TITLE

HEART RATE DRIVEN UNSUPERVISED TECHNIQUES FOR CONTINUOUS MONITORING OF AROUSAL TREND OF USERS

CROSS-REFERENCE TO RELATED APPLICATIONS AND PRIORITY

[001] The present application claims priority to Indian complete specification titled "HEART RATE DRIVEN UNSUPERVISED TECHNIQUES FOR CONTINUOUS MONITORING OF AROUSAL TREND OF USERS" Application No. 201821024337, filed in India on June 29, 2018.

TECHNICAL FIELD

[002] The disclosure herein generally relates to emotional state detection, and, more particularly, to heart rate driven unsupervised techniques for continuous monitoring of arousal trend of users.

BACKGROUND

[003] Neuroscientific and psychological research suggest that emotion plays an important role in influencing intelligence, rationality, behavior and inter-personal interaction. Emotional or mental well-being of each individual is essential for their efficient work performance and for maintaining a comfortable social or work environment. Over the world, depression, stress and related emotional disorders are having adverse effects on any organization's productivity, thereby incurring significant monetary losses. Mental health monitoring can detect these disorders in time and proper interventions can be provided thereafter.

[004] A vital step of any mental health monitoring process is to identify the emotional state of a person dynamically over time. According to the widely used Circumplex Model of Affect (e.g., refer 'J. A. Russell, "A circumplex model of affect," Journal of Personality and Social Psychology, vol. 39, no. 06, pp. 1161–1178, 1980.'), each emotion can be mapped to a two-dimensional space defined by two axes, valence and arousal. Valence signifies the pleasantness and

arousal represents the intensity of a particular emotion. The center point of this space, having neutral valence and medium arousal, caters to the baseline condition. The emotional state of a person moves on this emotion space over the course of time.

[005] The changes in the emotional state are spontaneously reflected in different physiological signals, e.g., cardio-vascular activation, skin conductance, skin temperature, etc. Since these signals are driven by the involuntary Autonomic Nervous System, these can be considered as more reliable modes of emotion recognition than audio-visual emotion channels or questionnaire-based surveys.

[006] Several attempts have been made to determine the emotional state or the valence-arousal level of an individual from physiological data. In traditional research, emotion analyses have been formulated as classification problems. On the other hand, few other researches have treated it as a regression problem where they have designed a real-time continuous arousal monitoring algorithm using features derived from physiological signals. Most of the aforementioned techniques have used supervised approaches for analyzing emotions. The primary challenge in using such approaches is that it is very difficult to obtain a sufficiently large physiological database with 'correct' valence-arousal annotations. This is because the idea of what should be considered as 'correct' in the emotional context is highly person-dependent. Moreover, for a particular person the emotional attributes tend to vary with time, situation and physical health and thus making it even more challenging to estimate the emotional state of a person.

SUMMARY

[007] Embodiments of the present disclosure present technological improvements as solutions to one or more of the above-mentioned technical problems recognized by the inventors in conventional systems. For example, in one aspect, there is provided a processor implemented method for continuous monitoring of arousal trend of users using Heart Rate (HR) driven unsupervised

techniques. The processor implemented method, comprising: obtaining instantaneous HR from a resampled time series of RR intervals of an ECG signal specific to user, wherein the instantaneous HR is obtained for a plurality of time windows from the resampled time series; computing, for each time window of the plurality of time windows, a measured average heart rate (a measured $\overline{HR}$) from the instantaneous HR specific to the user; inputting, to a Kalman Filter, the measured average HR (the measured $\overline{HR}$) and an initial estimate of an ideal $\overline{HR}$, wherein the Kalman Filter comprises a state space model that is designed based on baseline dynamics of an average heart rate ($\overline{HR}$); for each time window until a last time window of the plurality of time windows, performing: estimating an apriori state based on a last instance of an aposteriori state being initialized, wherein the initialized aposteriori state is based on the initial estimate of the ideal $\overline{HR}$ fed to the Kalman filter; determining, using the apriori state, an observation of the state space model of the Kalman filter; estimating an error based on the observation and the measured average heart rate (the measured $\overline{HR}$), and normalizing the error thereof; performing a comparison of the normalized error with a pre-defined threshold; and continually monitoring an arousal of the user based on the comparison to obtain an arousal trend.

[008] In an embodiment, the method may further comprise determining, for each time window of the plurality of time windows, a measurement noise of the state space model of the Kalman Filter based on a change in the arousal; and estimating a Kalman gain of the Kalman Filter using the measurement noise.

[009] In an embodiment, the method may further comprise updating, for each time window of the plurality of time windows, the aposteriori state using the Kalman gain. In an embodiment, the measurement noise is adjusted to a value based on the arousal detected.

[010] In an embodiment, the resampled time series is obtained by: receiving the ECG signal pertaining to the user, the ECG signal is captured for a pre-defined time interval; detecting a plurality of R peaks from the ECG signal; determining a time series of RR intervals based on a difference between positions of consecutive R peaks from the plurality of R peaks; filtering outliers

from the time series of RR intervals to obtain a corrected time series of RR intervals; and resample, at a predefined sampling rate, the corrected time series of RR intervals using a cubic spline interpolation technique.

[011] In another aspect, there is provided a system for continuous monitoring of arousal trend of users using Heart Rate (HR) driven unsupervised techniques. The system comprise a memory storing instructions; one or more communication interfaces; and one or more hardware processors coupled to the memory via the one or more communication interfaces, wherein the one or more hardware processors are configured by the instructions to: obtain instantaneous HR from a resampled time series of RR intervals of an ECG signal specific to user, wherein the instantaneous HR is obtained for a plurality of time windows from the resampled time series; compute, for each time window of the plurality of time windows, a measured average heart rate (a measured $\overline{HR}$) from the instantaneous HR specific to the user; input, to a Kalman Filter, the measured average heart rate (the measured $\overline{HR}$) and an initial estimate of an ideal average heart rate ($\overline{HR}$), wherein the Kalman Filter comprises a state space model that is designed based on baseline dynamics of an average heart rate ($\overline{HR}$); estimate, for each time window until a last time window of the plurality of time windows, an apriori state based on a last instance of an aposteriori state being initialized, wherein the initialized aposteriori state is based on the initial estimate of the ideal average heart rate ($\overline{HR}$) fed to the Kalman filter; determine for each time window until a last time window of the plurality of time windows, using the apriori state, an observation of the state space model of the Kalman filter; estimate, for each time window until a last time window of the plurality of time windows, an error based on the observation and the measured average heart rate (the measured $\overline{HR}$), and normalizing the error thereof; perform, for each time window until a last time window of the plurality of time windows, a comparison of the normalized error with a pre-defined threshold; and continually monitor, for each time window until a last time window of the plurality of time windows, an arousal of the user based on the comparison to obtain an arousal trend.

[012] In an embodiment, the one or more hardware processors are further configured by the instruction to: determine for each time window until a last time window of the plurality of time windows, based on the estimated arousal trend, a measurement noise of the state space model; and estimate a Kalman gain of the Kalman Filter using the measurement noise.

[013] In an embodiment, the one or more hardware processors are further configured by the instruction to: update, for each time window of the plurality of time windows, the aposteriori state using the Kalman gain.

[014] In an embodiment, the measurement noise is adjusted to a value based on the arousal detected.

[015] In an embodiment, the resampled time series is obtained by: receiving the ECG signal pertaining to the user, the ECG signal is captured for a pre-defined time interval; detecting a plurality of R peaks from the ECG signal; determining a time series of RR intervals based on a difference between positions of consecutive R peaks from the plurality of R peaks; filtering outliers from the time series of RR intervals to obtain a corrected time series of RR intervals; and resample, at a predefined sampling rate, the corrected time series of RR intervals using a cubic spline interpolation technique.

[016] In yet another aspect, there is provided one or more non-transitory machine readable information storage mediums comprising one or more instructions which when executed by one or more hardware processors causes for continuous monitoring of arousal trend of users using Heart Rate (HR) driven unsupervised techniques by obtaining instantaneous HR from a resampled time series of RR intervals of an ECG signal specific to user, wherein the instantaneous HR is obtained for a plurality of time windows from the resampled time series; computing, for each time window of the plurality of time windows, a measured average heart rate (a measured $\overline{HR}$) from the instantaneous HR specific to the user; inputting, to a Kalman Filter, the measured average heart rate (the measured $\overline{HR}$) and an initial estimate of an ideal $\overline{HR}$, wherein the Kalman Filter comprises a state space model that is designed based on baseline dynamics of an average heart rate ($\overline{HR}$); for each time window until a last time window of the

plurality of time windows, performing: estimating an apriori state based on a last instance of an aposteriori state being initialized, wherein the initialized aposteriori state is based on the initial estimate of the ideal $\overline{HR}$ fed to the Kalman filter; determining, using the apriori state, an observation of the state space model of the Kalman filter; estimating an error based on the observation and the measured average heart rate (the measured $\overline{HR}$), and normalizing the error thereof; performing a comparison of the normalized error with a pre-defined threshold; and continually monitoring an arousal of the user based on the comparison to obtain an arousal trend.

[017] In an embodiment, the instructions which when executed by the one or more hardware processors may further cause determining, for each time window of the plurality of time windows, a measurement noise of the state space model of the Kalman Filter based on a change in the arousal; and estimating a Kalman gain of the Kalman Filter using the measurement noise.

[018] In an embodiment, the instructions which when executed by the one or more hardware processors may further cause updating, for each time window of the plurality of time windows, the aposteriori state using the Kalman gain. In an embodiment, the measurement noise is adjusted to a value based on the arousal detected.

[019] In an embodiment, the resampled time series is obtained by: receiving the ECG signal pertaining to the user, the ECG signal is captured for a pre-defined time interval; detecting a plurality of R peaks from the ECG signal; determining a time series of RR intervals based on a difference between positions of consecutive R peaks from the plurality of R peaks; filtering outliers from the time series of RR intervals to obtain a corrected time series of RR intervals; and resample, at a predefined sampling rate, the corrected time series of RR intervals using a cubic spline interpolation technique.

[020] It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the invention, as claimed.

BRIEF DESCRIPTION OF THE DRAWINGS

[021] The accompanying drawings, which are incorporated in and constitute a part of this disclosure, illustrate exemplary embodiments and, together with the description, serve to explain the disclosed principles.

[022] FIG. 1 illustrates an exemplary block diagram of a system for continuous monitoring of arousal trend of users using Heart Rate (HR) driven unsupervised techniques in accordance with an embodiment of the present disclosure.

[023] FIG. 2 illustrates an exemplary flow diagram illustrating a method for continuous monitoring of arousal trend of users using Heart Rate (HR) driven unsupervised techniques using the system of FIG. 1 in accordance to an embodiment of the present disclosure.

[024] FIGS. 3A-3B depict correlation analysis of ground truth arousal annotation with (a) $\overline{HR}$, (b) arousal time-series A obtained using Kalman Filter (KF) approach implemented by the system of FIG. 1 in accordance with an example embodiment of the present disclosure.

[025] FIGS. 4A-4B depict Dynamic Time Warping (DTW) analysis of ground truth arousal annotation with (a) $\overline{HR}$, (b) arousal time-series A obtained using the KF approach implemented by the system of FIG. 1 in accordance with an example embodiment of the present disclosure.

[026] FIGS. 5A through 5H depict graphical representations illustrating ground truth annotations and corresponding arousal level time-series obtained using the KF approach implemented by the system of FIG. 1 in accordance with an example embodiment of the present disclosure.

DETAILED DESCRIPTION OF EMBODIMENTS

[027] Exemplary embodiments are described with reference to the accompanying drawings. In the figures, the left-most digit(s) of a reference number identifies the figure in which the reference number first appears. Wherever convenient, the same reference numbers are used throughout the

drawings to refer to the same or like parts. While examples and features of disclosed principles are described herein, modifications, adaptations, and other implementations are possible without departing from the spirit and scope of the disclosed embodiments. It is intended that the following detailed description be considered as exemplary only, with the true scope and spirit being indicated by the following claims.

[028] Existing techniques have been classifying broad levels of arousal (low/medium/high or baseline/arousal) but have been insufficient to provide any information about how the arousal level of a person is changing over time. This continuous arousal trend information is essential to have a proper understanding about the dynamics of the affective behavior of the person over time. Further, many of these existing techniques have used supervised approaches for analyzing emotions. The primary challenge in using such approaches is that it is very difficult to obtain a sufficiently large physiological database with 'correct' valence-arousal annotations. This is because the idea of what should be considered as 'correct' in the emotional context is highly person-dependent. Moreover, for a particular person the emotional attributes tend to vary with time, situation and physical health. Though using self-reports as the ground truth annotations seems to be a viable choice but different persons may perceive the same emotion differently. Thus, self-reports tend to be biased and there is no universally accepted method/technique of removing this bias. Considering these difficulties, an unsupervised emotion monitoring system seems to be a more realistic choice.

[029] Under normal condition, heart rate (HR) of a healthy individual varies between 60 and 100 bpm. Any increase in the arousal level leads to an increasing trend in HR. Thus by an accurate study of the temporal variations of the HR, an overall estimate of the arousal trend can be determined. Though it is very difficult to detect an arbitrary person's absolute level of arousal, the direction of arousal change (i.e., increasing/decreasing/no change) can be identified by the proposed method.

[030] Embodiments of the present disclosure propose a continuous and

unsupervised approach of monitoring the arousal trend of an individual from his heart rate using Kalman Filter. State-space model of the Kalman filter characterizes the baseline arousal condition. Deviations from this baseline model are used to recognize the arousal trend. The primary assumption behind the Kalman Filter is that the state α_t should follow a single linear Gaussian distribution. However, the distribution of $\overline{HR}$ (i.e., its mean and standard deviation) changes with any change in the arousal level. Observing this, the state-space model of the Kalman Filter is constructed to capture only the baseline condition of average heart rate ($\overline{HR}$). In the proposed framework, at any time t, the state α_t signifies the ideal $\overline{HR}$ for the baseline condition and the model observation β_t denotes the measured average heart rate (the measured $\overline{HR}$). During baseline, $\overline{HR}$ is not expected to vary significantly over consecutive time instants. The state model, given by (1), implements this fact by assuming $\overline{HR}$ to be constant over time with the only perturbation from a zero-mean, white, Gaussian noise called the process noise. This noise models the variations in HR associated with normal sinus rhythm. The measurement model, denoted by (2), relates the ideal $\overline{HR}$ estimated from the state model to the measured average heart rate (or the measured $\overline{HR}$) which again involves a zero-mean, white, Gaussian noise called the measurement noise. The measurement noise represents the errors incurred during measurement.

$$\alpha_t = \alpha_{t-1} + \text{process noise} \qquad (1)$$

$$\beta_t = \alpha_t + \text{measurement noise} \qquad (2)$$

[031] The deviation of the physically measured average heart rate (the measured $\overline{HR}$) from the baseline value estimated adaptively by the state-space model of the Kalman Filter is used to infer the arousal trend. Since the state model, given by (1), represents the condition of no arousal change, the $\overline{HR}$ values, for which arousal changes are detected, are treated as unreliable measurements. Hence their effect in refining the model estimates are to be reduced. This is done by increasing the variance of the measurement noise when any change in arousal is detected.

[032] The proposed methods and systems were implemented by the present disclosure for validation by using a publicly available dataset, DECAF (DECoding AFfective), comprising the physiological responses of 30 subjects (or users) while watching 'x' number of video clips (36 movie clips) inducing different emotions. For each clip, annotations of arousal given by experts per second are used to quantify the ground truth of arousal change. Experimental results suggest that the proposed method achieves a median correlation of 0.53 between the computed and expected arousal levels which is significantly higher than that achievable by the state-of-the-art technique(s).

[033] Referring now to the drawings, and more particularly to FIGS. 1 through 5H, where similar reference characters denote corresponding features consistently throughout the figures, there are shown preferred embodiments and these embodiments are described in the context of the following exemplary system and/or method.

[034] FIG. 1 illustrates an exemplary block diagram of a system 100 for continuous monitoring of arousal trend of users using Heart Rate (HR) driven unsupervised techniques in accordance with an embodiment of the present disclosure. In an embodiment, the system 100 may also be referred as an Arousal Detection System (ADS), and interchangeably used hereinafter. In an embodiment, the system 100 includes one or more processors 104, communication interface device(s) or input/output (I/O) interface(s) 106, and one or more data storage devices or memory 102 operatively coupled to the one or more processors 104. The memory 102 comprises a database 108. The one or more processors 104 that are hardware processors can be implemented as one or more microprocessors, microcomputers, microcontrollers, digital signal processors, central processing units, state machines, logic circuitries, and/or any devices that manipulate signals based on operational instructions. Among other capabilities, the processor(s) is configured to fetch and execute computer-readable instructions stored in the memory. In an embodiment, the system 100 can be implemented in a variety of computing systems, such as laptop computers, notebooks, hand-held devices, workstations, mainframe computers, servers, a

network cloud and the like.

[035] The I/O interface device(s) 106 can include a variety of software and hardware interfaces, for example, a web interface, a graphical user interface, and the like and can facilitate multiple communications within a wide variety of networks N/W and protocol types, including wired networks, for example, LAN, cable, etc., and wireless networks, such as WLAN, cellular, or satellite. In an embodiment, the I/O interface device(s) can include one or more ports for connecting a number of devices to one another or to another server.

[036] The memory 102 may include any computer-readable medium known in the art including, for example, volatile memory, such as static random access memory (SRAM) and dynamic random access memory (DRAM), and/or non-volatile memory, such as read only memory (ROM), erasable programmable ROM, flash memories, hard disks, optical disks, and magnetic tapes.

[037] The database 108 may store information but are not limited to, ECG signal (or Electrocardiography signal or Electrocardiogram signal)_and information from the ECG signal pertaining to user(s), and the like. Further, the database 108 stores information pertaining to inputs fed to the system 100 and/or outputs generated by the system (e.g., at each stage), specific to the methodology described herein. More specifically, the database 108 stores information being processed at each step of the proposed methodology.

[038] FIG. 2, with reference to FIG. 1, illustrates an exemplary flow diagram illustrating a method for continuous monitoring of arousal trend of users using Heart Rate (HR) driven unsupervised techniques using the system 100 of FIG. 1 according to an embodiment of the present disclosure. In an embodiment, the system 100 comprises one or more data storage devices or the memory 102 operatively coupled to the one or more hardware processors 104 and is configured to store instructions for execution of steps of the method by the one or more processors 104. The steps of the method of the present disclosure will now be explained with reference to the components of the system 100 as depicted in FIG. 1, and the flow diagram of FIG. 2. In an embodiment of the present disclosure, at step 202, the one or more hardware processors 104 obtain

instantaneous HR from a resampled time series of RR intervals of an ECG signal specific to user. In an embodiment, the resampled time series is obtained by: receiving the ECG signal pertaining to the user, the ECG signal is captured for a pre-defined time interval; detecting a plurality of R peaks from the ECG signal; and determining a time series of RR intervals based on 2 consecutive R peaks. The time series of RR intervals is determined by using a difference between positions of 2 consecutive R peaks from the plurality of peaks detected, in one example embodiment. In an embodiment, the R peak(s) are identified using modified Pan Tompkin's algorithm/technique. Further, outliers from the time series of RR intervals are filtered to obtain a corrected time series of RR intervals and the corrected time series of RR intervals is resampled using a cubic spline interpolation technique. Assuming that mean and standard deviation of the obtained RR intervals are μ_{RR} and σ_{RR} respectively. Samples of the time series lying outside a range (e.g., say $\mu_{RR} \pm 2\sigma_{RR}$) may be treated as outliers and, hence, discarded and the remaining samples are resampled (e.g., using the cubic spline interpolation) at a uniform sampling frequency of say 1 Hertz to obtain a resampled time series. Inverse of the RR intervals in the resampled time series gives the instantaneous HR.

[039] In an embodiment of the present disclosure, at step 204 the one or more hardware processors 104 compute a measured average heart rate ($\overline{HR}$) from the instantaneous HR specific to the user, and the measured average heart rate (the measured $\overline{HR}$) is inputted along with an initial estimate of an ideal $\overline{HR}$ to a Kalman Filter (KF) (not shown in FIG. 1) at step 206. In an embodiment of the present disclosure, the expression "average heart rate" is denoted by ($\overline{HR}$). Therefore measured average heart rate may also be referred as 'measured $\overline{HR}$', in one example embodiment of the present disclosure, and may be interchangeably used hereinafter. In an embodiment of the present disclosure, the measured $\overline{HR}$ is computed for each time window until a last time window from the plurality of time windows. The Kalman Filter comprises a state space model, wherein the state space model is designed based on baseline dynamics of an average heart rate ($\overline{HR}$), in an example embodiment. In an embodiment of the present disclosure,

the measure average heart rate (the measured $\overline{HR}$) is computed by averaging instantaneous HR values over a period of time windows with a certain percentage of overlap. In the present disclosure, the systems and methods of the present disclosure has computed the measured $\overline{HR}$ by averaging instantaneous HR values over 10 seconds time windows with a 90% overlap. This window-wise averaging operation performs smoothing over the normal heart rate variability and thereby the measured averaged heart rate (the measured $\overline{HR}$) provides the slow-varying HR trend.

[040] In an embodiment of the present disclosure, the Kalman Filter is comprised in the memory 102 and executed to perform the steps of the present disclosure and methodology described herein.

[041] The primary assumption behind the KF is that the state α_t should follow a single linear Gaussian distribution. However, the distribution of $\overline{HR}$ (i.e., its mean and standard deviation) changes with any change in the arousal level. Keeping this in mind, the state-space model of the KF is constructed to capture only the baseline $\overline{HR}$ condition. The deviation of the physically measured $\overline{HR}$ from the baseline value estimated adaptively by the KF is then used to infer the arousal trend as described below herein.

[042] In an embodiment of the present disclosure, at step 208 the one or more hardware processors 104 estimate an apriori state based on a last instance of an initialized aposteriori state, wherein the initialized aposteriori state is based on the initial estimate of the ideal $\overline{HR}$ fed to the Kalman filter. The apriori state is estimated for each time window until the last time window of the plurality of time windows.

[043] In an embodiment of the present disclosure, at step 208 the one or more hardware processors 104 determine, using the apriori state, an observation of the state space model of the Kalman Filter and an error is computed based on the observation and the measured average heart rate (or the measured $\overline{HR}$), and the error is normalized thereof at step 210 to obtain a normalized error. In other words, the error is computed and normalized at each time window from the

plurality of time windows.

[044] In an embodiment of the present disclosure, at step 212 the one or more hardware processors 104 perform a comparison of the normalized error with a pre-defined threshold at step 214 and determine an arousal trend of the user based on the comparison at step 216.

[045] In an embodiment of the present disclosure, a measurement noise of the state space model of the Kalman Filter is determined for each time window of the plurality of time windows based on a change in the arousal and a Kalman gain of the Kalman Filter is estimated using the measurement noise respectively. Based on the Kalman gain, the system 100 or the hardware processors 104 further update, for each time window of the plurality of time windows, the aposteriori state. In an embodiment, the measurement noise is adjusted to a value based on the arousal detected. Once the aposteriori state is updated the steps 208 till 216 are repeated for detecting arousal of user and obtaining arousal trend thereof. In other words, during the repetition of steps 208 till 216, each time the measurement noise is determined for computing Kalman gain at each time window and the aposterior state updation happens accordingly.

[046] In the present disclosure, at any time t, the state α_t signifies the ideal $\overline{HR}$ for the baseline condition and observation state β_t denotes the measured $\overline{HR}$. During baseline, $\overline{HR}$ is not expected to vary significantly over consecutive time instants. The state model, given by equation (1), implements this fact by assuming $\overline{HR}$ to be constant over time with the only perturbation from a zero-mean, white, Gaussian noise ω_t (process noise) of variance Υ_t. Here ω_t models the $\overline{HR}$ variations associated with normal sinus rhythm. The measurement model, denoted by equation (2), relates the ideal $\overline{HR}$ estimated from the state space model to the measured $\overline{HR}$ which again involves a zero-mean, white, Gaussian noise ν_t (measurement noise) representing error(s) incurred during measurement. The variance of ν_t is represented by Γ_t.

$$\alpha_t = \alpha_{t-1} + \omega_t, \text{ where, } \omega_t \sim \mathcal{N}(0, \Upsilon_t) \qquad (1)$$

$$\beta_t = \alpha_t + \nu_t, \text{ where, } \nu_t \sim \mathcal{N}(0, \Gamma_t) \qquad (2)$$

where $\mathcal{N}(m, s^2)$ denotes a Gaussian distribution with mean m and variance s^2

[047] At each time instant t, the following steps are performed to detect the arousal trend of the subject.

[048] The steps 208 till 216 and measurement noise, Kalman gain and updation/correction of the aposteriori state are described by of example below.

[049] **Prediction:** As described in step 108, first, the apriori state α_t^- and corresponding variance Λ_t^- are computed as below:

$$\alpha_t^- = \alpha_{t-1}^+ \qquad (3)$$

$$\Lambda_t^- = \Lambda_{t-1}^+ + \Upsilon_t \qquad (4)$$

where α_{t-1}^+ and Λ_{t-1}^+ denote the aposteriori state and variance at time $t-1$ respectively. Since normal HR lies in [60, 100] bpm, the initial state estimate, α_0^+, is set to the average of this range, i.e., 80 bpm. The variance estimate is initialized to a value (e.g., say small value) given by $\Lambda_0^+ = 0.01$. The observation state β_t is then obtained as follows:

$$\beta_t = \alpha_t^- \qquad (5)$$

[050] **Error estimation:** The error Δ_t between the HR trend $\overline{HR}_t$ obtained from ECG at time t and the observation β_t estimated by the Kalman Filter is obtained by equation (6). The present disclosure introduces equation (7) in order to model the normalized error, $\widetilde{\Delta_t}$, where μ_t and σ_t denote the mean and standard deviation of Δ_i for $i = 1, \ldots . (t-1)$.

$$\Delta_t = \overline{HR}_t - \beta_t \qquad (6)$$

$$\widetilde{\Delta_t} = \frac{\Delta_t - \mu_t}{\sigma_t} \qquad (7)$$

[051] It is to be noted that, as the state estimate is initialized to 80 bpm and the averaging of instantaneous HR over overlapping windows is done for the measurement only, the Markov assumption of KF is not violated.

[052] **Arousal detection:** With an increase (or decrease) in arousal, $\overline{HR}_t$ is assumed to rise (or fall) significantly, which in turn, changes Δ_t accordingly. Thus if Δ_t is sufficiently higher (or lower) than the error trend till that instant, i.e., $\widetilde{\Delta_t}$ assumes a significant positive (or negative) value, an increase (or decrease) in arousal level as compared to that of previous instant is inferred. If

absolute value of Δ_t remains within a threshold λ then it is decided that the arousal level has not changed. The present disclosure propose to quantify the direction of arousal change as ξ_t in equation (8). Here the values 0, 1 and -1 of ξ_t represent no change, an increase and a decrease in arousal level respectively.

$$\xi_t = \begin{cases} 0, |\widetilde{\Delta_t}| \leq \lambda \\ 1, \widetilde{\Delta_t} > \lambda \\ -1, \widetilde{\Delta_t} < -\lambda \end{cases} \quad (8)$$

[053] Thus, the arousal level A_t at time t is defined as,

$$A_t = A_{t-1} + \xi_t \quad (9)$$

[054] Here it is noteworthy that the estimated numerical value of A_t does not carry much significance, rather the pattern or trend of the A_t time-series is important.

[055] **Updation/Correction:** Since the state space model, given by (1), represents the condition of no arousal change, the $\overline{HR}_t$ values, for which an arousal change is detected, are treated as unreliable measurements. Hence their effect in refining the model estimates are to be reduced. This is done by increasing the variance Γ_t of the measurement noise as follows:

$$\Gamma_t = (1 - |\xi_t|)\psi_t + |\xi_t| P \psi_t \quad (10)$$

[056] Here, ψ_t denotes an estimate of the variance of measurement noise during baseline condition and P is a positive multiplier that helps in increasing Γ_t when ξ_t is ± 1.

[057] Now the Kalman gain G_t is determined by equation (11). Finally the refined (or updated/corrected) aposteriori state α_t^+ and corresponding variance Λ_t^+ are obtained by equation (12) and (13), respectively, as shown below by way of example expressions:

$$G_t = \frac{\Lambda_t^-}{\Lambda_t^- + \Gamma_t} \quad (11)$$

$$\alpha_t^+ = \alpha_t^- + G_t \Delta_t \quad (12)$$

$$\Lambda_t^+ = (1 - G_t)\Lambda_t^- \quad (13)$$

[058] The process noise variance Υ_t models the small changes in the ideally constant baseline $\overline{HR}$ due to sinus rhythm. Whereas, ψ_t involves

relatively larger variations in the physically measured baseline $\overline{HR}$ due to other physiological and instrumental perturbations. Hence, Υ_t should be considerably smaller than ψ_t. Here the system 100 heuristically choose Υ_t = 0.001 and ψ_t = 0.01. Further, on detection of any arousal change, the measurement noise variance is increased (Γ_t, as shown in equation (10)) significantly. Hence, the multiplier P is set to a high value of 10. Finally, threshold λ is experimentally set as 0.5.

[059] **Dataset:**

[060] The proposed methodology was validated on the movie portion of a publicly available dataset DECAF - a multimodal database for DECoding AFfective user responses (e.g., refer 'M. K. Abadi, R. Subramanian, S. M. Kia, P. Avesani, I. Patras, and N. Sebe, "DECAF: MEG-based multimodal database for decoding affective physiological responses," IEEE Transactions on Affective Computing, vol. 6, no. 3, pp. 209–222, 2015.'). It contains the affective responses of 30 subjects (16 male) aged 27.3 $\pm$ 4.3 years. Each subject watched 36 movie clips spread across nine different emotional genres including happy, sad, exciting, angry etc. The duration of a clip was 79.97 $\pm$ 20.37 seconds. The clips were shown in two sessions each comprising 18 videos. A random order was maintained in each session such that two videos of similar emotional characteristics were not shown successively. During these sessions, ECG (along with other physiological signals) of the subjects were recorded at a sampling frequency of 1 kHz. Seven expert annotators provided per-second arousal ratings ranging between [-1, 1] for each movie clip. The median of the seven annotations at each time instant is taken as the ground truth (GT) for all analysis as suggested by dataset creators (e.g., refer 'M. K. Abadi, R. Subramanian, S. M. Kia, P. Avesani, I. Patras, and N. Sebe, "DECAF: MEG-based multimodal database for decoding affective physiological responses," IEEE Transactions on Affective Computing, vol. 6, no. 3, pp. 209–222, 2015.').

[061] **Results and Discussion:**

[062] Results were obtained by applying the proposed KF based arousal trend monitoring technique on DECAF movie dataset and a thorough discussion

on the advantages of this technique are discussed. It is shown that the predicted arousal time-series A can be used as a direct measure of arousal, compared to the traditional $\overline{HR}$ time-series. The proposed is also compared with the supervised regressive approach given in traditional research (e.g., refer 'B. Grundlehner, L. Brown, J. Penders, and B. Gyselinckx, "The design and analysis of a real-time, continuous arousal monitor," in Wearable and Implantable Body Sensor Networks, BSN. Sixth International Workshop on. IEEE, 2009, pp. 156–161.'). The efficacy of these candidate arousal measures are compared based on the following performance metrics.

1. Pearson's correlation coefficient (ρ) with GT.
2. Minimum Euclidean distance (δ) from GT computed by Dynamic Time Warping (DTW).

[063] While computing δ, the two candidate time-series (measure of arousal and GT) are mapped to [0, 1]. The obtained metric was divided by the length of the aligned time-series in order to bring all the δ values on the same ground for comparison. High value of ρ and low value of δ indicate good trend monitoring performance.

[064] $\overline{HR}$ versus A - performance as an arousal measure:

[065] As mentioned in the above description, a rise in the HR trend corresponds to an increasing arousal level. The formulation of the proposed technique also incorporates this supposition. In order to examine the benefit of the KF based arousal monitoring approach, the present disclosure first considers using the $\overline{HR}$ directly as a measure of arousal. FIG. 3A shows that though a relation between $\overline{HR}$ and arousal is reported, there is no significant correlation (low median and high interquartile range) present in the current case; whereas, in case of A obtained by the KF, as shown in FIG. 3B, the correlation measures are significantly better with an increase of 0.47 in the median value. Moreover, the values of ρ for $\overline{HR}$ spans evenly over the entire range, whereas, in case of the arousal measure A the ρ values are mainly concentrated in the high correlation region. The second performance metric δ reported in FIG. 4A-4B, also reflects

the superiority of A over $\overline{HR}$ as a measure of arousal. More specifically, FIGS. 3A-3B, with reference to FIGS. 1 through 2, depict correlation analysis of ground truth arousal annotation with (a) $\overline{HR}$, (b) arousal time-series A obtained using KF approach in accordance with an example embodiment of the present disclosure. FIGS. 4A-4B, with reference to FIGS. 1 through 3B, depict Dynamic Time Warping (DTW) analysis of ground truth arousal annotation with (a) $\overline{HR}$, (b) arousal time-series A obtained using KF approach in accordance with an example embodiment of the present disclosure.

[066] Moreover, both FIGS. 3A through 4B show that there exists a pattern in the values of the performance metrics obtained in case of A unlike that for $\overline{HR}$, i.e., the KF model consistently performs well for the majority of movie clips, whereas the performance of $\overline{HR}$ is completely random. This substantiates the feasibility of the proposed model. The reason behind the not good performance of $\overline{HR}$, in spite of its reported relation with arousal, might be the presence of noise in the ECG signal from which $\overline{HR}$ is derived. This noise is reduced using the KF, thereby making A a more robust measure of arousal. Moreover the KF takes into account the slow variations of the baseline $\overline{HR}$ over time, which again facilitates obtaining a better estimate of arousal.

[067] **Analysis of trend predictions:**

[068] FIGS. 5A through 5H, with reference to FIGS. 1 through 4B, depict graphical representations illustrating ground truth annotations and corresponding arousal level time-series obtained using the KF approach in accordance with an example embodiment of the present disclosure. More specifically, FIGS. 5A through 5H present representative cases of the DECAF dataset. In the graphical representations depicted in FIGS. 5A through 5H, dotted line representation indicates ground truth annotations and solid line representation indicates corresponding arousal level time-series, in one example embodiment. Here the arousal time-series A obtained by the proposed KF approach and the corresponding GT are plotted, after mapping both the series to the range [0, 1]. It can be seen that the trend of A resembles that of GT

significantly in most of the cases, whereas, the resemblance is less prominent in some others. Even in some cases, as shown in Fig. 5(d), the KF may not perform as required where it predicts an opposite trend with respect to the GT. An in-depth analysis is performed to study the performance of the proposed technique as below:

[069] **Segment-wise arousal analysis:**

[070] It is observed that in the cases shown in FIGS. 5A, 5E and 5G, the arousal time-series, A, obtained using the KF approach matches with the GT more during the first half of the movie clip than the last half. In case of Fig. 5F, the resemblance is more during the last half of the clip. Finally, for FIGS. 5B and 5C, the entire sequence is at par with the annotated GT. Table 1 reports the performance metrics ρ and δ for both A and $\overline{HR}$ considering different segments of the movie clips. More specifically, Table 1 depicts performance analysis for different video segments.

Table 1

Segment of video clip	$\overline{HR}$ versus GT		A versus GT	
	ρ	δ	ρ	δ
First half	-0.11 [-0.66 0.51]	0.17 [0.08 0.31]	0.38 [-0.41 0.88]	0.10 [0.03 0.21]
Second half	0.12 [-0.51 0.63]	0.14 [0.08 0.25]	0.50 [-0.45 0.89]	0.06 [0.02 0.24]
Total	0.06 [-0.43 0.49]	0.14 [0.09 0.23]	0.53 [-0.36 0.83]	0.08][0.02 0.19]

[071] It is found that the later segments of the clips have slightly better values of the metrics. This is expected since the HR trend in the ECG may not instantaneously reflect the feeling a clip tends to induce.

[072] Table 2 depicts a comparison of the proposed algorithm with traditional systems and methods:

Table 2

Framework	ρ	δ
Traditional systems and methods	0.06 [-0.24 0.35]	0.22 [0.13 0.30]
Proposed systems and methods (Kalman Filter)	0.53 [-0.36 0.83]	0.08 [0.02 0.19]

[073] The embodiments of the present disclosure provide systems and methods for continuous arousal trend monitoring that employs the KF technique. Experimental validation on the DECAF movie dataset establishes the efficiency of the proposed method. Compared to $\overline{HR}$, the measure of arousal trend obtained using this proposed technique is found to relate more closely to the GT. This method outperforms the state-of-the-art technique with an increase of 0.47 in the median correlation between the computed and expected arousal levels. It is also observed that the computed arousal trend matches with the GT more during the second half, than the first, of the clips.

[074] The written description describes the subject matter herein to enable any person skilled in the art to make and use the embodiments. The scope of the subject matter embodiments is defined by the claims and may include other modifications that occur to those skilled in the art. Such other modifications are intended to be within the scope of the claims if they have similar elements that do not differ from the literal language of the claims or if they include equivalent elements with insubstantial differences from the literal language of the claims.

[075] It is to be understood that the scope of the protection is extended to such a program and in addition to a computer-readable means having a message therein; such computer-readable storage means contain program-code means for implementation of one or more steps of the method, when the program runs on a server or mobile device or any suitable programmable device. The hardware device can be any kind of device which can be programmed including

e.g. any kind of computer like a server or a personal computer, or the like, or any combination thereof. The device may also include means which could be e.g. hardware means like e.g. an application-specific integrated circuit (ASIC), a field-programmable gate array (FPGA), or a combination of hardware and software means, e.g. an ASIC and an FPGA, or at least one microprocessor and at least one memory with software modules located therein. Thus, the means can include both hardware means and software means. The method embodiments described herein could be implemented in hardware and software. The device may also include software means. Alternatively, the embodiments may be implemented on different hardware devices, e.g. using a plurality of CPUs.

[076] The embodiments herein can comprise hardware and software elements. The embodiments that are implemented in software include but are not limited to, firmware, resident software, microcode, etc. The functions performed by various modules described herein may be implemented in other modules or combinations of other modules. For the purposes of this description, a computer-usable or computer readable medium can be any apparatus that can comprise, store, communicate, propagate, or transport the program for use by or in connection with the instruction execution system, apparatus, or device.

[077] The illustrated steps are set out to explain the exemplary embodiments shown, and it should be anticipated that ongoing technological development will change the manner in which particular functions are performed. These examples are presented herein for purposes of illustration, and not limitation. Further, the boundaries of the functional building blocks have been arbitrarily defined herein for the convenience of the description. Alternative boundaries can be defined so long as the specified functions and relationships thereof are appropriately performed. Alternatives (including equivalents, extensions, variations, deviations, etc., of those described herein) will be apparent to persons skilled in the relevant art(s) based on the teachings contained herein. Such alternatives fall within the scope and spirit of the disclosed embodiments. Also, the words “comprising,” “having,” “containing,” and “including,” and other similar forms are intended to be equivalent in meaning and be open ended in that

an item or items following any one of these words is not meant to be an exhaustive listing of such item or items, or meant to be limited to only the listed item or items. It must also be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural references unless the context clearly dictates otherwise.

[078] Furthermore, one or more computer-readable storage media may be utilized in implementing embodiments consistent with the present disclosure. A computer-readable storage medium refers to any type of physical memory on which information or data readable by a processor may be stored. Thus, a computer-readable storage medium may store instructions for execution by one or more processors, including instructions for causing the processor(s) to perform steps or stages consistent with the embodiments described herein. The term "computer-readable medium" should be understood to include tangible items and exclude carrier waves and transient signals, i.e., be non-transitory. Examples include random access memory (RAM), read-only memory (ROM), volatile memory, nonvolatile memory, hard drives, CD ROMs, DVDs, flash drives, disks, and any other known physical storage media.

[079] It is intended that the disclosure and examples be considered as exemplary only, with a true scope and spirit of disclosed embodiments being indicated by the following claims.

CLAIMS:

1. A processor implemented method, comprising:

obtaining by one or more hardware processors, instantaneous HR from a resampled time series of RR intervals of an ECG signal specific to user, wherein the instantaneous HR is obtained for a plurality of time windows from the resampled time series;

computing by the one or more hardware processors, for each time window of the plurality of time windows, a measured average heart rate (a measured $\overline{HR}$) from the instantaneous HR specific to the user;

inputting by the one or more hardware processors, to a Kalman Filter, the measured average heart rate (the measured $\overline{HR}$) and an initial estimate of an ideal $\overline{HR}$, wherein the Kalman Filter comprises a state space model that is designed based on baseline dynamics of an average heart rate ($\overline{HR}$);

for each time window until a last time window of the plurality of time windows, performing:

estimating by the one or more hardware processors, an apriori state based on a last instance of an aposteriori state being initialized, wherein the initialized aposteriori state is based on the initial estimate of the ideal $\overline{HR}$ fed to the Kalman filter;

determining by the one or more hardware processors, using the apriori state, an observation of the state space model of the Kalman filter;

estimating by the one or more hardware processors, an error based on the observation and the measured average heart rate (the measured $\overline{HR}$), and normalizing the error thereof;

performing by the one or more hardware processors, a comparison of the normalized error with a pre-defined threshold; and

continually monitoring by the one or more hardware processors, an arousal of the user based on the comparison to obtain an arousal trend; wherein the processor implemented method further comprises:

determining by the one or more hardware processors, for each time window of the plurality of time windows, a measurement noise of the state space model of the Kalman Filter based on a change in the arousal;

estimating by the one or more hardware processors, a Kalman gain of the Kalman Filter using the measurement noise, wherein the measurement noise is adjusted to a value based on the arousal detected, wherein the measurement noise models variations in the average heart rate associated with sinus rhythm and wherein the state space model of the Kalman filter infers the arousal trend by increasing variance of the measurement noise when a change in arousal is detected, and wherein deviation of the measured ($\overline{HR}$) from the baseline value estimated adaptively by the state space model of the Kalman filter is used to infer the arousal trend;

performing window-wise averaging operation for smoothing over a normal heart rate variability and wherein the measured ($\overline{HR}$) performs a slow-varying heart rate trend; and

updating by the one or more hardware processors, for each time window of the plurality of time windows, the aposteriori state using the Kalman gain.

2. The processor implemented method of claim 1, wherein the resampled time series is obtained by:

receiving the ECG signal pertaining to the user, the ECG signal is captured for a pre-defined time interval;

detecting a plurality of R peaks from the ECG signal;

determining a time series of the RR intervals based on a difference between the positions of consecutive R peaks from the plurality of R peaks;

filtering outliers from the time series of the RR intervals to obtain a corrected time series of RR intervals; and

resampling, at a predefined sampling rate, the corrected time series of RR intervals using a cubic spline interpolation technique.

3. A system, comprising:

a memory storing instructions;

one or more communication interfaces; and

one or more hardware processors coupled to the memory via the one or more communication interfaces, wherein the one or more hardware processors are configured by the instructions to:

obtain instantaneous HR from a resampled time series of RR intervals of an ECG signal specific to user, wherein the instantaneous HR is obtained for a plurality of time windows from the resampled time series;

compute, for each time window of the plurality of time windows, a measured average heart rate (a measured $\overline{HR}$) from the instantaneous HR specific to the user;

input, to a Kalman Filter, the measured $\overline{HR}$ and an initial estimate of an ideal $\overline{HR}$, wherein the Kalman Filter comprises a state space model that is designed based on baseline dynamics of an average heart rate ($\overline{HR}$);

estimate, for each time window until a last time window of the plurality of time windows, an apriori state based on a last instance of an aposteriori state being initialized, wherein the initialized aposteriori state is based on the initial estimate of the ideal $\overline{HR}$ fed to the Kalman filter;

determine for each time window until a last time window of the plurality of time windows, using the apriori state, an observation of the state space model of the Kalman filter;

estimate, for each time window until a last time window of the plurality of time windows, an error based on the observation and the measured $\overline{HR}$, and normalizing the error thereof;

perform, for each time window until a last time window of the plurality of time windows, a comparison of the normalized error with a pre-defined threshold; and

continually monitor, for each time window until a last time window of the plurality of time windows, an arousal of the user based on the comparison to obtain an arousal trend; wherein the one or more hardware processors are further configured by the instructions to:

determine for each time window until a last time window of the plurality of time windows, based on the estimated arousal trend, a measurement noise of the state space model;

estimate a Kalman gain of the Kalman Filter using the measurement noise, wherein the measurement noise is adjusted to a value based on the arousal detected, wherein the measurement noise models variations in the average heart rate associated with sinus rhythm and wherein the state space model of the Kalman filter infers the arousal trend by increasing variance of the measurement noise when a change in arousal is detected, and wherein deviation of the measured ($\overline{HR}$) from the baseline value estimated adaptively by the state space model of the Kalman filter is used to infer the arousal trend;

perform window-wise averaging operation for smoothing over a normal heart rate variability and wherein the measured ($\overline{HR}$) performs a slow-varying heart rate trend; and

update for each time window of the plurality of time windows, the aposteriori state using the Kalman gain.

4. The system of claim 3, wherein the resampled time series is obtained by:

receiving the ECG signal pertaining to the user, the ECG signal is captured for a pre-defined time interval;

detecting a plurality of R peaks from the ECG signal;

determining a time series of the RR intervals based on a difference between positions of consecutive R peaks from the plurality of R peaks detected;

filtering outliers from the time series of the RR intervals to obtain a corrected time series of RR intervals; and

resampling, at a predefined sampling rate, the corrected time series of RR intervals using a cubic spline interpolation technique.

5. One or more non-transitory machine readable information storage mediums comprising one or more instructions which when executed by one or more hardware processors cause:

obtaining instantaneous HR from a resampled time series of RR intervals of an ECG signal specific to user, wherein the instantaneous HR is obtained for a plurality of time windows from the resampled time series;

computing, for each time window of the plurality of time windows, a measured average heart rate (a measured $\overline{HR}$) from the instantaneous HR specific to the user;

inputting, to a Kalman Filter, the measured average heart rate (the measured $\overline{HR}$) and an initial estimate of an ideal $\overline{HR}$, wherein the Kalman Filter comprises a state space model that is designed based on baseline dynamics of an average heart rate ($\overline{HR}$);

for each time window until a last time window of the plurality of time windows, performing:

estimating an apriori state based on a last instance of an aposteriori state being initialized, wherein the initialized aposteriori state is based on the initial estimate of the ideal $\overline{HR}$ fed to the Kalman filter;

determining, using the apriori state, an observation of the state space model of the Kalman filter;

estimating an error based on the observation and the measured average heart rate (the measured $\overline{HR}$), and normalizing the error thereof;

performing a comparison of the normalized error with a pre-defined threshold; and

continually monitoring an arousal of the user based on the comparison to obtain an arousal trend; wherein the instructions further cause:

determining for each time window of the plurality of time windows, a measurement noise of the state space model of the Kalman Filter based on a change in the arousal;

estimating a Kalman gain of the Kalman Filter using the measurement noise, wherein the measurement noise is adjusted to a value based on the arousal detected, wherein the measurement noise models variations in the average heart rate associated with sinus rhythm and wherein the state space model of the Kalman filter infers the arousal trend by increasing variance of the measurement noise when a change in arousal is detected, and wherein deviation of the measured ($\overline{HR}$) from the baseline value estimated adaptively by the state space model of the Kalman filter is used to infer the arousal trend;

performing window-wise averaging operation for smoothing over a normal heart rate variability and wherein the measured ($\overline{HR}$) performs a slow-varying heart rate trend; and

updating for each time window of the plurality of time windows, the aposteriori state using the Kalman gain.

6. The one or more non-transitory machine readable information storage mediums of claim 5, wherein the resampled time series is obtained by:

receiving the ECG signal pertaining to the user, the ECG signal is captured for a pre-defined time interval;

detecting a plurality of R peaks from the ECG signal;

determining a time series of the RR intervals based on a difference between the positions of consecutive R peaks from the plurality of R peaks;

filtering outliers from the time series of the RR intervals to obtain a corrected time series of RR intervals; and

resampling, at a predefined sampling rate, the corrected time series of RR intervals using a cubic spline interpolation technique.