the nascent field of emotion recognition. Furthermore, 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.

References

1. Microsoft project oxford emotion api. https://www.projectoxford.ai/ emotion.
2. Noise robust differentiators for second derivative estimation. http://www.holoborodko. com/pae/dحدود/NoiseRobustSecondDerivative.
3. Acharya, U.R., Joseph, K.P., Kannathal, N, Lim, C.M., Suri, J.S. Heart rate variability: a review. Medical and Biological Engineering and Computing 44, 12 (2006), 1033–1051.
4. Adib, F., Katabi, D. See through walls with wifi. Proceedings of the ACM SIGCOMM 2013 (New York, NY, USA, 2013), ACM, 75–86.
5. Adib, F., Kabelac, Z., Katabi, D., Miller, R.C. 3d tracking via body radio reflections. 11th USENSYMP on Networked Systems Design and Implementation (2014).
6. Adib, F., Mao, H., Kabelac, Z., Katabi, D., Miller, R.C. Smart homes that monitor breathing and heart rate. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (2015), ACM, 837–846.
7. Agrafioti, F., Hatzinakos, D., Anderson, A.K. Ecg pattern analysis for emotion detection. Affective Computing, IEEE Transactions on 3, 1 (2012), 102–115.
8. Ali, K., Liu, A.X., Wang, W., Shahzad, M. Keystroke recognition using wifi signals. Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (2015), ACM, 90–102.
9. Altiok, L., de Jong, A., Nijmeijer, H. Fmrcw radar for life-sign detection. Radar Conference, 2009 IEEE (2009), IEEE, 1–6.
10. Aspethans, E.M., Luecken, L.J. Heart rate variability as an index of regulated emotional responding. Review of General Psychology 10, 3 (2006), 229.
11. Boric-Lubecke, O., Massagram, W., Lubecke, V.M., Host-Madsen, A., Jokanovic, B. Heart rate variability assessment using doppler radar with linear demodulation. Microwave Conference, 2008 EuMC 2008. 38th European (2008), IEEE, 420–423.
12. Bocca, M., D’Mello, S. Affect detection: An interdisciplinary review of models, methods, and their applications. Affective Computing, IEEE Transactions on 1, 1 (2010), 18–37.
13. Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Feilenz, W., Taylor, J.G. Emotion recognition in human-computer interaction. Signal Processing Magazine, IEEE 18, 1 (2001), 32–80.
14. Devijver, P.A., Kittler, J. Pattern Recognition: A Statistical Approach, volume 761. Prentice-Hall, London, 1982.
15. Droitcour, A.D., Boric-Lubecke, O., Kovacs, G.T. Signal-to-noise ratio in doppler radar system for heart and respiratory rate measurements. Microwave Theory and Techniques, IEEE Transactions on 57, 10 (2009), 2488–2507.
16. Ekman, P., Friesen, W.V., Ellsworth, P. Emotion in the Human Face: A Guide ‘n for Research and an Integration of Findings. Elsevier, 2012.
17. El Ayadi, M., Kamel, M.S., Karray, F. Survey on speech emotion recognition: Features, classification schemes, and databases. Pattern Recognition 44, 3 (2011), 572–587.
18. Ellken, H.A., Ambady, N. On the universality and cultural specificity of emotion recognition: a meta-analysis. Psychological Bulletin 128, 2 (2002), 203.
19. Fletcher, R., Han, J. Low-cost differential front-end for doppler radar vital sign monitoring. Microwave Symposium Digest, 2009. MTT’09. IEEE MTT-S International (2009), IEEE, 1326–1328.
20. Hu, W., Zhao, Z., Wang, Y., Zhan, H., Lin, F. Noncontact accurate measurement of cardiopulmonary activity using a compact quadrature doppler radar sensor. Biomedical Engineering, IEEE Transactions on 61, 3 (2014), 725–735.
21. Jerritta, S., Murugappan, M., Nagarajan, R., Wan, K. Physiological signals based human emotion recognition: a review. Signal Processing and its Applications (CSPA), 2011 IEEE 7th International Colloquium on (2011), IEEE, 410–415.
22. Kohou, S.E., Bouthiller, X., Lambin, P., Gulcader, C., Michalski, V., Konda, K., Jean, S., Frumently, P., Dauvergne, B., Boulanger-Levesque, N., et al. Emonets: Multimodal deep learning approaches for emotion recognition in video. Journal on Multimodal User Interfaces (2015), 1–13.
23. Kaltiokallio, D., Yigitler, H., Jantti, R., Piispanen, N. Non-invasive respiration rate monitoring using a single cts tx-rx pair. Information Processing in Sensor Networks, IPSN-14 Proceedings of the 13th International Symposium on (2014), IEEE, 59–69.
24. Kamen, P.W., Kron, H., Tonkin, A.M. Poincare plot of heart rate variability allows quantitative display of parasympathetic nervous activity in humans. Clinical Science 92, 2 (1996), 201–208.
25. Keshan, TB, Mishra, A., Breen, WE, Froh, J.J. Gender differences in gratitude: Examining appraisals, narratives, the willingness to express emotions, and gratitude in psychological needs. J. Pers, 77, 3 (2009), 691–730.
26. Kim, J., Andreu, E. Emotion recognition based on physiological changes in music listening. Pattern Analysis and Machine Intelligence, IEEE Transactions on 30, 12 (2008), 2067–2083.
27. Kreibig, S.D. Autonomic nervous system activity in emotion: A review. Biological Psychology 84, 3 (2010), 394–421.
28. Lake, O.E., Richman, J.S., Griffin, MP, Moore, JR. Sample entropy analysis of neonatal heart rate variability. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology 283, 3 (2002), R789–R797.
29. Lane, R.D., McRae, K., Reiman, E.M., Chen, K., Ahern, GL, Thayer, J.F. Neural correlates of heart rate variability during emotion. Neuroimage 44, 4 (2009), 213–222.
30. Lang, P.J. The emotion probe studies of motivation and attention. American Journal Psycholog 92, 5 (1996), 372.
31. Massagram, W., Lubecke, VM, Host-Madsen, A., Boric-Lubecke, O. Assessment of heart rate variability and respiratory sinus arrhythmia via doppler radar. Microwave Theory and Techniques, IEEE Transactions on 57, 10 (2009), 2542–2549.
32. McKinley, S., Levine, M. Cubic spline interpolation. College of the Redwoods 43, 1 (1998), 1049–1060.
33. Patwardi, N., Brewer, L., Tate, G., Kaltiokallio, O., Bocca, M. Breatheinfd: A wireless network that monitors and locates breathing in a home. Selected Topics in Signal Processing, IEEE Journal of 6, 1 (2014), 30–42.
34. Penzel, T., Kantelhardt, JW., Groe, L., Peter, J.-H., Bunde, A. Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea. Biomedical Engineering, IEEE Transactions on 50, 10 (2003), 1143–1151.
35. Piccard, R.W. Affective computing: Toward machine emotional intelligence. Artificial Intelligence: Analysis of affective...
physiological state. Pattern Analysis and Machine Intelligence, IEEE Transactions on 23, 10 (2001), 1175–1191.

37. Postolache, O., Grão, P.S., Postolache, G., Gabriel, J. Cardio-respiratory and daily activity monitor based on fmcw doppler radar embedded in a wheelchair. Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE (2011), IEEE, 1917–1920.

38. Pu, Q., Gupta, S., Gollakota, S., Patel, S. Whole-home gesture recognition using wireless signals. Proceedings of the 19th Annual International Conference on Mobile Computing & Networking (2013), ACM, 27–38.

39. Quintana, D.S., Guastella, A.J., Guthred, T., Hickie, I.B., Kemp, A.H. 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 (2012), 188–172.

40. Saha, D.P., Martin, T.L., Knapp, R.B. Towards incorporating affective feedback into context-aware intelligent environments. Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on (2015), ACM, 49–55.

41. Sakamoto, T., Imaasaki, R., Taki, H., Sato, T., Yoshioka, M., Inoue, K., Fukuda, T., Sakai, H. Feature-based correlation and topological similarity for interbeat interval estimation using ultra-wideband radar. Biomedical Engineering, IEEE Transactions on (2015).

42. Scherer, K.R. Vocal communication of emotion: A review of research paradigms. Speech Communication 40, 1 (2003), 227–256.

43. Sun, L., Sen, S., Koutsourikolas, D., Kim, K.-H. Widraw: Enabling hands-free drawing in the air on commodity wifi devices. Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (2015), ACM, 77–88.

44. Szajdajt, J., et al. Heart rate variability: noninvasive electrocardiographic method to measure the autonomic nervous system. Swiss Medical Weekly 134 (2004), 514–522.

45. Tse, D., Viswanath, P. Fundamentals of Wireless Communication. Cambridge university press, 2005.

46. Wang, W., Liu, A.X., Shahzad, M., Ling, K., Lu, S. Understanding and modeling of wifi signal based human activity recognition. Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (2015), ACM, 65–76.

47. Wiens, A.D., Etemadi, M., Roy, S., Klein, L., Inan, O.T. Toward continuous, noninvasive assessment of ventricular function and hemodynamics: Wearable ballistocardiography. Biomedical and Health Informatics, IEEE Journal of 19, 4 (2015), 1435–1442.

48. Zeng, Z., Panico, M., Raisman, G.I., Huang, T.S. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. Pattern Analysis and Machine Intelligence, IEEE Transactions on 31, 1 (2009), 39–58.

49. Zhao, M., Adib, F., Katabi, D. Emotion recognition using wireless signals. Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking (2016), ACM, 95–108.

50. Zhu, J., Rosset, S., Hastie, T., Tibshirani, R. 1-norm support vector machines. Advances in Neural Information Processing Systems 16, 1 (2004), 49–56.

Mingmin Zhao, Fadel Adib, and Dina Katabi (mingmin, fadel, dk}@mit.edu), Massachusetts Institute of Technology, Cambridge, MA, USA.