

Article

Cognitive and Affective Assessment of Navigation and Mobility Tasks for the Visually Impaired via Electroencephalography and Behavioral Signals

Robert-Gabriel Lupu ¹, Oana Mitruț ^{2,*} , Andrei Stan ¹, Florina Ungureanu ¹, Kyriaki Kalimeri ³ and Alin Moldoveanu ² 

¹ Faculty of Automatic Control and Computer Engineering, Gheorghe Asachi Technical University of Iași, 700050 Iași, Romania; robert-gabriel.lupu@academic.tuiasi.ro (R.-G.L.); andrei.stan@academic.tuiasi.ro (A.S.); fungurea@tuiasi.ro (F.U.)

² Department of Computers, Faculty of Automatic Control and Computers, University Politehnica of Bucharest, 060042 Bucharest, Romania; alin.moldoveanu@cs.pub.ro

³ Institute of Scientific Interchange, Via Chisola 5, 10126 Torino, Italy; kyriaki.kalimeri@isi.it

* Correspondence: oana.balan@cs.pub.ro

Received: 15 September 2020; Accepted: 13 October 2020; Published: 15 October 2020



Abstract: This paper presented the assessment of cognitive load (as an effective real-time index of task difficulty) and the level of brain activation during an experiment in which eight visually impaired subjects performed two types of tasks while using the white cane and the Sound of Vision assistive device with three types of sensory input—audio, haptic, and multimodal (audio and haptic simultaneously). The first task was to identify object properties and the second to navigate and avoid obstacles in both the virtual environment and real-world settings. The results showed that the haptic stimuli were less intuitive than the audio ones and that the navigation with the Sound of Vision device increased cognitive load and working memory. Visual cortex asymmetry was lower in the case of multimodal stimulation than in the case of separate stimulation (audio or haptic). There was no correlation between visual cortical activity and the number of collisions during navigation, regardless of the type of navigation or sensory input. The visual cortex was activated when using the device, but only for the late-blind users. For all the subjects, the navigation with the Sound of Vision device induced a low negative valence, in contrast with the white cane navigation.

Keywords: sensory substitution; cognitive load; brain activity; navigation; multimodal; audio; haptic

1. Introduction

At the world level, approximately 2.2 billion people have a vision impairment or suffer from blindness, caused primarily by uncorrected refractive errors, cataracts, age-related macular degeneration, and glaucoma. The majority of people with vision impairments are over 50 years old, originating especially from low and middle-income countries [1]. The purpose of the Sound of Vision project (SoV) [2] was to develop an assistive system for the blind and visually impaired users that would facilitate navigation and obstacle detection. In this paper, we presented a study of cognitive load assessment and brain activation evaluation during an experiment in which eight visually impaired subjects performed various object detection and navigation activities while using both the white cane (a navigation aid they use on a daily basis) and the SoV device, which provided three types of sensory input—*audio cues* delivered through headphones, *haptic cues* delivered as vibrations applied on a vest that was placed on the user's abdomen, and a combination of both audio and haptic information, called the *multimodal sensory input*. We performed a metrics analysis, cognitive load,

working memory assessment, brain activity, visual cortex evaluation, and the identification of emotions during navigation in the real-world environment.

Section 2 presents an overview of mobility assistive devices, Section 3 introduces the biophysical signals and cognitive load, Section 4 describes the Sound of Vision device, Section 5 details the method, Section 6 presents the results and a discussion, and finally, Section 7 provides the final conclusions and future research directions.

2. Overview of Mobility Assistive Aids

The absence of visual information in the case of blind individuals can be substituted by conveying auditory and tactile stimuli, separately or simultaneously, through specialized assistive devices.

2.1. Auditory Vision Sensory Substitution

Auditory vision sensory substitution (AVSS) devices [3] map the image “seen” by the camera into a matrix of active audio sources. The diversity of AVSSs is very large, ranging from optophone-like systems [4] to devices that use spatialized 3D sounds. The *optophone* (or the piano transform device) scans the image from left to right and converts the detected input into sound cues. The most well-known optophone is the vOICe [5], where the pixels’ vertical position is mapped to frequency, and their brightness is mapped to sound intensity. In other AVSSs [6,7], height is correlated to frequency distribution in the horizontal plane to binaural loudness, and brightness is encoded as sound intensity. In *pitch transform systems* [7,8], distance is related to sound frequency, while in *verbal transform systems* [7,9], objects are assigned to synthetic voice recordings. A problem of most optophone-like modern systems is that they overwhelm the users with too much output information, increasing cognitive load, effort, and concentration. This drawback can be overcome by reducing scene complexity, by maintaining only the salient characteristics and relevant objects, and by applying an effective sonification technique in order to provide the users a natural, effective, and easy to understand environmental representation. Modern AVSSs use binaural 3D sounds spatialized with generic (prerecorded, stored in large databases) or individualized head-related transfer functions (HRTFs). Individualized HRTFs are preferable for creating 3D sound as they are more accurate and fit the user’s auditory characteristics.

2.2. Tactile Visual Sensory Substitution

Tactile visual sensory substitution (TVSS) systems use a matrix of controllable elements that provide spatial and temporal environmental information on the skin, either through kinesthetic or cutaneous sensations. In this type of device, a camera is used to acquire visual input that is consequently transformed into a tactile rendering via the multi-dimensional pin array, facilitating reading, shape, and face recognition [10–12]. One advantage of using TVSS devices is that the tactile sense, contrary to the auditory one, is less used and demanded in everyday activities. Thus, the user can receive cutaneous (awareness and stimulation of the outer surface of the body [13]) and kinesthetic (awareness of the limb position and displacement [10]) cues, without hampering locomotion or auditory perception at all [14]. On the other hand, a drawback lies in the fact that the capacity of the tactile channel is restrained to a limited maximum number of actuators and patterns to be applied. In addition, such devices cannot be used to a large extent because they are tiring and uncomfortable.

2.3. Auditory Tactile Visual Substitution Devices

When the scene is too difficult to be mapped onto the tactile array, the auditory channel is additionally recruited in order to enhance environmental representation, creating **auditory tactile visual substitution (ATVS)** devices. The first multi-sensory device was Nomad [15]. Tactile cues are delivered through a touch-sensitive tablet, and the auditory information consists of synthesized voice recordings. The Heard and Felt Vision Effects (HiFiVE) [16,17] system uses moving speech-like sounds

(tracers—area tracers and shape tracers), binaural panning, and tactile effects in order to map visual images to an audio-tactile representation.

3. Biophysical Signals and Cognitive Load

3.1. Electroencephalography

Electroencephalography (EEG) can provide neurophysiological markers of cognitive-emotional processes induced by stress and indicated by changes in brain rhythmic activity [18,19]. EEG signal processing techniques play a significant role in quantifying cognitive load [20–24]. Bos et al. [20] showed that cognitive load was an indicator of the learning progress. Berka et al. [22] extracted features from EEG signals for monitoring cognitive workload and task engagement. Nilsson et al. [23] showed learning outcomes from the subjects when they navigated a hypermedia environment. Scott et al. [24] also showed that a navigational map could create significantly more germane or extraneous cognitive load. Therefore, EEG/ Electrodermal Activity (EDA) signals are used to measure cognitive load and affective responses, and the overall process is explained in the following section.

Cognitive load and affective responses may impact the learning progress. The detection of reliable cognitive load and affective responses would improve the design of emotional intelligent mobility systems for the visually impaired people (VIPs). The complexity of the tasks is quantified in terms of cognitive load index and affective index, considering two well-established metrics in the scientific literature—the *event-related desynchronization (ERD)/event-related synchronization (ERS) index* and the *left-right asymmetry index*.

Affective responses directly influence the processes of cognitive learning. However, the challenges of learning can evoke negative affective responses [25]. Emotion assessment is a challenging and demanding task because people are not always able to express their emotions verbally [26]. Bos [20] showed that cognitive load could indicate changes in the learning process. He proposed an approach to determine the optimal placement of a limited number of electrodes, and then these electrodes were placed in an experiment aimed at determining arousal and valence. Left frontal inactivation is an indicator of a withdrawal response, which is often linked to negative emotion. On the other hand, right frontal inactivation is a sign of an approach response or positive emotion. High alpha activity (8–12 Hz in the EEG frequency band) is known to be an indicator of low brain activity.

Researchers have addressed the intertwining role of affective responses, learning, and cognitive load. Bower et al. [27] introduced the following hypothesis to study learning patterns: (1) a positive emotion usually increases the learning process through attention and motivation, (2) a positive emotion improves learning by enhancing cognitive load, and (3) a negative emotion decreases the learning process. Cattaneo et al. [28] employed the cognitive load theory for the understanding of the perceptual and neurocognitive mechanisms; however, there are still many open questions on how emotion and cognitive load can ease the learning process of the visually impaired people.

3.2. Electrodermal Activity and Heart Rate

Electrodermal activity (EDA) is a well-known indicator of physiological arousal and stress activation in affective computing [29,30]. It is more sensitive to emotion-related arousal variations as opposed to physical stressors, which can be better reflected by heart rate (HR) measurements. Blood volume pulse (BVP) patterns can also reflect transient arousal and cognition processes [31]. Two outdoor mobility studies from the early 1970s have suggested that some form of psychological rather than physical stress is responsible for visually impaired people's increased HR versus sighted pedestrians [32,33]. However, certain mobility tasks (for example, stairs climbing) may result in an interactive psychological stress effect and momentary physical workload; thus, cardiovascular measures may be less suitable than EDA.

4. The Sound of Vision Device

Spatial navigation is a category of spatial cognition related to performing tasks, such as following paths, detecting obstacles, and reaching targets. It is based on developing, maintaining, and recalling an internal representation of the environment [34,35]. This internal representation depends on the spatial relation between entities and on the subject's position, being classified into two categories: egocentric—the navigator is in the center of the coordinate system, and allocentric—the reference external to the navigator.

4.1. Technical Description

Sound of Vision is a wearable device that allows a visually impaired user to perceive and navigate the environment. It works by permanently scanning the environment, extracting essential features, and rendering them to the user through audio and haptic means.

The Sound of Vision final prototype includes an integrated custom hardware solution and a complex software solution, supporting the real-time operation of the device, as well as training tools and materials.

The hardware components of the system are:

- a headgear, including a 3D acquisition unit (depth camera for indoor or low light outdoor conditions, stereo camera for outdoor or bright light conditions, head and body inertial measurement unit (IMU) for body orientation) and an audio rendering unit (mounted on the head);
- a haptic belt with a matrix of 60 vibrating motors (six rows and 10 columns, placed on the abdomen);
- a processing unit: a small laptop with powerful CPU and GPU units (in a backpack);
- a wireless remote control (in the pocket).

When scanning the environment, the user can select from two different models for both audio and haptic: the *discrete model* (which renders the objects sequentially) and the *continuous model* (which provides real-time information at once about all the objects in the field of view). They are divided into sub-models and have different variations, as well as additional features for safe and reliable navigation: *Danger mode*—alerts about proximity objects and prevents collisions, *Flashlight*—enables the rendering of an object's distance in front of the user, *texts and special signs detection* and *best free space*—which indicates a navigable opening between surrounding objects.

The user movement is guided solely by the Sensory Substitution Device (SSD) with no additional feedback from the assistant or from other sources (i.e., maps from Google or GPS coordinates from a GPS device). The SSD device generates audio and haptic signals that are an encoded representation of the environment in the proximity of the user. Through intensive training, the user gains fluency in understanding the audio and haptic encoded feedback issued by the SSD device, and then he/she can make proper decisions for further movement in the environment.

Like any other person, the VIP wants to walk in the direction of the sound source. The SoV device scans the environment, detecting the obstacles and their features, and sends audio or/and haptic stimuli to users. These stimuli help the VIP to avoid obstacles and to find a secure path to the sound source. Obviously, the real scene and the stimulation are changed/updated according to the VIP's route, like in a maze. Depending on their perception, the VIPs may choose different paths.

4.2. The Focus of the Study

In line with the SoV project overall goals, the aim of this study was to evaluate the VIPs' cognitive load and emotional stress in real-world mobility experiments, in two cases: navigation with white cane and navigation relying on the SoV prototype with audio, haptic, or audio and haptic (multimodal) codification. The research questions we pursued were based on the following comparative assessments of cognitive load:

- when using the Sound of Vision device with audio vs. haptic vs. multimodal input;
- when using the Sound of Vision device vs. white cane during a navigation task in the real-world environment.

Based on some achievements presented in scientific papers and on the valuable previous experience [36–39], the experiments were oriented to collect EEG and physiological (EDA and HR) signals in five different mobility tasks in order to highlight the VIPs' cognitive load and stress in correlation with some events (collisions or total confusion) captured from the recorded videos.

As presented above, it has been proven by many studies that EEG is a promising and common approach to measuring cognitive load (CL), working memory load, emotional states, and any cerebral signals denoting cortex responses to specific stimuli. CL was an effective real-time index of task complexity backed up by behavioral evidence. Complementary, the peripheral physiological measurements reflected arousal and stress activation (EDA) and transient processes in arousal and cognition (HR). The mobility tasks did not include a consistent physical effort, so HR could be also considered in a multimodal approach.

An important aspect of this work was related to the VIPs' preference and long-term accommodation to navigating using the white cane, as they were educated to use it from the moment they lost sight or from childhood. Although the SoV device offered much more information about the environment (the number of nearby objects, their position and properties, the presence of specific objects, and so on), it was expected that for the first tests, the VIPs would have lower cognitive load and better performances during the benchmark task with the white cane than with audio or haptic stimuli. Obviously, the length of the accommodation period with the SoV device depended on each VIP's education and ability to learn. It was important for this study to understand how easily the SoV stimuli were perceived and processed by a VIP and which navigation modality (audio, haptic, or audio and haptic) was less stressful and more quickly accepted. We expected that audio mobility would outperform the haptic and fusion mobilities, knowing that blind people generally have a well-developed hearing sense.

Additionally, the VIPs' visual cortex excitation by audio and haptic stimuli during navigation was investigated, expanding the existing literature [40,41] that has reported brain activity in the visual cortex during EEG measurements for blind people who have received visual information through sensory substitution devices (SSDs).

To our best knowledge so far, this paper was the first one to present a comparative study regarding VIPs' stress, cognitive load, and visual cortex excitation while navigating in the real-world using a common white cane vs. a sensory substitution device.

5. Materials and Methods

This experiment has been carried out using the Acticap EEG device with 16 electrodes, provided by Brain Products GmbH from Germany and Shimmers Multisensory provided by the Shimmer Sensing company from Dublin, Ireland.

5.1. Experimental Setup

The aim of the experiment was to obtain a dataset, as large as possible, with EEG and physiological signals during the trials designed for traveling in fixed scenarios with the help of the SoV device [42]. There were two user setups. The first one was the **virtual training environment (VTE)** setup that was used to train the subjects and accommodate them with the audio and haptic encoding models prior to using the system in real-life scenarios [43,44]. The second setup (Figure 1) was the **real-world (RW)** setup used in an indoor controlled environment and outdoors.

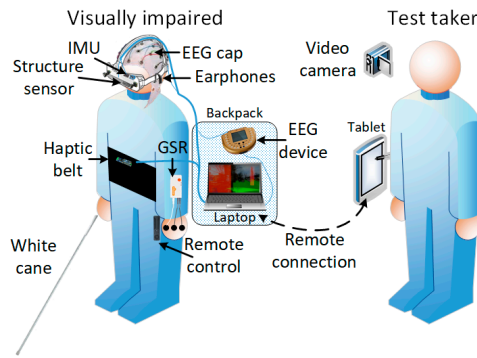


Figure 1. Real-world setup. IMU- Inertial Motion Unit, EEG–Electroencephalography, GSR–Galvanic Skin Response.

The difference between the VTE and RW setups consisted mainly of the video streaming sources that feed the processing and control unit. For VTE, the video stream was provided by a virtual reality serious game in which the VIP navigated using the keyboard or a joystick. In the VTE tests, the VIP wore the headset consisting of a structure sensor stereo video camera and an IMU sensor [45,46], but only the IMU signal was used in order to orient in the virtual environment by head movements. For the RW setup, the video stream was provided by the structure sensor (indoor use of the system) or stereo video camera (outdoor use of the system).

The SoV system used in the RW experiments is depicted in Figure 2 and consisted of two computing systems: the first one was the processing and control unit (PCU) attached to the VIP, and the second one (the tablet from Figure 2) was used by the assistant who controlled how the trial was performed. Via a remote connection with the PCU, the assistant could adjust parameters, select scenarios, and enable physiological signals recording. The PCU ran the SoV runtime application, which sensed the environment and provided audio and haptic stimuli. It also recorded physiological data from the user who performed the navigation tasks. The IMU signals were used to determine the user’s body and head orientation that was further used to render the audio and haptic output in accordance with the RW scene. The VIP had the opportunity to select the most appropriate audio or haptic encoding by using a remote control connected to the PCU. During navigation, there was no communication between the VIP and the assistant. The VIP walked autonomously based only on the stimulation provided by the SSD. The test taker’s task during tests was to ensure that the system was working and that the VIP received correct clues. He did this by using a tablet connected wirelessly to the VIP’s laptop.

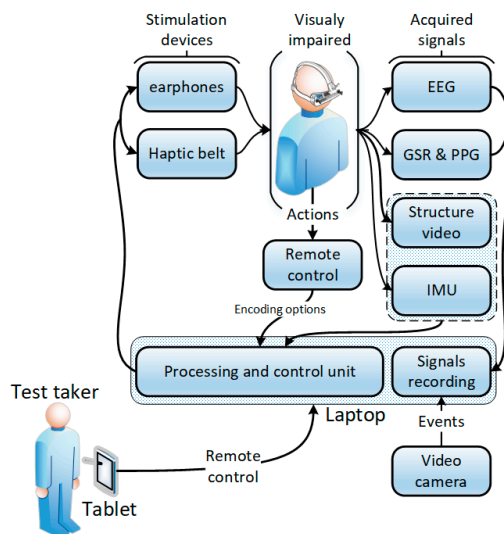


Figure 2. The system architecture for the real-world (RW) setup. PPG-plethysmography.

We used a lightweight laptop (Dell XPS) and two cooling fans in order to cool down the laptop, which are part of the SoV prototype for experimental purposes. The experimental system (laptop, battery, fans, and EEG device) weighed no more than 3.5 kg, and the users did not feel uncomfortable due to the weight and/or heat. Actually, a lighter version of the SoV system—more energy-efficient and low cost—is under development.

A print screen from the serious game called “**treasure hunt**” (TH) is presented in Figure 3. The VIP had to use the joystick in order to position himself in the virtual scene, exactly where the sound source originates. In RW, the user had to navigate through an indoor environment in order to reach a target sound source while avoiding cardboard box obstacles of various dimensions (Figure 4).

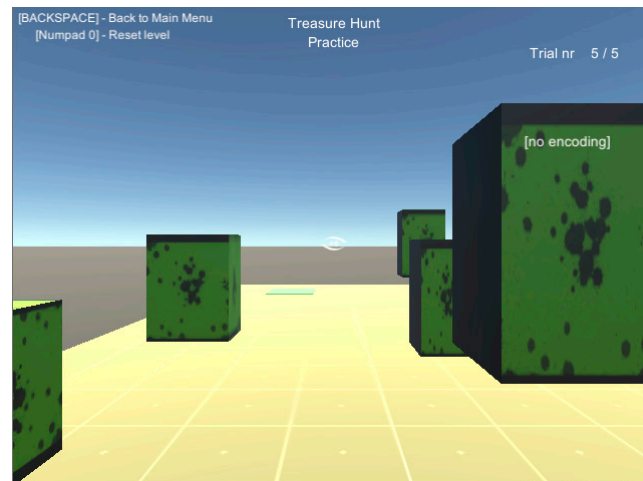


Figure 3. Treasure hunt in a virtual training environment (VTE).



Figure 4. Treasure hunt in RW.

We conducted EEG and EDA/HR recordings during experimentation in RW, under two conditions: **white cane only** (for the participants who used this mobility aid on a regular basis) and **SoV device only**. The scenes were tested using the *audio encoding*, the *haptic encoding*, and *both the audio and haptic encodings (multimodal)* with the SoV device, 5 trials each. In order to minimize the required testing resources, the users had the opportunity to choose the sonification model and tactile stimulation [47–49] that best suited their level of perception and understanding.

The **static scenes (1R)** were tested with the *discrete model*, while the **dynamic scenes (TH)** were tested with the *continuous model* in both virtual and real-world environments.

In the *discrete (or iterative) model*, the scene was rendered in a loop, one by one. A sphere was constantly expanding its radius until 5.25 m with a speed of 2 m/sec. Auditory and tactile stimuli were provided when this sphere intersected scene objects, allowing distance detection, as well as comparing the distance between objects. The *continuous model* rendered an entire scene at once, providing instantaneous information via audio and haptic.

VTE tests were recorded automatically by the SoV system, while RW tests were recorded manually by testing assistants. Furthermore, each trial in a test was videotaped for annotation purposes. The time needed to finish every trial and the accuracy were saved: number of collisions between user and obstacles, number of cane contacts with objects, time duration, together with path length followed.

Each test consisted of 5 trials, and each one was assigned to a fixed path/boxes arrangement for TH, as it is presented in Figures 5–9.

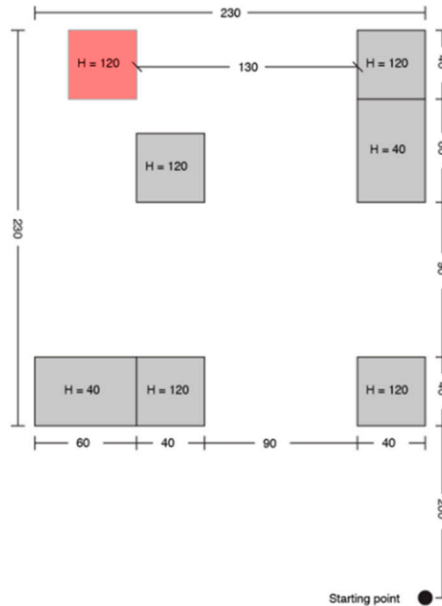


Figure 5. Treasure hunt (TH) configuration A.

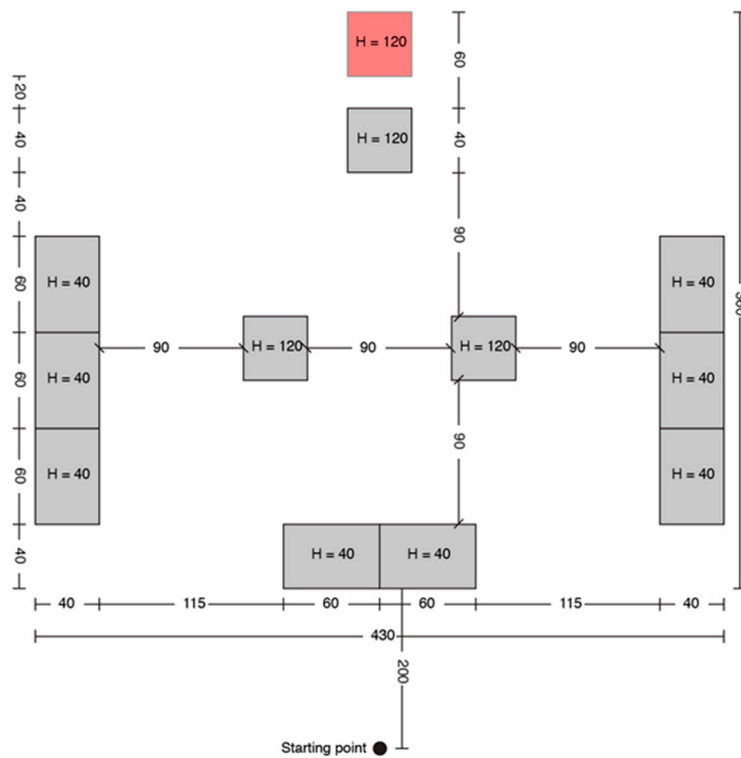


Figure 6. TH configuration B.

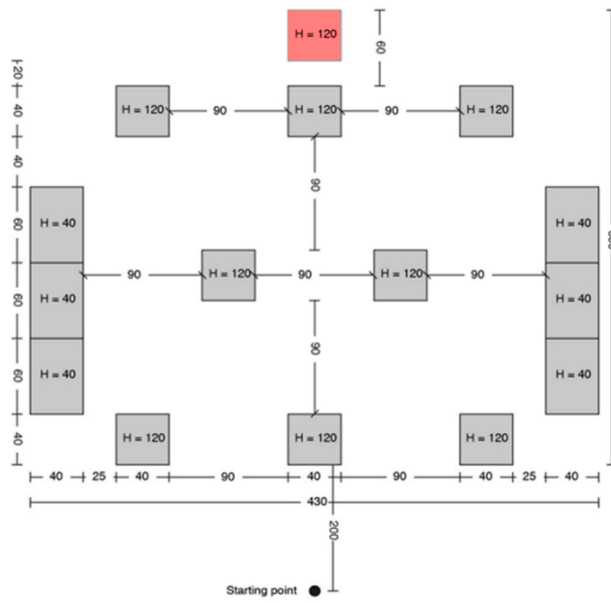


Figure 7. Treasure hunt configuration C.

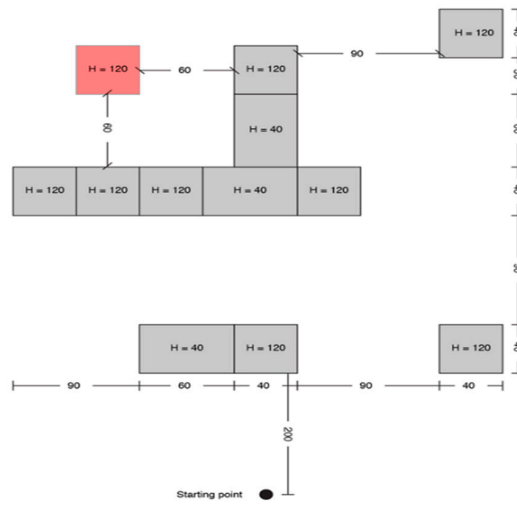


Figure 8. Treasure hunt configuration D.

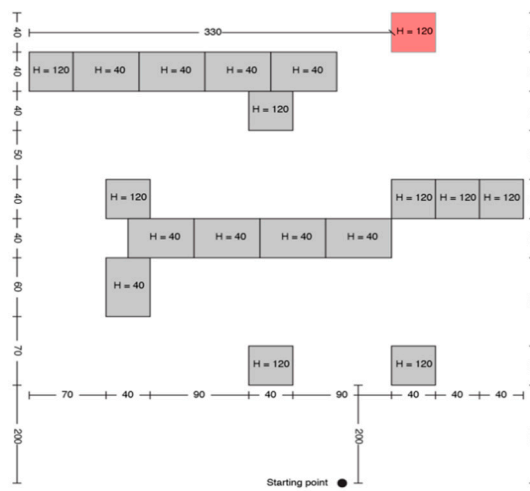


Figure 9. Treasure hunt configuration E.

5.2. Data Collection

In this study, 15 VIPs were involved in training and testing tasks in the virtual environment and real-world settings, using the SoV prototype and different releases of the SoV runtime. After each testing stage, important improvements were made to the hardware and software resources, including audio and haptic encodings based on the VIPs' feedback. Finally, only 8 complete datasets (corresponding to 3 females and 5 males, aged 20–42) with fully validated data were retained and subsequently analyzed. All the participants provided informed consent approved by the research ethics committee of the institutions involved in the project (Approval number: 9083/15.05.2017). One hour before performing the tests, the VIPs did not drink coffee nor black tea, and also smoking was forbidden before or during the experiments.

For EEG and additional physiological measurements, the following equipment was used:

- a BrainProducts V-Amp 16 amplifier and an EasyCap helmet with 19 sintered Ag/AgCl miniaturized passive electrodes for EEG signal acquisition with a sampling rate of 512 Hz;
- a Shimmer3 GSR+ unit sensor for measuring electrodermal activity/galvanic skin response (EDA/GSR) and continuous HR;
- a video camera or smartphone for video recording in real-time.

The acquisition procedure used 16 electrodes, namely Fp1, F7, F3, C3, P3, P7, C4, O1, O2, Fp2, F8, F4, C4, P4, T8, P8, and an ear reference, placed according to the 10-20 international system. The sampling rate was 512 Hz, and the AFz electrode was connected to the ground. To ensure reliable EEG raw data, the impedance of each electrode was maintained below 5 k Ω , by using a good abrasive gel. The OpenVibe open source software was used for EEG acquisition. The OpenVibe server acquired the EEG signals, and the OpenVibe client saved or sent the raw data as a stream. The Shimmer GSR unit sent the acquired data via Bluetooth.

The data acquisition process is outlined in Figure 10. An important part of the acquisition process is the usage of the lab streaming layer (LSL) protocol [50] so that each data component had to provide a stream of data as output. For the components that do not natively provide LSL output streams, simple adaptors had to be designed, as in the case of the data provided by the EDA/GSR device. The EEG data was available as an LSL stream provided by the OpenVibe application. The applications developed in this project (VTE and SoV runtime) provided LSL streams to source events that were internally generated and of interest for later analysis. The data was stored in a tabular format inside hdf5 files [51], together with a timestamp provided by the LSL.

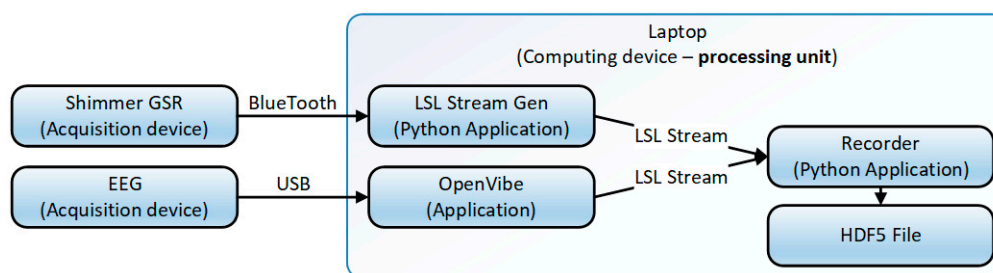


Figure 10. The data acquisition process. LSL, lab streaming layer.

In what concerns video recording and annotations, the mp4 video files recorded during the tests were annotated with Chronoviz [52]. The events: *Start* (beginning of the recording), *Collision*, *Find* (find the source/treasure), *Lost* (lost control), *TouchCane* (only for the test with the white cane), *Stop* (end of the recording) were considered. For each mp4 file, a CSV file with annotated events and corresponding Unix timestamps was generated.

Alternatively, if Chronoviz could not be used due to system constraints (Chronoviz needs a system with Mac OS X 10.6 or later), a Python script was designed to synchronize the data streams acquired by

the processing unit with the video recordings of the experiment. The script aligned the timestamps of the samples in the acquired data streams (provided as csv files) with the timing information found in the video recording. The application ExifTool [53] was used for gathering timing information from the movie files. The data streams were trimmed or padded in order to fit the movie length. The script detected and reported any timing misalignments and provided means to fine-tune the synchronization process. The resulting adjusted data streams and movies could be annotated later in a similar way as in Chronoviz. The synchronization process is presented in Figure 11.

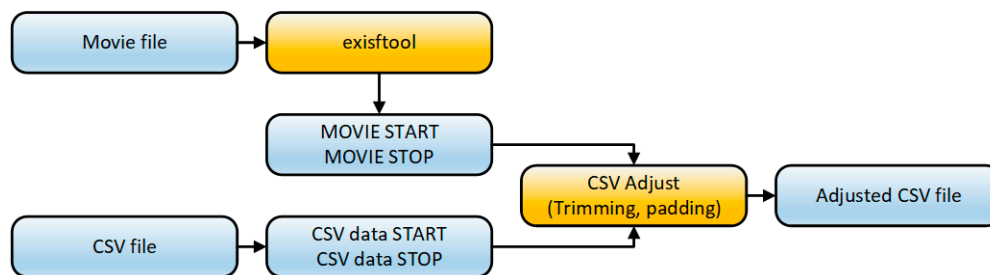


Figure 11. RW events synchronization process.

5.3. Data Acquisition and Preprocessing

The acquired brain waves were pre-processed. We applied a band-pass filter for 0.5–100 Hz, a notch filter to remove power line contamination at 50 Hz, and a band-pass filter to obtain frequency bands of interest (delta, theta, alpha, beta, and gamma). The artifacts (involuntary eye blinks, muscle movements, brief amplifier saturations) presented in the EEG signals were removed using an online Savitzky-Golay filter. The EEG data obtained after pre-processing were baseline-normalized by subtracting for each participant and for each channel the mean of the resting state recordings (recorded in the laboratory during the VTE sessions).

The Shimmer software for EDA acquisition could not be efficiently used in the SoV setup (Figure 1), and therefore the streams from the sensor were acquired over Bluetooth at 16 Hz. The skin resistance values (y , μS) were computed from the Shimmer ADC values with the following linear function:

$$y = p1 \times x + p2 \quad (1)$$

where $p1$ and $p2$ are parameters specific to the range setting and can be selected from the datasheet of the sensor. If the electrodes are not tightly attached and lose contact with the skin, motion artifacts (high-frequency noise) can be present in the acquired signals. A low pass filter was applied to remove high-frequency noise, which can be attributed to movement artifact and other noise components. A cutoff frequency of as low as 1–5 Hz could be used without affecting the data of interest due to the slowly varying nature of the EDA responses.

5.4. Data Analysis

Within the broader framework of the SoV project, the aim of this study was to explore the VIPs' brain activity during navigation tasks with the help of an SSD based on audio, haptic, and multimodal encoding, compared to white cane navigation. The research was focused on assessing cognitive load, visual cortex excitation, and emotions evaluation during RW navigation. For each exploration, the EEG signals were selected according to the analyzed brain lobes and the *power spectrum*, and the *asymmetry between the two cortex hemispheres* was calculated.

Usually, CL is investigated in the channels corresponding to the frontal lobe, which reflect the activity of short-term memory and consists of calculating **frontal-asymmetry**, meaning the difference between the logarithms of the power spectrum of the left and right hemispheres divided by the logarithm of the total power spectrum of both hemispheres. There is no single standard way to

calculate asymmetry, and some authors use the difference or the ratio between the spectral powers of the signals on the right and the correspondents in the left hemisphere. Anyhow, *higher asymmetry reflects a strong workload, while lower asymmetry reflects avoidance and relaxation* [54].

CL is strongly related to emotional well-being states. The “feeling good” aspect of well-being deals with the balance of positive emotions vs. negative emotions. Well-being reflects a person’s ability to identify and respond to the challenges of everyday life, even painful and unpleasant events [55]. Hawthorne presented an extensive study on how feeling good might contribute to cognitive load in different ways [55].

Certain states can be more accurately investigated if the EEG waves are analyzed in the five specific bands: *delta*, *theta*, *alpha*, *beta*, and *gamma*. The *delta* waves reflect the unconscious states, and it is usually recorded in deep dreamless sleep. The *theta* waves are typically associated with the subconscious mind, sleeping, dreaming, meditation, or even artistic creation. The *alpha* waves are visible in all the cortex lobes and give valuable information regarding brain activation and the relaxed (but yet aware) mental state. High alpha activity has been correlated to brain inactivation. The *beta* waves are correlated to high mental activity, more prominent in the frontal cortex but visible over other lobes as well. The alpha and beta waves are the most used to classify workload using EEG. The *gamma* waves (>30 Hz) reflect hyper brain activity and have become more and more studied as the sampling frequency of the acquisition systems has increased [56].

Regarding **visual cortex (VC) excitation**, it must be specified that it was not known before the year 2000 whether the visual cortex could receive input from other sensory modalities besides the eyes through the lateral geniculate nuclei. Afterward, the EEG measurements have revealed that the VC activity is higher for blind subjects during rest or auditory/tactile tasks than in normal control. Without a certain demonstration, Sadato et al. suggested that in blind subjects, the cortical areas normally reserved for vision might be activated by other sensory modalities [57]. In 2003, Burton reviewed various brain imaging studies, which investigated the visual cortex activity of VIPs during nonvisual tasks, such as hearing messages, Braille reading, or even sensory discriminations of tactile or auditory stimuli, and concluded that the loss of vision did not lead to a permanent inactivation of the visual cortex [58]. A scientific report from Georgetown University Medical Center concluded in 2010 that “people who have been blind from birth make use of the visual parts of their brain to refine their sensation of sound and touch” [59]. In recent years, several studies have highlighted enhanced auditory processing in blind persons to partially compensate their impairment, with greater sensitivity of the other senses. It has been proved that the VC plasticity allows this cortical lobe to be colonized by the auditory and somatosensory systems in the case of congenitally blind persons. The study conducted by Campus et al. revealed that the occipital activation to sound was strong in sighted persons and much lower in blind persons [60]. Another valuable conclusion was that the occipital lobe of sighted subjects played a major role in the reconstruction of the environmental spatial metrics and that vision loss blocked this process. Obviously, it is expected to remark differences in VC excitation between the people who are blind from birth and those who lost their sight later and know what color, distance, or shape mean. For this analysis, the O1 and O2 electrodes are the most important, but also the Oz and the electrodes from parietal lobes should be considered in an extensive study.

In terms of evaluating emotions, it is well known that the **amygdala** is responsible for the perception of emotions, such as anger, fear, and sadness. The pre-frontal cortex and the hippocampus (located in the medial region of the temporal lobe) are highly correlated to emotional activity [56,61]. Because the right hemisphere is associated with negative emotions (i.e., fear or disgust), and the left hemisphere is highly activated by positive emotions and motivation (i.e., happiness and satisfaction), the EEG asymmetries in the frontal and parietal lobes are relevant for valence and arousal assessment [56]. A thorough evaluation can be performed if the signal analysis is performed on the EEG frequency bands of alpha, beta, and gamma. According to these findings and based on some other studies related to efficient EEG channels selection for emotion recognition, Zhang and his coworkers recommended the following set of electrodes: Fp1, Fp2, F7, F8, C3, FC5, FC2, AF4 (frontal lobe), T7, T8 (temporal

lobe), O1, Oz (occipital lobe), and P3, P4, Pz, PO4 (parietal lobe) [62]. For emotions assessment in this study, only the channels C3, C4, T7, T8, P3, P4, F3, and F4 were considered due to the limited number of electrodes of the EasyCap helmet. The asymmetry in the pre-frontal lobe was presented in CL evaluation, and O1 was not considered because the standard list refers to sighted people, and, in our approach, the visual cortex was subjected to special attention.

6. Results and Discussion

6.1. Navigation Metrics Analysis

As mentioned above, the EEG, HR, and EDA (GSR) signals were acquired for the treasure hunt tests, using the white cane or the SoV device with three spatial information encodings—audio, haptic, and multimodal (audio and haptic). Besides the video recordings and the files containing the data obtained during the experiments, important metrics regarding navigation were collected for each user involved in the study: the *time required to accomplish a trial*, the *length of the path*, the *number of major or minor collisions*, and also the *numbers of white cane contacts with the obstacles*. All these data are summarized in Table 1 and reflect the cumulative performance of all the users for each scenario type.

Table 1. Cumulative experimental data for the treasure hunt (TH) tasks—navigation with the Sound of Vision (SoV) device and white cane.

Codification	Scenario Type	Collisions Total Number	Path Total Distance (m)	Total Time (s)
Audio	A	12	33.69	293
	B	10	44.36	299
	C	22	42.8	261
	D	18	41.8	306
	E	20	62.05	409
Haptic	A	13	28.3	179
	B	11	39.1	221
	C	17	41	261
	D	20	44.1	255
	E	34	54	440
Audio and Haptic	A	7	33.35	236
	B	14	39	267
	C	15	47.3	327
	D	12	51.6	271
	E	24	56.5	345
White cane	A	4	23.27	204
	B	5	28.8	205
	C	6	27.8	213
	D	2	30.16	206
	E	6	37.77	255

Figures 12–14 present the averages of time duration, number of collisions, and traveled distances for RW navigation with the help of the white cane and SoV device, in case of all the five obstacle arrangements (A to E). A+H stands for audio and haptic (multimodal). In the case of the short and easy routes, the walking durations were very similar for audio stimulation and cane traveling, while the haptic and multimodal stimulation required less time than the white cane. Only for the most complicated test scenario (E), the cane and multimodal tasks were performed in a shorter time than with haptic and audio input. It is known that the VIPs usually walk slowly, and it was encouraging that the SoV device did not slow down the movement of the users.

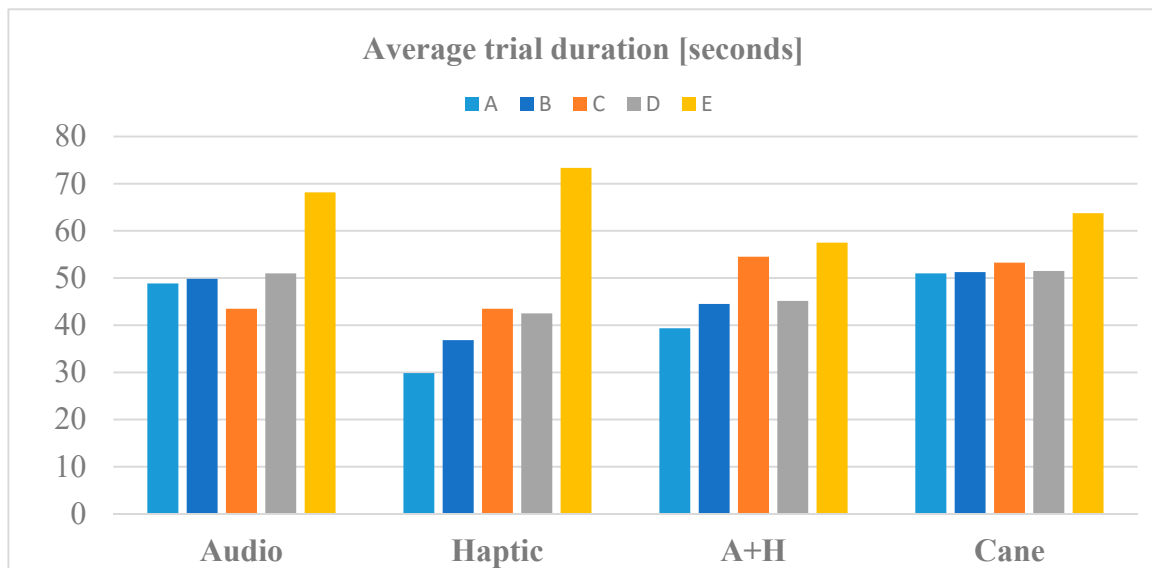


Figure 12. The average time to perform navigation tasks.

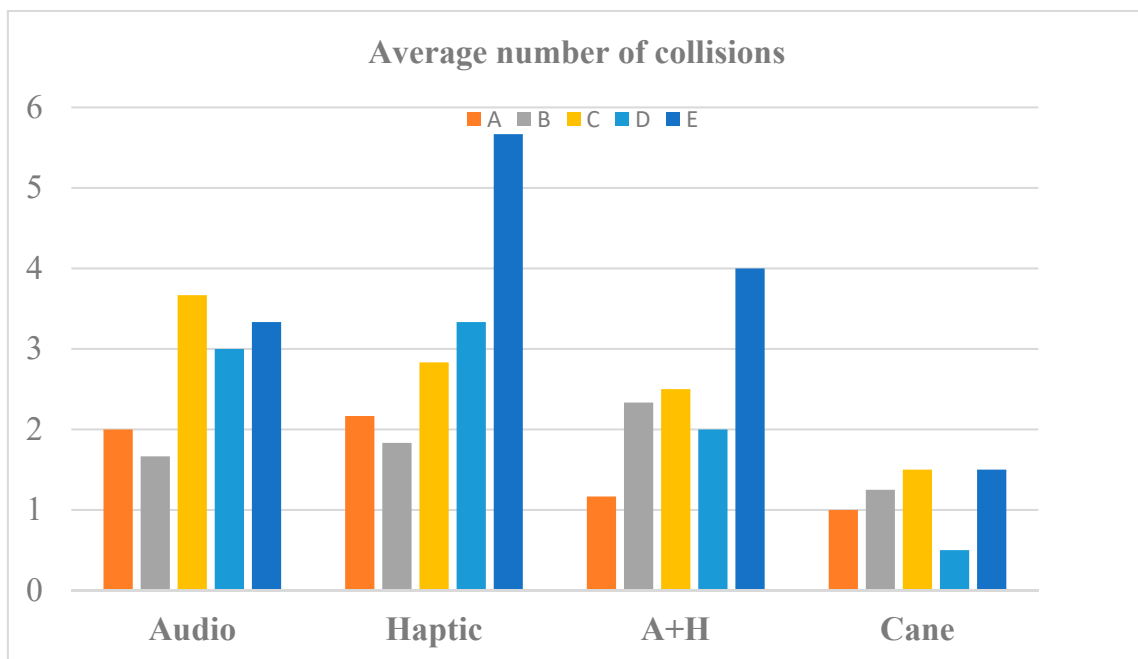


Figure 13. The average number of collisions during navigation tasks.

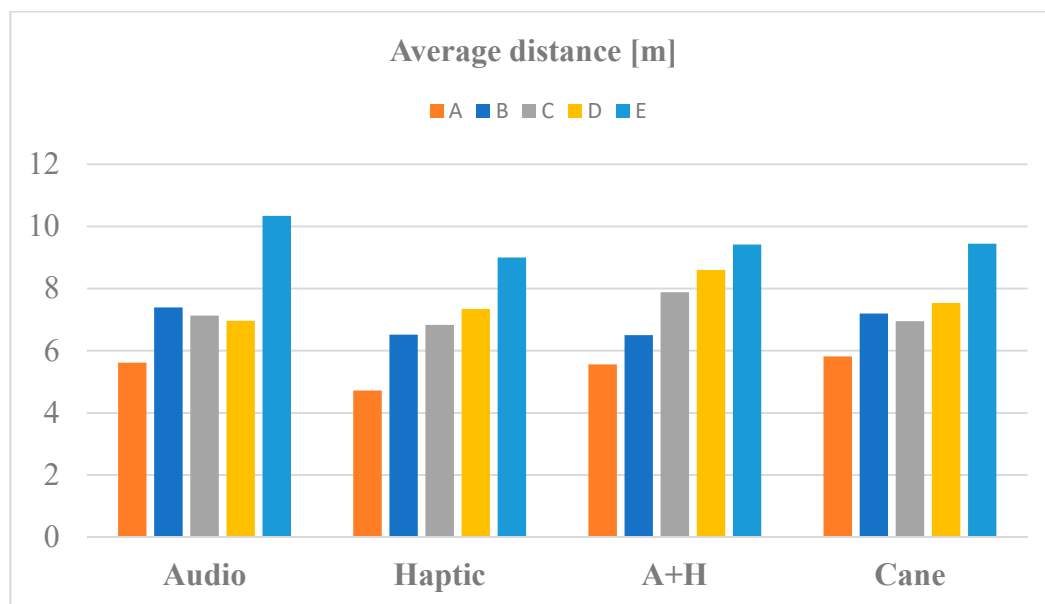


Figure 14. The average distances of navigation tasks.

We noticed a higher number of collisions when using the SoV device in comparison to the white cane navigation. This fact was expected because usually, the VIPs touch the objects with the cane along their path, avoiding the great majority of collisions. The average distances did not differ too much between cane and SoV navigation, except for the SoV audio mode, for which the routes were significantly longer regardless of the testing scenario.

From Table 1 and Figures 12–14, it can be concluded that the required time, the length of the path chosen by each VIP according to his/her perception of the SoV stimulation, and the number of collisions depended on the complexity of the scene and on the user's training and ability to adapt to a new navigation aid. Obviously, the time, length of the path, and the number of collisions were much higher for the scenes C, D, and F. Some VIPs had better results with the audio mode and others with the haptic mode, but the number of collisions was higher for the haptic mode. As expected, the metrics for the white cane navigation were better because the VIPs were accustomed to using it on a daily basis. As a particular conclusion, *the VIPs' navigation performance with the SoV device was better in the case of the multimodal encoding, in terms of duration and number of collisions*. On the other hand, no general conclusion could be drawn because the number of VIPs involved in the experiments was small, and also a VIP could have learned the scenes during the first trials and performed better during the last trial, even if the experiments were randomly conducted.

6.2. Cognitive Load Analysis

Figures 15 and 16 present the total cognitive load for all the validated experiments and the five test scenarios, in the case of both the SoV device and white cane navigation. In the case of SoV navigation, we computed the average of the audio, haptic, and multimodal stimulations. Regardless of the difficulty of the test scenario (A is the easiest, and E is the most difficult), high values of CL were observed for the electrodes related to the frontal cortex (O1 especially, in the vision area) if the SoV device was used. The increase of frontal cortical activity was expected, but the activity of the visual cortex (VC) was worth being investigated because it supports some previous opinions about VC activation in the case of the VIPs who received various environmental sensory stimulation. It should be noted that the brain activity corresponding to the O1 channel was significantly higher than for the O2 channel for both types of navigation, resulting in an increased emotional state.

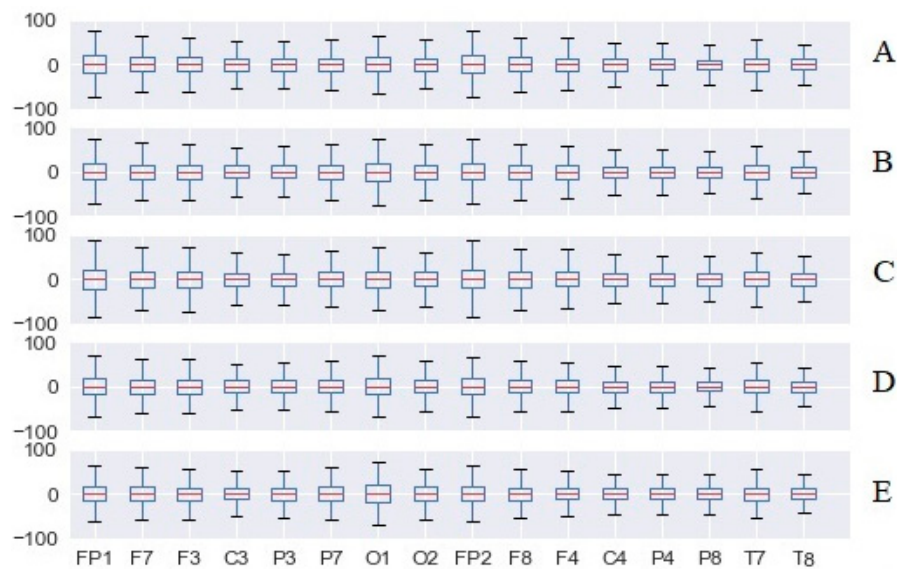


Figure 15. The total cognitive load of all electrodes for white cane navigation.

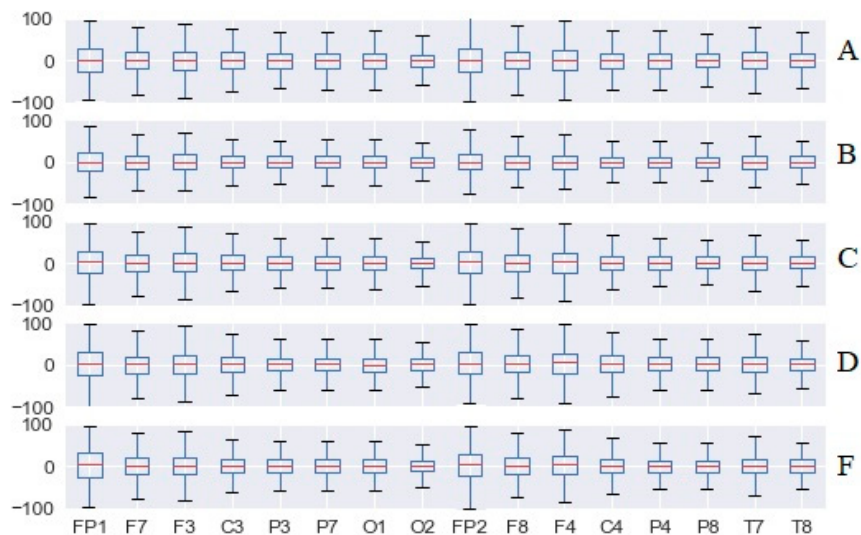


Figure 16. The total cognitive load of all electrodes for Sound of Vision (SoV) navigation.

In Figure 17, we present the total CL for the scenario treasure hunt (TH), configuration C. The main conclusion was that there was a *significant increase in the CL index* (indicated as a negative fluctuation according to the CL index definition) *in the case of using the SoV device with audio, haptic, and multimodal stimulation in comparison to white cane navigation*. The conclusion was similar in the case of the other testing scenarios (TH, configurations A, B, D, and E).

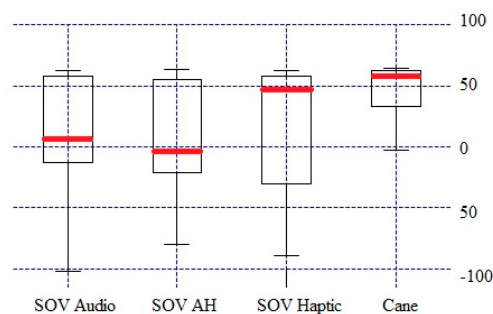


Figure 17. The total cognitive load (CL) for the TH—configuration C task.

In contrast to the CL values presented in Figures 15 and 16 for each electrode, the global CL index was calculated on average for all brain waves for all users, aiming to have a general representation of the brain activity. The short box related to cane traveling meant that the data consistently hovered around the center value, denoting a similar effort for all users, and the whiskers indicated a quite limited distribution as well. In the case of using the SoV device with audio, haptic, and multimodal stimulation, the taller boxes indicated more variable data, and the whiskers showed a wider distribution, namely more scattered data. The different ways in which the users perceived the haptic and sonification models could explain this conclusion, which anyway was in accordance with the plots depicted in Figures 12–14. The consistency of the experimental data and the preprocessing accuracy were proved by the lack of outliers in the total CL indexes. Although the median value for haptic stimulation was closer to the median for cane walking, the distribution of global CL was the widest one. It could be observed that multimodal stimulation had the effect of reducing the spread of the global CL index. The tactile and auditory stimuli were processed by distinct lobes of the cerebral cortex with significant differences in CL, and this should explain the negative skewness in the case of cane and haptic stimulation and the positive skewness in the case of audio and multimodal stimulation. The conclusion was similar for the other testing configurations (A, B, D, and F) of the TH scenario.

6.3. Brain Activity Analysis

Besides this general evaluation of cognitive load, it was relevant to explore how the VIPs' cortical lobes were activated during walking on certain routes with the white cane or guided by the SoV device using the three input encodings. Particular reactions were expected, depending on the type of visual impairment and on the users' training or education. For this, the analysis of the individual frequency bands was performed according to the literature guidelines. First of all, *the alpha waves* were investigated, especially in the *frontal lobe*, taking into account that there is an *inverse relationship between alpha power and cortical activity*; namely, more brain activity (engagement) means less alpha power [63]. A more detailed analysis should be done if the alpha-1 (lower alpha, 7–10 Hz) and the alpha-2 (higher alpha, 10–13 Hz) frequencies were considered because it is well known that *alpha-1 is related to response inhibition and attentional demands*, and *alpha-2 reflects task performance in terms of speed, relevance, and difficulty* [64]. It has been proved that *people with relatively increased left-frontal alpha activity are more motivated and focused in a positive way*, and their related emotions are joy or anger. In contrast, *the increase of right-frontal activity denotes a more negative motivation accompanied by fear, sadness, and disgust* [63,65]. The asymmetry was calculated based on the difference between the logarithms of the spectral powers from the left and right brain hemispheres.

For a more accurate assessment of brain activity related to users with different perceptions and visual impairments, the envelopes of the alpha1 and alpha2 bands asymmetries were depicted for the considered navigation tasks, and the collisions annotated with Chronoviz were marked with black dots. It must be mentioned that the acquired signals for navigation with the cane or with the SoV device had different lengths, according to the time required to perform the task and the path chosen by the user, as it is presented in Table 1. The brain activity exploration was oriented towards analyzing the late visual impaired users in a group and the subjects who were born blind in another group.

In Figures 18 and 19, the envelopes of the asymmetries depicted are related to a user who was born blind (early-blind). He usually navigates using the cane, and he took part in all the training sessions in the virtual environment and ego-static real-world tests. *A significant difference between navigation with the cane and the SoV device was observed only for the audio encoding, in terms of response inhibition and attentional demands*. Although SoV is a completely new device that implies a different way of navigation, however, the consistent training in the virtual environment and the ego-static real-world tests helped a lot the user to accommodate to the encodings. *The greater attentional demand (reflected by the alpha1 waves) was evident for audio stimulation*, and it could be assumed that this was due to the fact that the VIPs usually rely heavily on the environmental noise when they navigate. They also try to perceive natural noises when the SSD sonification is conveyed to them. From the perspective of

alpha2 frequencies, meaning speed and task difficulty, for this user, for all the encodings, the values obtained for navigation with the SoV device were significantly higher than those obtained in the case of using the white cane. Anyhow, this conclusion was expected, considering the novelty of the SoV system for the users and the fact that the VIPs walked relying on the white cane in a natural style for a long time. The collisions (marked with dark dots) were well correlated with the inflection points of the envelopes' variations.

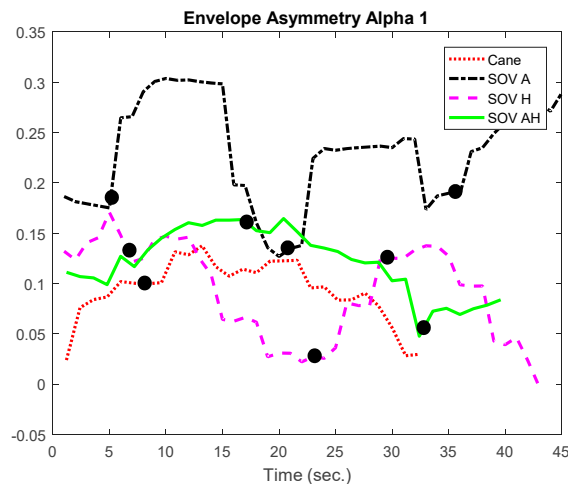


Figure 18. The envelopes of the alpha1 asymmetries for a user who was born blind.

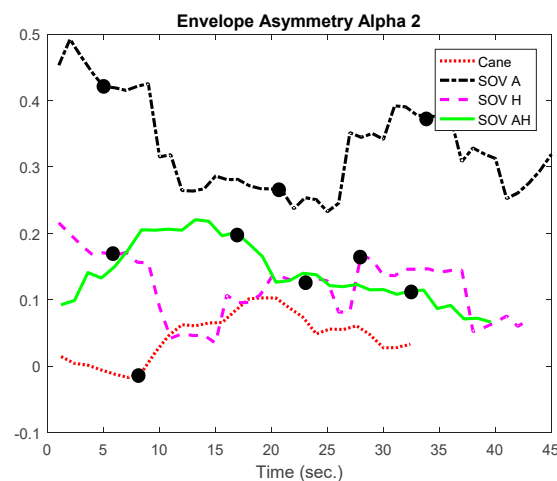


Figure 19. The envelopes of the alpha2 asymmetries for a user who was born blind.

In Figures 20 and 21, the envelopes presented correspond to a user from the late-blind group. He lost his sight at 17, has a good education, and usually navigates accompanied by a family member, without using the white cane. He quickly got used to the SoV device and got good scores in the training sessions. In this study case, the alpha1 asymmetry values (Figure 20) were higher for the cane navigation (even the necessary time was shorter), compared to those obtained for SoV navigation, regardless of how the environmental information was encoded. This demonstrated a higher concentration for the cane navigation and good and fast accommodation with the SoV device. The alpha2 asymmetries (Figure 21) highlighted increasing difficulties for the audio and multimodal encoding tasks. However, the range variations of the alpha1 and alpha 2 asymmetries were similar for the two users considered. This suggested that an early and a late VIP had the same cognitive load, but there were differences between the navigation tasks: cane vs. SSD and between the different types of encodings (audio, haptic, and multimodal).

Besides the two particular cases presented above, an overview of the CL analysis is presented in Figures 22–25. By averaging the results for all the VIPs (early- and late-blind), in case of the most difficult trial (scenario E), it could be concluded that the alpha1 asymmetries for audio and multimodal codifications

were a little higher (~0.2) compared to the cane and haptic modality (~0.1). The variations within the whole asymmetries data set are displayed in the whisker plot from Figure 24.

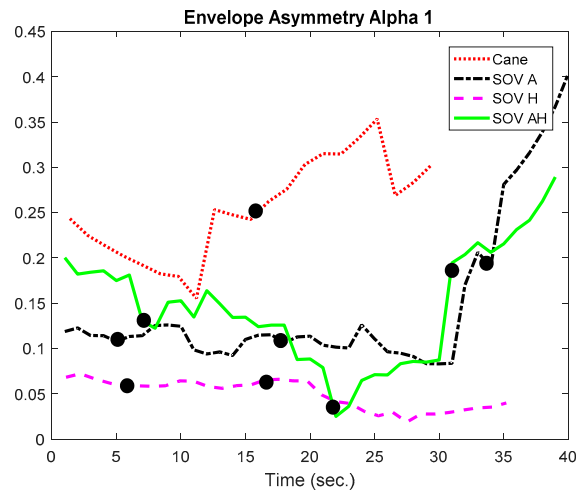


Figure 20. The envelopes of the alpha1 asymmetries for a late-blind user.

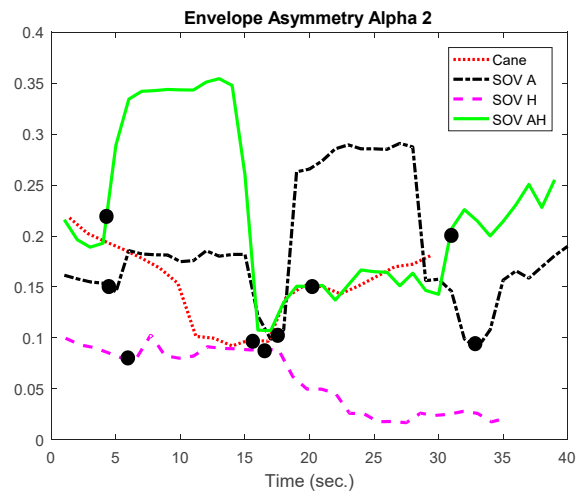


Figure 21. The envelopes of the alpha2 asymmetries for a late-blind user.

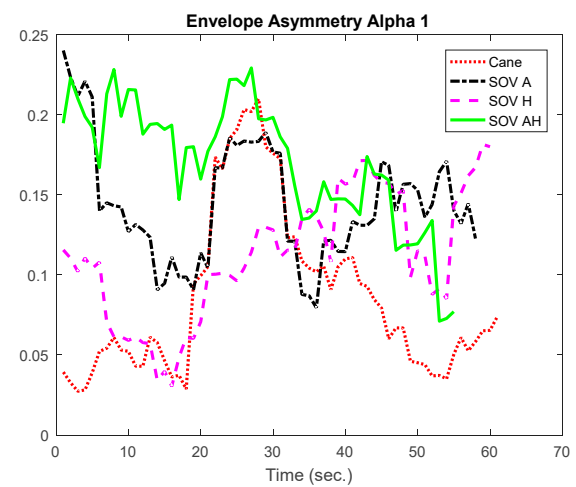


Figure 22. The average of the alpha1 asymmetry for all the users.

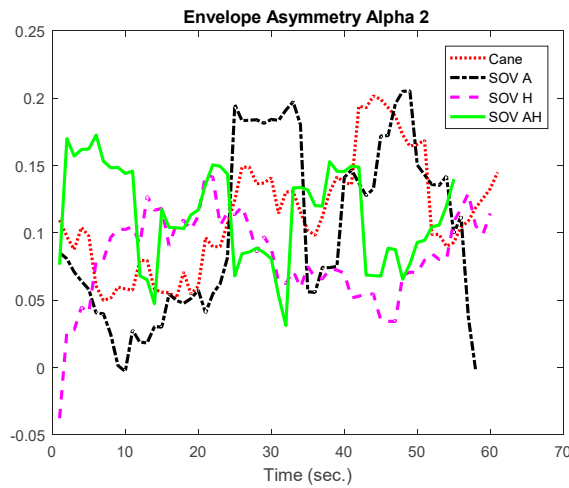


Figure 23. The average of the alpha2 asymmetry for all the users.

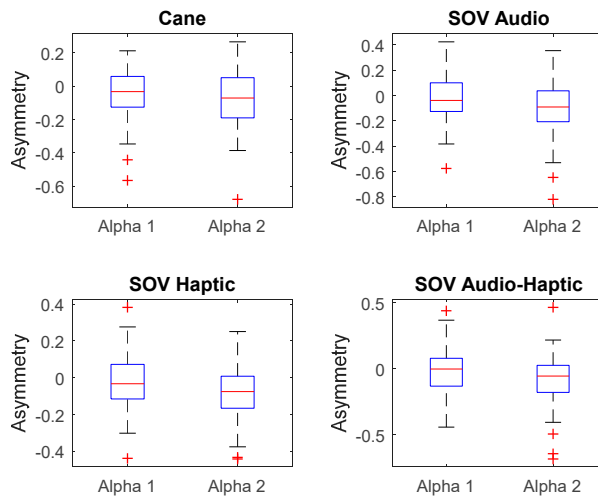


Figure 24. The envelopes for all the users, for scenario E.

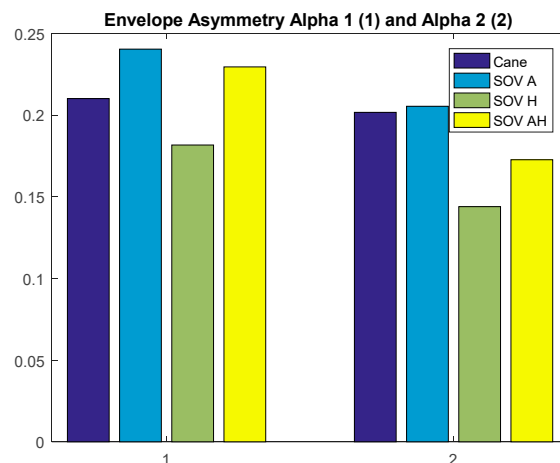


Figure 25. Average envelopes’ asymmetries for all users, for scenario E.

6.4. Visual Cortex Activation Analysis

Some previous studies have revealed the presence of visual cortex activity in the case of the VIPs if a sensory substitution system creates an “information map” of the environment. Therefore, the total cognitive load (TCL) for electrodes O1 and O2 was investigated. Preliminary investigations

of the experimental data showed that there was a major difference between the VC activity of the late-blind persons and of those who were born blind. Thus, a general conclusion regarding all VIPs could be drawn.

The asymmetries of TCL in the visual cortex for UserA (early-blind) and UserB (late-blind) were calculated. In Figures 26 and 27, the upper envelopes for TCL values are represented for UserA and UserB. *UserA was a born blind person, and, in his case, the asymmetry of the cane task was twice greater than the asymmetries of SoV tasks.* In line with some previous research, it is possible to associate the VC activity of the cane task with the fact that parts of his visual cortex were activated to refine his sensations and usual activities. In contrast, *the late-blind person's VC activity in SoV tasks was much higher than in the cane task (which was negative) and more than five times higher than UserA's visual activity.* The limited number of VIPs from each group (five early- and three late-blind users) did not permit to obtain valuable statistical results, but *for all the late-blind persons guided by audio and haptic stimuli, the average asymmetry of VC was around six times greater than that of the persons born blind, as can be seen in Figures 28 and 29.* Another important observation was that *VC asymmetry was lower in the case of multimodal stimulation than in the case of separate stimulation (audio or haptic).* It must be emphasized that there was *no correlation between visual cortical activity and the number of collisions during navigation, regardless of the type of navigation or sensory input.*

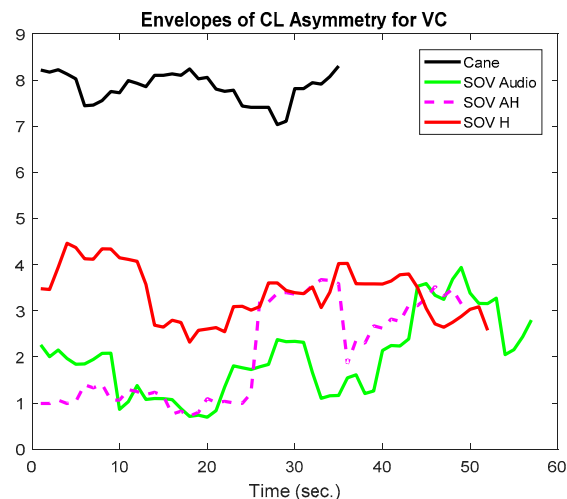


Figure 26. Visual cortex asymmetry for UserA (early-blind).

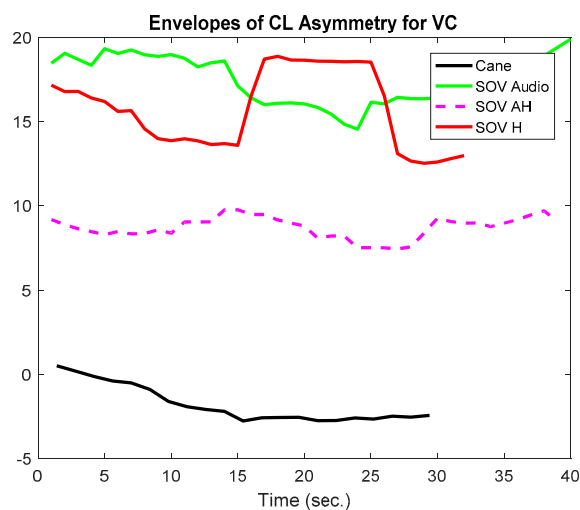


Figure 27. Visual cortex asymmetry for UserB (late-blind).

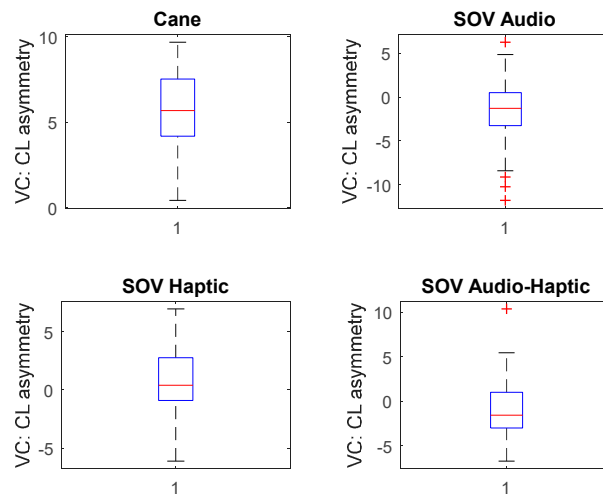


Figure 28. Visual cortex activity for UserA (early-blind).

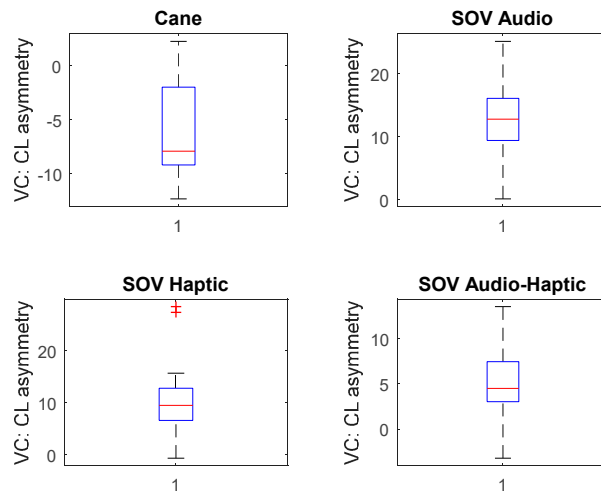


Figure 29. Visual cortex activity for UserB (late-blind).

6.5. Emotions Assessment During Real-World Navigation

As presented in the previous chapter, the emotional influence is a vast topic on which, from the dimensional perspective, the valence and arousal dimensions are advocated by Russell [66]. *Arousal* expresses calmness or excitement, whereas *valence* expresses a negative or positive effect. According to the comprehensive literature, *left frontal inactivation* is an indicator of a withdrawal response, which is often linked to a negative emotion, and *right frontal inactivation* is a sign of an approach response or positive emotion. Therefore, the **ratio of right and left asymmetry (valence state—VS)** was computed with the equation:

$$VS = \log\left(\frac{PS_R}{PS_L}\right) \quad (2)$$

where PS_R and PS_L are the power spectrum values of the right and the corresponding left hemisphere channels in a specific frequency band. The channels T7-T8, which are considered the most relevant for emotions assessment, but also C3, C4, P3, P4, F3, and F4, were considered based on the theoretical statements from the previous chapter [62].

In addition, the HR and EDA signals acquired during the tests were processed according to the standard approaches described in the literature [65]. The root mean square of the successive differences (RMSSD) values were calculated for the HR recorded using the Shimmer sensor. A low RMSSD value means a high HR, denoting a strong concentration, emotion, or physical effort, whereas a high RMSSD value corresponds to resting or to a relaxing activity. The HR values in the resting state for the involved

users were different according to their age and personal rhythm. Therefore, the percentage of variation of RMSSD compared to the resting state was calculated. By pre-processing the signals acquired with Shimmer, the EDA signals (μS) were obtained. Then, the deconvolution performed using the Ledalab software provided the phasic and tonic components and skin conductance responses (SCRs)—abrupt increases in the conductance of the skin, measured in μS , were calculated.

Firstly, the global VS of all the users and all the performed tests was calculated for the chosen pairs of electrodes. The result is graphically depicted in Figure 30. The C3 and C4 electrodes were considered because their waves could be associated with hippocampus activity, together with T8-T7, which obviously are the most relevant for assessing emotions [65]. The VS calculated using the T8-T7 pair indicated a *low positive valence for cane navigation and a negative valence for SoV navigation*, in accordance with the cumulative time, distance, and the number of collisions from Table 1. For the *parietal lobe*, the P4-P3 pair indicated a *moderate VS for cane navigation and low VS for SoV navigation*. The pair C4-C3 reflected a low positive valence, close to the neutral state. The same observation applied for the F4-F4 pair, for which the low negative valence in the SoV navigation using the audio encoding must be taken into account.

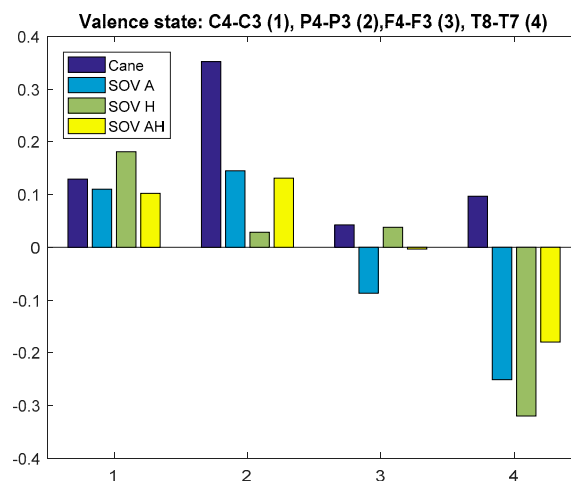


Figure 30. Average valence state (VS) for all visually impaired people (VIP).

The global percentage variation rate of RMSSD compared to the resting state decreased with: 21% for cane navigation, 39% for SoV navigation using the audio encoding, 44% for SoV navigation using the haptic encoding, and 41% for SoV navigation using the multimodal encoding. By computing the global SCR index, the following values were obtained: 0.18 for white cane navigation, 0.61 for SoV navigation using the audio encoding, 0.76 for SoV navigation using the haptic encoding, and 0.71 for SoV navigation using the multimodal encoding.

The trials were performed randomly within the same navigation type (cane or SoV) and for the same user, and no significant differences of the RMSSD values were remarked between the trials, even if they had different durations. Moreover, a slight increase (corresponding to an HR decrease) was observed towards the end of most of the tests. The users were not subjected to intense physical activity because they walked on the plain ground; however, the average of HR values was a little bit increased in comparison to the VTE tests. On the other hand, the EDA signals were sensitive to most of the collisions, especially in the case of the SoV navigation.

In Figures 31 and 32, we present the VS values for the two special users (UserA and UserB). In the case of UserA, who was intensively trained in the VTE and in the RW, the *valence was positive but very close to the neutral state for all the navigation types, with the remark that the valence for SoV navigation was a little bit higher than for cane navigation*. His RMSSD and SCR values did not differ significantly between the navigation modes.

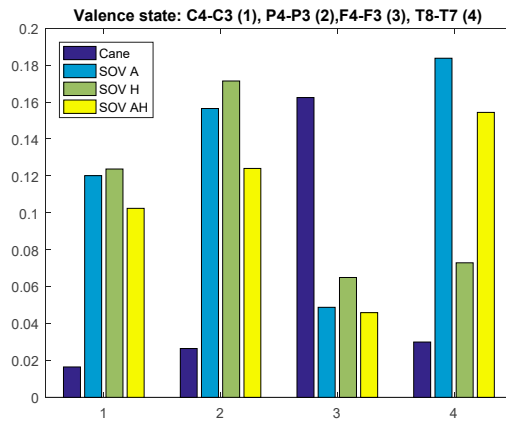


Figure 31. Average VS of UserA.

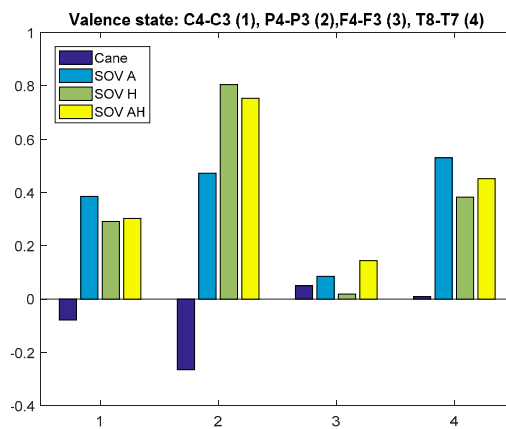


Figure 32. Average VS of UserB.

Analyzing the results for UserB, who is usually guided by a family person, it is obvious that for him, navigating using the white cane induced a neutral to low negative valence, and navigating using the SoV device gave him more security, satisfaction, and a more comfortable state. This was underlined by a decrease of the RMSSD percentages and an increase of SCRs for all navigation modes.

In Figures 33 and 34, the evolution in time of the VS for the T8-T7 pair is represented for UserA and UserB in order to highlight the slow evolution of valence during a trial. In general, the collisions did not essentially affect the valence changes, as in the case of the cognitive load assessment.

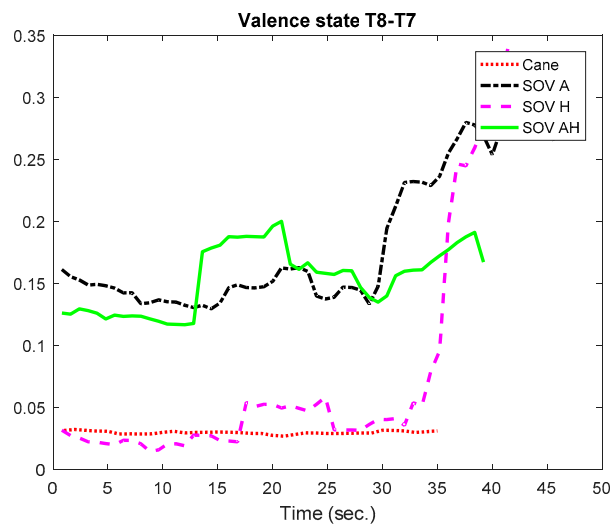


Figure 33. VS evolution in time for UserA.

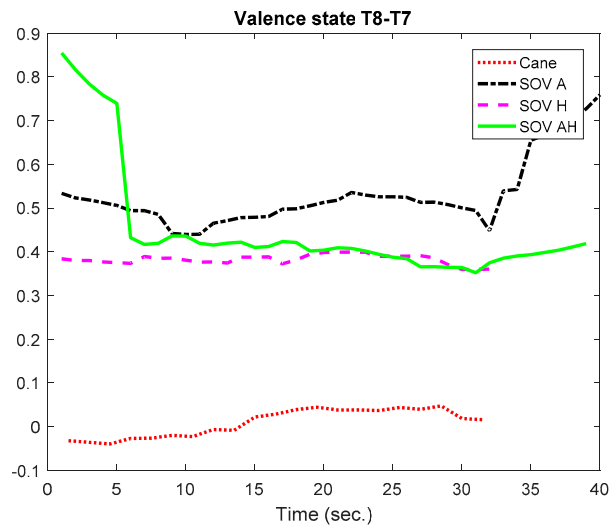


Figure 34. VS evolution in time for UserB.

6.6. Limitation of This Study

The limitations of the current study reside in the fact that we performed the tests with a small number of users for both categories (early-blind and late-blind). Although the results were interesting and in line with the existing literature, a more thorough evaluation should be realized. In addition, the EEG recordings were performed using a limited number of electrodes. As future directions, we plan to use a more advanced EEG recording device, with a higher number of electrodes, and to improve the SoV device so that it would be lighter and more comfortable to be worn.

The results of real-world experiments were strongly influenced by a consistent training period, similar to all users, which requires a great deal of time. Future work can extend the realistic scenarios of RW traveling for enhancing the impact of the study.

7. Conclusions

This paper presented an experimental framework and a study based on EEG, HR, and GSR signal analysis, aiming to assess the brain cortex activation and affective reactions of the visually impaired persons to the stimuli provided by a sensory substitution device used for navigation in real-world scenarios, compared to the white cane navigation. The study was focused on the evaluation of working memory load, visual cortex activation, and emotional experience when the VIPs perceived audio, haptic, and multimodal stimuli during a navigation task in five different types of scenarios.

The choice of the Brain-Computer Interface (BCI) equipment proved to be inspired because its characteristics allowed a good acquisition of EEG signals simultaneously with the use of the SoV device. The same BCI equipment has been employed successfully in other studies of our own concerning multimodal neuromotor rehabilitation [67,68]. An important feature of the experimental setup is the ability to synchronize the data streams and to align the acquired signals with the events extracted from the video recordings. The training performed in the VTE and the ego-static tests performed indoors had an essential role in preparing the users to perceive distances, positions, and object dimensions only by means of the audio, haptic, and multimodal stimuli, giving confidence to all of them in using the SoV device. The aim was to provide all users the ability to automatically understand the complexity of a scene. Besides, during the VTE training, multiple resting sessions were recorded for all volunteers, which had an important role in establishing a baseline.

The perception of audio and haptic stimuli using the SoV device was assessed in terms of cognitive load, pleasantness, excitement, and events, for all the visually impaired users, as well as for the specific categories (early-blind or late-blind). *All in all, the haptic stimuli appeared to be less intuitive than the audio stimuli.*

The analysis showed that *navigating with the SoV device increased the cognitive load and the working memory (lower accuracy and longer response times)*. The analysis of the EEG data revealed *the usage of verbal working memory in the posterior parietal cortices*. The obtained results indicated that the left-right asymmetry of the prefrontal cortex had distinguishable characteristics when the VIPs were navigating in real-world *environments with a wide range of obstacles*.

The visual cortex exploration revealed *a significant activation when using the SoV device, only for the late VIPs*. The low VC activity of congenitally blind persons during SoV navigation could be related to brain plasticity, which allows the auditory and somatosensory systems to extend their functionality in that part of the cortex.

Finally, we assessed the valence state of the users when navigating in unfamiliar indoor environments based on mobile monitoring and a fusion of EEG and physiological (EDA and HR) signals. *For the generic VIP population, the use of the SoV device induced a low negative valence in contrast with cane usage*. But the findings differed for the specific categories of sight loss (early- and late-blind), pointing out the particular needs/difficulties faced by each category of VIP.

This study proved once more that sensory substitution is an alternative method, which helps the blind people to acquire information about the surrounding space and to navigate independently in unknown real-world environments, safely and comfortably, after substantial training.

The findings hopefully empower the knowledge of how the visually impaired persons are stressed and emotionally affected by SSD navigation and contribute to the development of the intelligent navigation devices, aiming for the VIPs' safety and well-being. The results of our work can inspire researchers working in the field of IoT devices comprising sensors, antennas, and Bluetooth, which have created navigation rules based on a fuzzy controller [69], GPS embedded in a stick with voice recognition for obstacles detection [70], computer vision-based assistants [71], or assistive systems relying on wearable smart glasses and mobile applications [72].

Valuable research projects have investigated the efficiency of intelligent sensory substitution devices [73], and, in this context, our research brought an important contribution by analyzing EEG and physiological signals in order to assess the cognitive effort and emotional state of users in real-world navigation.

Author Contributions: Conceptualization, F.U. and R.-G.L.; methodology, A.S., F.U., K.K., O.M. and A.M.; software, R.-G.L. and A.S.; validation, O.M., F.U., K.K. and A.M.; formal analysis, F.U.; investigation, O.M., F.U. and K.K.; resources, R.-G.L., A.S. and O.M.; data curation, F.U.; writing—original draft preparation, R.-G.L., F.U., O.M. and A.M.; writing—review and editing, A.M.; visualization, F.U., O.M., R.-G.L., A.S. and K.K.; supervision, A.M.; project administration, A.M.; funding acquisition, A.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 643636 "Sound of Vision". Financial support from the Competitiveness Operational Program 2014–2020, Action 1.1.3: Creating synergies with RDI actions of the EU's HORIZON 2020 framework program and other international RDI programs, MySMIS Code 108792, Acronym project "UPB4H", financed by contract: 250/11.05.2020.

Acknowledgments: The authors wish to thank the volunteer visually impaired subjects for their collaboration and involvement. This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 643636 "Sound of Vision". Financial support from the Competitiveness Operational Program 2014–2020, Action 1.1.3: Creating synergies with RDI actions of the EU's HORIZON 2020 framework program and other international RDI programs, MySMIS Code 108792, Acronym project "UPB4H", financed by contract: 250/11.05.2020 is gratefully acknowledged.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. World Health Organization—Blindness and Vision Impairment. Available online: <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment> (accessed on 10 September 2020).
2. Sound of Vision. Available online: <https://soundofvision.net/> (accessed on 10 September 2020).

3. Väljamäe, A.; Kleiner, M. Spatial sound in auditory vision substitution systems. *Audio Eng. Soc. Conv.* **2006**, *120*, 6795.
4. Fournier d'Albe, E.E. The Optophone: An instrument for reading by ear. *Nature* **1920**, *105*, 295–296.
5. Meijer, P.B.L. The Voice Technology. Available online: <http://www.seeingwithsound.com> (accessed on 10 September 2020).
6. Fish, R.M. Audio Display for Blind. *IEEE Trans. Biomed. Eng.* **1976**, *23*, 144–154. [[CrossRef](#)]
7. Sanz, P.R.; Mezcuca, B.R.; Sánchez Pena, J.M. A sonification proposal for safe travels of blind people. In Proceedings of the 18th International Conference on Auditory Display, Atlanta, GA, USA, 18–21 June 2012.
8. Milios, E.; Kapralos, B.; Kopinska, A.; Stergiopoulos, S. Sonification of range information for 3-D space perception. *IEEE Trans. Neural Syst. Rehab. Eng.* **2003**, *11*, 416–421. [[CrossRef](#)]
9. Bestpluton World Cie. The “Mini-Radar”, Your Small Precious Companion that Warns You Obstacles in a Spoken Way, and that Helps You to Walk Straight. Available online: <http://aweb1208.free.fr/EnglishMiniRadar.htm> (accessed on 10 September 2020).
10. Pun, T.; Roth, P.; Bologna, G.; Moustakas, K.; Tzovaras, D. Image and video processing for visually handicapped people. *J. Image Video Process.* **2007**, *5*, 1–12.
11. Kaczmarek, K.A.; Bach-y-Rita, P. Tactile displays. In *Virtual Environments and Advanced Interface Design*; Barfield, W., Furness, T.A., Eds.; Oxford University Press Inc.: Oxford, UK, 1995.
12. Wall, S.A.; Brewster, S. Sensory substitution using tactile pin arrays: Human factors, technology, and applications. *Signal Process.* **2006**, *86*, 3674–3695. [[CrossRef](#)]
13. Loomis, J.M.; Lederman, S.J. Tactual perception. In *Handbook of Perception and Human Performance: Cognitive Processes and Performance*; Boff, K.R., Kaufman, L., Thomas, J.P., Eds.; John Wiley & Sons: New York, NY, USA, 1986.
14. Cancar, L.; Díaz, A.; Barrientos, A.; Travieso, D.; Jacobs, D.M. Tactile-sight: A sensory substitution device based on distance-related vibrotactile flow. *Int. J. Adv. Robot. Syst.* **2013**, *10*, 272. [[CrossRef](#)]
15. Parkes, D. Nomad: An audio-tactile tool for the acquisition, use and management of spatially distributed information by visually impaired people. In Proceedings of the 2nd International Symposium on Maps and Graphics for Visually Impaired People, London, UK, 20–22 April 1988; pp. 24–29.
16. Dewhurst, D. Creating and accessing audio-tactile images with HFVE vision substitution software. In *Proceedings of the Third Interactive Sonification Workshop*; KTH: Stockholm, Sweden, 2010; pp. 101–104.
17. Dewhurst, D. Using imprints to summarise accessible images. In Proceedings of the Interactive Sonification Workshop Ison, 4th Interactive Sonification Workshop, Fraunhofer IIS, Erlangen, Germany, 10 December 2013.
18. Hosseini, S.A.; Naghibi-Sistani, M.B. Classification of emotional stress using brain activity. In *Applied Biomedical Engineering*; Gargiulo, G.D., McEwan, A., Eds.; InTech: Rijeka, Croatia, 2011; pp. 313–336.
19. Jena, S.K. Examination stress and its effect on EEG. *Int. J. Med. Sci. Public Health* **2015**, *11*, 1493–1497. [[CrossRef](#)]
20. Bos, D.O. *EEG-Based Emotion Recognition: The Influence of Visual and Auditory Stimuli*; Department of Computer Science, University of Twente: Enschede, The Netherlands, 2006.
21. Antonenko, P.; Pass, F.; Grabner, R.; van Gog, T. Using Electroencephalography to Measure Cognitive Load. *Educ. Psychol. Rev.* **2010**, *22*, 425–438. [[CrossRef](#)]
22. Berka, C.; Levendowski, D.J.; Lumicao, M.N.; Yau, A.; Davis, G.; Zivkovic, V.T.; Olmstead, R.E.; Tremoulet, P.D.; Craven, P.L. EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviat. Space Environ. Med.* **2007**, *78*, B231–B244.
23. Nilsson, R.M.; Mayer, R.E. The effects of graphics organizers giving cues to the structure of a hypertext document on users' navigation strategies and performance. *Int. J. Hum. Comput. Stud.* **2002**, *57*, 1–26. [[CrossRef](#)]
24. Scott, B.M.; Schwartz, N.H. Navigational spatial displays: The role of metacognition as cognitive load. *Learn. Instr.* **2007**, *17*, 89–105. [[CrossRef](#)]
25. Pekrun, R. The impact of emotions on learning and achievement: Towards a theory of cognitive/motivational mediators. *Appl. Psychol.* **1992**, *41*, 359–376. [[CrossRef](#)]
26. Soleymani, M.; Pantic, M.; Pun, T. Multimodal emotion recognition in response to videos (Extended abstract). In Proceedings of the International Conference on Affective Computing and Intelligent Interaction (ACII), Xi'an, China, 21–24 September 2015; pp. 491–497.

27. Bower, G.H.; Forgas, J.P. *Mood and Social Memory*; Lawrence Erlbaum Associates Publishers: Mahwah, NJ, USA, 2001.
28. Cattaneo, Z.; Vecchi, T.; Cornoldi, C.; Mammarella, I.; Bonino, D.; Ricciardi, E.; Pietrini, P. Imagery and spatial processes in blindness and visual impairment. *Neurosci. Biobehav. Rev.* **2008**, *32*, 1346–1360. [[CrossRef](#)]
29. Healey, J.A.; Picard, R.W. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Trans. Intell. Transp. Syst.* **2005**, *6*, 156–166. [[CrossRef](#)]
30. Setz, C.; Arnrich, B.; Schumm, J.; La Marca, R.; Tröster, G. Discriminating stress from cognitive load using a wearable EDA device. *IEEE Trans. Inf. Technol. Biomed.* **2010**, *14*, 410–417. [[CrossRef](#)]
31. Peper, E.; Harvey, R.; Lin, I.M.; Tylova, H.; Moss, D. Is There More to Blood Volume Pulse Than Heart Rate Variability, Respiratory Sinus Arrhythmia, and Cardiorespiratory Synchrony? *Biofeedback* **2007**, *35*, 54–61.
32. Peake, P.; Leonard, J.A. The use of heart rate as an index of stress in blind pedestrians. *Ergonomics* **1971**, *14*, 189–204. [[CrossRef](#)]
33. Wycherley, R.J.; Nicklin, B.H. The heart rate of blind and sighted pedestrians on a town route. *Ergonomics* **1970**, *13*, 181–192. [[CrossRef](#)]
34. Urgen, B.A.; Plank, M.; Ishiguro, H.; Poizner, H.; Saygin, A. EEG theta and Mu oscillations during perception of human and robot actions. *Front. Neurobotics* **2013**, *7*, 19. [[CrossRef](#)]
35. Gallistel, C.R. *The Organization of Learning*; MIT Press: Cambridge, MA, USA, 1993.
36. Piquart, M.; Pfeiffer, J.P. Psychological well-being in visually impaired and unimpaired individuals: A meta-analysis. *Br. J. Vis. Impair.* **2011**, *29*, 27–45. [[CrossRef](#)]
37. Kalimeri, K.; Saitis, C. Exploring Multimodal Biosignal Features for Stress Detection during Indoor Mobility. In Proceedings of the 18th ACM International Conference on Multimodal Interaction, Tokyo, Japan, 12–16 November 2016.
38. Saitis, C.; Parvez, M.Z.; Kalimeri, K. Cognitive Load Assessment from EEG and Peripheral Biosignals for the Design of Visually Impaired Mobility Aids. *Wirel. Commun. Mob. Comput.* **2018**, *2018*. [[CrossRef](#)]
39. Saitis, C.; Kalimeri, K. Identifying urban mobility challenges for the visually impaired with mobile monitoring of multimodal biosignals. In Proceedings of the International Conference on Universal Access in Human-Computer Interaction, Toronto, ON, Canada, 17–22 July 2016.
40. Amedi, A.; Stern, W.; Camprodon, J.; Bermpohl, F.; Merabet, L.; Rotman, S.; Hemond, C.; Meijer, P.; Pascual-Leone, A. Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nat. Neurosci.* **2007**, *10*, 687–689. [[CrossRef](#)]
41. Naue, N.; Rach, S.; Struber, D.; Huster, R.J.; Zaehle, T.; Korner, U.; Herrmann, C.S. Auditory Event-Related Response in Visual Cortex Modulates Subsequent Visual Responses in Humans. *J. Neurosci.* **2011**, *31*, 7729–7736. [[CrossRef](#)] [[PubMed](#)]
42. Sound of Vision Outputs. Available online: <https://soundofvision.net/outputs/> (accessed on 10 September 2020).
43. Moldoveanu, A.D.B.; Ivascu, S.; Stanica, I.; Dascalu, M.-I.; Lupu, R.; Ivanica, G.; Morar, A. Mastering an advanced sensory substitution device for visually impaired through innovative virtual training. In Proceedings of the IEEE 7th International Conference on Consumer Electronics-Berlin (ICCE-Berlin), Berlin, Germany, 3–6 September 2017; pp. 120–125. [[CrossRef](#)]
44. Dascalu, M.; Moldoveanu, A.; Balan, O.; Lupu, R.G.; Ungureanu, F.; Caraiman, S. Usability assessment of assistive technology for blind and visually impaired. In Proceedings of the 2017 E-Health and Bioengineering Conference (EHB), Sinaia, Romania, 22–24 June 2017; pp. 523–526. [[CrossRef](#)]
45. Botezatu, N.; Caraiman, S.; Rzeszotarski, D.; Strumillo, P. Development of a versatile assistive system for the visually impaired based on sensor fusion. In Proceedings of the 2017 21st International Conference on System Theory, Control and Computing (ICSTCC), Sinaia, Romania, 19–21 October 2017; pp. 540–547. [[CrossRef](#)]
46. Kristjánsson, A.; Moldoveanu, A.; Jóhannesson, O.I.; Balan, O.; Spagnol, S.; Valgeirsdóttir, V.V.; Unnthorsson, R. Designing sensory-substitution devices: Principles, pitfalls and potential. *Restor. Neurol. Neurosci.* **2016**, *34*, 769–787. [[CrossRef](#)] [[PubMed](#)]
47. Moldoveanu, A.; Taloi, B.; Balan, O.; Stanica, I.; Flamaropol, D.; Dascalu, M.I.; Moldoveanu, F.; Mocanu, I. Virtual Mini-Games—A Serious Learning Tool for Sensory Substitution Devices. In Proceedings of the EduLearn Conference, Barcelona, Spain, 3–5 July 2017.

48. Spagnol, S.; Wersényi, G.; Bujacz, M.; Bălan, O.; Martínez, M.H.; Moldoveanu, A.; Unnthorsson, R. Current Use and Future Perspectives of Spatial Audio Technologies in Electronic Travel Aids. *Wirel. Commun. Mob. Comput.* **2018**, *2018*. [[CrossRef](#)]
49. Caraiman, S.; Morar, A.; Owczarek, M.; Burlacu, A.; Rzeszotarski, D.; Botezatu, N.; Moldoveanu, A. Computer Vision for the Visually Impaired: The Sound of Vision System. In Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVW), Venice, Italy, 22–29 October 2017; pp. 1480–1489. [[CrossRef](#)]
50. Lab Streaming Layer (LSL), Multi-Modal Time-Synched Data Transmission over Local Network. Available online: <https://github.com/scn/labstreaminglayer> (accessed on 10 September 2020).
51. HDF5 Data Model, Library, and File Format for Storing and Managing Data. Available online: <https://support.hdfgroup.org/HDF5/> (accessed on 10 September 2020).
52. ChronoViz, Visualization of Time-Based Multimodal Data. Available online: <https://chronoviz.com/> (accessed on 10 September 2020).
53. ExifTool. Available online: <https://exiftool.org/> (accessed on 10 September 2020).
54. iMotions. Available online: <https://imotions.com/> (accessed on 10 September 2020).
55. Hawthorne, B.; Vella-Brodrick, D.; Hattie, J. Well-Being as a Cognitive Load Reducing Agent: A Review of the Literature. *Front. Educ.* **2019**, *4*, 121. [[CrossRef](#)]
56. Alarcao, S.M.; Fonseca, M.J. Emotions Recognition Using EEG Signals: A Survey. *IEEE Trans. Affect. Comput.* **2019**, *10*, 374–393. [[CrossRef](#)]
57. Sadato, N.; Pascual-Leone, A.; Grafman, J.; Ibañez, V.; Deiber, M.P.; Dold, G.; Hallett, M. Activation of the primary visual cortex by Braille reading in blind subjects. *Nature* **1996**, *380*, 526–528. [[CrossRef](#)]
58. Burton, H. Visual cortex activity in early and late blind people. *J. Neurosci.* **2003**, *23*, 4005–4011. [[CrossRef](#)]
59. Georgetown University Medical Center. People Blind from Birth Use Visual Brain Area to Improve Other Senses: Can Hear and Feel with Greater Acuity. *ScienceDaily*. 10 October 2010. Available online: www.sciencedaily.com/releases/2010/10/101006131203.htm (accessed on 1 September 2020).
60. Campus, C.; Sandini, G.; Amadeo, M.B.; Gori, M. Stronger responses in the visual cortex of sighted compared to blind individuals during auditory space representation. *Sci. Rep.* **2019**, *9*, 1935. [[CrossRef](#)] [[PubMed](#)]
61. Bălan, O.; Moise, G.; Petrescu, L.; Moldoveanu, A.; Leordeanu, M.; Moldoveanu, F. Emotion Classification Based on Biophysical Signals and Machine Learning Techniques. *Symmetry* **2020**, *12*, 21. [[CrossRef](#)]
62. Zhang, J.; Chen, M.; Zhao, S.; Hu, S.; Shi, Z.; Cao, Y. Relief F-Based EEG Sensor Selection Methods for Emotion Recognition. *Sensors* **2016**, *16*, 1558. [[CrossRef](#)] [[PubMed](#)]
63. iMotions, Frontal Asymmetry 101—How to Get Insights on Motivation and Emotions from EEG. Available online: <https://imotions.com/blog/frontal-asymmetry-101-get-insights-motivation-emotions-eeeg/> (accessed on 10 September 2020).
64. Ray, W.J.; Cole, H.W. EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes. *Science* **1985**, *228*, 750–752. [[CrossRef](#)] [[PubMed](#)]
65. Saitis, C.; Kalimeri, K. Multimodal Classification of Stressful Environments in Visually Impaired Mobility Using EEG and Peripheral Biosignals. *IEEE Trans. Affect. Comput.* **2018**. [[CrossRef](#)]
66. Russell, J.A. Core affect and the psychological construction of emotion. *Psychol. Rev.* **2003**, *110*, 145–172. [[CrossRef](#)]
67. Ferche, O.; Moldoveanu, A.; Cintează, D.; Toader, C.; Moldoveanu, F.; Voinea, A.; Taslitchi, C. From neuromotor command to feedback: A survey of techniques for rehabilitation through altered perception. In Proceedings of the E-Health and Bioengineering Conference (EHB), Iasi, Romania, 19–21 November 2015; pp. 1–4.
68. Moldoveanu, A.; Ferche, O.; Moldoveanu, F.; Lupu, R.; Cintează, D.; Irimia, D.C.; Toader, C. The TRAVEE System for a Multimodal Neuromotor Rehabilitation. *IEEE Access* **2015**, *7*, 8151–8171. [[CrossRef](#)]
69. Tayyaba, S.; Ashraf, M.W.; Alquthami, T.; Ahmad, Z.; Manzoor, S. Fuzzy-Based Approach Using IoT Devices for Smart Home to Assist Blind People for Navigation. *Sensors* **2020**, *20*, 3674. [[CrossRef](#)]
70. Mala, N.S.; Thushara, S.; Subbiah, S. Navigation gadget for visually impaired based on IoT. In Proceedings of the 2nd International Conference on Computing and Communications Technologies (ICCT), Chennai, India, 23–24 February 2017; pp. 334–338.
71. Islam, M.M.; Sadi, M.S.; Zamli, K.Z.; Ahmed, M.M. Developing Walking Assistants for Visually Impaired People: A Review. *IEEE Sens. J.* **2019**, *19*, 2814–2828. [[CrossRef](#)]

72. Chang, W.; Chen, L.; Chen, M.; Su, J.; Sie, C.; Yang, C. Design and Implementation of an Intelligent Assistive System for Visually Impaired People for Aerial Obstacle Avoidance and Fall Detection. *IEEE Sens. J.* **2020**, *20*, 10199–10210. [[CrossRef](#)]
73. Jicol, C.; Lloyd-Esenkaya, T.; Proulx, M.J.; Lange-Smith, S.; Scheller, M.; O'Neill, E.; Petrini, K. Efficiency of Sensory Substitution Devices Alone and in Combination with Self-Motion for Spatial Navigation in Sighted and Visually Impaired. *Front. Psychol.* **2020**, *11*, 1443. [[CrossRef](#)] [[PubMed](#)]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).