June 2023

REAL-TIME AUDIENCE ANALYTICS SYSTEM FOR MEASURING ENGAGEMENT AND SENTIMENT DURING LIVE PRESENTATIONS

Peter Hraška

Lubo Drobny

Jozef Cechovsky

Gregor Gregor

Marek Šuppa

REAL-TIME AUDIENCE ANALYTICS SYSTEM FOR MEASURING ENGAGEMENT AND SENTIMENT DURING LIVE PRESENTATIONS

AUTHORS:
Peter Hraška
Lubo Drobny
Jozef Cechovsky
Gregor Karetka
Marek Šuppa

ABSTRACT

A presenter's ultimate goal is the effective communication of their message to their audience. However, it can be difficult ensuring that such a message resonates with and is understood by everyone in an audience, considering an audience's size, language makeup, and (e.g., possibly remote) location. Techniques are presented herein that leverage wearable technologies, which can provide valuable insights into the engagement, confusion, and amusement of an audience as a whole. Using biometric data (such as heart rate activity and gyroscope and accelerometer readings) that may be collected from such technologies, a deeper real-time understanding of how an audience is feeling may be developed. Aspects of the presented techniques may offer personalized suggestions based on collected data, allowing a presenter to make real-time adjustments to their presentation style to better facilitate an audience's understanding and engagement and thus connect with their audience like never before.

DETAILED DESCRIPTION

Delivering a presentation can be challenging when the presenter lacks real-time feedback from an audience, especially in virtual presentations or situations that involve a language barrier. Such a challenge makes it difficult for the presenter to know when it is time for them to adjust their presentation, cheer up the audience, or provide some space for questions. Even in a large conference room, it can be difficult for a presenter to read subtle shifts in a crowd's emotions, potentially leading to miscommunication and disengagement. As a result, a presenter needs a reliable and effective way of receiving real-time audience feedback to better engage their audience and deliver an effective presentation.

Techniques are presented herein that inform a presenter about the emotions of their entire audience (i.e., a crowd as a whole) during a meeting or a conference. Among other things, the presented techniques may provide real-time feedback and suggestions that can help a presenter improve their presentation.

Aspects of the presented techniques leverage wearable devices, such as smart watches and/or bracelets, that have integrated heart rate sensors, accelerometers, and gyroscopes. Those devices, which may be worn by a participant, may be connected to a smartphone application that is used by a participant and which collects and analyses data concerning heart rate activity. Such an application may establish a baseline using resting heart rate information and may process all of the biometric data, such as heart rate variability (HRV) data, using multi-modal machine learning techniques (which will be described below in connection with gyroscopic data) through pre-trained models that are contained inside of a worn device.

Importantly, according to the presented techniques the sensitive biometric data that is collected, as described above, does not need to leave a participant's device. Such a device would only send a message if a meaningful shift is detected in a device wearer’s emotion.

In connection with HRV data, the presented techniques support the monitoring of a root mean square of successive RR interval differences (RMSSD) value, which is a promising metric that is used in emotion detection, as well as a standard deviation of RR intervals (SDRR) value, a low-frequency (LF) power measure, and a high-frequency (HF) power measure, where an RR interval (also sometime referred to as R-R interval) is the time elapsed between two successive R-waves or successive heartbeats.

Figure 1, below, visually depicts elements of the relationship between RR intervals and the development of an RMSSD value.

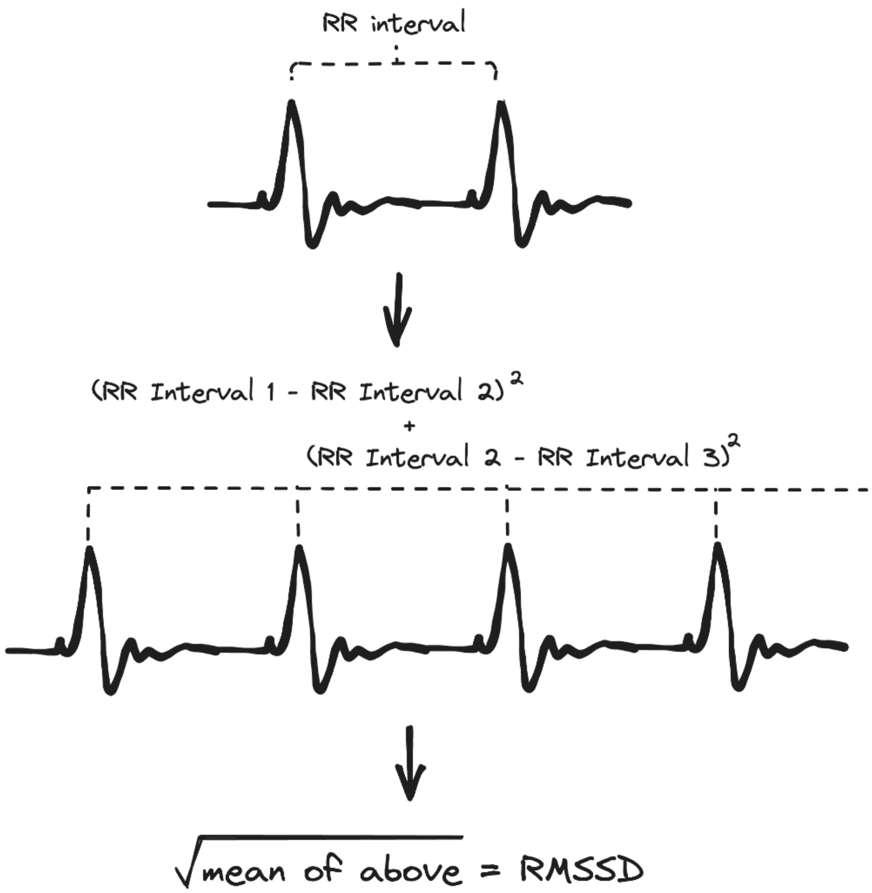

Figure 1: RR Intervals and RMSSD Value

The presented techniques may also analyze data from gyroscopes and accelerometers. Since the intended use of the presented techniques encompasses a wearable device, such a device's integrated gyroscope and accelerometer may be used to detect even small and low-body movements, such as, among others, a chuckle (even an inaudible one), a stirring, unease, or a relaxed state.

To detect a change in an emotional state, data may be collected from a three-axis gyroscope, a three-axis accelerometer, and a heart rate sensor while a wearer (in, for example, a laboratory condition) performs a series of activities that are designed to elicit different emotional states (such as watching a sad movie, listening to music, or engaging in a stressful task).

The collected data may be preprocessed by segmenting it into windows of a fixed duration (such as, for example, five seconds) and then extracting various features from each window. For example, the mean, standard deviation, and skewness of the signals in each window, or even the time-frequency representations, may be calculated. After those features have been extracted, a machine learning algorithm, such as a support vector machine (SVM) or a neural network, may be employed to classify the data into different emotional states. Such an algorithm may accept as input the feature vector for each window and may output a predicted emotional state (like, for example, happy, sad, stressed, etc.).

The following formula encompasses a simplified equation for a linear SVM that may be employed as described above:

$$f(x) = \text{sign}(\sum_{i=1}^{n} \alpha_i y_i k(x_i, x) + b)$$

The above equation develops a predicted emotional state f(x) for the input feature vector x using a linear SVM. The summation operation calculates the weighted sum of the inner products between the support vectors x_i and the input feature vector x, where the weights are given by the product α_i * y_i (with α_i being the Lagrange multiplier and y_i being the class label of x_i). The kernel function $k(x_i, x)$ measures the similarity between the input feature vectors, b is a bias term that shifts the decision boundary, and the sign function maps the output of the SVM to a decision (i.e., an emotional state).

A model as described above is capable of detecting a participant's emotions, such as neutral, amused, confused, angry, etc. Under the presented techniques, only when an emotion threshold is met, and the data is deemed to be significant, need the data be sent to a server. Accordingly, the number of devices that send data to a server can vary depending upon an event's size.

At a server, an event streaming technology may be employed to define a function that processes the incoming data. The following formula expresses one possible function:

$$\left|\left\{e \in E : \frac{|e|}{n} \geq 0.1\right\}\right| \geq m \implies T$$

In the above formula, E is the set of all of the events that are related to a certain emotion during the last five minutes; n is the total number of participants; m is the minimum number of participants that are required to meet a threshold, as defined by the product 0.1 * n (i.e., ten percent of the total number of participants); and T is the event that is to be triggered if the threshold is met.

In brief, the techniques presented herein work by triggering an event when a critical mass of participants experience a certain emotion. As a result, a presenter may receive suggestions in response to the triggering of such an event. For example, if an audience is confused, the presenter may be advised to offer a question and answer (Q&A) session to address their confusion; if the audience is amused, the presenter may be encouraged to empathize with their amusement; and if the audience shows signs of fatigue, a break may be recommended.

In addition to the above-described suggestions, the presented techniques may also display a real-time gauge or pie chart to inform a presenter about the current emotional distribution of an audience. Such a display may be non-distracting and may provide the presenter with valuable information before a critical threshold is met.

Figure 2, below, visually depicts elements of the techniques presented herein (as described and illustrated above) along with their various components.

The presented techniques, as described and illustrated below, may take advantage of a range of existing functionalities. For example, smartphone sensors, such as a gyroscope and an accelerometer, have been employed by some to detect user emotions; subtle changes, such as slight shift in the way that someone walks or types, that are measurable by a smartphone may be used to detect an emotion; the use of a huge camera system for facial recognition and attention detection can be used to discern a crowd's emotion; and movement, audio signals, and a body's temperature can indicate a crowd's emotion during a party or other high body activity events.

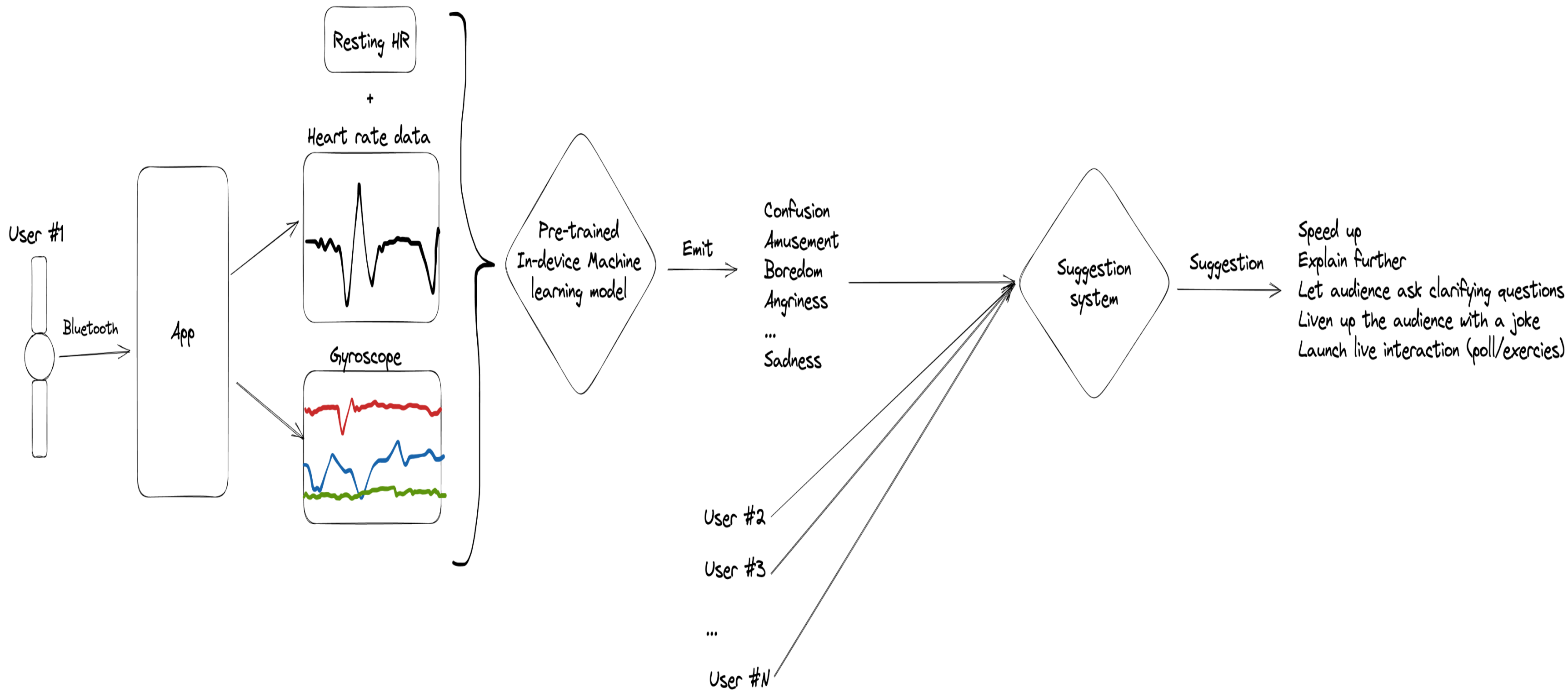

Figure 2: Illustrative Overview

Additionally, the presented techniques encompass a number of novel elements. First, the techniques will work in a live, hybrid, or fully-remote setting and do not require the use of expensive camera equipment. Next, the techniques will work even in a calm setting (such as a meeting or a conference) where a user typically does not move around. Further, during a presentation the techniques may provide a presenter with actionable, relevant, and real-time suggestions regarding how they might improve their presentation. Still further, as a benefit of aggregation the techniques are scalable, from supporting tens of participants to, if necessary, supporting hundreds of thousands of participants.

In summary, to aid a presenter as they deliver their material to an audience, techniques have been presented herein that leverage wearable technologies, which can provide valuable insights into the engagement, confusion, and amusement of an audience as a whole. Using biometric data (such as heart rate activity and gyroscope and accelerometer readings) that may be collected from such technologies, a deeper real-time understanding of how an audience is feeling may be developed. Aspects of the presented techniques may offer personalized suggestions based on collected data, allowing a presenter to make real-time adjustments to their presentation style to better facilitate an audience's understanding and engagement and thus connect with their audience like never before.