VErGE: A System for Collecting Voice, Eye Gaze, Gesture, and EEG Data for Experimental Studies

Ovidiu Gherman, Ovidiu Schipor, Bogdan-Florin Gheran
MintViz Lab / MANSiD Research Center
"Ştefan cel Mare" University of Suceava
Suceava, Romania
ovidiug@eed.usv.ro, schipor@eed.usv.ro, bogdan.gheran@gmail.com

Abstract—**Collecting multimodal data in controlled experiments is often challenging, especially when there is the need to record fine observations and experimental data over multiple input channels. Self-reporting and direct observation, including video recording, are often employed with good results, but in many cases user behavior may be elusive or difficult to quantify and record. This problem is particularly even more difficult to address in the case of multimodal systems, where input is elicited from users with various disabilities such as motor impairments, due to many factors like increased calibration time, constraints on body postures and limited range of movements, and interferences occurring at a physiological level, such as fatigue or involuntary spasms. In this work, we describe VErGE, a multimodal input system designed to collect voice, gestures, eye gaze and EEG (electroencephalography) data in experimental settings. We detail the architecture of our system and discuss a representative use case for collecting multimodal input from users with motor impairments.**

Keywords—touch gestures; voice; EEG; eye gaze; eye tracker; motor impairments; experiments; experimental data

I. INTRODUCTION

Behavioral experiments and user studies rely on the acquisition of quantifiable, well defined data. Besides recording user behavior, the timing between different types of behaviors must be also considered because of its relevance. Relevant input modalities frequently considered for experimental studies on user input are fine motor movements, such as gesture input [21] [22] [23], eye gaze (relevant to identify the focus of attention in subjects) [1] [11] [12] [28] [29], head/body movement [26] [27] and voice input [24] [25] [30] [31]. Multimodal data requires multiple sensors for acquisition, which demands specific calibration and synchronization.

In the last decade, new data acquisition devices have been developed enabling researchers and practitioners to collect a wide range of measurements, e.g., pulse, oxygenation, body heat, etc. Among these, a promising class of devices is represented by Brain-Computer Interfaces (BCI) [2]. The advent of affordable electroencephalography (EEG) headsets enables easy collection and integration of EEG data that can be synchronized with other input type, such as voice [24] [25], gesture [21] [22] [23] or eye gaze tracking [1] [11] [28] [29]. The resulted multimodal dataset can be aggregated in a reference database for researchers interested in understanding specific aspects of user behavior. Currently, there are some systems that allow recording of various parameters from subjects, such as BIOMET, MyIDea or Smartkom [3] [4] [5], mostly related to biometric operation, but none have integrated yet the above-mentioned modalities or offers a unified method of data acquisition.

In this paper, we describe VErGE, a system for collecting voice, gestures, eye gaze, and EEG (electroencephalography) data from multiple input devices. Our technical contributions are as follows:

- A multimodal input data collection system called VErGE that enables easy recording and storage of voice, eye gaze, touch gesture, and EEG data from users in various experimental settings. We present the modular software architecture of VErGE and its specific implementation in terms of input devices and software middleware.
- A discussion of the VErGE system for a specific context of use involving participants with motor impairments, for which we highlight specific features of our system, such as portability (to accommodate data collection in challenging contexts of use) and easiness of use.

VErGE records and stores multimodal measurements in the XML (Extensible Markup Language) format, ready for further processing and analysis. VErGE enables easy data connection, is flexible with respect to the configuration of input modalities, easy to configure and its design is based on four of the most frequent input modalities available today. Moreover, assisted by a touchscreen-capable laptop, VErGE was specifically designed to be portable, easy to deploy in the field and, consequently, suitable to collect multimodal data from users with motor impairments, as we illustrate in the next section of the paper with a specific context of use. VErGE was field-tested with several subjects with various types and severities of motor impairments to ascertain its effectiveness to acquire multimodal data under a variety of experimental conditions.

This work was supported by a grant of the Romanian National Authority for Scientific Research and Innovation, CNCS/CCCDI – UEFISCDI.

978-1-5386-1495-2/18/$31.00 ©2018 IEEE

II. THE EXPERIMENTAL SETUP

The experimental setup of the VErGE system is built around a Windows-based portable computer (Dell Inspiron, i7@2.7GHz, 16GB RAM) with touchscreen capabilities. All input devices are connected to the same computer and are time-synchronized. The latter is an important step in preserving the system's capability to offer relevant insights in user behavior. For a high quality of the recorded data, the laptop is further synchronized with an external source (i.e., a high-precision online clock) so that the timestamps can be compatible with the external devices that are not integrated, such as a professional video camera that takes footage of the entire experiment.

A. The Architecture of the System

The proposed architecture for the VErGE system is described in Fig. 1. The relevant data is acquired by the PC from its peripheral USB ports so that the system can be easily extended with new types of sensors to collect and process the recorded data. The file format used (XML) makes this approach easy to implement.

B. Software Use and Data Packaging

The VErGE system acquires relevant data from peripheral sensors and pre-processes it by adding timestamps, discarding unneeded values and filtering erroneous information. The result is then written on the PCs storage device (a SSD disk for increased data transfer performance) as XML files (Fig. 2, Fig. 3 and Fig. 4). The system generates one file for each subject, input channel and experimental trial. This structure makes it easy to understand and subsequently analyze the recorded data via dedicated software, such as graphics and/or statistical analysis packages.

C. Gesture Data Collection

Touch gesture data collection is achieved via the integrated touchscreen of the laptop, which can operate in various configuration modes, such as a tablet. Gestures are represented as a series of strokes, where each stroke is a time series of (x,y) points with associated timestamp information. Thus, both single-stroke [32] [33] and multi-stroke gestures [22] [34] can be acquired.

An important limitation in this setup is the inability of the recording equipment to detect the pressure level of the gesture drawn on the screen. This parameter would provide even more insight in the input characteristics for a given subject (especially for one with limited or abnormal motor capabilities).

D. Audio and Video Recording

The system features audio recording capability. Each subject's voice utterances can thus be recorded in conjunction with each gesture performed on the touch-sensitive screen with synchronized timestamps. The subsequent resulting files can be processed in terms of recognition, such as with the APIs provided by the Google Voice Recognition Cloud [35]. For more advanced analyses (e.g. the analysis of inflection or the strain on the user), a trained professional can listen to each recording and rate it with appropriate qualifiers.

The VErGE setup is accompanied by a video recording device (a professional, high resolution camera) that it is not directly integrated in the PC-based equipment. The camera is also synchronized at the beginning of the experiment and will record the relevant behavior for the entirety of the experiment, the video file being transferred at the end to the data storage and named appropriately to make easy identification of the subject id and the trials performed. This allows a post-facto

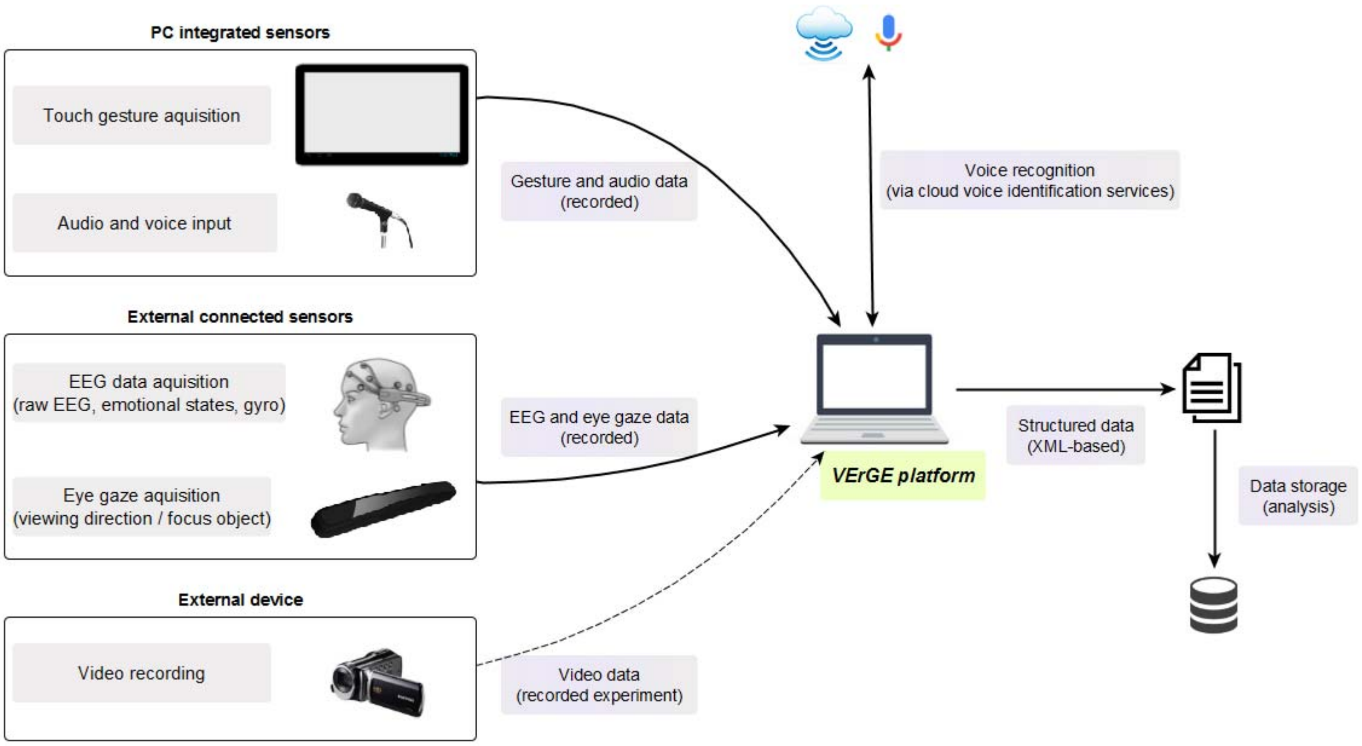

Fig. 1 - The architecture of the VErGE system for voice input, gesture, eye gaze, and EEG data collection.

```xml
<EEGEmoDataFrame timestamp="63646856672768"
        longTermExcitementScore="0.00000" shortTermExcitementScore="0.00000" meditationScore="0.00000"
        frustrationScore="0.00000" boredomScore="0.00000" valenceScore="0.00000" />
<EEGRawDataFrame timestamp="63646856672750"
        raw_cq="835.00000" af3="4156.92308" f7="3797.43590" f3="4139.48718" fc5="4141.02564"
        t7="4075.89744" p7="4161.53846" o1="4140.00000" o2="4152.30769" p8="4171.79487" t8="4085.64103"
        fc6="4090.76923" f4="4155.89744" f8="4076.41026" af4="4183.07692"
        gyrox="1937.00000" gyroy="2045.00000" />
<EEGRawDataFrame timestamp="63646856672758"
        raw_cq="803.00000" af3="4152.30769" f7="3804.10256" f3="4133.33333" fc5="4136.92308"
        t7="4077.94872" p7="4165.12821" o1="4141.53846" o2="4152.82051" p8="4169.23077" t8="4078.97436"
        fc6="4076.41026" f4="4148.20513" f8="4055.89744" af4="4168.71795"
        gyrox="1927.00000" gyroy="2046.00000" />
<EEGRawDataFrame timestamp="63646856672766"
        raw_cq="846.00000" af3="4150.25641" f7="3809.23077" f3="4133.84615" fc5="4137.94872"
        t7="4080.51282" p7="4170.76923" o1="4137.94872" o2="4145.64103" p8="4156.41026" t8="4083.07692"
        fc6="4078.46154" f4="4140.51282" f8="4063.58974" af4="4163.07692"
        gyrox="1921.00000" gyroy="2046.00000" />
...
<EEGEmoDataFrame timestamp="63646856672875"
        longTermExcitementScore="0.00000" shortTermExcitementScore="0.00000" meditationScore="0.00000"
        frustrationScore="0.00000" boredomScore="0.00000" valenceScore="0.00000" />
...
```


Fig. 2 - Data example output from the VErGE application for EEG headset (XML format).


```xml
<EyeTrackerDataFrame timestamp="63646856672661"
        left_gaze_x="220.37840" left_gaze_y="758.78240" right_gaze_x="220.37840" right_gaze_y="758.78240"
        diam_left_eye="0.00000" diam_right_eye="0.00000"
        left_eyePosition_X="3.96200" left_eyePosition_Y="-13.44200" left_eyePosition_Z="568.95400"
        right_eyePosition_X="63.92400" right_eyePosition_Y="-15.61000" right_eyePosition_Z="567.91600" />
<EyeTrackerFixationDataFrame timestamp="63646856672664"
        eye="l"
        time_start="84870490176" time_end="0"
        duration="0"
        Pos_X="286.65000" Pos_Y="285.40000" />
<EyeTrackerFixationDataFrame timestamp="63646856672664"
        eye="r"
        time_start="84870490176" time_end="0"
        duration="0"
        Pos_X="286.65000" Pos_Y="285.40000" />
<EyeTrackerDataFrame timestamp="63646856672691"
        left_gaze_x="220.37840" left_gaze_y="758.24240" right_gaze_x="220.37840" right_gaze_y="758.24240"
        diam_left_eye="0.00000" diam_right_eye="0.00000"
        left_eyePosition_X="4.19800" left_eyePosition_Y="-13.31500" left_eyePosition_Z="569.62900"
        right_eyePosition_X="64.16100" right_eyePosition_Y="-15.58100" right_eyePosition_Z="568.06000" />
...
```


Fig. 3 - Data example output from the VErGE application for eye gaze (XML format).

analysis of the experiment regarding the correctness of the acquired data, but also for offering additional insights regarding the subject's involvement during the experiment, such as their movements, gestures, gait, and general behavior.

E. Eye Gaze Tracking

Eye trackers can be either attached to a display (fixed) or mobile devices worn by users in the form of helmets or glasses (wearables), sometimes included in head-mounted displays. Most commonly, eye trackers include two components: a light source and a camera that captures the corneal reflection of the light source and some other visible ocular features. Data processing follows to determine the direction of the eye gaze [1]. Other information, such as eye blinks, blink frequency, left and right winks or modifications of the pupil diameter can be detected as well [6]. Eye trackers have been used for various purposes, such as to improve design and layout [7] [8], input for games [9], or assistive techniques and applications for people with disabilities [10] [11]. For instance, *Eye Draw* [10], is a software application that allows children with motor impairments to draw using eye movements.

Various eye gaze tracking equipment and software are available on the market [36] [37] [38] [39]. To implement the VErGE, we considered the *MyGaze* eye tracker that features a frame rate of 30Hz, accuracy for gaze position of 0.5° and space resolution of 0.1°. The software module for data acquisition captures screen coordinates for each individual eye during the user's interaction with the screen.

```xml
<Stroke id="0">
<Point T="63646856673207" X="422.30000" Y="243.90000" W="0.00000" H="0.00000" A="0.00000" />
<Point T="63646856673211" X="422.10000" Y="243.57000" W="0.00000" H="0.00000" A="0.00000" />
<Point T="63646856673217" X="422.30000" Y="243.90000" W="0.00000" H="0.00000" A="0.00000" />
<Point T="63646856673220" X="422.10000" Y="243.57000" W="0.00000" H="0.00000" A="0.00000" />
<Point T="63646856673224" X="422.10000" Y="243.57000" W="0.00000" H="0.00000" A="0.00000" />
<Point T="63646856673227" X="422.10000" Y="243.57000" W="0.00000" H="0.00000" A="0.00000" />
<Point T="63646856673231" X="422.30000" Y="243.90000" W="0.00000" H="0.00000" A="0.00000" />
<Point T="63646856673235" X="422.10000" Y="243.57000" W="0.00000" H="0.00000" A="0.00000" />
<Point T="63646856673240" X="422.10000" Y="243.57000" W="0.00000" H="0.00000" A="0.00000" />
<Point T="63646856673245" X="422.10000" Y="243.57000" W="0.00000" H="0.00000" A="0.00000" />
<Point T="63646856673250" X="422.30000" Y="243.90000" W="0.00000" H="0.00000" A="0.00000" />
<Point T="63646856673254" X="422.10000" Y="243.57000" W="0.00000" H="0.00000" A="0.00000" />
<Point T="63646856673258" X="422.10000" Y="243.57000" W="0.00000" H="0.00000" A="0.00000" />
<Point T="63646856673263" X="422.30000" Y="243.90000" W="0.00000" H="0.00000" A="0.00000" />
<Point T="63646856673267" X="422.10000" Y="243.57000" W="0.00000" H="0.00000" A="0.00000" />
...
<Point T="63646856673448" X="438.98000" Y="567.43000" W="0.00000" H="0.00000" A="0.00000" />
<Point T="63646856673453" X="439.10000" Y="567.10000" W="0.00000" H="0.00000" A="0.00000" />
</Stroke>
```


Fig. 4 - Data example output from the VErGE application for eye gaze (XML format).

F. The EEG Monitoring Subsystem

The EEG monitoring system can pick up the brain electrical activity from various cortical regions, clean it up, and store the data for further processing. By interpreting the signals generated by various brain areas, the EEG subsystem can identify specific mental patterns and allow identification of the user's mental state.

The human brain is a complex system with functions that are distributed in all its volume. However, for many practical goals, relevant information regarding this activity can be extracted via direct contact electrodes. The recorded information is used as input to reconstruct a source model of the most important regions of the brain. The model obtained can be queried to obtain various information regarding the underlaying mental state of the user. The raw data that describe the brain activity must be subsequently interpreted to assess the mental state via software algorithms based on empirical observation of neurological states [13] [14].

In this regard, various devices in use will provide algorithms to detect the neural patterns or even the emotional states of the person. This data (e.g., neural waves in various bands - delta, theta, beta, gamma and the mental states identified for various periods of time) can be correlated with the information acquired from other types of sensors, such as the gyroscopes and accelerometers integrated in the headset to create a bigger picture and thus infer the general state of the user more accurately [15] [16]. Given that the raw recorded data consist of electrical signals in the µV range, myoelectric activity (i.e., the electrical response of the musculature of the head in the mV range) is also recorded. Usually, EEG devices can separate the two types of signals so that an additional class of data (the muscular response of the head and head organs, such as eyes and mouth) related to facial expressions can be integrated in the analysis.

The EEG analysis is used in many fields of research to understand a person's subconscious response to various stimuli. Brain-computer interfaces are also used when the goal of the study is the control of various devices via mental state alone (or as assisting control) [17] [18] or in investigating the cognitive load of various operations [19]. Empirical studies have validated this approach [2] [20].

Traditionally, EEG analysis is performed with medical-grade equipment. The high precision of data acquisition for these kind of devices is offset by their high costs, the expertise required for applying the electrodes and interpreting the results, and the general bulkiness of the equipment. In contrast, in the last decade, a variety of low-cost, consumer-grade EEG devices have been released, many of them delivering very accurate results [2]. These devices connect to the PC and many expose a Software Development Kit (SDK) to allow data processing and analysis of the raw signals, making them attractive for researchers and practitioners.

The EEG device used with VErGE is an Emotiv EPOC+ Research Edition. This is a consumer-grade equipment with very good performance, manevrability and a source model of EEG activity almost on par with the entry-level medical devices. The model was validated via empirical tests and offers access to affective state of the subject via its EmoEngine subsystem [40] [41].

The Emotiv EPOC+ is a high-resolution, 14 channels wireless (via Bluetooth) EEG recording device, capable of acquiring and recording raw electrical data in the µV and mV range. It is backed up by an emotional detection suite capable of inferring various affective states, such as user frustration, meditation, engagement, boredom, and excitement (short term and long-term). The device records the electrical activity of the brain in certain predefined zones (from 14 active channels) via gold-plated electrodes that eliminate electric noise in the contact area, mediated by wet felt pads (saline solution). The easiness of use makes the headset ideal for studies with many participants and in various experimental conditions, such as outside the laboratory or with people with motor impairments.

Fig. 5 - The VErGE system employed in a practical experiment for collecting multimodal data (voice, touch gesture, eye gaze, and EEG) from participants with motor impairments (left) and without impairments (right).

The headset has a software suite that allows verification of electrodes placement via an electrode contact quality map and calibration for each user. The data acquired via the EEG headset for each subject and gesture is the raw EEG wave (for each of the 14 channels), the 5 emotional states, and the user's head movements (with the internal gyroscope).

Nevertheless, EEG devices have certain limitations, such as:

- The quality of contacts may vary in experimental use, mostly because the contact pads must be moisturized with a saline solution and placed directly on the scalp. The density of solution may vary with time due to evaporation.
- People with dry skin and/or high amount of head hair will cause low-quality raw EEG data acquisition.
- Interference with the surrounding equipment can affect the quality of data (e.g. electromagnetic noise).
- Affective state acquisition is limited in the first few data frames recorded because the EmoState algorithm is based on the previous data, i.e. it needs a certain amount of data frames to infer the emotional state of the user.

III. CONCLUSION

The VErGE platform was designed to collect, process, and integrate multimodal input from users, especially from users with motor impairments. As such, it is a valuable tool and enabling technology for research purposes, as it can capture data at multiple levels, helping researchers to understand and highlight ways to improve traditional user interfaces and make them more reliable and easy to use – especially for people with disabilities.

We validated VErGE in a practical experiment involving several participants with and without motor impairments (Fig. 5). The goal of the experiment was twofold: to test the viability of the setup (especially in the field, i.e., the system was transported to accommodate data collection from people with severe motor impairments) and to acquire relevant data regarding multimodal input. The VErGE system was proven to be quite capable at detecting and acquiring significant biological parameters regarding multimodal input. Fig. 2, Fig 3 and Fig. 4 illustrate examples of gesture, eye gaze, and EEG data in our XML format. We believe that the hardware and software platform presented in this work is a useful research tool for multimodal data collection in a variety of experimental conditions.

ACKNOWLEDGMENT

This work was supported by a grant of the Romanian National Authority for Scientific Research and Innovation, CNCS/CCCDI – UEFISCDI, project number PN-III-P2-2.1-PED-2016-0688 (209PED/2017) within PNCDI III.

REFERENCES

[1] D. W. Hansen and Q. Ji, “"In the eye of the beholder: A survey of models for eyes and gaze”, in IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, iss. 3, 2010, pp. 478–500.

[2] H. Ekanayake, “P300 and Emotiv EPOC: Does Emotiv EPOC capture real EEG?”, 2015, *retrieved from* http://neurofeedback.visaduma.info/EmotivResearch.htm (March 2018).

[3] J.Sutton and Z. Austin. “Qualitative research: Data collection, analysis, and management”, in The Canadian Journal of Hospital Pharmacy, vol. 68, no. 3, 2015, pp. 226-231.

[4] A. Corradini and P.R. Cohen, “On the relationships among speech, gestures, and object manipulation in virtual environments: Initial evidence”, in Van Kuppevelt J.C.J., Dybkjær L., Bernsen N.O. (eds) Advances in Natural Multimodal Dialogue Systems. Text, Speech and Language Technology, vol. 30, 2005, pp. 97-112.

[5] D. Petrovska-Delacrétaz, G.Chollet and B. Dorizzi (eds.), “Guide to Biometric Reference Systems and Performance Evaluation”, eISBN: 978-1-84800-292-0, 2009, Springer-Verlag – London, pp. 329-331.

[6] N. Riche, M. Duvinage, M. Mancaş, B. Gosselin and T. Dutoit, “Saliency and human fixations: State-of-the-art and study of comparison metrics”, in In Proceedings of ICCV 2013, 2013, pp. 1153-1160.

[7] N. Weibel, A. Fouse, C. Emmenegger, S. Kimmich and E. Hutchins, “Let's look at the cockpit: exploring mobile eye-tracking for observational research on the flight deck”, in Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '12), ACM, 2012, pp. 107-114.

[8] P. Jermann, D. Gergle, R. Bednarik and S. Brennan, “Duet 2012: Dual eye tracking in CSCW”, in Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work Companion (CSCW '12), ACM, 2012, pp. 23-24.

[9] E. Velosso and M. Carter, “The Emergence of EyePlay: A survey of eye interaction in games”, in Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '16), ACM, 2016, pp. 171-185.

[10] A.J. Hornof and A. Cavender, “EyeDraw: enabling children with severe motor impairments to draw with their eyes”, in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '05), ACM, 2005, pp. 161-170.

[11] C.A. Chin, A. Barreto, G. Cremades and M. Adjouadi, “Performance analysis of an integrated eye gaze tracking / electromyogram cursor control system”, in Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility (Assets '07), ACM, 2007, pp. 233-234.

[12] J.J. Cerrolaza, A. Villanueva, M. Villanueva and R. Cabeza, “Error characterization and compensation in eye tracking systems”, in Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '12), ACM, 2012, pp. 205-208.

[13] S. Villata et al., “Emotions and personality traits in argumentation: An empirical evaluation”, in Argument & Computation, vol. 8, no. 1, 2017, pp.61-87 (*extended version*).

[14] R.N. Khushaba et al., “Choice modeling and the brain: A study on the electroencephalogram (EEG) of preferences”, in Expert Systems with Applications, vol. 39, iss. 16, 2012, pp. 12378-12388.

[15] J. Mas I Manjon, “Theta rhythm (emotion) and the alpha rhythm (attention) EEG”, Fundacion J. Mas, Research Report 03, 2011, pp. 1-6, *unpublished*.

[16] P. Mavros, R. Coyne, J. Roe and P. Aspinall, “Engaging the brain: Implications of mobile EEG for spatial representation”, in Digital Physicality | Physical Digitality: Proceedings of the 30th eCAADe Conference, 2012, pp. 657.

[17] S. Dharmasena, K. Lalitharathne, K. Dissanayake, A. Sampath and A. Pasqual, "Online classification of imagined hand movement using a consumer grade EEG device", in IEEE 8th International Conference on Industrial and Information Systems ICIIS 2013 - Conference Proceedings, 2013, pp. 537-541.

[18] M.N. Fakhruzzaman, E. Riksakomara and H. Suryotrisongko, “EEG wave identification in human brain with Emotiv EPOC for motor imagery”, in Procedia Computer Science, vol. 72, 2015, pp. 269-276.

[19] E.W. Anderson, K.C. Potter, L.E. Matzen, J.F. Shepherd, G.A. Preston and C.T. Silva, “A user study of visualization effectiveness using EEG and cognitive load”, in Eurographics/IEEE Symposium on Visualization 2011 (EuroVis 2011), vol. 30, no. 3, 2011.

[20] G.S. Taylor and C. Schmidt, “Empirical evaluation of the Emotiv EPOC BCI headset for the detection of mental actions”, in Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 56, iss. 1, 2012, pp. 193-197.

[21] Radu-Daniel Vatavu, Gabriel Cramariuc, Doina Maria Schipor, “Touch Interaction for Children Aged 3 to 6 Years: Experimental Findings and Relationship to Motor Skills”, International Journal of Human-Computer Studies, 74, 2015, pp. 54-76.

[22] Radu-Daniel Vatavu, Lisa Anthony, Jacob O. Wobbrock, “Gestures as Point Clouds: A $P Recognizer for User Interface Prototypes”, in Proceedings of ICMI'12, the 14th ACM International Conference on Multimodal Interaction. ACM, 2012, pp. 273-280.

[23] Bogdan-Florin Gheran, Gabriel Cramariuc, Ionela Rusu, Elena-Gina Craciun, “Tools for Collecting Users' Touch and Free-Hand Gestures on Mobile Devices in Experimental Settings”, in Proceedings of the 13th International Conference on Development and Application Systems, 2016. IEEE Press, pp. 248-253.

[24] Khe Chai Sim, “Speak-as-you-swipe (SAYS): a multimodal interface combining speech and gesture keyboard synchronously for continuous mobile text entry” in Proceedings of the 14th ACM international conference on Multimodal interaction (ICMI '12). ACM, New York, NY, USA, 2012, 555-560.

[25] Lode Hoste, Beat Signer, “SpeeG2: a speech- and gesture-based interface for efficient controller-free text input”, in Proceedings of the 15th ACM on International conference on multimodal interaction (ICMI '13). ACM, New York, NY, USA, 2013, pp. 213-220.

[26] Radu-Daniel Vatavu, “User-defined gestures for free-hand TV control” in Proceedings of the 10th European Conference on Interactive TV and Video (EuroITV '12). ACM, New York, NY, USA, 2012, pp. 45-48.

[27] Radu-Daniel Vatavu, Ovidiu-Ciprian Ungurean, Ștefan-Gheorghe Pentiuc, “Body Gestures for Office Desk Scenarios”, in D. England (Ed.), Whole Body Interaction. Springer HumanComputer Interaction Series. London: Springer-Verlag, 2011, pp. 163-172.

[28] Radu-Daniel Vatavu, Matei Mancaș, “Visual attention measures for multi-screen TV”, in Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '14). ACM, New York, NY, USA, 2014, pp. 111-118.

[29] Raphael Menges, Chandan Kumar, Daniel Müller, Korok Sengupta. “GazeTheWeb: A Gaze-Controlled Web Browser”, in Proceedings of the 14th Web for All Conference on The Future of Accessible Work (W4A '17). ACM, New York, NY, USA, 2017, Article 25, 2 pages.

[30] Susumu Harada, Jacob O. Wobbrock, James A. Landay, “Voicedraw: a hands-free voice-driven drawing application for people with motor impairments”, in Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility (Assets '07). ACM, New York, NY, USA, 2007, pp. 27-34.

[31] Ovidiu Andrei Schipor, Stefan Gheorghe Pentiuc, Maria-Doina Schipor, “Improving Computer Based Speech Therapy Using a Fuzzy Expert System”, Computing and Informatics 29(2), 2010, pp. 303-318.

[32] Vittorio Fuccella, Gennaro Costagliola, “Unistroke Gesture Recognition Through Polyline Approximation and Alignment”, in Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15). ACM, New York, NY, USA, 2015, pp. 3351-3354.

[33] Radu-Daniel Vatavu, Daniel Vogel, Géry Casiez, Laurent Grisoni, “Estimating the Perceived Difficulty of Pen Gestures” in Proceedings of INTERACT'11, the 13th IFIP TC13 Conference on Human-Computer Interaction. LNCS 6947, Springer Berlin Heidelberg, 2011, pp. 89-106.

[34] Yosra Rekik, Laurent Grisoni, Nicolas Roussel, “Towards Many Gestures to One Command: A User Study for Tabletops”, in Proceedings of INTERACT (2), 2013, pp. 246-263.

[35] Speech API, Speech Recognition – Google Cloud Platform, https://cloud.google.com/speech/ , (March 2018).

[36] http://www.eyetracking.com/ (March 2018).

[37] http://www.umoove.me/ (March 2018).

[38] https://tobiigaming.com/ (March 2018).

[39] https://imotions.com/ eye-tracking/ (March 2018).

[40] S.S. Shankar and R. Rai, “Human factors study on the usage of BCI headset for 3D CAD modeling”, in Computer-Aided Design, vol. 54, Elseviere, 2014, pp. 51-55.

[41] A. R. Shahrestani, “Brain Computer Interfaces emotional state detection (EEG pattern recognition)”, master thesis presented to The Faculty of the Computer Science Program California State University Channel Islands, 2013, *unpublished*, *retrieved online* at http://repository.library.csuci.edu/bitstream/handle/10211.3/57093/Shahrestani%20Thesis~.pdf;sequence=1 (March 2018).