Exploring Multimodal Biosignal Features for Stress Detection during Indoor Mobility

Kyriaki Kalimeri
Data Science Lab
ISI Foundation, Turin, Italy
kalimeri@ieee.org

Charalampos Saitis
Audio Communication Group
Technical University Berlin, Germany
charalampos.saitis@campus.tu-berlin.de

ABSTRACT

This paper presents a multimodal framework for assessing the emotional and cognitive experience of blind and visually impaired people when navigating in unfamiliar indoor environments based on mobile monitoring and fusion of electroencephalography (EEG) and electrodermal activity (EDA) signals. The overall goal is to understand which environmental factors increase stress and cognitive load in order to help design emotionally intelligent mobility technologies that are able to adapt to stressful environments from real-time biosensor data. We propose a model based on a random forest classifier which successfully infers in an automatic way (weighted AUROC 79.3%) the correct environment among five predefined categories expressing generic everyday situations of varying complexity and difficulty, where different levels of stress are likely to occur. Time-locating the most predictive multimodal features that relate to cognitive load and stress, we provide further insights into the relationship of specific biomarkers with the environmental/situational factors that evoked them.

CCS Concepts

•Human-centered computing → Ubiquitous and mobile computing; Ubiquitous and mobile computing design and evaluation methods;

Keywords

Multimodal classification, data fusion, multimodal interaction, biosignals, mobile EEG, electrodermal activity, affective computing, visually impaired mobility

1. INTRODUCTION

Indoor mobility can be a challenging and emotionally stressful task for visually impaired people (VIP), especially when navigating in unfamiliar environments. Despite an increasing number of assistive technologies that help individuals with sight loss to augment their spatial awareness and wayfinding abilities when in move, very few systems provide a high degree of independence beyond known environments that would allow VIP to significantly achieve mobility and integrate in everyday active life [13, 21]. Placing the visually impaired in the center of attention and exploiting recent developments in pervasive physiological computing and wearable wireless sensor devices, a multimodal study was designed to better understand how people with sight loss perceive and interact with space as manifested in their management of cognitive load and stress.

Orientation and mobility (O&M) in humans heavily relies on sight, which provides instantaneous, effortless access to anticipatory (e.g., stairs, turns, signs) and proactive (e.g., moving people, pillars) information at various distances simultaneously [24]. Visually impaired people learn to obtain critical environmental information primarily through touch (sensing the ground surface with a white cane) and hearing (identifying and localizing events and landmarks through sound). Indoor mobility challenges can be summarized in three main problems: avoiding objects or obstacles (e.g., people moving or standing, pillars, doors opening outwards); detecting ground level changes (e.g., stairs, ramps); and finding entrance/exit points (e.g., automated doors, elevators). Although these problems generally diminish with increased experience of an environment (e.g., person's own home), they still make traveling in unfamiliar settings particularly challenging, often preventing VIP from visiting new spaces altogether.

Despite a significant amount of research on understanding the perceptual and neurocognitive mechanisms by which people with sight loss access and process wayfinding information [7], there is still little practical knowledge of how the management of mental load and stress relates to the wayfinding process itself. This is a critical aspect of designing mobility technologies that has only recently been considered essential in developing an understanding of how environmental factors affect the cognitive-emotional states of the visually impaired [33]. One way of detecting emotion and psychological stress is through identifying patterns in physiological modalities, the most common being electrodermal activity, cardiovascular activity (e.g., heart rate, blood volume pulse), and electroencephalography.

Electrodermal activity (EDA) is a well-known indicator of physiological arousal and stress activation in affective computing [14, 31]. It is more sensitive to emotion related variations in arousal as opposed to physical stressors, which can be better reflected in measurements of heart rate (HR). Blood volume pulse (BVP) patterns can also reflect transient

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

ICMI'16, November 12–16, 2016, Tokyo, Japan
© 2016 ACM. 978-1-4503-4556-9/16/11...$15.00
http://dx.doi.org/10.1145/2993148.2993159

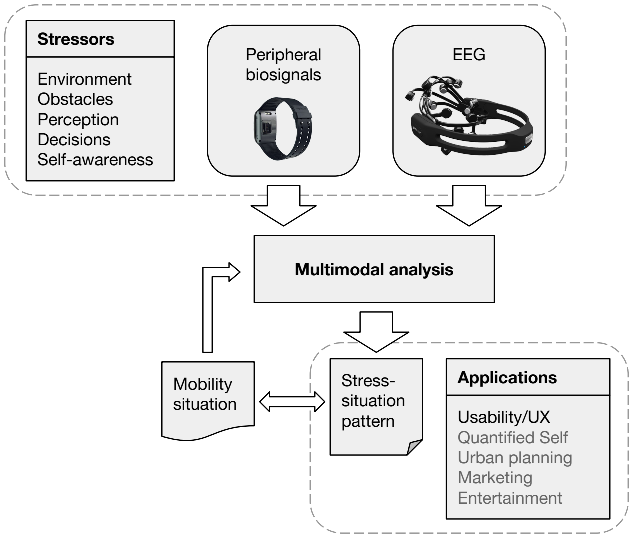


Figure 1: Multimodal biosignal data capture and analysis framework for detecting stress during mobility.

processes in arousal and cognitions [27]. Two outdoor mobility studies in the early 1970s suggested that some form of psychological rather than physical stress is responsible for increased HR in visually impaired versus sighted pedestrians [26, 34]. However, certain mobility tasks (e.g., stair climbing) may result in an interactive effect of psychological stress and momentary physical work load, thus cardiovascular measures may be less suitable than EDA. Electroencephalography (EEG), on the other hand, can provide neurophysiological markers of cognitive-emotional processes induced by stress and indicated by changes in brain rhythmic activity [15, 17].

In recent years, the advent of ubiquitous mobile and sensing technologies, consumer brain-computer interfaces (BCI), and the quantified self movement has driven the development of wireless wearable multi-sensor systems (from devices to smartphone apps) for easy and reliable automatic collection of brain and peripheral biosignal data streams, making it possible to monitor human affective states in virtually any real-world situation [9, 18]. Massot and colleagues [22] used a custom mobile biosensor to collect EDA from 27 blind pedestrians as they walked through urban scenes of varying complexity. Examination of arousal-relevant EDA features showed that VIP experience increased psychological stress when walking on busy shopping streets, passing through large open areas, and crossing junctions. In another study, analyses of EEG signals recorded from 12 VIP during outdoor travel using a commercial BCI headset [23] further indicated that busy streets, open spaces, and street crossings induce larger cognitive engagement than quieter and less complex urban settings.

Taking advantage of the inherent and complementary properties of the EEG, EDA, and BVP modalities, this paper presents a multimodal approach to automatic inference of stressful environmental conditions affecting VIP when moving in unfamiliar spaces using a random forest classifier and features extracted from the three signals. The aim was to discover biomarkers that can be used to detect shifts in emotional stress and cognitive load between different settings and situations. We have previously explored this method during an urban mobility study with 8 visually impaired pedestrians, where we obtained a high prediction rate (weighted AUROC 93%) [30]. Building on those results, here we report an indoor wayfinding study during which EEG and EDA signals were collected from the same group of VIP using two state-of-the-art portable biosensor devices and used in unimodal and multimodal classification experiments (Fig. 1). We further studied the relationship of the most predictive stress biomarkers with the environmental/situational factors that evoked them through assessing and visualizing their density distributions.

The overall goal of the study was to develop an understanding of which environmental factors increase cognitive load and stress during visually impaired mobility in order to help build *emotionally intelligent* O&M systems, which are capable of implicitly adapting not only to changing environments but also to shifts in the internal experience of the user in relation to environmental factors [28, 25, 33]. The proposed framework thus differs fundamentally from context-based approaches to environment recognition, for example, GPS-based geolocation. While the latter allow a certain degree of independence for VIP, identifying dynamic stressors in different or the same environments can lead to even more independent mobility systems that recognize, interpret, and adapt to the affective states of the user.

2. EXPERIMENTAL SETUP

2.1 Site and Route

The experiment was conducted inside a university building, which houses various service units for students, a bookstore, two restaurants, classrooms, and reading rooms. As such, the building provided sufficiently complex indoor environments for the purposes of our study. With the assistance of VIP caretakers and O&M instructors, we planned a route to take the VIP through circumstances where different levels of stress were likely to occur (i.e., of varying complexity and difficulty, see Fig. 2).

The route linked the entrance at the back of the building (START) to the main entrance at its front (END) and comprised five distinct environments representable of a variety of indoor mobility challenges (see Table 1). Indicatively, participants had to enter through automated doors (environments A), use an elevator (B), move across a busy open space (D–main entrance hall), and walk down a large spiral staircase (E). The route was approximately 200 meters in length and took on average 5 minutes to walk (range = 4–8 minutes).

2.2 Participants

Nine healthy visually impaired adults with different degrees of sight loss walked individually a complex route in an educational building (6 female; average age = 41 yrs, range = 22–53 yrs). To help make them feel comfortable and safe, they were encouraged to walk as usual using their white canes if they wished so, and were accompanied by their familiar O&M instructor. Participants were instructed to avoid smoking normal or e-cigarettes and consuming caffeine or sugar (e.g., coffee, coke, chocolate) approximately 1 hour prior to the walk. Recruitment was based on volunteering and all VIP were capable of giving free and informed consent. The

Table 1: **Descriptions and mobility challenges of the different indoor environments.**

ID	Description	Challenges
A	Entering through automated doors (two hinged and one rotating)	finding the pushbutton (hinged doors only), finding where and when to enter the rotating door, other people going through the door at the same time
B	Using an elevator to move between floors	finding the pushbuttons (calling the elevator, selecting floor), other people exiting or entering the elevator
C	Walking along a narrow corridor	Moving people, noise, classroom doors opening suddenly
D	Moving across an open space (in front of the classrooms and at the main entrance hall)	Moving people, standing people, tables, chairs, trash bins, pillars, people talking, loud noises
E	Using stairs to move between floors	Finding start of stairs, people using the stairs in the opposite direction, walk down a large spiral flight of stairs

study was approved by the National Bioethics Committee of Iceland. All data was anonymized before analysis.

All participants actively experience indoor environments other than where they reside on a daily basis: three work full-time, three are part-time employees, and three attend educational/vocational establishments. Two participants, who reported regular use of a white cane when mobile, felt safe enough to walk the indoor route without any aid. When asked to describe their feelings regarding the ease of mobility over the previous year, four VIP believed that it has not changed, three thought that it has become easier, while two considered it to have become less easy.

2.3 Data Collection

EEG was recorded using the Emotiv EPOC+, a mobile headset with 16 dry electrodes registering over the 10-20 system locations AF3, F7, F3, FC5, T7, P3 (CMS), P7, O1, O2, P8, P4 (DRL), T8, FC6, F4, F8, and FC4 (sampling rate $f_s = 128$ Hz).[1] Given the practical constraints involved in a mobility study, EPOC+ was chosen as it provides a good compromise between performance (i.e., number of channels and quality of the acquired EEG signals) and usability (i.e., portability, preparation time and user comfort) with respect to other commercial wireless EEG systems [1, 8, 9, 10].

Along with the Emotiv headset, participants were asked to wear the Empatica E4 wristband[2] [11]. E4 measures EDA through 2 ventral (inner) wrist electrodes ($f_s = 4$ Hz) and BVP through a dorsal (outer) wrist photoplethysmography (PPG) sensor ($f_s = 64$ Hz). The wristband also includes an infrared thermopile sensor and a 3-axis accelerometer. E4 is currently the only commercial multi-sensor device developed based on extended scientific research in the areas of psychophysiology and physiological computing. Additionally, it has a cable-free, watch-like design, which makes it easier and more aesthetically pleasant to wear, and thus better fitted to use in ambulatory measurements compared to other wearable biosignal devices. Participants were asked to wear the wristband on the non-dominant hand to minimize motion artifacts related to handling the white cane [4].

Participants walked the charted route three times for training purposes. Directions were only provided during the first walk to help the VIP familiarize with the route. They were instructed to avoid unnecessary head movements and hand gestures as well as talking to their O&M instructor unless there was an emergency. Video and audio were registered by means of a smartphone camera to facilitate data annotation (observing behaviors across the different indoor environments) and synchronization (start/end of walk, environments and obstacles). At the end of the third walk, participants were asked to describe stressful moments along the route.

[1]http://emotiv.com/epoc/

[2]https://www.empatica.com/e4-wristband

3. MULTIMODAL DATA ANALYSIS

3.1 Biosignal Feature Extraction

Blood volume pulse data obtained via the E4 wristband were not included in the analysis as they were found to be highly dependent on the more physically demanding parts of the walk, especially stair climbing. The analysis was therefore focused on the EEG and EDA modalities. Incomplete data recorded from two participants during the first walk as a result of temporary dysfunctions of the equipment were discarded.

3.1.1 Electroencephalography

The Emotiv EPOC+ system involves a number of internal signal conditioning steps. Analogue signals are first high-pass filtered with a 0.16 Hz cut-off, pre-amplified, low-pass filtered with a 83 Hz cut-off, and sampled at 2048 Hz. Digital signals are then notch-filtered at 50/60 Hz and down-sampled to 128 Hz prior to transmission. For approximately half of the participants, EEG data obtained from the headset was first time-domain interpolated using the Fast Fourier Transform (FFT) to account for missing samples due to connectivity issues. Subsequently, all signals were baseline-normalized by subtracting for each participant and for each channel the mean of resting state registrations. These were obtained during a laboratory study with the same participants [30].

Brain activity is characterized by rhythmic patterns across distinct frequency bands, the definition of which can vary somewhat among studies. Here we analyzed EEG in six bands, namely delta (0.5–4 Hz), theta (4–7 Hz), alpha-1 (7–10 Hz), alpha-2 (10–13 Hz), beta (13–30 Hz), and gamma (30–60 Hz). Beta activity is associated with psychological and physical stress, whereas theta and alpha-1 frequencies reflect response inhibition and attentional demands such as phasic alertness [29]. Alpha-2 is related to task performance in terms of speed, relevance, and difficulty [20]. Gamma waves are involved in more complex cognitive functions such as multimodal processing or object representation [19].

For each of the 14 EEG channels, the absolute and relative spectral power in each frequency band was first computed using the Power Spectral Intensity (PSI) and Relative Intensity Ratio (RIR) functions of the PyEEG open source Python module [2]. For the kth band, these are defined as

$$\mathrm{PSI}_k = \sum_{i=\lfloor N(f_k/f_s) \rfloor}^{\lfloor N(f_{k+1}/f_s) \rfloor} |X_i| \quad \text{and}$$

$$\mathrm{RIR}_k = \frac{\mathrm{PSI}_k}{\sum_{j=1}^{K-1} \mathrm{PSI}_j}, \quad k = 1, 2, \ldots, K-1$$

where f_s is the sampling rate, N is the time series length, $|X_1, X_2, \ldots, X_N|$ is the FFT of the series, and K is the total number of bands. Furthermore, the irregularity (or uniformity) of the structure of the total EEG power spectrum, that is the peaks or plateaus in the distribution of power across the different frequencies, was quantified by computing the information measure of entropy [16]. Using the RIR function of PyEEG, entropy is defined as

$$H = -\frac{1}{log\,(K)} \sum_{i=1}^{K} \mathrm{RIR}_i \; log\,(\mathrm{RIR}_i)\,.$$

In total, 182 EEG features were extracted from each individual EEG recording.

3.1.2 *Electrodermal Activity*

A measurement of skin conductance is traditionally characterized by two types of behavior: short-lasting phasic responses (can be thought of as rapidly changing peaks in EDA) and long-term tonic level in the absence of phasic responses (can be thought of as the underlying slow-changing level of EDA). In terms of physiology, a skin conductance response (SCR) is a sudden rise in the electrical conductance of the skin due to secretion from the skin's sweat glands (sweat contains electrolytes) in response to sympathetic nervous activation. Another characteristic of the signal is the superposition of subsequent SCRs (i.e., one SCR emerges on top of the preceding one), typically observed in states of high arousal.

Skin conductance data obtained from the E4 was first low-pass filtered (1st order Butterworth, $f_c = 0.6$ Hz) to remove steep peaks stemming from artifacts and subsequently min-max normalized to reduce inter-individual variance [6]. Conditioned SC signals were then decomposed into two continuous components of phasic and tonic EDA using a deconvolution-based method implemented in Ledalab, a Matlab based toolbox [3]. Six features were extracted: mean tonic EDA and the number of "spontaneous" SCRs (i.e., phasic changes not traceable to specific stimulation), which are known to be particularly suitable for longitudinal monitoring of emotional stress-elicited EDA (i.e., tonic arousal); sum of amplitudes of registered SCRs and average, maximum, and cumulative phasic EDA, which provide varying indicators of instantaneous phasic arousal [4].

3.2 Classification Design

In order to identify automatically the affective meaning of an indoor environment based on biosignals recorded from VIP walking through it, we postulated the study as a supervised classification process. A widely-used ensemble learning method for classification was employed, namely Random Forest (RF) classifier [5], selected due to its ability to deal with possibly correlated predictor variables as well as because it provides a straightforward assess of the variable importances. For each of the distinct environments described in Table 1, hereinafter called classes, each time point of the corresponding biosignal data was annotated based on a binary schema per second, where "1" signaled the presence of the participant in the given environment at the given time point and "0" otherwise.

A series of experiments were designed to assess and compare the predictive power of each modality (EEG or EDA) as well as of their fusion in a feature-level basis, in both

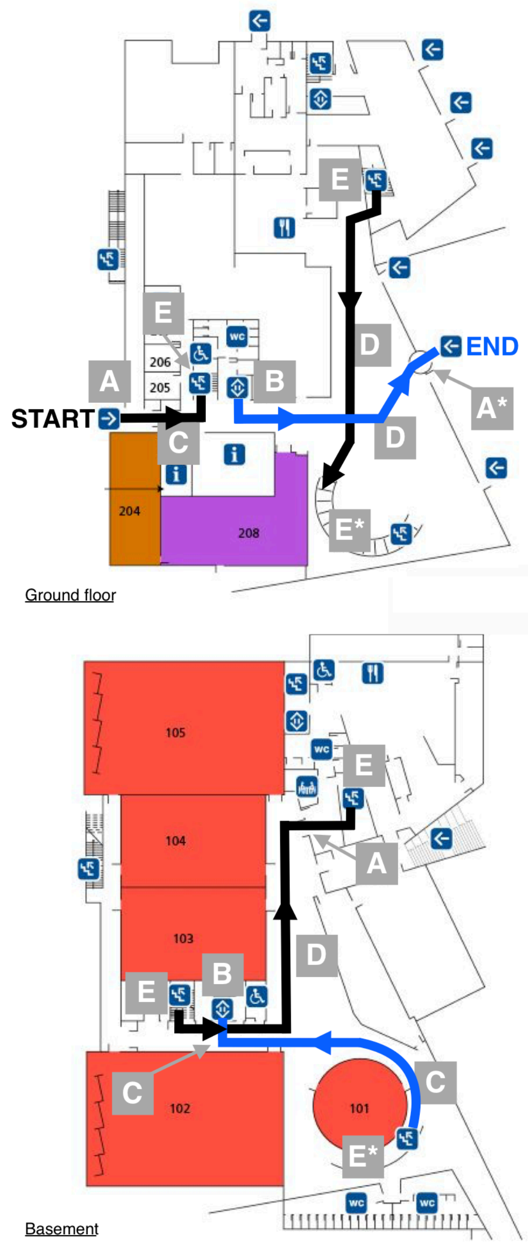


Figure 2: Layout of the site and charted route. Letters depict the different indoor environments reported in Table 1, with A* indicating the rotating door and E* the large spiral stairs.

single-class and multi-class scenarios:

- Exp. I: Single-class classification using as predictors the unimodal features extracted from the EEG signals ($N = 182$).
- Exp. II: Single-class classification using as predictors the unimodal features extracted from the EDA signals ($N = 6$).
- Exp. III: Single-class classification using feature-level multi-modal fusion of EEG and EDA features ($N = 188$). The prediction is made on single-class binary target variables as before.
- Exp. IV: Multi-class classification using as predictors the unimodal features extracted from the EEG signals ($N = 182$).
- Exp. V: Multi-class classification using as predictors the unimodal features extracted from the EDA signals ($N = 6$).
- Exp. VI: Multi-class classification using feature-level multi-modal fusion of EEG and EDA features ($N = 188$).

The number of estimators was set to 300 and the maximum number of features was set equal to the total number of features for each experiment. For each experiment we estimated the relative rank (i.e. depth), as emerged from the "Gini" impurity function, of each feature in order to assess the relative importance of that feature to the predictability of the target variable [5].

All data from the three times each participant walked the route were employed for the analysis. While overall familiarity with the route might have gradually increased, individual environments still retained a dynamic complexity due to "new" stressors such as people coming out of the elevator, classroom doors opening, or chairs and tables being displaced. There were approximately 7k data points, with classes ranging from about 2k (most frequent) to 570 points (less frequent). Sequential data points were split randomly in training and testing subsets (which, as a result, no longer contain sequential points). We trained one model for each of the single-class cases and one for the multi-class experiment following a five folds cross-validation schema, where the 80% of the data points were used for training and the 20% for testing. The best model is chosen as the one that maximized the area under of the receiver operating characteristic (AUROC) weighted statistic, taking into account the lack of balance between the labels.

4. RESULTS AND DISCUSSION

Table 2 summarizes the AUROC weighted metric for all experiments. In the one-versus-all scenario (Exps. I–III), both modalities (Exps. I and II) are predictive of the distinct environments, with EEG performing considerably better than EDA. Fusing the two modalities (Exp. III) gave marginally improved results with the exception of the elevator environment, the detection of which was substantially better than when using only EEG features. Importantly, all participants reported the elevator as being one of the most stressful parts of the walk. In the multi-class scenario (Exps. IV–VI), EEG spectral power (absolute and relative) and entropy (Exp. IV) showed much higher discriminative ability than tonic and phasic EDA (Exp. V), but results were slightly improved with the multimodal fusion model (Exp. VI). Figure 3b depicts the weighted ROC curves of the latter in an one-against-all binary scenario, assessing the qualitative performance of each class. Interestingly, we note that the model performs equally well for all classes showing proof of its stability.

Table 2: Classification AUROC weighted metric for all the environments across the various experiments. The reported numbers refer to the mean AUROC over all folds in percentile and in parenthesis the standard deviation is reported.

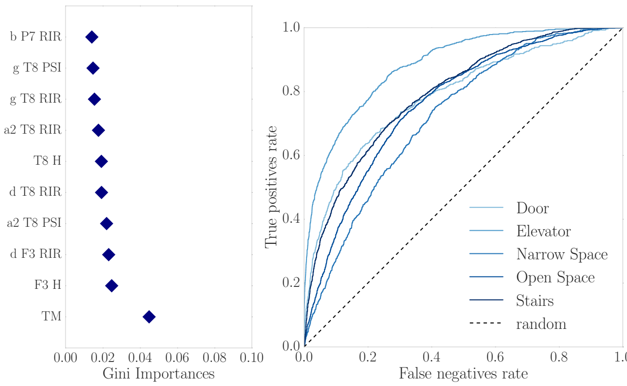
Environment	EEG	EDA	FUSION
Single-class classification			
	Exp. I	Exp. II	Exp. III
Door	76.7 (0.8)	75.8 (2.1)	76.7 (1.5)
Elevator	82.4 (0.9)	78.4 (0.8)	90.1 (0.9)
Corridor	70.7 (2.3)	66.9 (1.4)	72.5 (2.6)
Open space	74.9 (1.2)	67.0 (1.7)	75.6 (1.2)
Stairs	77.7 (0.9)	69.6 (1.8)	78.7 (1.3)
Multi-class classification			
	Exp. IV	Exp. V	Exp. VI
All	77.3 (0.4)	53.3 (1.0)	79.3 (0.3)

Figure 3a depicts the nine most predictive features of Exp. VI. The feature importances were estimated also for all experiments and the most predictive ones appear always with the highest ranks. Interestingly, the most predictive feature was the mean tonic EDA despite the overall low discriminative performance of this modality. This result confirms previous research in that skin conductance tonic level is a highly relevant index of stress-induced physiological arousal. However, predictions were dominated by spectral power and entropy of the F3 and T8 EEG channels. Although real-time EEG acquisition may be subject to very noisy signals, this finding is in line with the neuroscientific literature where the 10-20 system locations F3/4, F7/8 and T7/8 are often considered suitable enough to monitor brain activity under emotional stress [35].

We had previously applied the same multimodal approach to studying mental load and stress in people with sight loss during an outdoor mobility task with the same participants [30].

Extracting the same features from the EDA and EEG signals, the former outperformed the latter in predicting stressful urban environments, with four of the ten most predictive features coming from the EDA modality (mean tonic EDA was again found to be the most predictive biomarker). This difference between the two studies further supports the need for a multimodal approach to stress detection in visually impaired mobility: both brain and peripheral biosignals are important for assessing stress under different conditions (e.g., indoor vs. outdoor) and tasks (e.g., active wayfinding vs. passive walking).

To better understand the properties of the most predictive features that emerged from the classification experiments (see Fig. 3a), as well as the intensity of the cognitive-emotional responses they express, we assigned feature values to 1-second steps from the start point based on recorded timestamps and



(a) (b)

Figure 3: (a) Feature importances in Exp. VI. Mean tonic EDA (TM) emerged as the most predictive biomarker but predictions were dominated by spectral power and entropy of the F3 and T8 EEG channels. (b) One against all ROC curves for each one of the classes in Exp. VI. The overall AUROC weighted metric for the multi-class classification of environments is 79.3% and, importantly, the trained model seems able to learn equally well all the different environments.

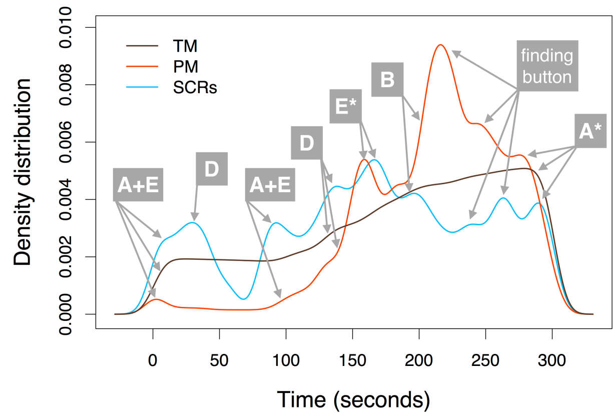
assessed their temporal distributions by means of weighted kernel density estimation. Since mean tonic EDA (TM) was the only EDA feature among the nine most predictive features of Exp. VI, we also examined the density distribution of maximum phasic EDA (PM) and number of phasic EDA peaks (SCRs).

Due to different walking speeds and behaviors, individual walk times varied between participants and trials, ranging from 4 to 8 minutes with an average length of 5 minutes. To temporary align all features so that same times corresponded to same environments we performed dynamic time warping [12], postulating that a certain environment induced similar biomarker patterns. Each feature vector was warped to a reference vector that was 300 seconds (5 minutes) long. Let $(x_1, x_2, \ldots, x_n)$ be an independent random sample drawn from some distribution with density function $f(x)$ defined on $\mathbb{R}$. The (univariate) weighted kernel density estimate of f is defined in [32] as:

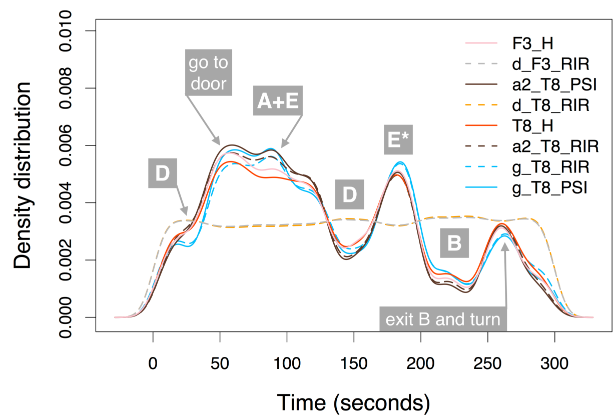
$$\hat{f}_h(x) = \frac{1}{n}\sum_{i=1}^{n} W(x_i, w)\, K_h(x - x_i)$$

where K is a kernel function, $h > 0$ is a parameters which controls the bandwidth (or smoothing) of the estimate, $K_h(x) = 1/h\, K(x/h)$, and W is a function weighting each data point in the sample with a value from $w \in \mathbb{R}$. A popular choice for K is the Gaussian (or normal) kernel, which was also applied here.

For each of the EDA and EEG features, using the values as weights (w) for 1-second time steps (x) and a bandwidth of $h = 7.99$ helped estimate the feature-weighted density of time points (temporal distances) on a 400-point grid, and based on this generate a density function across all participants and walks. Figure 4a shows the resulting density distributions for the TM, PM and SCRs features plotted together. Tonic EDA gradually increases towards the second half of the walk where



(a)



(b)

Figure 4: Temporal biomarker density distribution plots. Time points were weighted according to (a) the most predictive feature of mean tonic EDA (TM) as well as the features of tonic (SCRs) and phasic EDA (PM); (b) the 8 most predictive EEG features reported in Fig. 3a (H = entropy; a2 = alpha-2; d = delta; g = gamma; PSI = absolute power; RIR = relative power). Letters depict the different indoor environments reported in Table 1, with A* indicating the rotating door and E* the large spiral stairs. EDA features depicted stressful situations whereas EEG markers traced differences in cognitive load.

the most stressful parts were according to reports from all participants at the end of the study. The estimated density of SCRs generally followed the trend of TM, while allowing to observe transitory increases in tonic skin conductance. The high density of PM in the second half of the walk suggested the presence of a higher number of instantaneous stressors, for example: finding the elevator buttons and door, maintaining direction amidst loud noises and people moving in the main entrance hall, and safely passing through the rotating door.

Estimated temporal densities of the EEG features are plotted in Fig. 4b. Spectral entropy as well as alpha-2 and gamma power (absolute and relative) in the F3 and T8 channels all followed very similar distributions, with peaks present when participants had to solve a wayfinding task, for example: turning and going towards the door, finding where stairs begin, negotiating orientation while climbing down the spiral stairs and after exiting the elevator. The density distribution of relative delta power in the same two channels showed much weaker fluctuations along the route, but it appeared to decrease when power in the other bands increased and vice versa. A possible explanation in line with the neuroscientific literature is that delta waves, known to dominate EEG spectra during "deep sleep," were suppressed during increased cognitive activity. Therefore, whereas EDA depicted stressful situations across different environments and situations, EEG traced changes in cognitive load, thus providing further evidence of the complementary nature of the two modalities in assessing human affective states.

5. CONCLUSIONS

To the best of the authors' knowledge, this is the first study that presents a framework for assessing the cognitive-emotional experience of people with sight impairment while navigating in unfamiliar indoor environments based on ambulatory monitoring and fusion of multimodal biosignal data. Different scenarios were compared, aiming to address the robustness of the model as well as emerging differences in the perception and interaction of the VIP with their surroundings. Even if the chosen site does not represent all possible different indoor environments and situations in terms of complexity and difficulty, the charted route was designed so as to combine most of the mobility challenges faced by VIP. The high prediction rate (79.3% AUROC weighted) is highly encouraging of this approach and, interestingly, the most predictive biomarkers indicate as stressful and cognitively loaded "hotspots" (see Fig. 4) spaces and situations that coincide perfectly with participants' self-reported experience.

Reported classification results, despite being promising, should be considered with caution due to the limited number of participants. Furthermore, the well-established Emotiv EPOC+ EEG headset has certain limitations with respect to the quality of the recorded signal during experiments involving physical activity such as the one presented in the study. The number of the provided electrodes is limited and hence the EEG markers discussed above are meant to provide only insights on the most predictive features and their connection to specific tasks and conditions.

A rich multimodal dataset has been collected, which can be made available upon request in order to maximize the impact of the work and encourage further investigations. Future steps of the present study include refining the predictive model through exploring novel multimodal biosignal features, comparing different classifiers, implementing leave-one-subject-out cross-validation to account for user-to-user variation in physiological arousal, as well as performing an in-depth analysis of specific psychological stressors in each category of vision impairment (e.g., totally blind vs. partially sighted individuals). Such findings hopefully pave the way to emotionally intelligent mobile technologies that take the concept of navigation one step further, accounting not only for the shortest path but also for the most effortless, least stressful and safest one.

6. ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Union's Horizon 2020 Research and Innovation program under grant agreement No 643636 "Sound of Vision." The authors wish to thank the administration and O&M instructors at the National Institute for the Blind, Visually Impaired, and Deafblind in Iceland for their valuable input and generous assistance.

7. REFERENCES

[1] N. A. Badcock, P. Mousikou, Y. Mahajan, P. de Lissa, J. Thie, and G. McArthur. Validation of the Emotiv EPOC EEG gaming system for measuring research quality auditory ERPs. *PeerJ*, 19, 2013.

[2] F. S. Bao, X. Liu, and C. Zhang. PyEEG: An open source Python module for EEG/MEG feature extraction. *Comp. Intel. Neurosc.*, 2011:Art. ID 406391, 2011.

[3] M. Benedek and C. Kaernbach. A continuous measure of phasic electrodermal activity. *J. Neurosc. Methods*, 190:80–91, 2010.

[4] W. Boucsein. *Electrodermal Activity*. Springer, New York, 2nd edition, 2012.

[5] L. Breiman. Random forests. *Mach. Learn.*, 45(1):5–32, 2001.

[6] J. T. Cacioppo and L. G. Tassinary. Inferring psychological significance from physiological signals. *American Psychologist*, 45(1):16–28, 1990.

[7] Z. Cattaneo, T. Vecchi, C. Cornoldi, I. Mammarella, D. Bonino, E. Ricciardi, and P. Pietrini. Imagery and spatial processes in blindness and visual impairement. *Neurosc. Biobehav. Reviews*, 32:1346–1360, 2008.

[8] W. David Hairston, K. W. Whitaker, A. J. Ries, J. M. Vettel, J. Cortney Bradford, S. E. Kerick, and K. McDowell. Usability of four commercially-oriented EEG systems. *J. Neural Eng.*, 11:046018, 2014.

[9] S. Debener, F. Minow, et al. How about taking a low-cost, small, and wireless eeg for a walk? *Psychophysiology*, 49:1449–1453, 2012.

[10] J. I. Ekandem, T. A. Davis, I. Alvarez, M. T. James, and J. E. Gilbert. Evaluating the ergonomics of BCI devices for research and experimentation. *Ergonomics*, 55:592–598, 2012.

[11] M. Garbarino, M. Lai, D. Bender, R. W. Picard, and S. Tognetti. Empatica E3 - A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition. In *EAI 4th Int. Conf. Wirel. Mob. Commun. Healthcare (Mobihealth)*, pages 39–42, 2014.

[12] T. Giorgino. Computing and visualizing dynamic time warping alignments in R: The dtw package. *J. Stat. Soft.*, 31(7):1–24, 2009.

[13] N. A. Giudice and G. E. Legge. Blind navigation and the role of technology. In A. Helal, M. Mokhtari, and B. Abdulrazak, editors, *The Engineering Handbook of Smart Technology for Aging, Disability, and Independence*, pages 479–500. John Willey & Sons, 2008.

[14] J. A. Healey and R. W. Picard. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Trans. Intel. Transport. Sys.*, 6(2):156–166, 2005.

[15] S. A. Hosseini and M. B. Naghibi-Sistani. Classification of emotional stress using brain activity. In G. D. Gargiulo and A. McEwan, editors, *Applied Biomedical*

Engineering, pages 313–336. InTech, Rijeka, Croatia, 2011.

[16] T. Inouye, K. Shinosaki, H. Sakamoto, S. Toi, S. Ukai, A. Iyama, Y. Katsuda, and M. Hirano. Quantification of EEG irregularity by use of the entropy of the power spectrum. *Electroenc. Clin. Neurophys.*, 79(3):204–210, 1991.

[17] S. K. Jena. Examination stress and its effect on eeg. *Int J Med Sci Pub Health*, 11(4):1493–1497, 2015.

[18] E. Kanjo, L. Al-Husain, and A. Chamberlain. Emotions in context: examining pervasive affective sensing systems, applications, and analyses. *Pers. Ubiquit. Comput.*, 19(7):1197–1212, 2015.

[19] A. Keil, M. M. Müller, W. J. Ray, T. Gruber, and T. Elbert. Human gamma band activity and perception of a gestalt. *J. Neurosc.*, 19(16):7152–7161, 1999.

[20] W. Klimesch. Eeg alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Rev.*, 29:169–195, 1999.

[21] J. R. Marston and R. G. Golledge. The hidden demand for participation in activities and travel by persons who are visually impaired. *J. Vis. Imp. Blind.*, 97(8):475–488, 2003.

[22] B. Massot, N. Baltenneck, C. Gehin, A. Dittmar, and E. McAdams. EmoSense: An ambulatory device for the assessment of ANS activity—application in the objective evaluation of stress with the blind. *IEEE Sensors J.*, 12(3):543–551, 2012.

[23] P. Mavros, K. Skroumpelou, and A. H. Smith. Understanding the urban experience of people with visual impairments. In *Proc. GIS Research UK 2015*, pages 401–406, Leeds, UK, April 15–17 2015.

[24] S. Millar. *Understanding and Representing Space: Theory and Evidence from Studies with Blind and Sighted Children*. Oxford: Clarendon, New York, 1994.

[25] D. A. Norman. *The Design of Future Things*. Basic Books, New York, 2007.

[26] P. Peake and J. A. Leonard. The use of heart rate as an index of stress in blind pedestrians. *Ergonomics*, 14(2):189–204, 1971.

[27] E. Peper, R. Harvey, I.-M. Lin, H. Tylova, and D. Moss. Is there more to blood volume pulse than heart rate variability, respiratory sinus arrhythmia, and cardiorespiratory synchrony? *Biofeedback*, 35(2):54–61, 2007.

[28] R. W. Picard. *Affective Computing*. The MIT Press, Cambridge, MA, 1997.

[29] W. J. Ray and H. W. Cole. EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes. *Science*, 228(750–752), 1985.

[30] C. Saitis and K. Kalimeri. Identifying urban mobility challenges for the visually impaired with mobile monitoring of multimodal biosignals. In M. Antona and C. Stephanidis, editors, *Universal Access in Human-Computer Interaction. Users and Context Diversity*, pages 616–627. Springer, 2016.

[31] C. Setz, B. Arnrich, J. Schumm, R. L. Marca, and G. Tröster. Discriminating stress from cognitive load using a wearable EDA device. *IEEE Trans. Inform. Tech. Biomed.*, 14(2):410–417, 2010.

[32] M. P. Wand and M. C. Jones. *Kernel Smoothing*. Chapman & Hall, London, 1995.

[33] R. L. Welsh. Improving psychosocial functioning for orientation and mobility. In W. R. Wiener, R. L. Welsh, and B. B. Blasch, editors, *Foundations of Orientation and Mobility*, volume 2: Instructional Strategies and Practical Applications. AFB Press, New York, 3rd edition, 2010.

[34] R. J. Wycherley and B. H. Nicklin. The heart rate of blind and sighted pedestrians on a town route. *Ergonomics*, 13(2):181–192, 1970.

[35] W.-L. Zheng and B.-L. Lu. Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. *IEEE Trans. Auton. Mental Devel.*, 7(3):162–175, 2015.