An improved privacy-aware system for objective and subjective ecological momentary assessment

ULRIK KOWALK*, SVEN KISSNER, PETRA VON GABLENZ, INGA HOLUBE, AND JOERG BITZER

Institute of Hearing Technology and Audiology, Jade University of Applied Sciences, Oldenburg, Germany

The technical components and software features of a new hearing-aid compatible smartphone-based ecological momentary assessment (EMA) system are presented in this paper. EMA is an assessment strategy that seeks to minimise instrumental infliction on the measured entity while data is gathered at multiple points of time. This work builds upon an already developed and deployed smartphone-based system. Objective data is gathered in the form of acoustical features, while subjective data is collected via automatised questionnaires. Since linking objective acoustical measures to subjective assessments is particularly promising with regard to the hearing rehabilitation process, our system has been specially tailored for hearing aid users. The introduction of wireless data transfer has eliminated cable clutter, a customisable questionnaire allows for subjective assessment, and a streamlined user interface complements the design. Like the former version, the current revision ensures the privacy of both participants and third parties. To facilitate cooperative research, source code and custom-built hardware will be released under open source licenses. All additional components are commercially available.

ECOLOGICAL MOMENTARY ASSESSMENT

When conducting a study, practised procedures involve the completion of one or more questionnaires, usually taken in retrospective. The answers given rely heavily on memory, but because memory is of transient nature its distorting effects can strongly influence the results. According to Shiffman *et al.* (2008), ecological momentary assessment (EMA) uses a different approach. Data is recorded at several points during the study, often on sub-hour intervals, resulting in more reliable answers and evaluable evolution of parameters over time. To ensure that measurement does not interfere with the entity being measured, data collection needs to be as ecological as possible. EMA therefore aims at gathering data at the time it is generated without having an influence on the data. For practical reasons, repeated surveys are nowadays usually conducted using digital devices. Galvez *et al.* (2012) incorporated a personal digital assistant-based (PDA) EMA application to explore when and how hearing problems occur throughout the day. As opposed to strictly subjective assessments, a number of

*Corresponding author: ulrik.kowalk@jade-hs.de

Proceedings of the International Symposium on Auditory and Audiological Research (Proc. ISAAR), Vol. 6: Adaptive Processes in Hearing, August 2017, Nyborg, Denmark. Edited by S. Santurette, T. Dau, J. C.-Dalsgaard, L. Tranebjærg, T. Andersen, and T. Poulsen. The Danavox Jubilee Foundation, 2017.
 ISBN: 978-87-990013-6-1.

audiological studies have implemented recordings of physical variables to correlate with survey results. Banerjee (2011a) used logged data of manual multimemory and volume control adjustments by hearing aid users together with broadband input levels of the hearing aids combined with EMA in order to successfully identify situations in which users desire to modify settings. In a second study, Banerjee (2011b) investigated the correlation between automatic behaviour of hearing aids and EMA surveys to learn more about parameter decisions made by the hearing aids. Timmer *et al.* (2017) verified the validity of EMA according to hearing experiences, using data from an environmental classifier bound to a smartphone-based EMA system by showing high correlation between objective and subjective results as well as an overall high compliance rate. In a recent study published by Wu *et al.* (2017) the authors complemented randomly timed smartphone-based surveys with audio recordings from a portable device the test subjects carried around their neck in order to classify different listening situations.

Our method, in contrast, implements a single application that extracts acoustic features from a live binaural audio stream in real time and combines them with data from randomly or fixedly timed questionnaires. Implementing bluetooth audio transmission from ear-level microphones to a smartphone, it is a singular open-source experiment device that can be used without programming knowledge. High-level privacy-awareness allows for legally unconstrained use in everyday situations.

DESCRIPTION OF THE PREVIOUS SYSTEM

In Kissner *et al.* (2015), a smartphone-based EMA system was presented that included microphones for the purpose of extracting audio features. These microphones were worn like behind-the-ear (BTE) hearing aids and signals were transmitted via cables to an external USB audio adapter connected to the smartphone. Two applications ran simultaneously – one performed extraction of acoustical features and one conducted questionnaires without any internal communication channel between the two applications. Audio features were archived as blocks of bundled data over short periods of time, creating a separate series of blocks for each feature.

ATTRIBUTES OF THE NEW SYSTEM

The technical components and properties of the new system as well as advantages over the old system are described in this section.

Hardware

For the newly developed system, different choices with regard to hardware have been made. The microphones are no longer mounted behind the ears but attached to glasses, right above the ears (see Fig. 1). They are connected to a pocket-sized transmitter box (weight: ca. 36 g) by means of audio cables (length: ca. 0.5 m). The manufactured compartment contains a set of two A/D converters with preamplifiers and voltage offset filters, a lithium-ion polymer (LiPo) battery, and a Bluetooth transmitter for wireless audio transfer to the smartphone. Signals are sent in high resolution via the A2DP protocol. An integrated safeguard circuit monitors the voltage of the LiPo battery and performs automatic shutdown in order to preserve battery life. The transmitter unit has a runtime of more than 8 hours and is charged by either USB, external power supply, or induction coil. RGB-LEDs indicate the current state of the device and transmission power is regulated dynamically when necessary. It can be easily attached to any clothing by an external clip. An outline of the system schematics is shown in Fig. 2.

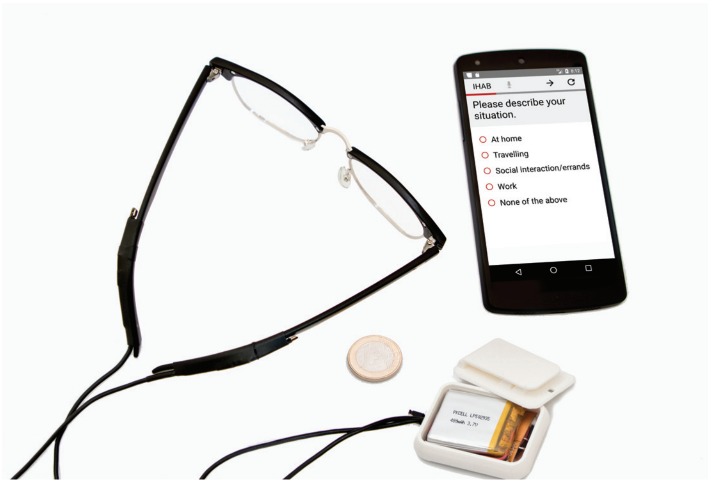

Fig. 1: Prototype of EMA system: Microphones attached to glasses, pocket-sized Bluetooth transmitter, and smartphone with questionnaire

A challenge during development has been the transformation of an Android device to act as a dual channel audio receiver. Since this feature is not provided by the conventional Android system, the device (LG Nexus 5) has been equipped with a modified Android Automotive operating system.

Signal processing

At the current stage, three acoustical features are extracted in realtime by the smartphone from signal blocks $x_m[n] = x[n + m \cdot N]$ of size N, m being the block index. The first measure is the RMS

$$\mathrm{RMS}_m = \sqrt{\sum_{n=0}^{N-1} x_m^2[n]}, \qquad \text{(Eq. 1)}$$

which is calculated in order to obtain binaural loudness levels. As a basic input to voice activity detection, the zero-crossing rate (ZCR) of the signal $x[n]$ and its first derivative $\Delta x[n] = x[n] - x[n-1]$ are recorded as well. The ZCR is calculated

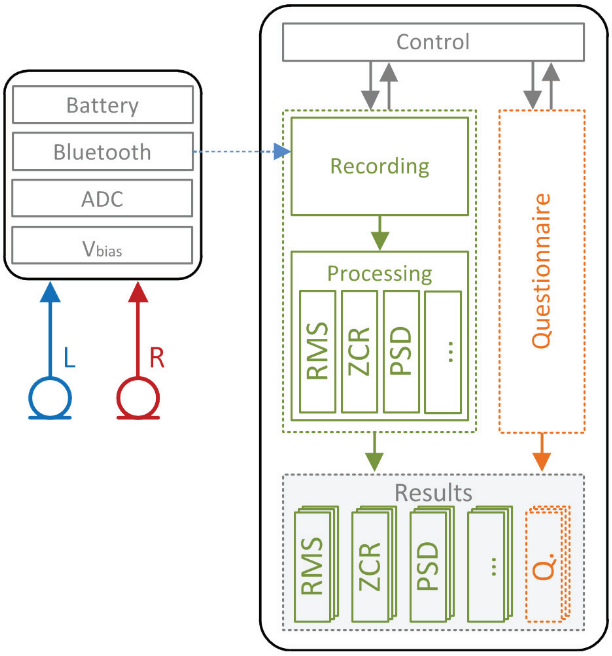

Fig. 2: System layout – dual microphone signals sent to smartphone via Bluetooth transmitter, extraction of acoustical features, and automatic conduction of questionnaire.

according to

$$\mathrm{ZCR}_m = \frac{1}{N-1}\sum_{n=0}^{N-1} s_m[n] \qquad \text{(Eq. 2)}$$

$$\text{with } s_m[n] = \begin{cases} 1 & \text{if } x_m[n]\cdot x_m[n-1] < 0 \\ 0 & \text{else,} \end{cases} \qquad \text{(Eq. 3)}$$

counting the number of zero crossings per signal block. The third feature are power spectral densities Φ, that are calculated for both channels separately, as well as cross-power spectral densities (Φ_{LR}) between left and right channel. In Eq. 4 and Eq. 5, X_m is the Fourier transform of the N-point Hann-windowed signal x_m:

$$\Phi_m[n] = X_m[n] \cdot X_m^*[n] \qquad \text{(Eq. 4)}$$

$$\Phi_{\mathrm{LR},m}[n] = X_{\mathrm{L},m}[n] \cdot X_{\mathrm{R},m}^*[n] \qquad \text{(Eq. 5)}$$

These features were chosen to give an acoustical overview of the participants' daily routine (Bitzer *et al.*, 2016) and will in future be complemented by other objective metrics. To simplify further extension of functionality, the system uses a defined plug-in architecture. Feature extraction is performed on chunks of fixed length (e.g., 60 s) resulting in one time-stamped series per feature.

Privacy-awareness

Two requirements are met by the system: 1) No audio data is stored; 2) No content can be reconstructed from the extracted data. The first criterion is met by the design of the processing engine. The second is implemented in each feature respectively. Special consideration is taken for the PSD feature. In order to prevent reconstruction, PSD time series are smoothed with a time constant of $\tau = 125$ ms and certain frames are omitted. As shown by Kissner *et al.* (2015), no content can thus be retrieved while acoustical information is still usable for study.

Questionnaires

A simplified interface has been developed for adaptation and creation of new questionnaires without the restriction of programming skills. Two software components are essential: the main application in the form of an installable .apk file and at least one questionnaire. The questionnaire is implemented as a formatted, human-readable .xml file with intuitive attributes and optional comments. Time scheduling values are specified as mean interval and randomness margin represented by seconds, the usual case being periodically recurring questionnaires. A margin of 0 seconds yields a steady sampling interval and different questionnaires are selected via a preferences menu. Dynamic structuring helps extract a maximal amount of data through tailored questionnaires by only stating relevant questions, true to a filter attribute. A question is only visible if its criteria are met based on a system of unique answer identifiers (IDs) that come with every answer. Once an answer has been selected, the corresponding ID is saved to memory. Two possible restrictions exist. Either a predefined ID must exist in the memory (positive criterion) or it explicitly must not exist (negative criterion). While a question is shown if one or more positive criteria are satisfied, it will be hidden if one or more negative criteria are met, negative overriding positive. For intuitive assessment, answer formats include radio buttons, checkboxes, emojis, sliders with fixed or arbitrary scales, and free text.

Open source

In order to allow for collaboration, all source code and construction plans will be published under open source licenses. The system uses Nexus 5 smartphones, which are commercially available. The questionnaire management application has been localised across English and German. Future releases will include further languages.

Performance

Several acoustical parameters have been measured in order to asses the technical properties of the current system. For reproduction of test signals, an NTi Audio Talk-Box is used and reference measurements are taken by a G.R.A.S. 40AF free-field microphone pre-amplified by a Brüel & Kjær 2829 type 26TK system. Sound pressure levels are calibrated using a Norsonic Sound Calibrator type 1251 and the experiments are conducted in an anechoic chamber with a volume of approximately $43\,m^3$.

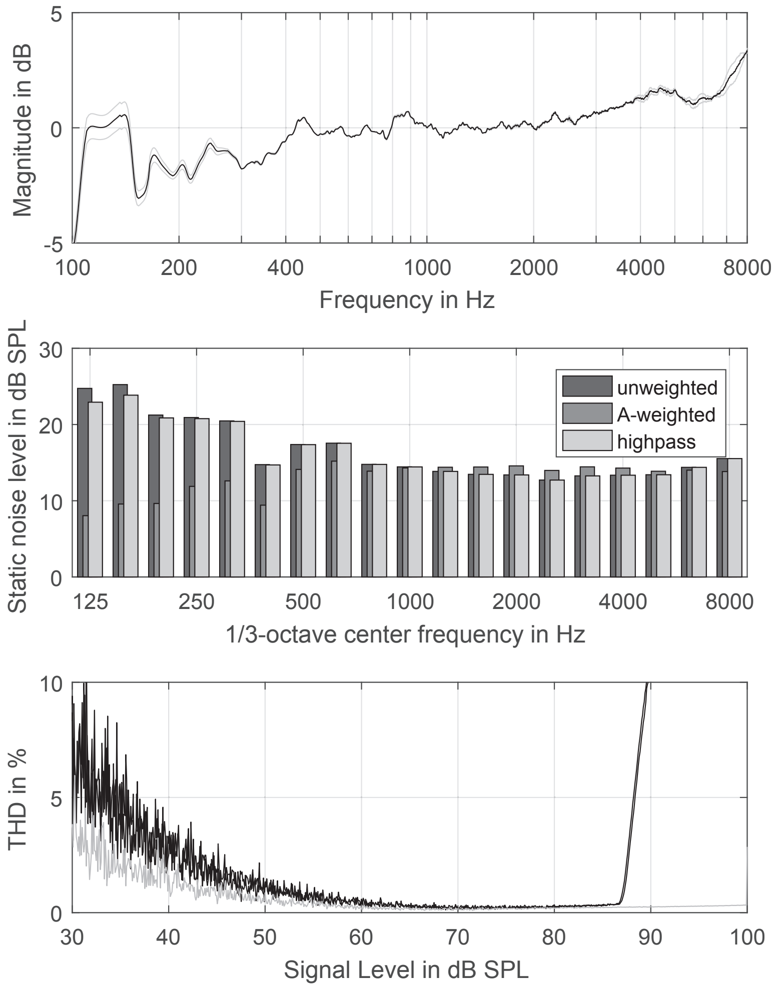

Fig. 3: *Top:* Frequency response of the current system, smoothed in ERB bands. Individual microphone responses are drawn as grey lines, the averaged response is drawn in black. *Middle:* Third-octave noise levels for different pre-processing filters. *Bottom:* Average percentage of total harmonic distortion by the system (black line) and reference microphone (grey line).

Frequency response is measured with calibrated, zero-centred broadband noise played back at a level of 60 dB SPL. The distance between speaker and system microphones is 1.60 m. As shown in Fig. 3 (top), we find a reasonably flat response between 100 Hz and 8 kHz with a slight upward tendency towards higher frequencies. Tolerances between multiple microphones are less than ±0.5 dB between 100 Hz and 200 Hz, less than ±0.1 dB between 200 Hz and 4 kHz and a little higher beyond.

Equivalent input noise levels are examined under the same external conditions as the frequency response. Three different states of internal processing are evaluated: unweighted, A-weighted, and filtered by a second order Butterworth high pass filter with f_0=100 Hz. As depicted in Fig. 3 (middle), unweighted noise levels are relatively equally distributed with emphasis on frequencies lower than 400 Hz.

Total harmonic distortion determines the dynamic range of the system. Measurements are conducted using a Fostex 6301B loudspeaker generating an amplitude sweep signal with a sinusoidal carrier at a frequency of 1 kHz and level ranging from 20 to 100 dB SPL. For a distortion acceptance limit of 2 %, results show dynamic validity ranging from approximately 45 up to 88 dB SPL yielding an estimated dynamic of 43 dB (see Fig. 3, bottom). This renders the system applicable for one of the main acoustical situations of everyday life – conversational speech – which usually lies in the range above 50 dB SPL according to Bitzer *et al.* (2016).

Advantages over the previous system

While being fully functional, the preceding system by Kissner *et al.* (2015) included certain attributes that were updated for the new system. Because a line connection was used to transmit microphone signals to the smartphone, a third party USB audio interface was required at the receiving end. This implied the need for proprietary device drivers, counteracting complete open source publication, and also leading to mechanical instabilities. This system uses wireless transmission, thus eliminating the need for additional hardware. Another advantage is reduced ecological influence on measurements due to omitted inhibitory attributes (e.g., cable clutter). Because the microphones are no longer mounted behind the ears, but are attached to the temple, the concurrent use of behind-the-ear hearing aids is now possible. Integration of signal processing and questionnaire management into one single application has introduced process supervision on a high level, increasing functional security and simplifying usability. New answer formats and scheduling options within the questionnaire editing interface have increased flexibility and have lead to more intuitive assessment.

SUMMARY AND OUTLOOK

A system has been presented that incorporates all instruments to perform privacy-aware EMA of both subjective and objective parameters, while granting the experimenter a high degree of freedom, flexibility, and simplicity. Currently in development is a shared database for swift data exchange, pooling and comparison, which would facilitate international collaboration. Wireless signal transmission over a serial protocol to loosen the constraint on phone brand and model are also being investigated as well as a modulation-based blind estimator of speech quality, the speech to reverberation modulation energy ratio (SRMR, see Falk *et al.*, 2010) to supplement the current feature set. Further options regarding questionnaires are event-based and fixed scheduling. The system will be integrated in a field study scheduled for 2018 (see Meis *et al.*, 2017).

ACKNOWLEDGEMENTS

The authors would like to thank Holger Groenewold and Sven Franz for hardware development and modification of the operating system. This work was supported by the Hearing Industry Research Consortium (IRC). Further support was granted by the governmental funding initiative "Niedersaechsisches Vorab" of the Lower Saxony Ministry for Science and Culture, research focus "Hoeren im Alltag Oldenburg (HALLO)".

REFERENCES

Banerjee, S. (**2011a**). "Hearing Aids in the real world: Use of multimemory and volume controls," J. Am. Acad. Audiol., **22**, 359-374. doi: 10.3766/jaaa.22.6.5

Banerjee, S. (**2011b**). "Hearing aids in the real world: Typical automatic behavior of expansion, directionality, and noise management," J. Am. Acad. Audiol., **22**, 34-48. doi: 10.3766/jaaa.22.1.5

Bitzer, J., Kissner, S., and Holube, I. (**2016**). "Privacy-aware acoustic assessments of everyday life," J. Audio Eng. Soc., **64**, 395-404. doi. 10.17743/jaes.2016.0020

Falk, T.H., Zheng, C., and Chan, W.Y. (**2010**). "A non-intrusive quality and intelligibility masure of reverberant and dereverberated speech," IEEE T. Audio Speech, **18**, 1766-1774. doi: 10.1109/TASL.2010.2052247

Galvez, G., Turbin, M.B., Thielman, E.J., Istvan, J.A., Andrews, J.A., Henry, J.A. (**2012**). "Feasibility of ecological momentary assessment of hearing difficulties encountered by hearing aid users," Ear Hearing, **33**, 497-507. doi: 10.1097/AUD.0b013e3182498c41

Kissner, S., Holube, I., and Bitzer, J. (**2015**). "A smartphone-based, privacy-aware recording system for the assessment of everyday listening situations," Proc. ISAAR, **5**, 445-452.

Meis, M., Krueger, M., Gebhard, M., von Gablenz, P., Holube, I., Grimm, G., and Paluch, R. (**2017**). "Overview of new outcome tools addressing auditory ecological validity: Analyses of behavior in real life listening environments," Proc. ISAAR, **6**, 31-38.

Shiffman, S., Stone, A.A., and Hufford, M.R. (**2008**). "Ecological momentary assessment," Annu. Rev. Clin. Psycho., **4**, 1-32. doi: 10.1146/annurev.clinpsy.3.022806.091415

Timmer, B.H.B., Hickson, L., and Launer, S. (**2017**). "Ecological momentary assessment: Feasibility, construct validity, and future applications," Am. J. Audiol., **26.3S**, 436-442. doi: 10.1044/2017_AJA-16-0126

Wu, Y.H., Stangl, E., Chipara, O., Shabih Hasan, S., Welhaven, A., and Oleson, J. (**2017**). "Characteristics of real-world signal to noise ratios and speech listening situations of older adults with mild to moderate hearing loss," Ear Hearing, in press. doi: 10.1097/AUD.0000000000000486