

A Point-Process Approach for Tracking Valence using a Respiration Belt

Revanth Reddy, *Student Member, IEEE*, Saman Khazaei, *Student Member, IEEE*,
and Rose T. Faghih, *Senior Member, IEEE*.

Abstract—Emotional valence is difficult to be inferred since it is related to several psychological factors and is affected by inter- and intra-subject variability. Changes in emotional valence have been found to cause a physiological response in respiration signals. In this study, we propose a state-space model and decode the valence by analyzing a person’s respiration pattern. Particularly, we generate a binary point process based on features that are indicative of changes in respiration pattern as a result of an emotional valence response. High valence is typically associated with faster and deeper breathing. As a result, (i)depth of breath, (ii)rate of respiration, and (iii)breathing cycle time are indicators of high valence and used to generate the binary point process representing underlying neural stimuli associated with changes in valence. We utilize an expectation-maximization (EM) framework to decode a hidden valence state and the associated valence index. This predicted valence state is compared to self-reported valence ratings to optimize the parameters and determine the accuracy of the model. The accuracy of the model in predicting high and low valence events is found to be 77% and 73%, respectively. Our study can be applied towards the long term analysis of valence. Additionally, it has applications in a closed-loop system procedures and wearable design paradigm to track and regulate the emotional valence.

I. INTRODUCTION

Human emotions are complex and difficult to interpret because, unlike physiological signals such as Electrocardiogram (ECG) and Electroencephalogram (EEG), emotions are psychological and immeasurable. In addition, emotion is affected by a number of psychological, physiological, and environmental factors that lead to high subject to subject variability and a lack of an objective measure of emotion [1]. Human emotion recognition is a crucial field of study for understanding the mechanisms that drive emotion. It has potential applications in developing actuators to regulate emotions or collecting data in a study where objective measures of emotion are required.

To simplify the diverse spectrum of human emotion, emotions can be modeled on a two-dimensional plane consisting of a measure of valence and arousal for each unique emotion [2]–[4]. Valence refers to the degree of pleasantness or unpleasantness associated with an emotion, while arousal refers to the intensity of the emotion. [3]. In this study, the emotional valence is the affective quality of interest. Previous studies have found that certain physiological changes occur

in response to emotional changes [5]. These changes can be physiologically measured with common devices such as the EEG, ECG, and respiration belts. Many studies have found success using EEG signals to recognize emotions since an EEG is a direct measure of central nervous system activity that is immediately affected by emotional changes [6]. However, EEGs are impractical for continuous monitoring because of their susceptibility to artifacts and lack of portability [7]. In comparison, respiration activity is more practical for continuous measurement and can contain information indicative of emotional changes.

Respiration activity is typically measured through a respiration belt around the abdomen. The belt measures changes in the circumference of the abdomen in response to the expansion and contraction of the lungs during a regular breathing cycle. The resulting signal consists of repeating waves representing the inhalation, exhalation, and breathing cycle of the subject. Respiration differs from other physiological signals since, despite respiration typically occurring involuntarily, it can also be controlled voluntarily. During involuntary respiration, the respiration waveforms are more uniform, where certain qualities of respiration, such as respiration rate and respiration depth, follow a Gaussian distribution around an individual-specific baseline. Voluntary changes in respiration patterns occur in response to an external stimulus. In the absence of confounding factors, an emotional stimulus has been shown to cause changes in respiration patterns [8]. For example, deep and fast breathing has been shown to occur in response to excitement, while slow and shallow breathing has been shown to occur in response to negative emotion. These qualities of respiration in addition to the feasibility of data collection make respiration activity an invaluable predictor of emotional state.

In the emotion recognition paradigm, clinical studies and available datasets collect physiological signals from subjects while emotions are elicited using visual and auditory stimuli. In this study, we utilize the Multimodal Database for Affect Recognition and Implicit Tagging (MAHNOB-HCI) database [9]. The dataset consists of physiological measurements taken from subjects in response to 20 movie and video clips meant to elicit specific emotions. The valence and arousal levels of subjects were self-reported after each clip is shown. The authors of the dataset were able to achieve a valence and arousal classification accuracy of 45.5% and 46.2%, respectively, using a support vector machine classifier with features derived from physiological signals including galvanic skin response, ECG, respiration, and skin temperature. Similarly, other studies have achieved high accuracy

This work was supported in part by NSF under Grant 2226123 - CAREER: MINDWATCH: Multimodal Intelligent Noninvasive brain state Decoder for Wearable Adaptive Closed-loop architectures and in part by NYU start-up funds.

Revanth Reddy, Saman Khazaei, and Rose T. Faghih are with Department of Biomedical Engineering, New York University, USA (e-mail: rfaghih@nyu.edu). Rose T. Faghih served as the senior author.

in emotion recognition using machine learning and deep learning methods. However, these studies do not provide a continuous measure of valence or consider underlying neural stimuli [5], [8].

In this study, we develop a state-space model similar to [10]–[27] to relate emotional valence to deviations from involuntary respiration patterns. For each respiration cycle, we measure the depth of breath, derivative of respiration amplitude (i.e. rate of respiration), and total respiration cycle time. We represent the deviations from a Gaussian distribution in these features as binary point processes for each feature, in which impulses represent events indicative of high valence over time. The point process approach is a common method to model neural spiking activity and represents the underlying neural stimuli driving the involuntary changes in respiration immediately after an emotional stimulus. Several studies have used a similar approach to relate underlying neural stimuli to an unobserved process in the brain [3], [22], [26]. A point process approach is ideal for respiration signals since features are measured at discrete times corresponding to the temporal location of each breath. In particular for each breath, we make a point process by assigning a value of one if both the respiration amplitude and respiration rate probabilities or the respiration cycle time probabilities are below a specified threshold probability value, or zero otherwise. We use an Expectation-Maximization (EM) algorithm to decode the valence state in relation to the high valence indicative events. Eventually, we evaluate the decoded valence state given the self-reported valence ratings from subjects.

II. METHODS

A. Dataset

The MAHNOB-HCI dataset consists of an emotion recognition experiment and an implicit tagging experiment. For the purposes of this study, we will only be analyzing data from the first experiment. The experiment involved data collection from 27 subjects. Subjects were shown 20 video clips meant to elicit different emotional responses. During each trial period where a video clip is shown, subjects are shown a neutral clip meant to induce no emotion, the emotional clip, and then self assess their emotions. During the experiment, several physiological signals are measured such as ECG, EEG, galvanic skin response, eye gaze, respiration, and skin temperature. For each subject, the physiological measurements are recorded over an approximately 50 minute period in which the video clips are shown in a randomized order. From the 27 subjects, we exclude subjects with incomplete recordings from technical difficulties, significant measurement noise, and those who did not consent to publishing. We perform our analysis on the 14 remaining subjects.

B. Pre-Processing

The respiration signal is susceptible to noise, motion artifacts, and wandering baselines that can impact analysis. Similar to previous methods, noise and wandering baselines are corrected using a Butterworth bandpass filter with cutoff frequencies of 0.2 and 2 Hz [28]. This filter corrects the

baseline of the signal to zero and removes polynomial trends from the respiration signal. The cutoff frequencies are selected to remove high frequency sensor noise as well as potential low frequency artifacts. Motion artifacts are present in the respiration signals as irregular increases in respiration amplitude in the order of 5-10 times the magnitude of a typical breath in the same subject. Since the amplitude of the normal signal differs greatly from the signal impacted by motion artifacts, a moving average filter is implemented on the signal to determine the locations at which a motion artifact is likely occurring based on the magnitude of the signal exceeding the feasible range of amplitudes. Once these locations are detected, they are removed from the original signal to limit the impact of motion artifacts on our analysis.

C. Feature Extraction and Point Process Generation

Features from the respiration signal should be extracted based on the underlying physiology such that they can be used to infer the valence level. Based on empirical evidence and prior work [5], [8], we choose to monitor the behavior of respiration amplitude, rate of respiration, and total cycle time of respiration. We use the MATLAB *findpeaks* function to detect the local maxima throughout the respiration signal. The local minima found before and after each peak correspond to the beginning of inhalation and ending of exhalation, respectively. Respiration amplitude is measured as the difference between the amplitude of the peak and the amplitude of respiration when inhalation begins. The total cycle time of respiration is defined as the length of time between the beginning of inhalation and the end of exhalation for a single breath. Here, the rate of respiration refers to the derivative of the respiration amplitude signal. The derivative of the signal represents the rate at which inhalation and exhalation is occurring. For each respiration wave, the respiration amplitude, cycle time, and number of times the derivative of respiration exceeds a threshold within a 2.5 second window before the wave are measured. We also measure the number of times the derivative crossed the threshold to remove extraneous derivative peaks from noise. A 2.5 second interval was selected to allow us to consider the rate of exhalation for the previous breath and the rate of inhalation for the current breath while ignoring the rate of inhalation of the previous breath, which has already been considered. In other words, a 2.5 second interval is the largest time interval in which no more than one pulse can occur.

A binary point process is created based on the features measured at each respiration wave. Assuming that respiration amplitude, rate of respiration, and cycle time are all Gaussian distributed in the absence of an emotional stimulus, significant deviations from the Gaussian distribution can be seen as a response to a stimulus. As a result, for each respiration wave, the probability of a measured value occurring for each feature is calculated using the Gaussian cumulative distribution function. For each respiration wave, we generate a binary impulse if both the respiration amplitude and derivative probabilities or the cycle time probabilities are below a specified threshold probability value. The amplitude and

derivative features are considered together since they contain overlapping information. The derivative feature consideration removes peaks that are impacted by artifacts or noise, which would not contain the rise in amplitude associated with inhalation before. The cycle time is considered separately from these features because it contains temporal information measured from the raw respiration signal. The thresholds are unique for each feature, are subject-specific, and have been determined empirically. Specific thresholds are required to consider the highly subjective emotional responses of each individual. For example, a subject who is more susceptible to physiological changes to emotion may show a 50% increase in respiration amplitude, whereas a less susceptible subject may only show a 30% increase. In other words, the skewness from the Gaussian distribution in the presence of an emotional stimulus likely depends on the intensity of the subject's specific physiological response.

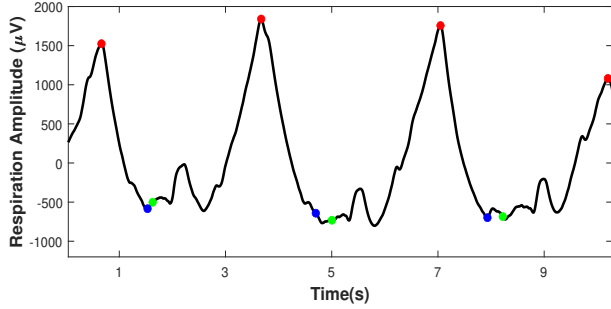


Fig. 1: An example of detected locations in a sample respiration signal. Red points represent the detected respiration peaks at the end of inhalation. Green points represent the start of inhalation. Blue points represent the end of exhalation.

D. State-Space Model

The state-space model relates the binary point process of high valence indicative impulses to a predicted valence state of the subject [29]. We first bin the time axis of the respiration signal into 2.5 second duration intervals (please see II-C). Using the point process generated based on the extracted features, each bin is assigned a value of 1 or 0 based on the presence of a peak. Using the state-space model for point-processes [19], we estimate a continuous valence level throughout the duration of the signal. The state-space model consists of two equations: A state equation describing the latent valence state and an observation equation relating the observed binary pulses to the state equation. We assume that a Gaussian state equation best describes the latent state and a Bernoulli probability model describes the presence of peaks in individual bins as shown in [3], [4], [15]–[18]. We analyze the data from the perspective of an ideal observer, having knowledge of the entire signal, to estimate the valence state at each bin. The valence state equation is described as a random walk through time as follows:

$$z_k = z_{k-1} + \epsilon_k \quad (1)$$

where $\epsilon_j \sim \mathcal{N}(0, \sigma_\epsilon^2)$ is the process noise.

We define the total number of bins as K and index each bin as $k = 1, 2, 3, \dots, K$, where K is the duration of the

entire respiration signal divided by the bin size. For the observation equation, we use s_k to represent the presence of a peak in each bin. s_k follows a Bernoulli distribution where $s_k = 1$ if bin k contains a peak and $s_k = 0$ in the absence of a peak. We define q_k as the probability of a peak occurring in bin k . For the state equation, we define z_k as the latent valence state at bin k . Since the changes in respiration pattern are indicative of changes in valence state, the probability q_k depends on the state z_k . Given an arbitrary value of valence state z_k , the observation equation describes the probability of observing s_k using a Bernoulli distribution as such $P(s_k|q_k) = q_k^{s_k}(1 - q_k)^{1-s_k}$, where q_k is defined by the logistic equation [3], [4], [29]:

$$q_k = \frac{1}{1 + e^{-(\alpha + z_k)}} \quad (2)$$

The value of α depends on the random probability that a peak occurs in a bin at the start of the experiment. It can be estimated using (2) by assuming $z_0 = 0$ and finding q_0 empirically by calculating the proportion of high valence events to all breaths.

Based on the binned point process, we observe $S_{1:K} = \{s_1, s_2, \dots, s_K\}$, which indicates the presence of a peak at each bin. The goal is to estimate the valence state $Z = \{z_1, z_2, \dots, z_K\}$ and σ_ϵ^2 in order to estimate q_k for each bin. Since z is unobserved and σ_ϵ^2 is a parameter, we can use the expectation maximization algorithm to estimate both z and σ_ϵ^2 .

E. Valence State Estimation

The estimation step contains a filter algorithm that calculates the valence state estimate of the subject and a fixed interval smoothing algorithm to refine the estimate based on the ideal observer's perspective [20], [30]. The filter algorithm estimates the valence state $z_{k|k}$ given $S_{1:k}$, with σ_ϵ^2 being replaced by its maximum likelihood estimate. The fixed interval smoothing estimates $z_{k|K}$ given $S_{1:K}$, again with the maximum likelihood estimate of σ_ϵ^2 . We use a Gaussian approximation for the point process observations to design the filter [20], [30]. We define $q_{k|k}$ as the probability of a peak occurring given observations until bin k and $q_{k|K}$ as the probability of a peak occurring given all observations throughout the duration of the signal. The expectation maximization algorithm is further detailed in the following sections.

F. Expectation Step

Given $S_{1:K}$, z_k , $\sigma_\epsilon^{2(l)}$, and z_0^l , we compute the expectation of all data log likelihood at iteration $l + 1$, where l is the iteration number of the algorithm.

Forward Filter: We estimate the state $z_{k|k}$ and the variance $\sigma_{k|k}^2$, given $\sigma_\epsilon^{2(l)}$ and $z_0^{(l)}$, using a recursive filter. This algorithm is derived from the maximizing posterior probability of the Kalman filtering algorithm as described in [31], [32]. The algorithm is described as:

$$z_{k|k-1} = z_{k-1|k-1} \quad (3)$$

$$\sigma_{k|k-1}^2 = \sigma_{k-1|k-1}^2 + \sigma_\epsilon^2 \quad (4)$$

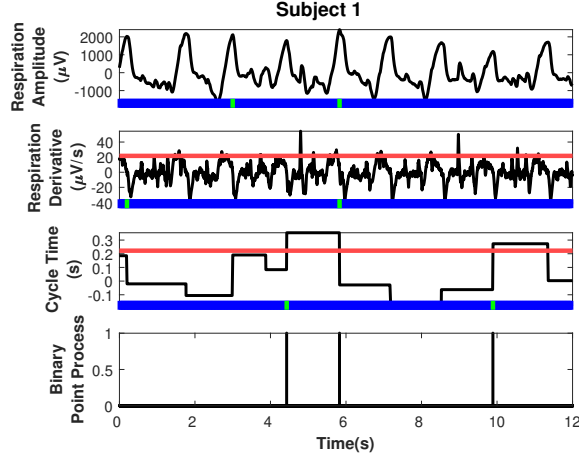


Fig. 2: A sample binary process generated from a respiration signal. Respiration amplitude, respiration derivative, and cycle time are shown. The green markers indicate the presence of a high valence indicative event. The red lines indicate the thresholds for the features to be considered a high valence indicative event. The last plot depicts the generated binary point process based on the events of the individual features. A point is generated if an event is detected in both the amplitude and derivative features or the cycle time feature.

$$z_{k|k} = z_{k|k-1} + \sigma_{k|k-1}^2 \left[s_k - \frac{1}{1 + e^{-(\alpha + z_{k|k})}} \right] \quad (5)$$

$$\sigma_{k|k}^2 = \left[\frac{1}{\sigma_{k|k-1}^2} + \frac{e^{(\alpha + z_{k|k})}}{[1 + e^{-(\alpha + z_{k|k})}]^2} \right]^{-1} \quad (6)$$

for $k = 1, 2, \dots, K$. The initial conditions for the algorithm are $z_0 = z_0^{(l)}$ and $\sigma_{0|0}^2 = \sigma_\epsilon^{2(l)}$. Since the term $z_{k|k}$ appears on either side of (6) and the equation is non-linear, it can be solved using Newton's method.

Backward Filter: Based on the posterior mode estimate of $z_{k|k}$ and the variance $\sigma_{k|k}^2$, we use a fixed-interval smoothing algorithm to compute $z_{k|K}$ and $\sigma_{k|K}^2$. The fixed interval smoother is described as follows:

$$A_k = \frac{\sigma_{k|k}^2}{\sigma_{k+1|k}^2} \quad (7)$$

$$z_{k|K} = z_{k|k} + A_k(z_{k+1|K} - z_{k+1|k}) \quad (8)$$

$$\sigma_{k|K}^2 = \sigma_{k|k}^2 + A_k^2(\sigma_{k+1|K}^2 - \sigma_{k+1|k}^2) \quad (9)$$

For $k = K-1, K-2, \dots, 1$, where the initial conditions are $z_{K|K}$ and $\sigma_{K|K}^2$.

G. Maximization Step

In the maximization step, the expected value of the complete data log likelihood is maximized according to the following:

$$\begin{aligned} \sigma_\epsilon^{2(l+1)} = & \frac{2}{K+1} \left[\sum_{k=2}^K (\sigma_{k|K}^2 + z_{k|K}^2) - \sum_{k=2}^K (A\sigma_{k|K} + z_{k|K}z_{k-1|K}) \right] \\ & + \frac{1}{K+1} \left[\frac{3}{2}(\sigma_{1|K}^2 + z_{1|K}^2) - (\sigma_{K|K}^2 + z_{K|K}^2) \right] \quad (10) \end{aligned}$$

$$z_0^{(l+1)} = \frac{1}{2}z_{1|K} \quad (11)$$

The expectation maximization algorithm iterates between the E and M steps until the convergence of parameters. A change of variables formula is used on the Gaussian probability density function with mean $z_{k|j}$ and variance $\sigma_{k|j}^2$ to determine the probability density for q_k , where $j = k$ for the forward filter and $j = K$ for the backward filter, as shown below [19]:

$$\begin{aligned} f(q|\alpha, z_{k|j}, \sigma_{k|j}^2) = & \frac{1}{\sqrt{2\pi\sigma_{k|j}^2}q(1-q)} \\ & \times \exp \left\{ \frac{-1}{2\sigma_{k|j}^2} \left[-z_{k|j} + \log \frac{q}{(1-q)e^\alpha} \right] \right\} \quad (12) \end{aligned}$$

From (12), we can compute confidence intervals to determine, from an ideal observer's perspective, if a given peak occurred due to more than just random chance.

H. Valence Index and Threshold Determination

During feature extraction and generation of the point process, we use thresholds to determine whether an observed feature is a significant deviation from a Gaussian distribution. These thresholds have to be selected for each subject to minimize the error between the predicted valence index and self-reported valence ratings of the subject throughout the respiration recording. We define the predicted valence index as $p(x_k > x_{median})$, which provides a continuous value for the valence index in terms of the state's deviation from its median value. In other words, the predicted valence index represents the probability that the events are occurring due to more than just random probability [3].

To compare the predicted valence index to the self-reported valence ratings, the self-reported valence ratings have to be converted into a continuous valence rating. First, for each subject, we normalize and rescale valence ratings on a scale of 0 to 1. We set the continuous valence rating equal to the value of the normalized valence rating at the intervals in which the corresponding video clip is shown. During periods between clips, the valence rating is linearly interpolated to create a continuous rating over time. Even though we perform interpolation, the true valence rating at these points is unknown, though the rating is expected to decrease in the absence of an emotional stimulus.

We consider the first 25% of the respiration signal for threshold determination and use the remaining 75% of the signal to test the algorithm. To determine optimal threshold values, mean-squared error is calculated between the continuous self-reported valence rating and the predicted valence index. For each unique configuration of threshold values, a unique point process is generated and valence state is estimated after the EM algorithm. We then calculate the valence index and associated error. Over all configurations of thresholds, we minimize the error to find the ideal thresholds that would cause the predicted valence index to best represent

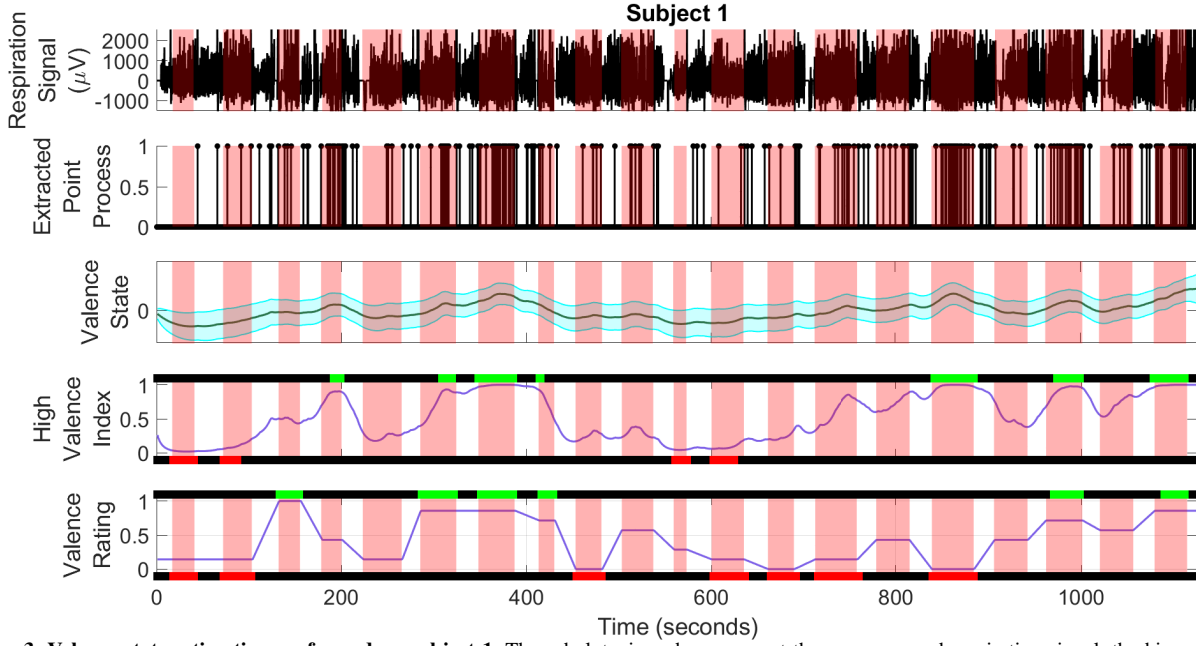


Fig. 3: Valence state estimation performed on subject 1. The subplots, in order, represent the pre-processed respiration signal, the binary point process, the estimated valence state, the predicted valence index, and the normalized and interpolated self-reported valence rating. The red intervals indicate the periods in which video clips are shown to subjects. High valence periods are shown in green for both valence index and self-reported rating. Low valence periods are indicated in red.

the self-reported valence rating. Using the thresholds found from the first portion of the signal, valence estimation is then performed for the entire signal.

I. Model Accuracy

The accuracy of the valence estimation is determined by identifying whether the estimation properly identified periods of high and low valence in the subjects. This is done by categorizing each bin in the valence index and self-reported valence rating as either high valence, low valence, or neither. In the self-reported valence rating, valence was considered to be high if the index was above 0.8, and low if the index was below 0.2. These thresholds were selected empirically to represent periods of very high or very low valence throughout the duration of the recordings. For the valence index, valence was considered to be high if the index was above 0.9, and low if the index was below 0.1. These thresholds are selected because the index represents the probability of belonging in either state. We use stricter definitions for high and low valence events to identify only the extremes of valence with high certainty, since only these periods would be significant enough to alert the user in the implementation of a control system for valence. Accuracy, sensitivity, and specificity are calculated based on whether each high valence period was properly identified by the estimation. For the purposes of accuracy calculations, the periods between clips are ignored. In addition, if the subject had a valence rating drastically conflicting with the expected valence response from a video (i.e. positive valence response to a horror movie clip) according to experimental setup, these portions are corrected to the expected response of the clip. These calculations are then repeated to determine if low valence periods were properly predicted in the valence index.

III. RESULTS

Sample results for peak detection, as well as detection of the beginning and end of breaths, are shown in Fig 1. The cycle time is the time interval between the local minimum detected on either side of a respiration peak. For this subject, the respiration pattern contains additional waveform prior to each breath. In this case, the start of inhalation is considered to be at the start of this smaller wave.

Fig 2 shows a binary point process generated for a given respiration signal based on the three features. A separate binary point process is generated for each feature by marking the points at which the feature exceeds the threshold value. These point processes are combined to generate the final binary point process. The binary point process for each individual feature is also shown. For the respiration amplitude feature, the depth of each breath is measured based on the difference between the amplitude at the start and end of inhalation, rather than just the amplitude at the end of inhalation. For the respiration derivative feature, a point is generated when the derivative exceeds the threshold a certain number of times (See Section IIH). This number limit is set to avoid extraneous derivative spikes not belonging to a breath that can be observed in the plot. The cycle time is shown as a continuous value of the change in cycle time corresponding to the most recent breath. A point is generated where the cycle time increases above the threshold. The resulting combined binary process contains the events that occurred in both the respiration amplitude and derivative features or only the cycle time feature.

Fig 3 shows the valence state estimation results for one subject. The pre-processed respiration signal, the binary point process, the estimated state, the estimated valence

index, and the self-reported valence ratings are shown. Confidence intervals for state estimation are also shown in blue. The red intervals within each plot denote the intervals in which video clips are shown to the subject. The valence rating is converted into a continuous value for comparison to the predicted valence index. This is done by setting the continuous valence rating to the self-reported scores during clips and linearly interpolating for valence values between clips. This transformation allows us to directly compare the valence index to the self-reported scores.

The accuracy of the model in predicting high valence and low valence events was found to be 77% and 73%, respectively. The sensitivity and specificity for high valence prediction was 0.47 and 0.85, while the sensitivity and specificity of low valence prediction was 0.21 and 0.92.

Subject	Participant	Low Valence Accuracy	High Valence Accuracy
1	1	72.58%	78.34%
2	3	92.15%	71.17%
3	4	86.91%	79.55%
4	5	63.19%	93.28%
5	6	76.84%	65.55%
6	8	58.51%	73.94%
7	9	77.38%	81.46%
8	10	90.48%	76.25%
9	15	52.50%	84.53%
10	16	99.70%	67.83%
11	18	83.33%	79.53%
12	19	72.48%	65.92%
13	20	59.27%	86.18%
14	21	58.55%	74.43%

TABLE I: **High and Low Valence Estimation Accuracy for all subjects.** Participant Number corresponds to the numbering of subjects in the MAHNOB-HCI database.

IV. DISCUSSION AND CONCLUSIONS

Respiration peak amplitude, rate of respiration, and respiration cycle time appear to be indicators of high valence. For most subjects, during high valence periods, there would be a very high density of pulses in the binary point process as shown in Fig 3. In some cases, there is also a high density of pulses found in the intervals between video clips being shown. This observation could be a result of the subject's valence being unreported during these intervals, and the subject could have had an emotional response during this resting period. We have also observed that the respiration signal is more prone to motion artifacts between clips that could attribute to this observation.

The valence index calculated based on the latent valence state appears to correlate to the self-reported valence ratings provided by most subjects. However, the quality of correlation varies greatly among different subjects. In more strongly correlated subjects, the predicted valence can capture faster changes in valence (i.e. a brief period of high valence between extended periods of low valence) and even depicts a drop in valence between clips being visible as shown in the estimation for Subject 1 in Fig 3. In moderately correlated subjects, the predicted valence follows general trends in valence (extended periods of high or low valence), but fails to capture faster changes in valence. In weakly correlated

subjects, the predicted valence appears to be slow changing and typically within the neutral valence range. This could be a result of these subjects providing a large number of neutral valence ratings, or the features not being good indicators of valence in these specific subjects. In all subjects, the predicted valence significantly deviates from the reported valence during fast changes in valence in some clips. During these fast changes, the incorrect estimate continues to follow the general trend in valence rather than an abrupt change.

The accuracy of the model in predicting high valence periods validates the potential applications of the model in monitoring the valence state. The high specificity values indicate that when the valence index is above 0.9, the subject is very likely to be experiencing high valence. In comparison, the low valence event prediction yielded lower accuracy values. However, this could be a result of the valence index reaching the lower threshold less frequently. This model seeks to predict high valence from increases in depth of breath and rate of respiration. These features could be less effective in determining the lower end of the valence spectrum since we are not measuring reductions in respiration amplitude that could be more indicative of low valence. The model also has a significantly lower sensitivity than specificity. This could be a result of stricter thresholds for high and low valence period determination in the valence index. These thresholds were stricter to only identify the extreme valence periods. However, the self-reported ratings contained longer and more frequent high and low valence periods than the index, resulting in a high specificity but low sensitivity. These periods could also be affected by the subject's interpretation of the valence rating or individual bias. It should also be noted that the self-reported rating could be inaccurate in representing the subject's true valence level since the ratings are interpolated from only 20 ratings.

The accuracy of the predicted valence also depends on the accuracy of the self-reported valence ratings. Self-reported ratings could be subject to bias and have high subject variability [1]. In some cases, subjects would provide only high or low valence ratings but fail to provide neutral ratings. In other cases, subjects provided mostly neutral ratings with few ratings in either extreme. Normalizing the self-reported ratings can correct for this issue in some subjects, but a true accuracy of the predicted valence cannot be determined without an objective measure of valence.

From a physiological perspective, respiration peak amplitude, rate of respiration, and respiration cycle time appear to be indicators of high valence. High valence is associated with deep and faster breathing [5], [8]. During deep breaths, the respiration amplitude would increase as the abdomen expands further due to the increase in air volume. During faster breathing, the rate of inhalation and exhalation would increase to compensate for the increased amplitude. This physiological response to an emotional stimulus is captured in our features as deviations from the typical physiological behavior. It should also be noted that these features could only be indicative of high valence, which would explain the decrease in quality of estimation as the subjects' valence

rating fluctuates.

The accuracy of the model makes it ideal for continuous monitoring of valence over extended periods. In a practical application to continuously estimate valence, the valence parameter estimation would have to be performed in batches of ~ 3 minute intervals. Threshold estimation would have to be performed for one batch per subject given that self-reported valence levels are also provided. In the future, this model can be improved by implementing a more complex filter that incorporates the amplitude of impulses. The amplitude values can also contain information that would make the estimate more robust to minor changes in valence. The model can be further improved by incorporating other physiological signals that have shown to be predictors of emotion to improve the quality of prediction. Future directions for this project can also include incorporating it with previous work performed with predicting arousal as in [3], [4] to estimate a two-dimensional emotion state including both valence and arousal.

REFERENCES

- [1] L. G. Eaton and D. C. Funder, "Emotional experience in daily life: valence, variability, and rate of change," *Emotion*, vol. 1, pp. 413–421, Dec. 2001.
- [2] J. A. Russell, "A circumplex model of affect," *Journal of personality and social psychology*, vol. 39, no. 6, p. 1161, 1980.
- [3] D. S. Wickramasuriya and R. T. Faghih, "A marked point process filtering approach for tracking sympathetic arousal from skin conductance," *IEEE Access*, vol. 8, pp. 68499–68513, 2020.
- [4] D. S. Wickramasuriya and R. T. Faghih, "A mixed filter algorithm for sympathetic arousal tracking from skin conductance and heart rate measurements in pavlovian fear conditioning," *PLoS one*, vol. 15, no. 4, p. e0231659, 2020.
- [5] W. Sato, T. Kochiyama, and S. Yoshikawa, "Physiological correlates of subjective emotional valence and arousal dynamics while viewing films," *Biological Psychology*, vol. 157, p. 107974, 2020.
- [6] L. A. Schmidt and L. J. Trainor, "Frontal brain electrical activity (eeg) distinguishes valence and intensity of musical emotions," *Cognition & Emotion*, vol. 15, no. 4, pp. 487–500, 2001.
- [7] V. Roy, P. K. Shukla, A. K. Gupta, V. Goel, P. K. Shukla, and S. Shukla, "Taxonomy on EEG artifacts removal methods, issues, and healthcare applications," *J. Organ. End User Comput.*, vol. 33, pp. 19–46, Jan. 2021.
- [8] Q. Zhang, X. Chen, Q. Zhan, T. Yang, and S. Xia, "Respiration-based emotion recognition with deep learning," *Comput. Ind.*, vol. 92–93, pp. 84–90, Nov. 2017.
- [9] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, "A multimodal database for affect recognition and implicit tagging," *IEEE Transactions on Affective Computing*, vol. 3, pp. 42–55, 2012.
- [10] R. Amin and R. T. Faghih, "Physiological characterization of electrodermal activity enables scalable near real-time autonomic nervous system activation inference," *PLoS computational biology*, vol. 18, no. 7, p. e1010275, 2022.
- [11] L. R. Branco, A. Ehteshami, H. F. Azgomi, and R. T. Faghih, "Closed-loop tracking and regulation of emotional valence state from facial electromyogram measurements," *Frontiers in computational neuroscience*, vol. 16, p. 747735, 2022.
- [12] D. D. Pednekar, M. R. Amin, H. F. Azgomi, K. Aschbacher, L. J. Crofford, and R. T. Faghih, "Characterization of cortisol dysregulation in fibromyalgia and chronic fatigue syndromes: a state-space approach," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 11, pp. 3163–3172, 2020.
- [13] T. Yadav, M. M. U. Atique, H. F. Azgomi, J. T. Francis, and R. T. Faghih, "Emotional valence tracking and classification via state-space analysis of facial electromyography," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, pp. 2116–2120, IEEE, 2019.
- [14] M. B. Ahmadi, A. Craik, H. F. Azgomi, J. T. Francis, J. L. Contreras-Vidal, and R. T. Faghih, "Real-time seizure state tracking using two channels: A mixed-filter approach," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, pp. 2033–2039, IEEE, 2019.
- [15] D. S. Wickramasuriya and R. T. Faghih, "A bayesian filtering approach for tracking arousal from binary and continuous skin conductance features," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 6, pp. 1749–1760, 2019.
- [16] D. S. Wickramasuriya and R. T. Faghih, "A novel filter for tracking real-world cognitive stress using multi-time-scale point process observations," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 599–602, IEEE, 2019.
- [17] D. S. Wickramasuriya, M. Amin, R. T. Faghih, et al., "Skin conductance as a viable alternative for closing the deep brain stimulation loop in neuropsychiatric disorders," *Frontiers in neuroscience*, vol. 13, p. 780, 2019.
- [18] D. S. Wickramasuriya, C. Qi, and R. T. Faghih, "A state-space approach for detecting stress from electrodermal activity," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3562–3567, IEEE, 2018.
- [19] A. C. Smith, L. M. Frank, S. Wirth, M. Yanike, D. Hu, Y. Kubota, A. M. Graybiel, W. A. Suzuki, and E. N. Brown, "Dynamic analysis of learning in behavioral experiments," *J. Neurosci.*, vol. 24, pp. 447–461, Jan. 2004.
- [20] A. C. Smith, M. R. Stefani, B. Moghaddam, and E. N. Brown, "Analysis and design of behavioral experiments to characterize population learning," *Journal of Neurophysiology*, vol. 93, no. 3, pp. 1776–1792, 2005.
- [21] T. P. Coleman, M. Yanike, W. A. Suzuki, and E. N. Brown, "A mixed-filter algorithm for dynamically tracking learning from multiple behavioral and neurophysiological measures," *The dynamic brain: An exploration of neuronal variability and its functional significance*, pp. 3–28, 2011.
- [22] M. M. Shanechi, R. C. Hu, M. Powers, G. W. Wornell, E. N. Brown, and Z. M. Williams, "Neural population partitioning and a concurrent brain-machine interface for sequential motor function," *Nature neuroscience*, vol. 15, no. 12, pp. 1715–1722, 2012.
- [23] H. F. Azgomi and R. T. Faghih, "A wearable brain machine interface architecture for regulation of energy in hypercortisolism," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, pp. 254–258, IEEE, 2019.
- [24] S. Santaniello, G. C. McConnell, J. T. Gale, R. T. Faghih, C. Kemere, J. D. Hilliard, and M. Han, "Towards the next generation of deep brain stimulation therapies: Technological advancements, computational methods, and new targets," *Frontiers in Neuroscience*, vol. 15, 2021.
- [25] A. G. Steele, S. Parekh, H. F. Azgomi, M. B. Ahmadi, A. Craik, S. Pati, J. T. Francis, J. L. Contreras-Vidal, and R. T. Faghih, "A mixed filtering approach for real-time seizure state tracking using multi-channel electroencephalography data," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 2037–2045, 2021.
- [26] M. M. Shanechi, R. C. Hu, and Z. M. Williams, "A cortical–spinal prosthesis for targeted limb movement in paralysed primate avatars," *Nature communications*, vol. 5, no. 1, pp. 1–9, 2014.
- [27] S. Khazaei, M. R. Amin, and R. T. Faghih, "Decoding a neurofeedback-modulated cognitive arousal state to investigate performance regulation by the yerkes-dodson law," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2021.
- [28] G. Castegnetti, A. Tzovara, M. Staib, S. Gerster, and D. R. Bach, "Assessing fear learning via conditioned respiratory amplitude responses," *Psychophysiology*, vol. 54, pp. 215–223, Feb. 2017.
- [29] D. S. Wickramasuriya, C. Qi, and R. T. Faghih, "A state-space approach for detecting stress from electrodermal activity," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, July 2018.
- [30] A. C. Smith, L. M. Frank, S. Wirth, M. Yanike, D. Hu, Y. Kubota, A. M. Graybiel, W. A. Suzuki, and E. N. Brown, "Dynamic analysis of learning in behavioral experiments," *J. Neuroscience*, vol. 24, no. 2, pp. 447–461, 2004.
- [31] J. M. Mendel, *Lessons in estimation theory for signal processing, communications, and control*. Prentice Hall, Mar. 1995.
- [32] E. N. Brown, L. M. Frank, D. Tang, M. C. Quirk, and M. A. Wilson, "A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells," *J. Neurosci.*, vol. 18, pp. 7411–7425, Sept. 1998.