SONEX: AN EVALUATION EXCHANGE FRAMEWORK FOR REPRODUCIBLE SONIFICATION

Norberto Degara, Frederik Nagel
Fraunhofer IIS
Am Wolfsmantel 33
91058 Erlangen, Germany
norberto.degara@iis.fraunhofer.de

Thomas Hermann
Bielefeld University
Ambient Intelligence Group
P.O.-Box 10 01 31, 33501 Bielefeld, Germany
thermann@techfak.uni-bielefeld.de

ABSTRACT

After 18 ICAD conferences, Auditory Display has become a mature research community. However, a robust evaluation and scientific comparison of sonification methods is often neglected by auditory display researchers. In the last ICAD 2012 conference, only one paper out of 53 makes a statistical comparison of several sonification methods and still no comparison with other state-of-the-art algorithms is provided. In this paper, we review profitable standards in other communities and transfer them to derive recommendations and best practices for auditory display research. We describe SonEX (Sonification Evaluation eXchange), a community-based framework for the formal evaluation of sonification methods. The goals, challenges and architecture of this evaluation platform are discussed. In addition, a simple example of a task definition according to the guidelines of SonEX is also introduced. This paper aims at starting a vivid discussion towards the establishment of thorough scientific methodologies for auditory display research and the definition of standardized sonification tasks.

1. INTRODUCTION

Auditory Display research takes place in a community that builds up on a wide range of disciplines [1]. Among others, disciplines that contribute to sonification technology are physics, acoustics, psychoacoustics, signal processing, statistics, computer science and musicology. We can find application examples that go from auditory display in assistive technology, bio-feedback and sonification of movement to navigation of data and process monitoring. Independently of the discipline and application, the primary goal of Auditory Display research is to exploit our complex and powerful listening system to enable the use of sound to understand the world on a similar level as with visual displays – or even exceeding the understanding enabled by visual displays. The aim of Auditory Display research is therefore to develop systems where audio is the main carrier of information in a broad sense.

After 18 ICAD conferences, sonification has become a mature research field. Still, the community is fighting against being considered just as an amusement experiment or a pseudo-science by the audio scientific community. This is partially due to the fact that in some cases it is difficult to make a clear distinction between sonification, music, art and science as Ballora discusses in his very inspiring TEDxPSU talk [2]. However, if we attend to Hermann's definition of sonification [3] and take sonification as data-dependent generation of sound using a systematic, objective and reproducible transform, then sonification is clearly defined as a well-defined scientific method.

Nevertheless, a robust evaluation and scientific comparison of sonification methods is often neglected by auditory display researchers. In a recent review of the evaluation works presented in the ICAD conferences between 1992 and 2009, Vogt [4] shows that there are only few quantitative examples that allow for the objective comparison of sonification methods. In such an interdisciplinary but small community, the list of applications where sonification is used is long and finding publicly available methods and datasets that attack a similar application to compare with is in many cases difficult. In addition, when we are facing the problem of sonifying a particular dataset, many sonification possibilities exist. Just think of all the possible mapping combinations we could think of when using a parameter mapping sonification technique. A specific sonification method must be then selected to enable the best performance and user experience. However, in many cases, this selection is not mathematically justified but based on the subjective assessment of the researcher.

To support this observation, we discuss next one of the most successful sonification examples developed in the last years, the online acoustic feedback system for on-water rowing training proposed by Schaffert et al. in [5]. The work shows a significant increase in the average speed of the boat when using sonification. In this study, acceleration values are mapped to a tone on the musical MIDI scale. This mapping is arbitrary and only a single sonification method is tested. In a following COST-SID workshop for sonification, six different methods for rowing data were proposed [6]. These alternatives were further discussed in [7]. However, as already pointed in [8], no formal evaluation concerning the accurate extraction of information and aesthetics has been performed in this COST-SID workshop. A robust evaluation on the selection of the best sonification method is missing. In his work on elite rowing [8], Dubus quantitatively compares four sonification methods in terms of function and aesthetics, but still it is not shown which method provides the best performance results in terms of increasing the average speed or final user performance.

If we analyze the contributions of the ICAD 2012 conference, we can observe that the majority of the literature where a sonification method is proposed shows a lack of proper statistical evaluation and comparison. Reading through the ICAD 2012 proceeding papers dealing with some form of sonification we observe the following distribution: of the total number of 53 papers, 8 presented artistic works and 18 introduced new sonification methods for applications such as process monitoring, medicine, movement sonification and navigation of data. These 18 papers reported some sort of quantitative results and, therefore, a proper statistical evaluation of the selected sonification method should be considered. Of these 18 papers only 5 works propose more than one sonification method in order to select the most appropriate ([9, 10, 11, 12, 13]). However only one work makes a statistical comparison to determine the best performing algorithm [11]. This study statistically evaluates 3 different sonification metaphors plus a control system of sound source distance sonification for virtual auditory display. Still, none of the reviewed papers includes a state-of-the-art reference system for comparison. This makes the advancement of proper sonification methods based on previous findings very difficult, and hence the development of successful commercial applications.

What should we do then to raise the quality of sonification systems and publications to a higher level and have sonification not be relegated to pseudo-science? Greg Kramer, the founder of the ICAD community, already answered this question in 2004 [14], stating that a methodical research approach is needed. Therefore, considering the lack of robust evaluation observed, we can argue that a benchmarking framework that allows for the comparison of sonification algorithms is required. Experiments must be carefully designed and reproducible. This would allow the ICAD community to build upon each other's work and invest more time developing new methods and combining them with the existing techniques than recreating existing methods. From a practical point of view, it has been shown that making research reproducible and comparable also increases its potential impact factor [15].

In this paper we define SonEX (Sonification Evaluation eXchange), a community-based framework for the formal evaluation of sonification methods. The platform allows for the definition of a number of standardized sonification tasks and their corresponding evaluation measures used to compare the sonification algorithms submitted to the system for comparison. In SonEX, the tasks are collaboratively defined by the members of the community and independently evaluated every year, ranking sonification techniques according to their statistical performance. This platform would therefore allow the comparison of algorithms along their different runs, overcoming the lack of formal analysis and comparison currently observed. If this evaluation exchange platform would be finally accepted by the ICAD community, it would bring the community a bit closer to achieve established and standardized sonification techniques, which is one of the main goals of sonification in the next years [1].

Section 2 describes the platform, related work, goals and challenges of SonEX. Then, Section 3 presents a first approach to the implantation of SonEX, introducing an example of task definition according to SonEX. Finally, some conclusions are drawn in Section 4.

2. SONEX: A SONIFICATION EVALUATION EXCHANGE PLATFORM

2.1. Related Work

Algorithm benchmarking is a common practice in many research communities. We can find international competitions that, for example, evaluate the performance of biometric algorithms for face detection and recognition or fingerprint resistance and verification which are generally associated with an international conference, such as the International Joint Conference on Biometrics (IJCB)[1] or the IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)[2]. In the area of audio technology, signal processing algorithms are benchmarked in the context of international conferences such as the Audio Engineering Society (AES)[3] and the IEEE Signal Processing Society (ICASSP)[4]. In the area of Sound Source Separation, algorithms are evaluated in the context of a Signal Separation Evaluation Campaign (SiSEC)[5] . For music signal processing and analysis, the International Society for Music Information Retrieval (ISMIR) holds the Music Information Retrieval Evaluation eXchange (MIREX)[6] which was inspired by the success of the Text Retrieval Conference (TREC) evaluation framework [16]. In MIREX, music information retrieval algorithms are regularly evaluated over a set of predefined tasks and databases. Examples of these evaluation competitions include audio artist, genre and mood classification, chord and melody extraction, music similarity and retrieval and beat tracking and tempo estimation. Although this is considered to be a very young research community, the evaluation contest has significantly contributed to the development of new and very competitive methods since its first run in 2005 [17, 18].

In the context of the ICAD, several competitions have been run. In 2004, the Listening to the Mind Listening concert asked for sonifications of 5 minutes EEG data, defining specific aspects such as that real-time playback (mapping time to sonification time). This created a comparability for patterns in a single given data set using many different sonifications. The criterion of the sonification being systematic in the mapping process was set relevant for the jury. In 2011, a contest on the sonification of head related transfer functions was proposed. However, the contest was focused on the aesthetics of the task instead of an objective measurable objective. A panel of experts was selected to decide the best sonification but no objective performance measure was defined. The competition held in the ICAD 2012 adopted the theme "Listening to the World Listening" and its aim was to explore what could be learnt by listening to the sonification of social media data. In the context of data exploration, what can be potentially learnt from a sonification is unknown, or at least not defined properly, and therefore it is very difficult to specify an objective performance measure. The winner of the context was also selected by a jury and no explicit and measurable performance indices were defined. In the present ICAD 2013 conference, a very interesting contest for the sonification of spatial data for visually impaired people is organized. Although the evaluation measures have not been yet specified, performance could be objectively measured in terms of accuracy, error rate, reaction speed and aesthetics. This constitutes a good example of a task that could be run every year to challenge ICAD researchers to advance on the development of proper sonification methods for assistive technology.

2.2. Goals

Although some isolated competitions have taken place in the context of the ICAD, our first aim is to bring to the attention of sonification researchers the need for comparative evaluations and to inspire a general discussion on this issue. We have already identified the lack of robust analysis in the comparison of sonification

[1] http://atvs.ii.uam.es/icb2013/

[2] http://www.btas2013.org/

[3] http://www.aes.org

[4] http://www.signalprocessingsociety.org

[5] http://sisec.wiki.irisa.fr

[6] http://www.music-ir.org/mirex/wiki/MIREX_HOME

methods and given examples of how formal benchmarking in other fields substantially contributes to the advancement of technology. Therefore, if we want to see sonification to be considered a mature and scientific research field, comparison with state-of-the-art methods should be generally provided in our research works.

Also, following the success in other communities and in particular the example of MIREX [17, 18], we aim at developing the idea of the sonification contest further by defining a Sonification Evaluation eXchange (SonEX) framework for reproducible sonification. This platform is designed to enable the comparison and interoperability of different sonification methods and the sharing of data, evaluation methods and results. More specifically, this infrastructure should be flexible enough to enable the following:

- Submission and evaluation of *community-based* defined tasks.
- Agreement on an established database or data model for the evaluation of methods / sonification approaches.
- Definition of a standardized evaluation method, facilitating user-based subject evaluation and interaction.
- Publication of results to compare algorithms.

Thus, SonEX aims at representing a community-based framework for the formal evaluation of sonification methods and algorithms. As future perspective, we could even envision SonEX to enable researchers to test their approaches against each defined challenge outside the annual evaluation since the limitation to only annual evaluations could hinder the advance of technologies.

Another ambitious objective is to support the interoperability of methods. This means that researchers could directly use the functionality provided via methods already available in SonEX and thereby avoid reinventing wheels – SonEX could offer a community-shared toolkit infrastructure.

2.3. Workflow

Following the example of the MIREX community, a possible workflow for the definition and evaluation of sonification tasks could be as follows: First, a call of interest for sonification tasks is submitted to the community. If there are at least three researchers willing to participate, then the task could be incorporated into the official list of tasks to be evaluated. Then, the potential participants should redefine the ideas for the task. A measurable and objective task should be defined so that the performance of the algorithms can be evaluated. So the participants should also agree on the metrics to be used. The participants should also accept a common database or data model for evaluating the sonification algorithms. Finally, the input/output interface of the methods must be specified in order to make the task definition independent of the programming language to be used.

Once the submission is completed and the algorithms evaluated, the final results should be posted in the SonEX collaborative working environment prior to the ICAD conference. As in MIREX, it is advisable for the ICAD conference to include a poster session to present the results of the evaluation. Then, the pros and cons of the evaluation and future improvements should be discussed in a meeting. In addition, researchers participating in the evaluated tasks have to submit a short paper (2 pages) describing their system. The workflow is depicted in Fig. 1.

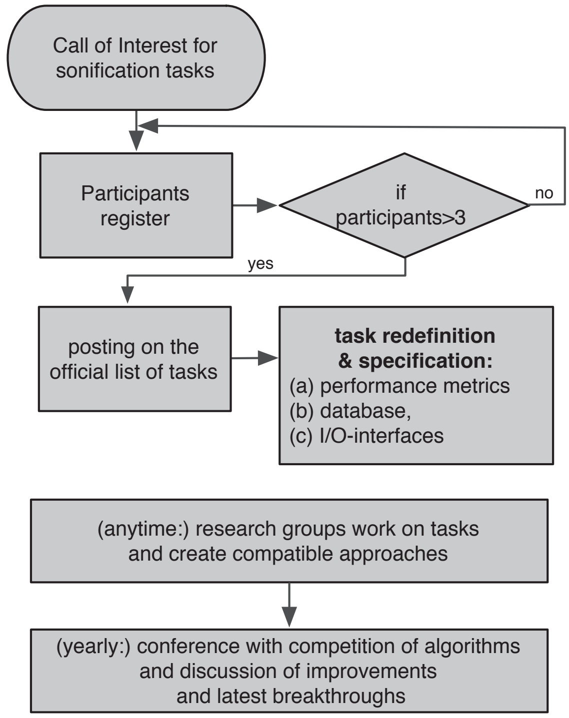

Figure 1: Block diagram of the proposed SonEX workflow.

2.4. Challenges

Facilitating the comparison of sonification methods is a big challenge. Some of the issues that we will have to face are the following:

1. *The definition of the sonification task must be objective and measurable in terms of performance:*

   According to [3], the data-dependent generation of sound in a sonification method reflects *objective* properties of the input data, furthermore there needs to be a precise definition on how the data and user interactions change the sound, and finally it has to be reproducible. This definition not only imposes constraints on the sound generation method but also on the knowledge we have about the data to be sonified. This means that we must be able to precisely define the properties of the input data and the possible interactions with the sonification system. This way, quantitative performance measures can be defined to evaluate how good the sonification method reflects objective properties of the data. Since sonification transforms are reproducible, this allows us to statistically characterize the performance. When defining a task, researchers should always consider how to quantitatively evaluate performance and how to compare their algorithms.

   We emphasize this point because the sonification problem is sometimes not well defined. For example, in data navigation and pattern discovery users aim at discovering previously unknown relations, arguing that our brain is a very good pattern recognition system. For such tasks, a clear objective can not be defined since we don't know a priori what we are looking for and performance can not be measured. Therefore, this task would not be well-defined for evaluation within the SonEX framework. However, the data navigation task could properly be redefined, for instance by creating a labelled database with multiple types of known data interactions. Then, we could evaluate which methods are best suited for the discovery of specific patterns hidden in the data, resulting probably in established standards for the discovery of certain patterns in complex data.

2. *Context reproducibility as precondition for proper evaluation:*

   SonEX only considers the online evaluation of sonification methods and for this, a computer framework is used. However, sonification is just the part of the whole auditory display that translates information to sound. The auditory display includes the application and usage context, interactions and the specific technical sound display [3].

   In some cases, such as in data exploration, a *computer equivalent*, i.e. the simulation of contextual conditions or even data generation within the computer, might be appropriate to evaluate sonification algorithms. However, in some other cases, such as in the rowing example describe above, the sonification evaluation results might not be completely valid. We are aware that the proposed evaluation does not reflect the details of a real system and that small factors change results when implementing a real auditory display. Still, the evaluation results obtained from SonEX can be used to select a reduced number of sonification approaches to evaluate in real systems.

   For this reason, computer equivalents to real problems must be carefully designed so that results can hold to some extent. Good existing examples of this computer equivalents for sonification research are the helicopter flying control experiment described in [19], the virtual space for blind people navigation introduced by Bujacz in [20] and the sonification experiment for situation awareness in surveillance operations presented in [21].

3. *Aesthetics is an additional criterium for evaluation:*

   Qualitative evaluations show that a balance between function and aesthetics is needed in interactive sonification design from a practical point of view [8]. Still, the subjective evaluation of the aesthetics of a sonification should not be the only measure of performance of a SonEX task, a common approach in current sonification research. Aesthetics should be combined with other objective evaluation measures, such as error rates or precision measures. For that, a weighted evaluation measure using for example a Multi-Criteria Decision Aid as the one suggested by Vogt in [4] could be used.

4. *Databases:*

   Sonification experiments are affected by extreme issues of data availability. In some cases, it is difficult to find quality ground-truth data which are widely accepted by the community for benchmarking algorithms. A significant challenge for the ICAD community is to build a large collection of data sets for all the defined tasks with their associated ground-truth. In some cases, the ground truth data could be defined by a data generation model. Good examples are the flying control system of [19] and the virtual space for moving around obstacles of [20]. This way, the expensive task of annotating a database is avoided.

   Sharing these databases is important because sonification algorithms can then be applied and evaluated over common data and thus their performance can be compared. However, participants should not have direct access to the data. Otherwise, researchers could fall into the error of designing sonification methods based on a very specific dataset what would lead to overfitting. For that reason, and as in many other communities, data should be provided to the sonification methods through SonEX, keeping the database secret but accessible.

5. *Subjective evaluation:*

   Contrary to most of the tasks evaluated in MIREX, sonification tasks must finally be evaluated with the active interaction of users, and finding volunteers for performing online evaluations will be difficult. Furthermore, users might be biased or not naive, for instance if they participated to a study several times. The undefined conditions on the side of the user create further problems: users might use headphones of different quality, or built-in loudspeakers in laptop that even fail to reproduce certain frequencies at all. Further more there is no control over environmental noises during the time of the participation to the evaluation. Instructions for adjusting audio equipment may help, but in the end, we need to trust in the cooperation of the participating users.

   However, with the collaboration of the whole community it will be possible to overcome most obstacles satisfactorily. For this to happen, evaluation tasks must be attractive and evaluation sessions should be also short in time to avoid discouraging volunteers. For this reason, for SonEX we propose the use of game-like evaluation approaches, which can be specified together with the task definition [22]. An example for this is given in Sec. 3 Finally, if researchers submit a task, they are also expected to collaborate in the evaluation of other tasks.

6. *Support a large number of programming languages and software toolkits:*

   ICAD is a multidisciplinary community where researchers use different programming languages and software toolkits. To extend and facilitate the proposed evaluation platform, SonEX should allow the submission of sonification methods programmed in different languages and this is a big challenge. As a starting point, SonEX is being currently developed to support Python, a flexible programming language which is able to integrate modules developed in other languages and includes an extensive collection of open access modules for statistical analysis, signal processing and graphical user interface development.

   In addition, there are Python libraries that support the Open Sound Control[7] (OSC) protocol, which is the format selected by SonEX for messaging. To separate the sonification from the evaluation platform, the database and the user interaction, and to make the integration of algorithms easier, an input/output interface must be implemented for the correct submission of the sonification method. This interface must be specified during the task definition step, describing the format of the OSC messages to allow the interaction of the user with the sonification method and how the method accesses the data for generating the sound. This interaction protocol could specify user control actions such as "stop", "play" and "go back", and parameters messaging as, for example, "sonification_method/parameter/frequency 440".

7. *Support a regular evaluation every year with only a small community:*

[7] http://opensoundcontrol.org

Another challenge is to keep the interest of the ICAD community on running these evaluation tasks every year together with the ICAD conference. We are aware that this requires a lot of effort and commitment. But we strongly believe that this is necessary if we want sonification to establish as scientific field.

With the collaboration of the ICAD community solutions for these (and probably more) challenges can certainly be found. A substantial part of the community needs to contribute if we want to see a significant scientific advance in sonification.

2.5. SonEX Architecture Summary

Figure 2 depicts the SonEX framework and summarizes the ideas discussed above.

As described in Section 2.3, researchers agree first on the "Task Definition", the "Task Evaluation" and the "Task Data Model" using a collaborative working interface. The sonification methods are submitted to the platform using a "Submission Interface" and the task is run in SonEX. Note that the sonification methods access the data and optionally user interaction, using a defined OSC protocol, and generate the sound which is presented to the users. Then, the resulting sound is subjectively evaluated using ratings and tests with users, allowing to measure and calculate a method's performance according to the agreed metrics. The sonification methods can potentially make use of other sonification libraries stored in a "Core Repository", and thereby more and more avoiding the "reinventing of wheels". Finally, the statistical analysis of the evaluation results are published in the platform.

3. "THE WALKING GAME" EXAMPLE

To provide the reader with a more tangible idea of SonEX tasks in practice, we introduce an example of task definition according to the workflow presented in Section 2.3. Therefore, we would first make a *call of interest* for ICAD researchers to participate in a "Walking Game" task experiment and agree on the database, evaluation and interface:

1. Task Description:

   In this game-based experiment, the player must guide an avatar to a target point in a virtual space avoiding obstacles and barriers as fast as possible. The virtual space is visually presented to the player together with the sonification of the position of the obstacles. For the first runs, the players may move the avatar in an audiovisual condition, allowing them to understand how sound and situation relate. Yet after some iterations (with always changing obstacle, target and initial avatar location), lights go off and the avatar must be guided just using the auditory information alone.

   Figure 3 presents an schematic representation of the proposed game. Delivering location-based information to support eyes-free navigation is a challenging task. However, it is also of great interest for the blind and visually impaired community since these sonification algorithms can be used for the development of Electronic Travel Aids (ETA) [23]. Therefore, the development of accurate and sophisticated sonification methods is of great importance.

2. User Interfaces / Interactions: For this navigation task, formatted for SonEX as a game, it must be defined how users actually control their avatar. Interfaces can range from cursor keys to using handheld smartphones equipped with a compass module to reorient and gestures to move forward. It must be defined whether rotation and translation proceeds in quantized units or continuously, and how long it takes for the game application to update the sound. Most importantly it must be determined how sound is displayed (e.g. using stereo headphones). Some methods (such as echolocation) will ask for the ability to inject a sound probe, e.g. a clicking sound, using a microphone on the client side and render a sonification on the basis of simulated spatial responses. This illustrates that even for a seemingly tiny task there are many questions to be addressed for defining a well-specified setup for reproducible research.

3. Database:

   In this task, the avatar moves in a virtual space where the position of obstacles and barriers are known. Multiple scenarios can be generated by placing obstacles in different positions or generating these positions at random. The benefit of this data generating model approach is that we avoid annotating the position of obstacles and using scene image segmentation algorithms if, for example, real video images were used. By using this virtual environment, the sonification task is also isolated from other problems (such as image segmentation) providing us with a very controlled experiment. Note also that we could also consider the definition of other similar subtasks where, instead of using a virtual model space, a database of real images or videos could be used.

4. Evaluation:

   For the evaluation of sonification methods, multiple evaluation measures can be proposed and agreed among the participants. We could for example consider to evaluate the total time for getting the avatar to the goal position. Obviously the number of obstacle collisions should also be considered. And finally, game players could rate their preferred sonification method so that preference and aesthetic quality can also be considered. These performance measures can be evaluated independently but, as introduced in Section 2.4, also a weighted average could be used to reduce the manifold features onto a single quality function.

   A web based interface should be implemented so as subjects of all over the world can participate in the evaluation. Also, a mobile-phone application could be provided to make the evaluation more attractive and accessible.

5. Software Interfaces:

   The software interface for communicating methods with the task must be also specified. The geometry, texture and positions of the obstacles could be defined in a configuration file in XML, for example. This information could be sent to the sonification algorithms using an OSC message such as:

   </SonEX/walking_game/method_X/cofiguration "path_to_the_configuration">

   The first part of the OSC message identifies the address of sonification method "X" and the second part the path to the configuration file. For the interaction with the player, the task user interface could send OSC messages to move left, right, forward or backward:

   </SonEX/walking_game/method_X/move_left 1.0>
   </SonEX/walking_game/method_X/move_right 1.0>
   </SonEX/walking_game/method_X/move_forward 1.0>
   </SonEX/walking_game/method_X/move_back 1.0>

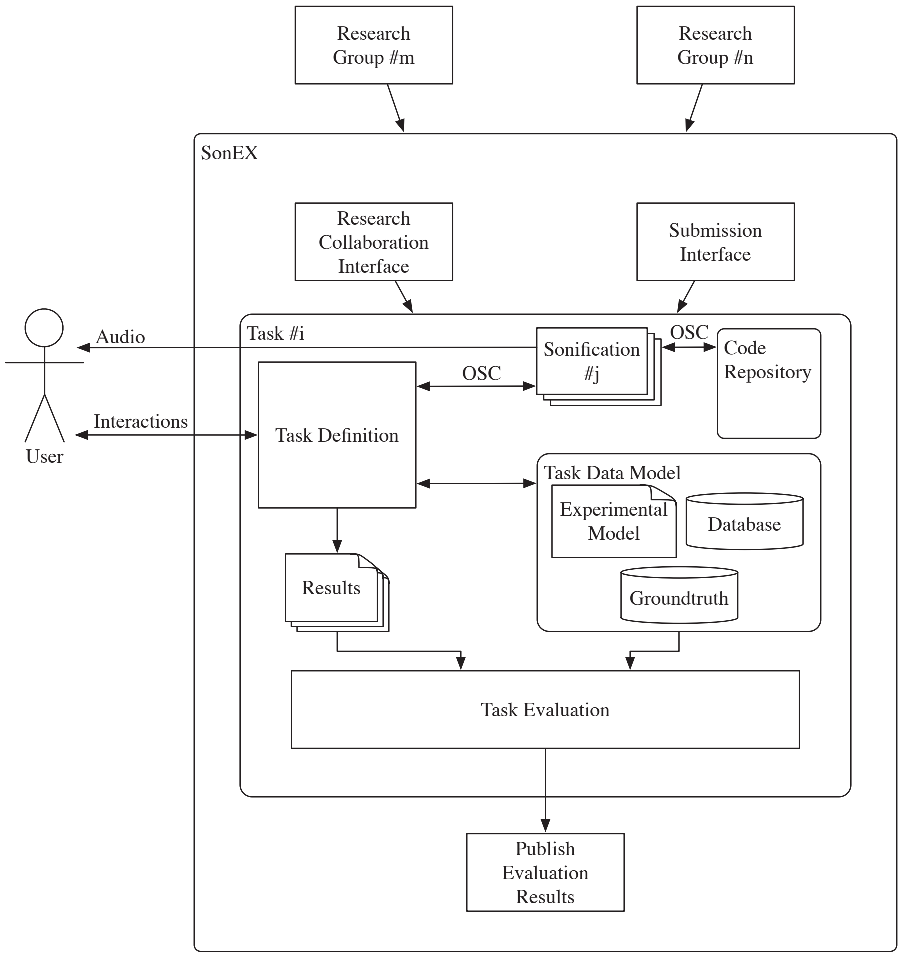

Figure 2: Architecture of the Sonification Evaluation eXchange framework (SonEX) for reproducible sonification research.

Depending on the task to be performed, a more detailed protocol could be needed.

After agreeing on the description, database and evaluation, researchers submit their contributions and the (public) evaluations of sonifications can start. The results should then be made publicly available and discussed in a panel discussion of the ICAD conference.

Note that the task has been defined so as researchers can independently develop their algorithms without any restriction on the sonification technique. They could, for example, use either a spherically expanding shock wave scan as in the data sonogram sonification model [24] or a radar mode for the sonification of the positions of the obstacles [25]. Also, there is a well defined set of evaluation measures that can be used to classify algorithms according to their performance. As already discussed in Section 2.4, this information can be used for discarding very bad performing algorithms when going for a real implementation of a ETA.

4. CONCLUSION

In this paper we have proposed and discussed a benchmarking framework to help the ICAD community to raise the quality of current research. SonEX (Sonification Evaluation eXchange) is a community-based platform that enables the definition and evaluation of standardized tasks for the formal comparison of sonification methods, supporting open science standards and reproducible research in the context of ICAD. This would allow the ICAD community to build upon each other's work and invest more time developing new methods and combining with the existing techniques than recreating existing methods.

The goals and architecture of this evaluation platform have been discussed. When building such a platform, many challenges have to be faced. In particular, the definition of objective tasks, computer sonification equivalents, database availability and subjective evaluation are specially important. To provide the reader with a clearer view of SonEX, a selected example of a task definition has been presented.

The submission web interface and evaluation has not been implemented yet, since this requires a lot of resources and first, the agreement and support of the ICAD community. The implementation of the submission system and its ongoing optimization for reproducible sonification research is proposed as our future work.

We believe SonEX could be a strong driver of research, encouraging the Auditory Display community to clearly define tasks, research goals and standardized evaluation measures that enable

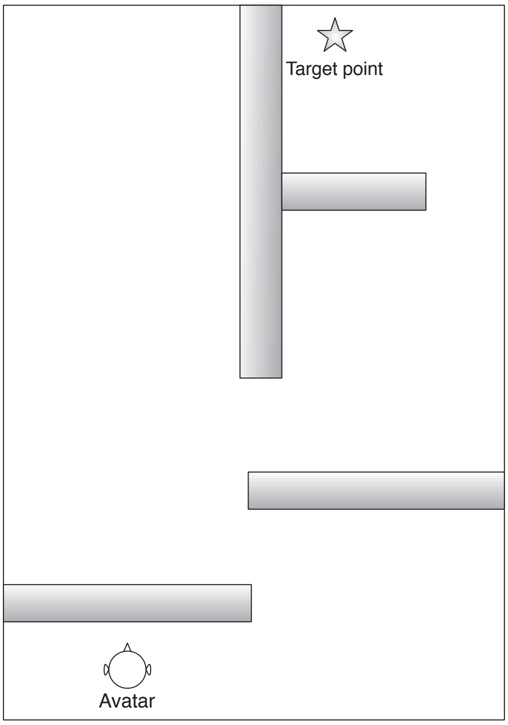

Figure 3: The Walking Game example.

formal and statistically based state-of-the-art comparison of algorithms. Therefore, to arrive at the best possible definition of tasks, standards and evaluation methods, we invite you to share your opinions, ideas, data and methods. We look forward to a fruitful discussion.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] T. Hermann, A. Hunt, and J. G. Neuhoff, "Introduction," in *The Sonification Handbook*, T. Hermann, A. Hunt, and J. G. Neuhoff, Eds. Berlin, Germany: Logos Publishing House, 2011, ch. 1, pp. 1–6.

[2] M. Ballora, "Opening your ears to data," in *TEDxPSU Talks at Penn State University*, November 2011. [Online]. Available: https://www.youtube.com/watch?v=aQJfQXGbWQ4

[3] T. Hermann, "Taxonomy and definitions for sonification and auditory display," in *Proceedings of the 14th International Conference on Auditory Display (ICAD 2008)*, P. Susini and O. Warusfel, Eds. IRCAM, 2008.

[4] K. Vogt, "A quantitative evaluation approach to sonifications," in *International Conference on Auditory Display*, 2011.

[5] N. Schaffert, R. Gehret, and K. Mattes, "Modeling the rowing stroke cycle acoustically," *J. Audio Eng. Soc*, vol. 60, no. 7/8, pp. 551–560, 2012.

[6] http://www.cost-sid.org/.

[7] N. Schaffert, K. Mattes, S. Barrass, and A. O. Effenberg, "Exploring function and aesthetics in sonifications for elite sports," in *The Second International Conference on Music Communication Science*, vol. 5954, December 2009, pp. 465–472.

[8] G. Dubus, "Evaluation of four models for the sonification of elite rowing," *Journal on Multimodal User Interfaces*, vol. 5, pp. 143–156, 2012.

[9] J. Gillard and M. Schutz, "Impoving the efficacy of auditory alarms in medical devices by exploring the effect of amplitude envelope on learning and retention," in *Proceedings of the International Conference on Auditory Display (ICAD)*, 2012.

[10] T. Hermann, B. Ungerechts, H. Toussaint, and M. Grote, "Sonification of pressure changes in swimming for analysis and optimization," in *Proceedings of the International Conference on Auditory Display (ICAD)*, 2012, pp. 60–67.

[11] G. Parseihian, B. F. Katz, and S. Conan, "Sound effect metaphors for near field distance sonification," in *Proceedings of the International Conference on Auditory Display (ICAD)*, 2012.

[12] D. Pirrò, A. Wankhammer, P. Schwingenschuh, R. Höldrich, and A. Sontacchi, "Acoustic interface for tremor analysis," in *Proceedings of the International Conference on Auditory Display (ICAD)*, 2012.

[13] H. Terasawa, J. Parvizi, and C. Chafe, "Sonifying ecog seizure data with overtone mapping: a strategy for creating auditory gestalt from correlated multichannel data," in *Proceedings of the International Conference on Auditory Display (ICAD)*, 2012.

[14] G. Kramer, "A letter from Greg Kramer: founder of ICAD," in *Proceedings of the International Workshop on Interactive Sonification*, January 2004.

[15] P. Vandewalle, J. Kovacevic, and M. Vetterli, "Reproducible research in signal processing," *Signal Processing Magazine, IEEE*, vol. 26, no. 3, pp. 37–47, may 2009.

[16] E. Voorhees and D. Harmon, "The text retrieval conference," in *TREC Experiment and Evaluation in Information Retrieval*. MIT Press, 2005.

[17] J. S. Downie, "The music information retrieval evaluation exchange (2005–2007): A window into music information retrieval research," *Acoustical Science and Technology*, vol. 29, no. 4, pp. 247–255, 2008.

[18] J. S. Downie, A. F. Ehmann, M. Bay, and M. C. Jones, "The music information retrieval evaluation exchange: Some observations and insights," in *Advances in Music Information Retrieval*, ser. Studies in Computational Intelligence, Z. W. Ras and A. Wieczorkowska, Eds. Springer, 2010, vol. 274, pp. 93–115.

[19] J. Williamson and R. Murray-Smith, "Sonification of probabilistic feedback through granular synthesis," *MultiMedia, IEEE*, vol. 12, no. 2, pp. 45 – 52, april-june 2005.

[20] M. Bujacz, P. Skulimowski, and P. Strumillo, "Navitona prototype mobility aid for auditory presentation of three-dimensional scenes to the visually impaired," *J. Audio Eng. Soc*, vol. 60, no. 9, pp. 696–708, 2012.

[21] B. Höferlin, M. Höferlin, B. Goloubets, G. Heidemann, and D. Weiskopf, "Auditory support for situation awareness in video surveillance," in *Proceedings of the International Conference on Auditory Display (ICAD)*, 2012.

[22] V. X. Nguyen, "Circosonic: a sonification of circos, a circular graph of pair wise table data," in *Proceedings of the International Conference on Auditory Display (ICAD)*, 2012.

[23] A. D. N. Edwards, "Auditory display in assistive technology," in *The Sonification Handbook*, T. Hermann, A. Hunt, and J. G. Neuhoff, Eds. Berlin, Germany: Logos Publishing House, 2011, ch. 17, pp. 431–453. [Online]. Available: http://sonification.de/handbook/chapters/chapter17/

[24] T. Hermann, "Sonification for exploratory data analysis," Ph.D. dissertation, Bielefeld University, Bielefeld, Germany, February 2002.

[25] D. El-Shimy, F. Grond, A. Olmos, and J. Cooperstock, "Eyes-free environmental awareness for navigation," *Journal on Multimodal User Interfaces*, vol. 5, pp. 131–141, 2012.