A Multimodal System for Assessing Alertness Levels due to Cognitive Loading

Anwesha Sengupta, Anirban Dasgupta, *Member, IEEE,* Aritra Chaudhuri, *Member, IEEE,* Anjith George, *Member, IEEE,* Aurobinda Routray, *Member, IEEE,* Rajlakshmi Guha

***Abstract*—This paper proposes a scheme for assessing the alertness levels of an individual using simultaneous acquisition of multimodal physiological signals and fusing the information into a single metric for quantification of alertness. The system takes Electroencephalogram (EEG), high-speed image sequence and speech data as inputs. Certain parameters are computed from each of these measures as indicators of alertness and a metric by a fusion of the parameters for indicating alertness level of an individual at an instant is proposed. The scheme has been validated experimentally using standard neuropsychological tests such as the Visual Response Test (VRT), Auditory Response Test (ART), a Letter Counting (LC), and the Stroop Test. The tests are used both as cognitive tasks to induce mental fatigue as well as a tool to gauge the present degree of alertness of the subject. Correlation between the measures has been studied and the experimental variables have been statistically analyzed using measures such as multivariate linear regression and Analysis of Variance (ANOVA). Correspondence of trends obtained from biomarkers and neuropsychological measures validate the usability of the proposed metric.**

***Index Terms*—Alertness; EEG; Eye Saccade; Speech; Cognitive loading**

I. Introduction

Alertness in human beings illustrates the link between brain and behavior of an individual [1]. The definition includes the basic set of mechanisms that underlie the awareness of the world and the voluntary regulation of thoughts and feelings [2],[3].

Alertness comprises of a state of general wakefulness (tonic arousal or tonic alertness) [4],[5] and the ability to increase response readiness for a short period of time subsequent to external cues or stimuli (phasic alertness) [6]. Neural systems underlying phasic and tonic alertness activation indicate that alertness activation is not unidimensional and highlights the necessity of studying these activation states utilizing multiple parameters of physiologic and behavioral origin [7]. Alertness may also refer to the state of paying close and continuous attention [8].

Increase in attentional effort is defined as the motivated activation of attentional systems in response to detrimental challenges on attentional performance, such as presentation of distractors, prolonged time-on-task, changing target stimulus characteristics and stimulus presentation parameters, circadian phase shifts, stress or sickness. Alertness is maintained by functional states of the neural systems. Research has focused on EEG studies of human arousal and alertness in terms of the sleep/wake cycle [9]. Brain networks based on EEG synchronization are being used recently for studying fatigue and loss of alertness [10],[11],[12].

A drop in alertness levels may be caused by a number of factors, such as chronic kidney disease, high or low blood sugar or thyroid hormone levels, liver failure, brain infection, disorders or injury, heart or breathing problems such as arrhythmia/hypoxia, exposure to toxins or use / abuse of drugs, and mental or physical fatigue. Mental fatigue may be caused by working on cognitively demanding tasks for a considerable span of time and is characterized by a state of decreased cognitive performance [13], [14], [15], [16] increased resistance against further effort [17],[18],[19],[20],[21] and an increased tendency towards less analytic information processing [22].

Extended periods of effort on a single task lead to decline in performance [16]. Maintenance of vigilance is essential in tasks that demand sustained attention over an extended period, including air traffic control, military surveillance, seaboard navigation, industrial process/quality control, medical systems and long-distance driving etc.[21]. Reduction in alertness level of operators involved in these tasks may lead to loss of lives or damage to property [23]. Lyznicki *et al.* have reported in [24] that reduced alertness in drivers is a causative factor in 1% to 3% of all US motor vehicle crashes and nearly 96% of sleep-related crashes involve passenger vehicle drivers and 3% involve drivers of large trucks. Chronic sleep restriction can also have a detrimental effect on neurobehavioral and physiological functioning [22]. In their review of accidents that have occurred due to sleep deprivation, Horne and Reyner [25] have discussed circadian effects and the duration of work to be a profound cause of loss of attention, mainly observed in night shift workers.

It is imperative that a drop in alertness be detected early, and appropriate countermeasures be designed and deployed. A human-machine system intended for monitoring the alertness level will, therefore, be helpful in safety-critical installations. Systems to gauge alertness levels of an individual have been widely reported in literature. EEG, eye saccades and speech signals have been reported as popular non-invasive physiological parameters for determination of alertness level. The velocity profile of eye saccades as a reliable indicator of alertness level has also been discussed in literature [26]. The

Manuscript received February, 2016, revised December, 2016, accepted January, 2017

A.Sengupta, A. Dasgupta, A. George and A. Routray are with the Department of Electrical Engineering, Indian Institute of Technology Kharagpur, India

A. Chaudhuri is with the Advanced Technology Development Center, Indian Institute of Technology Kharagpur, India

R. Guha is with the Centre for Educational Technology, Indian Institute of Technology Kharagpur, India

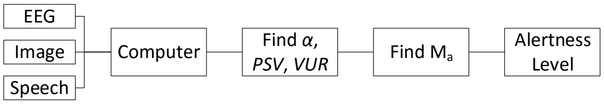

Fig. 1. The Multimodal Scheme

voiced-unvoiced ratio (VUR) of speech signals has also been pointed out as a valid indicator of the loss of attention [27]. A P300-based brain-computer interface using event-related brain potentials (ERP) as a medium for communication between human and computer has been presented by Serby *et al*. [28]. Blood biochemicals such as urea and creatinine have also been reported to be authentic indicators of alertness level [29]. Fletcher *et al*. have developed indices for development of fatigue based on self-assessment in [30], continuous monitoring of which can help to prevent road accidents occurring due to fatigue. Hockey *et al*. describes [31] an experimental procedure to track the changes in level of strain under a cyclic loading procedure using EEG power ratios and Heart Rate Variability (HRV), along with a task load index. The works in ([10], [20], [26], [32]) also attempt to track the changes in the level of alertness using various physiological signals. However, a noninvasive system for assessment of alertness levels remains a challenge on the whole.

A considerable body of work has been carried out on EEG-based fatigue detection and various methods have been reported to find the changes in EEG signal characteristics during the onset of fatigue. The use of EEG to analyze the alertness level of an individual during monotonous tasks, tasks demanding sustained attention or response to specific stimuli has been widely reported in literature [20], [32]. The study in [33] presents new insight into the mechanisms underlying brain waves and resultant EEG signals.

The study in [34] develops an algorithm for automatic recognition of alertness level using full-spectrum EEG recording. Most studies using EEG for estimating changes in alertness levels have revealed an increase in delta and theta frequency[35], and a drop in beta rhythms with increase in fatigue levels [10]. An increase in theta activity has also been correlated with a drop in cognitive task performance [36]. Alpha band frequency has been found to be the most important component for judging alertness level in an expectancy task [20]. Relative energy of different energy bands (alpha, beta, beta/alpha ratio and (alpha+theta)/beta ratio) has often been used as an indicator of fatigue [37]. Relative spectral amplitudes in alpha and theta bands, as well as the mean frequency of the EEG spectrum, have been used to predict alertness level in an auditory response test.

It is evident that most existing systems for determining the level of alertness rely on a single measure of alertness, or at best the combination of a couple of measures ([38], [39], [40], [41], [42]). There also seems to be a gap in combining physiological and neuropsychological measures for the detection of drop in alertness levels. Moreover, the development of systems for alertness assessment using physiological signals can give rise to issues such as[43]:

- Presence of artifacts in EEG signals [10]

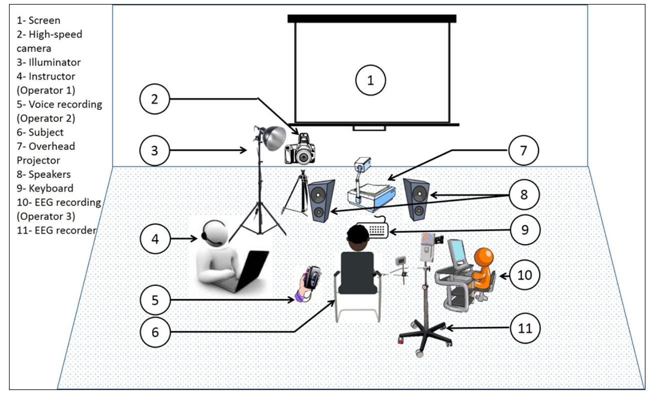

Fig. 2. The experimental setup

- Presence of noise in ocular images due to high-speed acquisition of image data [44]
- Difficulty in locating the iris position due to occlusion by the eyelids [45]
- Contamination of speech data with background noise [46]

A metric for quantifying the drop in alertness levels using a fusion of information from multiple cues is likely to nullify the problems arising out of using various measures individually for assessment of alertness levels. The present work proposes a scheme to gauge the alertness level of an individual, utilizing a number of noninvasive measures. The scheme has been validated experimentally and cross-correlated with neuropsychological measures. The overall scheme is shown in Fig. 1. The major contributions of the research work are:

- Introduction of a multi-modal scheme of simultaneous acquisition of high-frame-rate videos of eye images, speech signal, and EEG records
- Analysis of physiological and neuropsychological/behavioral measures for identifying changes in alertness levels
- Development of a unified metric for assessment of alertness levels by fusion of multi-modal cues

II. Experiment Design for validation

A. Selection of Participants

Thirty subjects (25.16 ± 5.78 years) in the age group $20-40$ years (6 female subjects, 23.67 ± 4.61 years, 24 male subjects, 25.36 ± 5.98 years) were selected randomly from graduate students and employees available on campus at IIT Kharagpur. The ratio of male and female participants was affected by the general gender ratio in the campus as well as by the difficulty of obtaining a clean EEG record with female subjects/subjects with long and/or dense hair.

B. Inclusion and Exclusion Criteria for Subjects

The subjects were required to provide an informed consent prior to their participation in the experiment. Appropriate certificate of approval was obtained from the Institute Ethical Committee at IIT Kharagpur. The suitability of the subject for undertaking the tests was judged and possibility of minor psychiatric disorders was examined using two separate

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2017.2672080, IEEE Transactions on Neural Systems and Rehabilitation Engineering

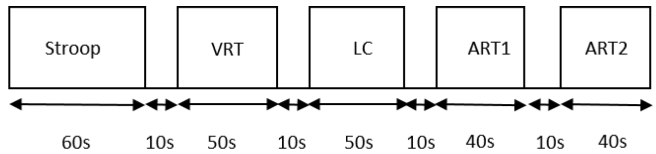

Fig. 3. Timing diagram of a single stage of the experiment

questionnaires. The subjects were also questioned about a possible history of major physical or mental illness and/or addiction. Subjects with a history of substance use or abuse, or subjects under antidepressants, antipsychotics, mood stabilizers or benzodiazepines were excluded from the scope of the experiment. Pre-questionnaire I (Table II) was used to collect information about the food habits, normal sleeping times, vision, and normal duty hours of the subjects, as well as psychological details. Pre-questionnaire II is the GHQ-28 [47]. Table I summarizes the inclusion and exclusion criteria for subjects.

C. Acquisition of Signals

The proposed system uses a computer with 2.60 GHz, core i5 processor having 6 GB of RAM. The speech data is transmitted as an analog signal via the 3.5 mm microphone jack port. The image signals are sent through Gigabit Ethernet protocol using the IEEE 802.3z standard. The EEG signals are sent through an RS-232 standard 9-pin serial port. The data transfer protocol has been shown in Fig. 4.

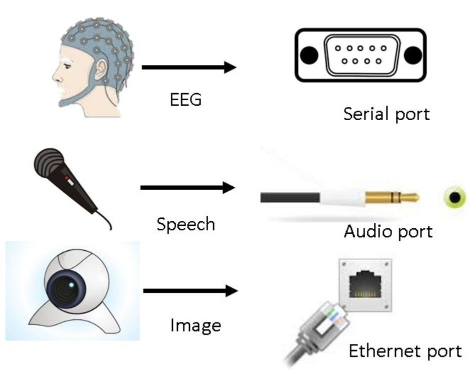

Fig. 4. The data transfer protocol

TABLE I
INCLUSION AND EXCLUSION CRITERIA OF SUBJECTS FOR THE EXPERIMENT

Inclusion criteria	**Exclusion criteria**
- Age range 20 to 40 years.	- Sleep disorders
- Normal/corrected vision and/or hearing ability.	- Psychological illness as assessed by a psychologist
- Right-handed	- History of use of any psychotropic drugs
- Adequately rested for past two hours	- History of head injury/ neuropsychological disorders
	- Color-blindness if any

TABLE II
PRE-QUESTIONNAIRE - I

Are you well-rested right now? - Yes / No / Somewhat
How many hours of sleep did you get last night?
Have you consumed coffee/tea in the past 4 hours? - Yes / No
Did you smoke/consume alcohol in the past 10 hours? - Yes / No
Do you have a history of mental illness? - Yes / No If yes, please provide details
Have you been taking any medication regularly? - Yes / No If yes, provide details of medicine and for how long you have been taking it

D. Experiment Design

The experiment was carried out in a $6m \times 10m \times 4m$ air-conditioned, sound-proof laboratory at IIT Kharagpur. The experimental environment was maintained to ensure that alertness level of subjects is minimally affected by heat, noise, and other environmental conditions. The room temperature was maintained at 22^o C. Fig. 2 shows the relative arrangement of the subject, operators, and instruments used in the experiment. The subject was made to sit in a chair facing a blank white screen of size $0.7m \times 0.9m$ kept at a distance of approximately $1.5m$ from the subject, on which the test contents were projected. The high-speed camera was placed on a tripod at an approximate distance of $0.5m$ from the subject. A filament bulb was used for lighting the face because of the low exposure time of the high-speed camera. The operators were positioned outside the scope of vision of the subject. Responses to the ART and VRT were recorded by an extended keyboard provided to the subject. EEG was recorded at a frequency of 256Hz using 64 scalp electrodes (Ag/AgCl, RMS,India) placed following the international 10-20 system. The electrodes corresponding to the earlobes are used as reference and the electrode at the forehead is used as ground. Speech was recorded at a frequency of 16 kHz using a recorder from Sony held at approximately $0.2m$ from the subject. An event marker was used to synchronize the time with the task.

Use of stimulants such as tea, coffee, etc. for the subjects was restricted for at least four hours before the experiment. The subjects were provided with a set of verbal and visual instructions about the experiment. The subjects were also allowed to undertake a sample test to remove the novelty effect.

The test battery comprised of:

- Stroop Task [48],
- Letter Counting Test [49],
- Psychomotor Vigilance Test (PVT) [50] comprising of an Auditory Response Test (ART) and a Visual Response Test (VRT)

A single stage of the experiment comprised of one complete test set, followed by a questionnaire regarding subjective assessment of the level of cognitive fatigue. The experiment was stopped when the subject complained of extreme discomfort, or when the performance of the subject revealed extreme fatigue. 21 of the 30 participants could continue the experiment till the 15^{th} stage, while the others withdrew earlier. It was

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2017.2672080, IEEE Transactions on Neural Systems and Rehabilitation Engineering

observed that all the subjects could complete 10 stages of the experiment.

E. Neurocognitive Tests

1) Stroop: The subject is required to name the color of the ink in which color-words are printed, instead of the name of the color. The performance is evaluated on the basis of the time taken to complete the task and the accuracy (number of correct responses). In the present experiment, the subject is asked to read the Stroop set row-wise.

2) Visual Response Test (VRT): In this test, the subject is required to press the direction keys on the keyboard according to the position (top, bottom, left, right) in which a green circle appears on the screen. The response time from the appearance of the circle is the marker of attention. The response is taken to be correct when the subject presses the key in response to the appearance of a green circle. Response to any other colored circle is considered as a wrong response.

3) Letter Counting (LC) Test: The subject is required to find out the number of occurrences of a particular alphabet in a 13×13 array of English letters placed on the screen. Evaluation criteria is the completion time (computed from the speech data as a difference between the onset of stimulus to the voice response of the subject) and accuracy.

4) Auditory Response Test (ART): A random recorded sequence of 1's and 2's (total 40 in number) is played. In the first round of the test, the subject is required to press the same digit as the sound played (1 for 1 and 2 for 2), whereas in the second round, the subject is required to press the opposite digit as the recording (that is, 2 for 1 and 1 for 2). The inter-stimulus-interval is kept between 1 and 2 seconds to minimize the chance of the subject remembering the sequence from the previous run.

III. Computation of Parameters and Development of Metric

The system captures EEG, high-frame-rate videos of eye images and speech signals using an EEG recorder, a high-speed camera, and a microphone respectively. Alpha energy, Peak Saccadic Velocity (PSV) and VUR are computed from these measures. Finally, a metric M_a is computed using (24).

A. Calculation of EEG Energy

The raw EEG signal is passed through a band-pass filter with cut-off frequencies of 0.5 Hz and 30 Hz followed by normalization of the data. This operation ensures the removal of power-line artifacts and any unwanted bias that might have been introduced during experimental recording. A Hanning window of length N as given in (1) has been used for windowing the data as

$$w(n) = \frac{1}{2}\left(1 - \cos\left(\frac{2\pi n}{N-1}\right)\right) \tag{1}$$

The windowed EEG signal is decomposed into Delta (δ),($0.5-4$ Hz), Theta (θ), ($4-8$ Hz), Alpha (α),($8-14$ Hz) and Beta (β) ($14-30$ Hz) bands [51] using the Discrete wavelet transform (DWT) technique using Daubechies wavelet 4[52]. The in-band ocular and muscular artifacts have been removed by a wavelet thresholding approach [53]. The energy at the jth level of decomposition has been computed from the wavelet coefficients as

$$E_j = \sum_{k=1}^{N} C_j^2(k) \tag{2}$$

Here $C_j(k)$ is the wavelet coefficient (approximate or detail) and N is the total number of wavelet coefficients at the jth level. Hence the relative energy p_j of a particular band represented at the jth level is given by

$$p_j = \frac{E_j}{\sum_j E_j} \tag{3}$$

The system computes the relative energy p_j for the alpha band ($8-13$ Hz) to determine the alertness level of the subject.

B. Calculation of Ocular Parameters from Video Data

The localization of the eye region[38] is the first step in the computation of PSV. The eyes are detected from the input images using Haar-like features.

1) Preprocessing: The preprocessing steps used for iris detection include correction for illumination variation and removal of glint.

a) Illumination Variation: The localized eye region is gamma corrected to make the algorithm illumination invariant [54]. The gamma correction enhances the local dynamic range of the eye image in dark regions while compressing it in bright areas. The transformation may be written as

$$I_{out} = cI_{in}^{\gamma} \tag{4}$$

Where I_{in} is the input image and I_{out} is the corresponding gamma corrected image(both non-negative real values). c and γ are constants. γ was selected adaptively in the range $(0, 1]$.

b) Glint Removal: Removal of glint is carried out by use of a structuring element B to create a morphological opening using a disk of radius 3 pixels (obtained empirically).

$$I_{op} = I_{out} \circ B = (I_{out} \ominus B) \oplus B; \tag{5}$$

where $\ominus$ and $\oplus$ denote erosion and dilation, respectively.

2) Iris Detection: The iris is detected using the Timm-Barth algorithm [55] based on image gradients. This method is found to be invariant to changes in scale, pose, contrast and to variation in illumination. The center of the iris is obtained from the detected circle as,

$$\theta_p(k) = [u_k v_k]^T \tag{6}$$

where u_k and v_k represent the x and y locations of the iris respectively, at the kth frame.

3) Tracking the iris: The position of the center of the iris is estimated using a Kalman Tracker when the iris gets occluded by the eyelids or any other object. States x_k are defined as

$$x_k = [u_k v_k]^T \tag{7}$$

The states are assumed to obey the linear stochastic difference equation

$$x_{k+1} = \phi_k x_k + \omega_k \tag{8}$$

ϕ_k is the 3×3 state transition matrix for the kth sample. ω_k is the process noise, assumed to be Additive White Gaussian Noise (AWGN) with covariance matrix $\mathbf{Q}$. The matrix ϕ_k is the 3×3 state transition matrix, for the kth sample. Considering a constant acceleration model, the matrix ϕ_k is found out to be

$$\phi_k = \begin{bmatrix} 1 & k & \frac{k^2}{2} \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix} \tag{9}$$

The filtering problem is to estimate x_k from the measurements z_k using

$$z_k = x_k + \epsilon_k \tag{10}$$

Here, ϵ_k is the measurement noise assumed to be AWGN with co-variance matrix $\mathbf{R}$. The measurements z_k are obtained from the iris detection algorithm as explained earlier. It is assumed that ω_k and ϵ_k are mutually uncorrelated. We carry out the updates recursively using (11) - (14) to find the state estimates.

$$\hat{x}_k^- = \phi_k \hat{x}_{k-1}^- \tag{11}$$

$$\mathbf{P}_k = \phi_k \mathbf{P}_{k-1} \phi_k^T + \mathbf{Q} \tag{12}$$

The Kalman Gain $\mathbf{K}_k$ is obtained as

$$\mathbf{K}_k = \mathbf{P}_k^- (\mathbf{P}_k^- + \mathbf{R}) \tag{13}$$

The error co-variance matrix $\mathbf{P}$ is updated as

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k)\mathbf{P}_k^- \tag{14}$$

The Kalman tracker not only provides a better estimate robust against AWGN, but also provides an optimal estimate of the iris center, with missing measurements which happens during occlusion.

4) Eye Corner Detection: Since the position of eye corners is invariant under variation of facial expressions, levels of eye closure and gaze, presence of eyelashes and eye makeup, these have been chosen as reference points for computing relative iris center position. The eye corners have been selected as reference positions because the position of eye corners does not vary with different facial expressions, levels of eye closure, gaze, eyelashes or makeup. The work in describes about detection of both the eye corners (temporal as well as nasal). Semantic Feature Extraction has been used to detect the eye corner [56]. In this method, the eyelids are fitted to construct an angle model. Two features are extracted from this model. The first feature characterizes the appearance difference between inner and outer canthus regions. The second feature emphasizes the role of the canthus angle bisector region. Finally, the eye corners are detected using a classifier obtained as a combination of the two features.

5) Computation of Peak Saccadic Velocity (PSV): Saccadic amplitude $\theta(k)$ at kth frame is computed as the normalized difference between the pupil centre position, $\theta_p(k)$ and the nasal eye corner position, $\theta_{c1}(k)$. The normalization is carried out by dividing the distance of the iris center from the eye corner as a fraction of the eye width $\theta_w(k)$ i.e. difference of nasal $\theta_{c1}(k)$ and temporal $\theta_{c2}(k)$ corner positions. This normalization is done to nullify the effect of the distance from the camera.

$$\theta_w(k) = \sqrt{|\theta_{c2}(k) - \theta_{c1}(k)|^2} \tag{15}$$

$$\theta(k) = \frac{\theta_p(k) - \theta_{c1}(k)}{\theta_w(k)} \tag{16}$$

The PSV ω_p at kth frame is hence computed as

$$\omega_p = max(\theta(k) - \theta(k-1)) \tag{17}$$

C. *Calculation of Voiced-to-Unvoiced Ratio from Speech Data*

Speech signals can be classified into three broad classes based on the mode of generation viz. silence, voiced and unvoiced speech [57]. Voiced speech is produced by the vibrations of vocal cords whereas unvoiced sounds are due to the turbulence of air in vocal tract (mouth, tongue, velum, etc.). Unvoiced speech has lower energy and higher zero crossing rates as compared to voiced speech. The speech signal in the proposed system is captured at 20 kHz [58]. The vocal fold vibrations may be assumed periodic if the signal is of short duration ($10 - 30$ms) [59]. For this reason, the speech data is processed in small frames of size $10 - 30$ ms. Singular Value Decomposition (SVD) is performed framewise to remove noise and redundant information.
Let $\mathbf{M}$ be the covariance matrix obtained from a frame of the speech data. $\mathbf{M}$ is decomposed as

$$\mathbf{M} = \mathbf{U\Sigma V}^* \tag{18}$$

Here, $\mathbf{U}$ is a $M \times M$, unitary matrix, $\mathbf{\Sigma}$ is a $M \times L$ diagonal matrix bearing non-negative real numbers on the diagonal, and $\mathbf{V}$ is an $L \times L$ unitary matrix. The diagonal entries of are the singular values of $\mathbf{M}$. The singular values are arranged in the descending order. The first five singular values are reconstructed to obtained the filtered speech signal. Voiced and unvoiced classification is carried out using a support vector machine (SVM) [60] with the Mel frequency cepstral coefficients (MFCC) as features. MFCC is a representation of the short-term power spectrum of the speech signal. First, the DCT coefficients of the filtered speech data x_n are obtained as

$$X_k = \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}\left(n + \frac{1}{2}\right)k\right] \qquad k = 0, \ldots, N-1. \tag{19}$$

A triangular window $w(n)$ of length L is used on the N DCT components to truncate the DCT spectra. The window is

$$w(n) = 1 - \left|\frac{n - \frac{N-1}{2}}{\frac{L}{2}}\right|, \tag{20}$$

The coefficients are then converted from f hertz into m mel using

$$m = 2595 \log_{10}(1 + \frac{f}{700}) \tag{21}$$

Once the MFCCs are obtained, the SVM returns the voiced speech $v_s(n)$ and unvoiced speech $u_s(n)$f lengths N_v and N_u

respectively. The VUR is finally obtained as the ratio of the energies as

$$VUR = \frac{\sum_0^{N_v - 1} v_s^2(n)}{\sum_0^{N_u - 1} u_s^2(n)} \tag{22}$$

D. Calculation of Correlation between Measures

Cross-correlation between different measures have been calculated using Pearson's Correlation coefficient as

$$R(i,j) = \frac{C(i,j)}{\sqrt{C(i,i)C(j,j)}} \tag{23}$$

C is the co-variance matrix of a given measure and R is the matrix of normalized correlation coefficients with i and j as the indices of row and column respectively.

E. Development of Metric

A combination of multiple cues is likely to be more accurate in indicating alertness levels than the use of individual cues. Literature [10] reports EEG to be a more authentic measure of alertness as compared to VUR and PSV. The present paper proposes a unified metric using a combination of information from the three cues. A metric M_a is proposed in (24) as a weighted combination of parameters obtained from EEG, eye saccades and voice.

$$M_a = 0.6M_\alpha + 0.25M_{sac} + 0.15M_{vur} \tag{24}$$

$$M_{sac} = 1 - PSV_n, 0 < M_{sac} \leq 1 \tag{25}$$

$$M_{vur} = 1 - VUR_n, 0 < M_{vur} \leq 1 \tag{26}$$

α_n, PSV_n and VUR_n are normalized alpha energy, PSV and VUR respectively. The weights have been obtained empirically using the experimental data and are assigned to the parameters based on the correlation of each to the level of alertness.

$$M_\alpha = \alpha_n = \frac{\alpha}{max(\alpha)}, 0 < M_\alpha \leq 1 \tag{27}$$

$$PSV_n = \frac{\omega_p}{max(\omega_p)} \tag{28}$$

$$VUR_n = \frac{VUR}{max(VUR)} \tag{29}$$

$max(.)$ indicates the maximum value of the sequence. $0 < M_a \leq 1$, with 0 corresponding to the maximum level of alertness.

IV. Results and Discussion

A. Standard error of the mean (SEM)

The variation of the biomarkers (Alpha energy, PSV and VUR) and subjective scores for all subjects and stages has been shown in Fig. 5, with the solid black lines indicating the mean trend. The non-uniformity observed may be explained as an attempt of the subjects to force attentiveness after a drop in performance levels in the corresponding previous stages. The trends reflected have been corroborated by p-values and z-scores corresponding to the performance of subjects in the tasks as shown in Table VI. Low p-values indicate a significant decline in the task performance with progression in stages. The overall variation in biomarkers and results of neuropsychological parameters is distinctly indicative of the onset of fatigue due to the inability of subjects to maintain sustained attention due to the monotonicity of the tasks and due to the tasks being administered continuously[7].

B. Cross Correlation

Correlation values between measures have been tabulated in a confusion matrix in Table III. A value of $R(i,j)$ closer to 1 indicates high correlation between corresponding measures and a negative value indicates an inverse variation between the corresponding pair of parameters. It may be noticed that both PSV and VUR decrease, while the remaining measures increase with a decrease in alertness.The high positive correlation of both SRT and VRT with PSV indicates the involvement of visual attention in these tasks.

C. Regression Analysis

A multi-linear regression analysis among the physiological data has been performed and the results presented in Fig. 6. It may be seen that alpha energy increases, while PSV and VUR decrease with progression in the number of stages. It is evident in from the trend in the regressors that corresponding changes occur in both the biomarkers and results of behavioural tests. The coefficient of determination for each measure is tabulated in Table IV.

D. Inferential Statistics

For the present experiment, it is important to observe simultaneous changes of the behavioral variables with progression in stages. The hypotheses for the analysis are defined as:
$H_0 =$ The null hypothesis: when all variables have same mean through the stages
$H_a =$ The alternate hypothesis: when all variables have different means through the stages
The z-scores and p-values for various cognitive tasks have been examined to study the significant decline in the variables with increase in stages and the results tabulated in Table VI. Bonferroni correction has been applied to the data.
It may be seen that results of all the other tests except SRT have high levels of significance, with results for VRT being the most significant. This may be attributed to the fact that response to a visual stimulus provides a better marker of alertness compared to other kinds of stimuli, while repeated performance of the Stroop task may show an improvement in outcome due to practice, in spite of a reduction in vigilance of the subjects.

E. Analysis of Variance (ANOVA) with Repeated Measures

The variation in different physiological and neuropsychological measures with respect to the stages has been analysed using standard repeated measures ANOVA [61] and results presented in Table V. The physiological parameters (alpha energy, PSV and VUR) and performances in neuropsychological tests (SRT, VRT, LC, ART1 and ART2) have been

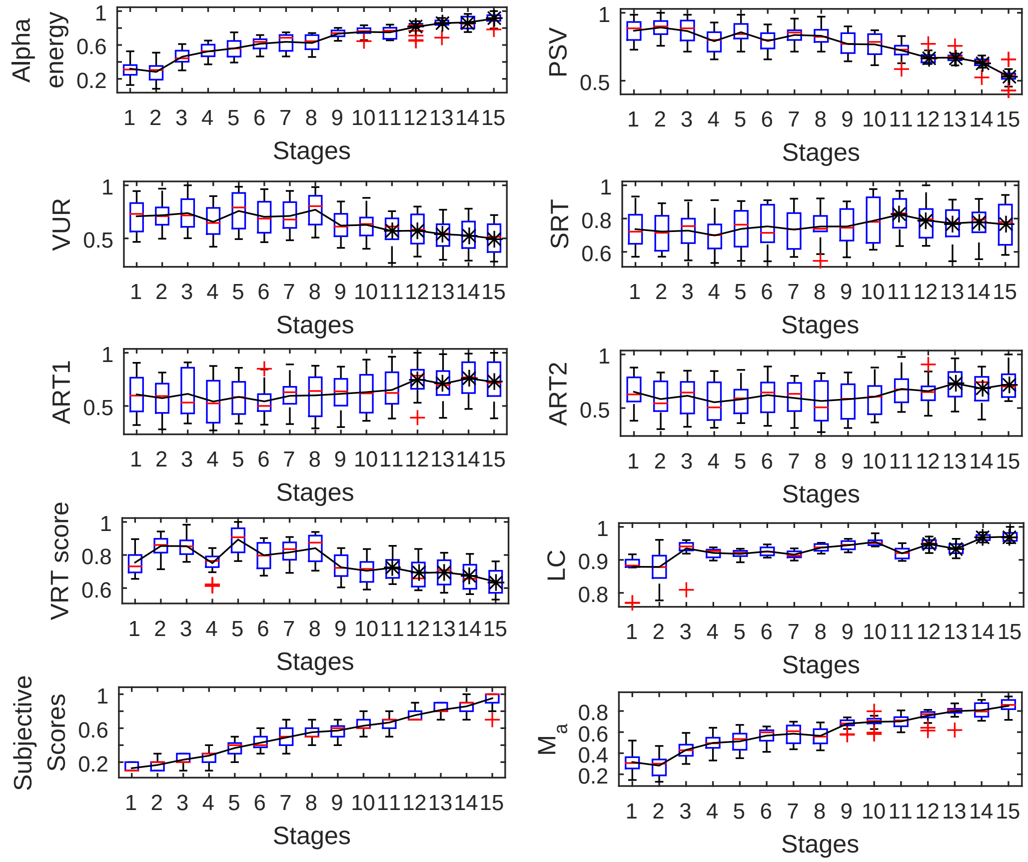

Fig. 5. Boxplots showing the variation of psychophysiological parameters with varying stages of alertness; stages corresponding to on-set of fatigue are marked with an asterisk and outliers are indicated using a red '+'.

TABLE III
THE PEARSON'S CORRELATION COEFFICIENTS SHOWN AS A CONFUSION MATRIX AMONG ALL THE MEASURES

	Alpha	**PSV**	**VUR**	**SRT**	**ART1**	**ART2**	**VRT**	**LC**	**Sub**	M_a
Alpha	1	-0.84	-0.74	0.30	0.71	0.32	-0.63	0.40	0.97	0.99
PSV	-0.84	1	0.92	-0.27	-0.78	-0.57	0.86	-0.43	-0.90	-0.86
VUR	-0.74	0.92	1	-0.32	-0.80	-0.61	0.92	-0.39	-0.82	-0.77
SRT	0.30	-0.27	-0.32	1	0.36	0.48	-0.26	0.11	0.31	0.31
ART1	0.71	-0.78	-0.80	0.36	1	0.58	-0.69	0.43	0.80	0.73
ART2	0.32	-0.57	-0.61	0.48	0.58	1	-0.54	-0.06	0.45	0.35
VRT	-0.63	0.86	0.92	-0.26	-0.69	-0.54	1	-0.25	-0.72	-0.67
LC	0.40	-0.43	-0.39	0.11	0.43	-0.06	-0.25	1	0.42	0.41
Sub	0.97	-0.90	-0.82	0.31	0.80	0.45	-0.72	0.42	1	0.98
M_a	0.99	-0.86	-0.77	0.31	0.73	0.35	-0.67	0.41	0.98	1

TABLE IV
COEFFICIENT OF DETERMINATION (R^2) VALUES FOR PHYSIOLOGICAL PARAMETERS AGAINST BEHAVIOURAL MEASURES

	Alpha	**PSV**	**VUR**
SRT	0.0958	0.0777	0.1070
LC	0.1637	0.1866	0.1578
VRT	0.4092	0.4814	0.8594
ART1	0.5078	0.6240	0.6467
ART2	0.1071	0.3279	0.3809

chosen as groups for the analysis.

The results obtained further emphasize the correspondence between the variation of physiological parameters and behavioral data, and establish the efficacy of the proposed system in capturing the drop in alertness levels of subjects during the continuous performance of cognitive tasks and hence in determining the state of alertness of an individual.

F. System Validation

The z-scores and p-values for the statistical correspondence between M_a and the subjective scores have been presented in Table VII. Bonferroni correction has been used to adjust for multiple comparisons. Predictability of the proposed metric has been established using a best fit line in the scatter plot of M_a with respect to the subjective scores (7) A strong

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2017.2672080, IEEE Transactions on Neural Systems and Rehabilitation Engineering

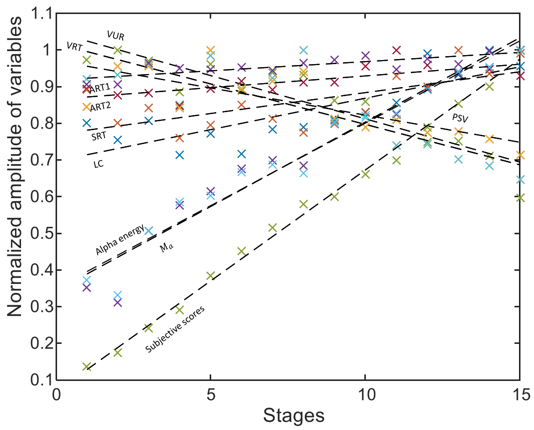

Fig. 6. Regression analysis of all variables

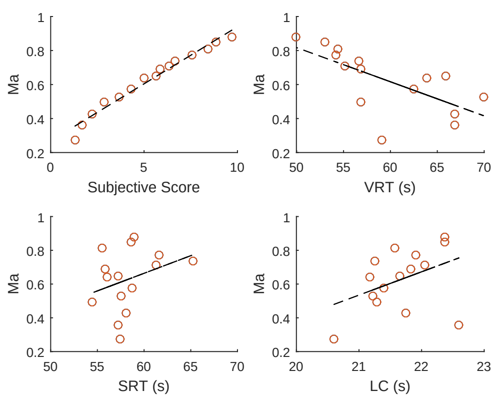

Fig. 7. Scatter plot with a best fit line showing the predictability of M_a against subjective scores, and time required to complete the VRT, SRT and LC tests

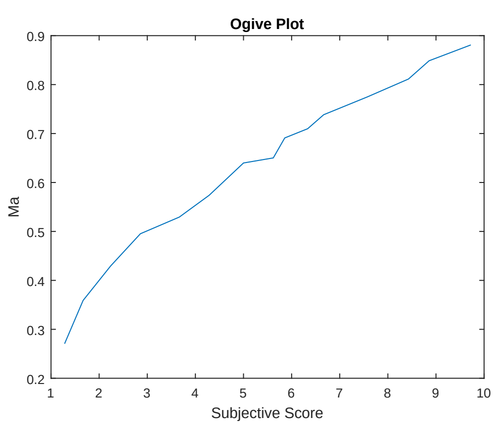

Fig. 8. Ogive plot showing the validation of the proposed metric

TABLE V
REPEATED MEASURES ANOVA

		Within group mean	**Between groups SSD**
Physiological Data	**Alpha energy**	0.6753	0.4565
	PSV	0.8634	0.0028
	VUR	0.8413	0.0011
Task Performance Data	**SRT**	0.8933	0.0284
	ART1	0.8342	0.0036
	ART2	0.8819	0.0155
	VRT	0.8503	0.0003
	LC	0.9588	0.1781
	Overall	0.8498	

Between group sum of squares $= 7.859$
Within group sum of squares $= 4.0354$
Between group $df = 7$
Within-group $df = 8\times 14 = 112$
F-ratio $= \frac{7.859}{4.0354} = 1.9475$
$F_{critical}(2, 15) = 1.7674$ at $\alpha = 0.1$ significance.

TABLE VI
SIGNIFICANT DECLINES IN TASK PERFORMANCE

Variables	**SRT**	**ART1**	**ART2**	**VRT**	**LC**
z-**score**	1.7642	2.0103	1.9856	2.7438	2.3415
p-**value**	0.0777	0.0444**	0.0102**	0.0061*	0.0192**

($**$ indicating $p < 0.05$,$*$ indicating $p < 0.01$)

correlation between the proposed metric and the subjective scores may be observed. However, the metric does not show satisfactory correlation with the time required to complete the SRT and LC tasks. The findings are further corroborated by the Ogive's plot for the variation of the proposed metric with respect to the subjective scores (Fig. 8).

V. CONCLUSION

The present work describes a system that combines information from multimodal data to determine the state of alertness of an individual. The system receives EEG, speech signals and high-speed image sequences through the Serial, Audio and Ethernet ports respectively, and calculates the alpha energy of EEG, voiced-unvoiced ratio of speech signals and peak saccadic velocity. The system has been cross-validated using neuropsychological measures of alertness in an experiment on sustained attention. The experiment makes use of a unique design of the auditory response test and uses the stimuli as both task and tool. To the best of knowledge of the authors, the present study is the first of its kind that uses a varied combination of biomarkers and neuropsychological parameters for analysis of alertness levels. Significant correspondence has been observed between the results of the biomarkers and results of neuropsychological parameters. The results of the

TABLE VII
SIGNIFICANCE OF M_a AND SUBJECTIVE SCORES

Variables	M_a	**Sub**
z-**score**	2.6233	2.4996
p-**value**	0.0087*	0.0124**

($**$ indicating $p < 0.05$,$*$ indicating $p < 0.01$)

present study would hence be instrumental in developing a multi-modal engine to determine the state of alertness of an individual.

ACKNOWLEDGEMENTS

The authors would like to express gratitude towards the subjects for participation in the experiment and I.I.T. Kharagpur for permission to use the facilities.

REFERENCES

[1] A. Raz and J. Buhle, "Typologies of attentional networks," *Nature Reviews Neuroscience*, vol. 7, no. 5, pp. 367–379, 2006.

[2] M. I. Posner and M. K. Rothbart, "Research on attention networks as a model for the integration of psychological science," *Annu. Rev. Psychol.*, vol. 58, pp. 1–23, 2007.

[3] M. I. Posner, M. R. Rueda, and P. Kanske, "18 probing the mechanisms of attention," *Handbook of psychophysiology*, p. 410, 2007.

[4] M. I. Posner and S. J. Boies, "Components of attention.," *Psychological review*, vol. 78, no. 5, p. 391, 1971.

[5] M. I. Posner and R. D. Rafal, "Cognitive theories of attention and the rehabilitation of attentional deficits.," 1987.

[6] W. Sturm and K. Willmes, "On the functional neuroanatomy of intrinsic and phasic alertness," *Neuroimage*, vol. 14, no. 1, pp. S76–S84, 2001.

[7] B. S. Oken, M. C. Salinsky, and S. Elsas, "Vigilance, alertness, or sustained attention: physiological basis and measurement," *Clinical Neurophysiology*, vol. 117, no. 9, pp. 1885–1901, 2006.

[8] P. Tassi and A. Muzet, "Defining the states of consciousness," *Neuroscience & Biobehavioral Reviews*, vol. 25, no. 2, pp. 175–191, 2001.

[9] J. T. Coull, "Neural correlates of attention and arousal: insights from electrophysiology, functional neuroimaging and psychopharmacology," *Progress in neurobiology*, vol. 55, no. 4, pp. 343–361, 1998.

[10] S. Kar, M. Bhagat, and A. Routray, "Eeg signal analysis for the assessment and quantification of drivers fatigue," *Transportation research part F: traffic psychology and behaviour*, vol. 13, no. 5, pp. 297–306, 2010.

[11] S. Kar and A. Routray, "Effect of sleep deprivation on functional connectivity of eeg channels," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 43, no. 3, pp. 666–672, 2013.

[12] A. Sengupta, A. Routray, and S. Kar, "Complex brain networks using visibility graph synchronization," in *2013 Annual IEEE India Conference (INDICON)*, pp. 1–4, IEEE, 2013.

[13] G. Matthews, *Human performance: Cognition, stress, and individual differences*. Psychology Press, 2000.

[14] T. Meijman, "The theory of the stop-emotion: On the functionality of fatigue," *Ergonomics and safety for global business quality and production*, pp. 45–50, 2000.

[15] J. S. Warm, R. Parasuraman, and G. Matthews, "Vigilance requires hard mental work and is stressful," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 50, no. 3, pp. 433–441, 2008.

[16] A. Gevins and M. E. Smith, "Neurophysiological measures of cognitive workload during human-computer interaction," *Theoretical Issues in Ergonomics Science*, vol. 4, no. 1-2, pp. 113–131, 2003.

[17] C. Marzano, F. Fratello, F. Moroni, M. C. Pellicciari, G. Curcio, M. Ferrara, F. Ferlazzo, and L. De Gennaro, "Slow eye movements and subjective estimates of sleepiness predict eeg power changes during sleep deprivation," *SLEEP-NEW YORK THEN WESTCHESTER-*, vol. 30, no. 5, p. 610, 2007.

[18] C. Cajochen, J. M. Zeitzer, C. A. Czeisler, and D.-J. Dijk, "Dose-response relationship for light intensity and ocular and electroencephalographic correlates of human alertness," *Behavioural brain research*, vol. 115, no. 1, pp. 75–83, 2000.

[19] L. Bi, R. Zhang, and Z. Chen, "Study on real-time detection of alertness based on eeg," in *Complex Medical Engineering, 2007. CME 2007. IEEE/ICME International Conference on*, pp. 1490–1493, IEEE, 2007.

[20] T.-P. Jung, S. Makeig, M. Stensmo, and T. J. Sejnowski, "Estimating alertness from the eeg power spectrum," *Biomedical Engineering, IEEE Transactions on*, vol. 44, no. 1, pp. 60–69, 1997.

[21] L. E. Reinerman-Jones, G. Matthews, L. K. Langheim, and J. S. Warm, "Selection for vigilance assignments: A review and proposed new direction," *Theoretical Issues in Ergonomics Science*, vol. 12, no. 4, pp. 273–296, 2011.

[22] S. Banks and D. F. Dinges, "Behavioral and physiological consequences of sleep restriction," *J Clin Sleep Med*, vol. 3, no. 5, pp. 519–528, 2007.

[23] D. F. Dinges, "An overview of sleepiness and accidents," *Journal of sleep research*, vol. 4, no. s2, pp. 4–14, 1995.

[24] J. M. Lyznicki, T. C. Doege, R. M. Davis, M. A. Williams, *et al.*, "Sleepiness, driving, and motor vehicle crashes," *Jama*, vol. 279, no. 23, pp. 1908–1913, 1998.

[25] J. Horne and L. Reyner, "Vehicle accidents related to sleep: a review.," *Occupational and environmental medicine*, vol. 56, no. 5, pp. 289–294, 1999.

[26] A. Ueno and Y. Uchikawa, "Relation between human alertness, velocity wave profile of saccade, and performance of visual activities," in *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*, vol. 1, pp. 933–935, IEEE, 2004.

[27] L. S. Dhupati, S. Kar, A. Rajaguru, and A. Routray, "A novel drowsiness detection scheme based on speech analysis with validation using simultaneous eeg recordings," in *Automation Science and Engineering (CASE), 2010 IEEE Conference on*, pp. 917–921, IEEE, 2010.

[28] H. Serby, E. Yom-Tov, and G. F. Inbar, "An improved p300-based brain-computer interface," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 13, no. 1, pp. 89–98, 2005.

[29] S. Kar, A. Routray, and B. P. Nayak, "Functional network changes associated with sleep deprivation and fatigue during simulated driving: validation using blood biomarkers," *Clinical Neurophysiology*, vol. 122, no. 5, pp. 966–974, 2011.

[30] A. Fletcher, K. McCulloch, S. D. Baulk, and D. Dawson, "Countermeasures to driver fatigue: a review of public awareness campaigns and legal approaches," *Australian and New Zealand Journal of Public Health*, vol. 29, no. 5, pp. 471–476, 2005.

[31] G. R. J. Hockey, P. Nickel, A. C. Roberts, and M. H. Roberts, "Sensitivity of candidate markers of psychophysiological strain to cyclical changes in manual control load during simulated process control," *Applied ergonomics*, vol. 40, no. 6, pp. 1011–1018, 2009.

[32] R. R. Johnson, D. P. Popovic, R. E. Olmstead, M. Stikic, D. J. Levendowski, and C. Berka, "Drowsiness/alertness algorithm development and validation using synchronized eeg and cognitive performance to individualize a generalized model," *Biological psychology*, vol. 87, no. 2, pp. 241–250, 2011.

[33] F. L. da Silva, "Neural mechanisms underlying brain waves: from neural membranes to networks," *Electroencephalography and clinical neurophysiology*, vol. 79, no. 2, pp. 81–93, 1991.

[34] M. K. Kiymik, M. Akin, and A. Subasi, "Automatic recognition of alertness level by using wavelet transform and artificial neural network," *Journal of neuroscience methods*, vol. 139, no. 2, pp. 231–240, 2004.

[35] S. K. Lal and A. Craig, "Driver fatigue: electroencephalography and psychological assessment," *Psychophysiology*, vol. 39, no. 3, pp. 313–321, 2002.

[36] W. Klimesch, "Eeg alpha and theta oscillations reflect cognitive and memory performance: a review and analysis," *Brain research reviews*, vol. 29, no. 2, pp. 169–195, 1999.

[37] A. Sengupta, S. Kar, and A. Routray, "Study of loss of alertness and driver fatigue using visibility graph synchronization," *Innovative Research in Attention Modeling and Computer Vision Applications*, p. 171, 2015.

[38] A. Dasgupta, A. George, S. Happy, and A. Routray, "A vision-based system for monitoring the loss of attention in automotive drivers," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 14, no. 4, pp. 1825–1838, 2013.

[39] P. Smith, M. Shah, and N. da Vitoria Lobo, "Monitoring head/eye motion for driver alertness with one camera," in *icpr*, p. 4636, IEEE, 2000.

[40] R. F. Kaplan and K. A. Loparo, "Alertness and drowsiness detection and tracking system," Sept. 29 1998. US Patent 5,813,993.

[41] A. Dasgupta, A. George, S. Happy, A. Routray, and T. Shanker, "An on-board vision based system for drowsiness detection in automotive drivers," *International Journal of Advances in Engineering Sciences and Applied Mathematics*, vol. 5, no. 2-3, pp. 94–103, 2013.

[42] P. Artaud, S. Planque, C. Lavergne, H. Cara, C. Tarriere, B. Gueguen, *et al.*, "An on-board system for detecting lapses of alertness in car driving," in *Proceedings: International Technical Conference on the Enhanced Safety of Vehicles*, vol. 1995, pp. 350–359, National Highway Traffic Safety Administration, 1995.

[43] R. M. Vick and C. S. Ikehara, "Methodological issues of real time data acquisition from multiple sources of physiological data," in *System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on*, pp. 7–pp, IEEE, 2003.

[44] S. Gupta and A. Routray, "Estimation of saccadic ratio from eye image sequences to detect human alertness," in *Intelligent Human Computer*

Interaction (IHCI), 2012 4th International Conference on, pp. 1–6, IEEE, 2012.

[45] Z. He, T. Tan, Z. Sun, and X. Qiu, “Toward accurate and fast iris segmentation for iris biometrics,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, no. 9, pp. 1670–1684, 2009.

[46] H. K. Solvang, K. Ishizuka, and M. Fujimoto, “Voice activity detection based on adjustable linear prediction and garch models,” *Speech Communication*, vol. 50, no. 6, pp. 476–486, 2008.

[47] D. Goldberg and P. Williams, “General health questionnaire (ghq),” *Swindon, Wiltshire, UK: nferNelson*, 1988.

[48] C. Golden, “Stroop colour and word test,” *age*, vol. 15, p. 90, 1978.

[49] K. Tham, “The baking tray task: a test of spatial neglect,” *Neuropsychological Rehabilitation*, vol. 6, no. 1, pp. 19–26, 1996.

[50] D. F. Dinges and J. W. Powell, “Microcomputer analyses of performance on a portable, simple visual rt task during sustained operations,” *Behavior research methods, instruments, & computers*, vol. 17, no. 6, pp. 652–655, 1985.

[51] J. T. Cacioppo, L. G. Tassinary, and G. Berntson, *Handbook of psychophysiology*. Cambridge University Press, 2007.

[52] R. Croft and R. Barry, “Removal of ocular artifact from the eeg: a review,” *Neurophysiologie Clinique/Clinical Neurophysiology*, vol. 30, no. 1, pp. 5–19, 2000.

[53] V. Krishnaveni, S. Jayaraman, L. Anitha, and K. Ramadoss, “Removal of ocular artifacts from eeg using adaptive thresholding of wavelet coefficients,” *Journal of Neural Engineering*, vol. 3, no. 4, p. 338, 2006.

[54] X. Tan and B. Triggs, “Enhanced local texture feature sets for face recognition under difficult lighting conditions,” *Image Processing, IEEE Transactions on*, vol. 19, no. 6, pp. 1635–1650, 2010.

[55] F. Timm and E. Barth, “Accurate eye centre localisation by means of gradients.,” *VISAPP*, vol. 11, pp. 125–130, 2011.

[56] C. Xu, Y. Zheng, and Z. Wang, “Semantic feature extraction for accurate eye corner detection,” in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pp. 1–4, IEEE, 2008.

[57] B. S. Atal and L. R. Rabiner, “A pattern recognition approach to voiced-unvoiced-silence classification with applications to speech recognition,” *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 24, no. 3, pp. 201–212, 1976.

[58] B. C. Moore, C. Füllgrabe, and M. A. Stone, “Effect of spatial separation, extended bandwidth, and compression speed on intelligibility in a competing-speech task,” *The Journal of the Acoustical Society of America*, vol. 128, no. 1, pp. 360–371, 2010.

[59] P. C. Loizou, *Speech enhancement: theory and practice*. CRC press, 2013.

[60] F. Qi, C. Bao, and Y. Liu, “A novel two-step svm classifier for voiced/unvoiced/silence classification of speech,” in *International Symposium on Chinese Spoken Language Processing*, pp. 77–80, Citeseer, 2004.

[61] R. J. Little and D. B. Rubin, *Statistical analysis with missing data*. John Wiley & Sons, 2014.

Anwesha Sengupta has received her B. E. degree from Jadavpur University, India in 2005 and M.S. degree from Indian Institute of Technology, Kharagpur, India in 2011. Presently, she is pursuing her PhD. from Indian Institute of Technology, Kharagpur. Her research interests include signal processing and EEG signal analysis.

Anirban Dasgupta has received his B.Tech. (Hons.) degree in Electrical Engineering from the National Institute of Technology (NIT), Rourkela, India in 2010 and MS degree in Electrical Engineering from the Indian Institute of Technology (IIT) Kharagpur, India. He is presently pursuing his PhD from IIT Kharagpur. His current research interests include computer vision, embedded systems, signal and image processing. He has been a student member of IEEE since 2011. He became a member of IEEE EMBS in 2014.

Aritra Chaudhuri has received the B.E. degree from BESU, India in 2010 and M.Tech. degree from Indian Institute of Technology (IIT) Kharagpur, India. His research interests include bio-medical signal processing and embedded systems. He has been a student member of IEEE since 2014. He became a member of IEEE EMBS in 2015. He has an experience of over 4 years on EEG-based signal processing and information extraction.

Anjith George has received his B.Tech. (Hons.) Degree from Calicut University, India in 2010 and M-Tech degree from Indian Institute of Technology (IIT) Kharagpur, India in 2012. Presently he is pursuing Ph.D. from IIT Kharagpur, India. His current research interests include embedded systems, computer vision and machine learning.

Aurobinda Routray is a professor in the Department of Electrical Engineering, Indian Institute of Technology, Kharagpur. His research interest includes non-linear and statistical signal processing, signal based fault detection and diagnosis, real time and embedded signal processing, numerical linear algebra, and data driven diagnostics. More information is available at $http://aroutray.org/$.

Rajlakshmi Guha is an Assistant Professor in the Centre for Educational Technology, IIT, Kharagpur. Her research interest includes cognitive psychology, affective psychology, human behaviour analysis. She is involved in several clinical and experimental research studies.