# Using Mobile Serious Games in the Context of Chronic Disorders
## A Mobile Game Concept for the Treatment of Tinnitus


Marc Schickler, Rüdiger Pryss, Manfred Reichert, Johannes Schobel
*Ulm University*
*Institute of Databases and Information Systems*
*{marc.schickler,ruediger.pryss,manfred.reichert}@uni-ulm.de, johannes.schobel@uni-ulm.de*

Berthold Langguth, Winfried Schlee
*University of Regensburg*
*Clinic and Policlinic for Psychiatry and Psychotherapy*
*berthold.langguth@medbo.de, winfried.schlee@gmail.com*



***Abstract*—Tinnitus ("ringing in the ear") is characterized by the perception of a sound in the absence of a corresponding acoustic stimulus. While many affected people habituate to the phantom sound, others are severely bothered and impaired in their quality of life. It is assumed that the latter group is characterized by a deficient noise cancelling mechanism in the brain. To train tinnitus patients to focus on target sounds and hence to suppress irrelevant background sounds, we developed a mobile serious game application, which is presented in this paper. The application runs on three mobile operating systems. We describe its goals and architecture as well as results from an evaluation study. Study results indicate that the gaming approach is feasible for training affected patients in focusing on directional hearing and, thereby, to suppress their tinnitus. Compared to traditional hearing training, advances of this approach are anytime availability, higher enjoyment, immediate feedback, and the option to stepwise increase game difficulty. From this, we expected an increased patient motivation and adherence as well as improved training and learning effects.**



***Keywords*-Mobile Serious Game, Personalized Healthcare, Patient Feedback, Tinnitus, Mobile Healthcare Assistance**


## I. Introduction

Brain disorders characterized by neuroplastic changes can be potentially treated with training procedures that either reduce pathological brain changes or enhance compensatory mechanisms. Example of such training procedures is rehabilitative training after stroke or logopedic treatment in stuttering. Chronic tinnitus, the perception of a sound in the absence of a corresponding acoustic stimulus, is a frequent disorder, which is also characterized by neuroplastic alterations in the brain [1], [2], [3]. While many affected people learn to suppress their tinnitus, others continuously perceive its sound and are hence severly impaired in their quality of life. It is assumed that in the latter group an inhibitory mechanism in the brain, the so-called noise cancelling system, is dysfunctional [4]. Hearing training approaches, therefore, have been developed to train patients to better suppress their tinnitus. Recent developments include training procedures for the localisation and selective attention to sounds [5]. An improved ability to selectively pay attention to localized sounds as well as to filter out irrelevant background noise should be an appropriate training procedure for strengthening a deficient "noise cancellation system".

Further requirements for an effective training device include anytime availability and patient enjoyment to increase motivation and learning effects. To meet these requirements, we developed a novel kind of mobile application taking principles from serious gaming into account. In particular, we adopted the concept presented in *Audio Defence*[1], where players must fight against opponents without having any visual information. Instead, they acoustically determine the direction of the opponent based on sounds generated by *Audio Defence*. Using these acoustic information, players have to take countermeasures to defend themselves.

To train the directional hearing ability of patients, in the newly realized game patients act as photographers instead of fighters (cf. Fig. 1). More precisely, patients must detect animals based on corresponding sounds. In this context, the following procedure will be applied: The application generates a virtual environment (e.g., farm) for the patient. For each animal to be detected, its position is randomly calculated (i.e., angle and distance; cf. Fig. 1 ①&②).

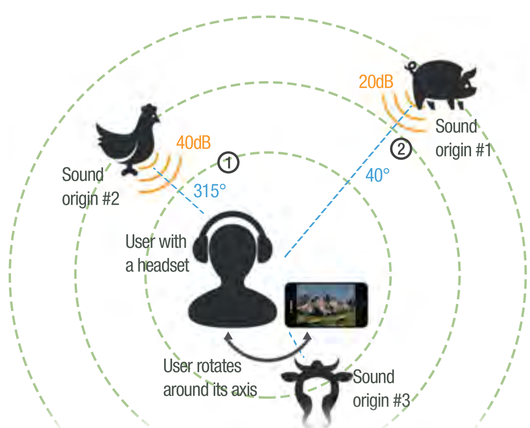



Figure 1: Audio Defence Principle

[1] http://www.audiodefence.com/

Patients must now relate their own position to the one of the animal. Therefore, they change the heading of their smart mobile device as long as they assume animals being directly in front of them. To verify the animal's position, users take a virtual picture. If the latter shows the animal, they succeed, otherwise they fail.

Using this kind of mobile healthcare application in practice, we unveiled several new aspects (cf. Fig. 2). First, different mobile operating systems must be supported (cf. Fig. 2③). Second, a flexible and smart collection of patient sensor data is crucial (cf. Fig. 2①). On one hand, it should consider requests from medical experts in a flexible way (cf. Fig. 2①); e.g., the latter requested more information about the detection procedure (e.g., on the time required for detecting an animal; cf. Fig. 2②). On the other, the different mobile operating systems revealed specific pecularities to be taken into account. For example, *Android* provides no built-in support for required 3D audio features. Instead, external libraries have to be used. In turn, the latter must be carefully considered in the context of collected data.

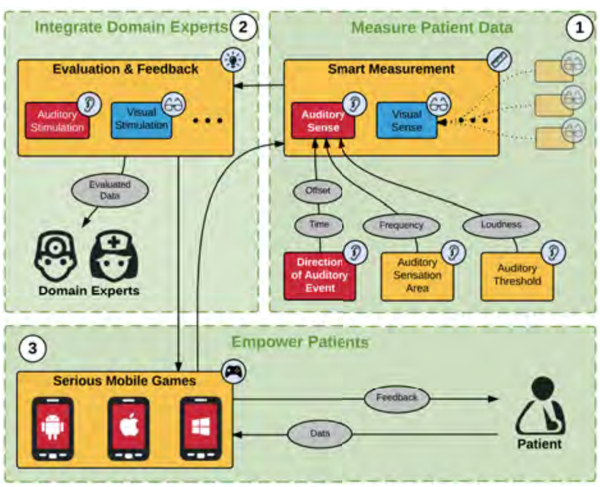



Figure 2: Overall Goal

Based on the lessons learned, we contribute fundamental technical aspects[2] that are crucial for the development of a generic platform empowering patients with the the help of serious games to foster their collaboration with clinicians. The remainder of this paper is organized as follows: Section II discusses fundamental technical aspects of the realized serious game. In particular, binaural audio and the head-related transfer function are presented. In Section III, we illustrate the developed architecture and discuss the lessons learned when implementing the mobile applications. Section IV illustrates study results evaluating the serious game in practice. Section V discusses related work, while Section VI concludes with a summary and an outlook.

[2]Fig. 2 highlights discussed aspects in *red*

## II. Game Fundamentals

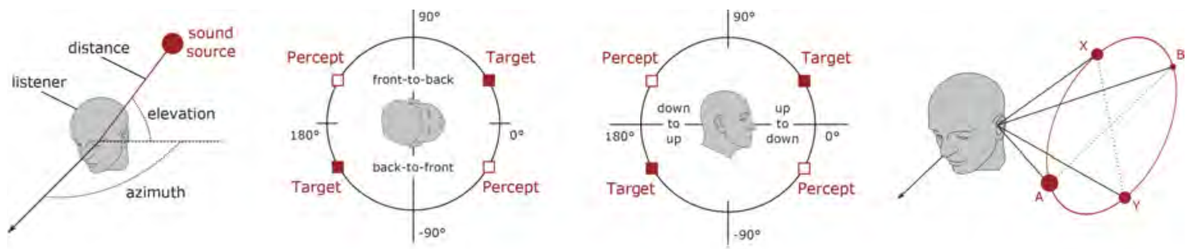
Two fundamentals related to 3D audio are required when realizing the mobile serious game: *binaural listening* and the *head-related transfer function*. The latter is relevant when providing the mobile application on the different mobile operating systems.

### A. Binaural Listening

In order to localize the source of sound-emitting virtual objects in the game, we must handle complex calculations with respect to binaural listening. The latter provides humans with the ability to localize sound signals. In this context, our brain analyzes time and intensity variations caused by the different positions of our ears. The overall analysis procedure, in turn, is denoted as binaural listening. In the following, we discuss the notion of time and intensity difference, which are both crucial regarding the technical implementation of the serious game.

Time difference is denoted as *interaural time difference* (ITD). Sound signals generated by a particular object will be perceived by each ear. Since our ears have different positions, there is a delay between the two points in time each ear detects a respective sound signal. Accordingly, intensity difference is denoted as *interaural intensity difference* (IID). Again, the different positions of our ears cause a difference in the detected intensity of the sound signal. For example, the ear being next to the sound source, detects a higher intensity of the signal. In this context, a spherical coordinate system needs to be established between the user and the virtual sound objects. In general, the position of each virtual sound object is defined by its *azimuth*, *elevation*, and *distance* to the user (cf. Fig. 3a). Azimuth and elevation constitute angular measurements inside the spherical coordinate system between the player and the virtual sound object. Azimuth is related to the horizontal field of view, while elevation is related to the vertical field of view (cf. Fig. 3a).

Interaural time difference as well as interaural intensity difference might cause localization errors (cf. Figs. 3b+c). Errors, in turn, are categorized as Azimuth and Elevation confusion. Both error types are solely related to one hemisphere of the user. In case of azimuth confusion, players cannot distinguish whether objects are in front of their view or in their back. Confusion is caused by the fact that ITD and IID are equal for both objects, which have the same distance to the player (cf. Fig. 3b). Accordingly, elevation confusion is related to the same phenomenon regarding the vertical field of view of the user; i.e., objects above or below the user have the same ITD and IID (cf. Fig. 3c). Azimuth and elevation confusion are summarized to cone of confusion (cf. Fig. 3d). Note that all points at the edge of the cone depicted in Fig. 3d have the same ITD and IID. To deal with the cone of confusion, in turn, our brain determines numerous additional parameters. For example, frequency variations between our ears that are caused by sound reflections.




(a) Spherical coordinate system (b) Azimuth confusion (c) Elevation confusion (d) Cone of confusion

Figure 3: Fundamentals of Binaural Audio

Examples of sources for reflections include our shoulders, the outer ear, or the head. Due to lack of space, we omit a detailed discussion of other parameters.

### B. Head-Related Transfer Function

To simulate interaural time difference, interaural intensity difference as well as other parameters (e.g., frequency variations for virtual objects), a digital filter called Head-Related Transfer Function (HRTF) is used. Each ear has its own HRTF. The latter, in turn, relies on two convolution operations applied to sound signals. In particular, the operations apply modifications to the sound signal of virtual objects before the signal is perceived by the ears. Modifications, for example, consider the outer ear, the head of the player, and the relative position of the latter to the sound source.

Since the physiognomy of the head varies for individuals, the development of a universal HRTF is quite complex. Ideally, each player has an individual HRTF. In practice, however, such an approach is not feasible [6], i.e., a more general approach is required. The MIT Media Lab, for example, developed a feasible and generic HRTF. Note that many other HRTF realizations exist [7].

As alternative HRTF realizations may be used, it must be determined which one is suitable in a given scenario. The HRTF used in the mobile serious game is a salient factor regarding overall game experience. Note that an important goal of our approach is to support different mobile operating systems. Therefore, amongst other important implemention aspects, we analyzed the built-in HRTF support for the considered mobile operating systems.

## III. Game Architecture

Fig. 5 illustrates the realized version of the game. Recall that users must detect animals solely based on their sounds. We restrict our discussion to several important technical as well as practical issues:

1) Players must configure two parameters: the number of animals and the existence or absence of background sounds (e.g., a farm). In the clinical application version we are developing for tinnitus patients, the latter option will allow for background sounds fitting to the individual's tinnitus.
2) When playing, the user cannot see the animals. More precisely, the application shows a black screen to the player. The target of the game is to detect all configured animals by taking pictures.
3) If players assume an animal being in their virtual field of view (cf. Fig. 5), they may take a picture. To change their field of view, players must rotate their body like in real-life holding the smart mobile device in front of them.
4) If the animal was in their field of view when taking the picture, the latter will show the animal. Otherwise, only a panorama is displayed. Note that with every picture taken, two additional details are provided to the user: (1) The angle offset between the centre of the field of view (i.e., $0°$) and the position of the animal.[3] (2) The duration required to take the picture.
5) The players succeeds if all animals are detected.

To realize the game on different mobile platforms, a sophisticated architecture was required (cf. Fig. 4): The *repository* component enables scenario and level management. The *feedback & evaluation* component, in turn, allows configuring feedback options for players and patients respectively. This enables IT experts, for example, to integrate evaluation algorithms more easily. Medical experts, in turn, may configure which parameters shall be included for the given feedback (e.g., duration to take a picture). The *data logger* component stores all collected data in a database. Furthermore, the *sensor and device management* component covers technical issues of required mobile sensors with respect to the pecularities of the different mobile platforms. Finally, to realize the *3D audio* features of the game, a separate component became necessary.

Technically, three particular challenges need to be tackled to cope with the requirements of the medical experts. The first one is related to the mobile sensors, while the second and third challenges are related to 3D audio aspects.

*First*, data collected by the mobile device sensors must be comparable (i.e., similar precision) among the different mobile operating systems. To be able to ensure this, two

[3]Note that the difference is $45°$ in Fig. 5

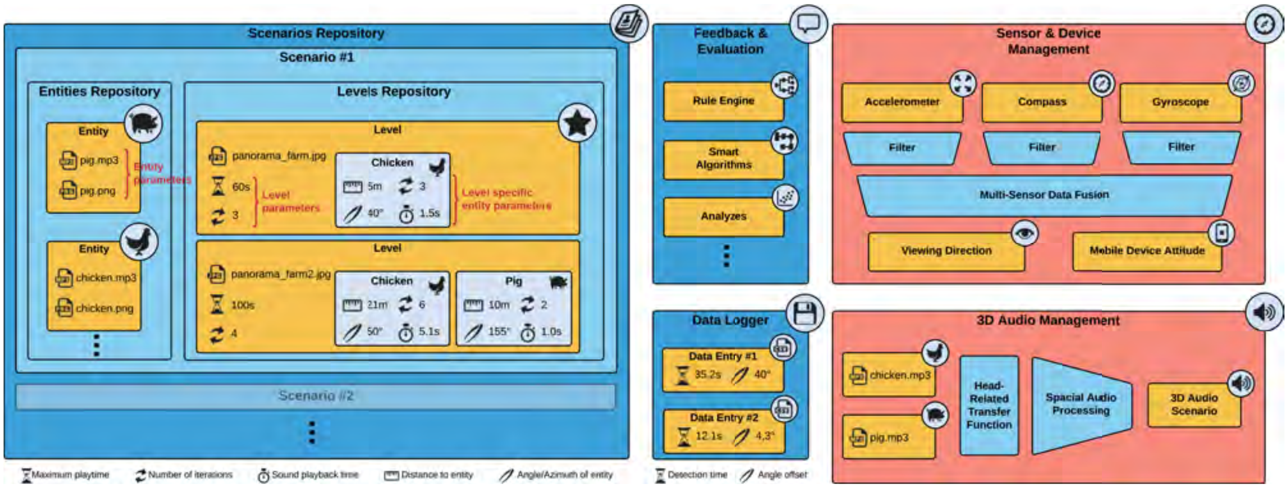



Figure 4: Game Architecture

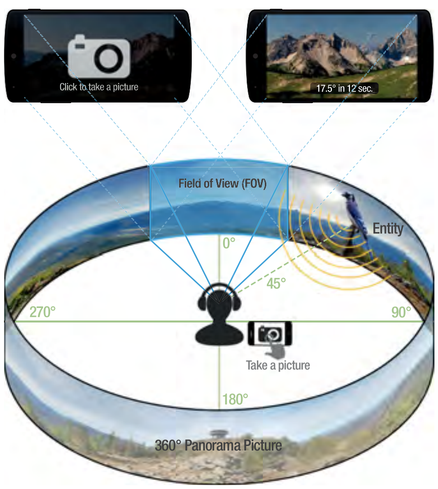



Figure 5: Game Concept

aspects had to be addressed:

1) We had to apply several filters to gathered sensor data (e.g., a low-pass and kalman filter). Most importantly, one of these filters mitigates sensor noise and, therefore, enables players to precisely point the device to the animal target.
2) The built-in support of the sensor APIs must be evaluated to provide similar game behaviour on all mobile operating systems.[4] Interestingly, all operating system vendors rely on an advanced *multi-sensor data fusion* approach. However, they revealed several important differences. For example, Android and Windows Phone enable developers to manually select required sensors, while iOS does not offer such feature.

[4]Note, that this is a crucial requirement for using the data collected with the mobile serious game in the context of clinical studies.

*Second*, to provide appropriate 3D audio features on different mobile operating systens, complex considerations became necessary. For example, Windows Phone has an easy-to-use 3D audio built-in framework, whereas, several other features cannot be simply adjusted; e.g., the head-related transfer function (HRTF) could not be manipulated.

HRTF constitutes the *third* challenge. In this context, iOS provides a 3D audio built-in framework. However, iOS enables adjustments of the head-related transfer function with a limited extent compared to Windows Phone. Conversely, Android provides no appropriate 3D audio built-in framework. Instead, external libraries must be used (e.g., OpenAL-MOB). In turn, only Android enabled us to flexibly change the head-related transfer function. On the other hand, to establish 3D audio on Android, the required efforts are higher compared to Windows Phone and iOS. Facing this heterogeneity of built-in functions, the goal to ensure the same 3D audio experience on all mobile operationg systems was challenging.

Altogether, all mobile platforms are appropriate for the game concept. However, the efforts that need to be spent in the context of the various mobile operation systems vary significantly (cf. Fig. 6).

## IV. Game Evaluation

To demonstrate the applicability of the developed mobile serious game, we conducted a user study. The latter included 24 subjects with different experiences. Note that two of the subjects indicated that they were affected by tinnitus.

Each subject had to accomplish 3 different levels with all mobile operating systems (i.e., 9 games in total). In the first level, a bird (cf. Fig. 7) had to be detected and no background noise was used. In the second level, a bird and a frog (cf. Fig. 7) were targets, again with no background

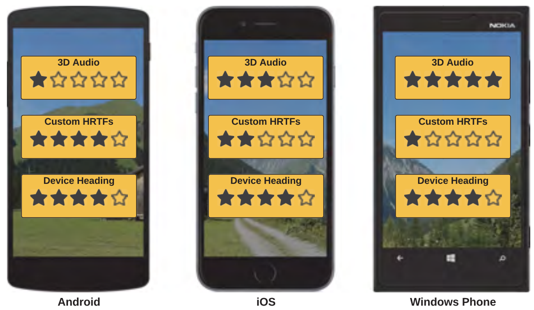



Figure 6: Built-in Support of Mobile Platforms

noise.[5] The third level is similar to the second one, except that background noise was used. Note that players wore stereo headphones while playing. To conduct the study, we recorded the duration required to take a picture as well as the angle offset of the user to the actual position of the target after taking the pictures.

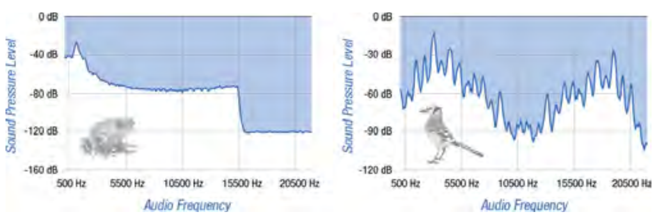



Figure 7: Frequency Spectrums of Frog and Bird

In the study, we applied a linear mixed effect model to evaluate differences between the three mobile platforms. We parameterized the study along the linear mixed effect model as follows:

- Fixed effects are the three levels, the targets, and the mobile platforms. Note that the fixed effects were evaluated in combination with each other as well.
- Random effects are limited to the parameter subject.

Based on the linear mixed effect model, we calculated an analysis of variance (ANOVA) with respect to gathered data (cf. Table I) along the dimensions *duration to take a picture* and the *angle offset to the target*. For the first dimension, our calculations revealed a strong significance for both the mobile platforms ($F(2, 1082) = 44.05$, $p < 0.0001$) and the target type ($F(1, 1082) = 20.71$, $p < 0.0001$). Furthermore, while the played levels are not significant ($F(2, 1082) = 1.60$, $p = 0.203$), the interaction between mobile platform and played levels are significant ($F(4, 1082) = 3.01$, $p = 0.017$).

Regarding the second dimension, again the mobile platforms revealed a strong significance ($F(2, 1082) = 8.90$, $p = 0.0001$). In turn, the played levels ($F(2, 1082) = 0.91$, $p = 0.401$), the targets ($F(1, 1082) = 0.02$, $p = 0.882$), and the interaction between mobile platform and played levels ($F(4, 1082) = 1.74$, $p = 0.138$) are not significant.

| Effect | DF | Time | | Offset | |
|---|---|---|---|---|---|
| | | $F$ | $p$ | $F$ | $p$ |
| OS | $2, 85$ | $33.55$ | $< 0.0001$ | $10.67$ | $< 0.0001$ |
| Level | $1, 85$ | $0.60$ | $0.438$ | $0.11$ | $0.745$ |
| Target | $1, 85$ | $24.76$ | $< 0.0001$ | $0.020$ | $0.889$ |
| OS×Level | $2, 85$ | $4.28$ | $0.014$ | $1.46$ | $0.233$ |
| OS×Target | $2, 85$ | $0.90$ | $0.407$ | $9.24$ | $0.0001$ |
| Level×Target | $1, 85$ | $1.54$ | $0.214$ | $1.52$ | $0.218$ |
| OS×Level×Target | $2, 85$ | $3.88$ | $0.021$ | $0.23$ | $0.796$ |

Table I: Analysis of Variance

Overall, the analysis revealed that users show different results with respect to the different mobile platforms. To elaborate the significance, consider the following two observations. First, Fig. 8 shows the average time to take a picture by the user. On iOS, the average playing time was 9.8 s with a standard deviation of 6.1 s, while on Android the playing time was 12.1 s and the standard deviation 7.4 s. Finally, on Windows Phone the average playing time was 15.8 s with a standard deviation of 9.0 s. Despite these playing time differences, all mobile platforms were appropriate for the clinical context.

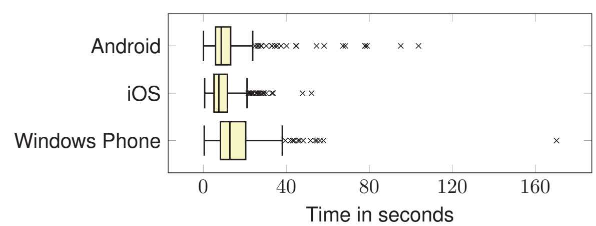



Figure 8: Measured Playing Time

Second, Fig. 9 shows the offset angles after taking pictures. On iOS, the best results with an average of 13.3° and a standard deviation of 18.8° were achieved. Android revealed comparable results with an average of 20.6° and a standard deviation of 19.3°. Finally, Windows Phone revealed an average of 21.1° and a standard deviation of 18.9°.

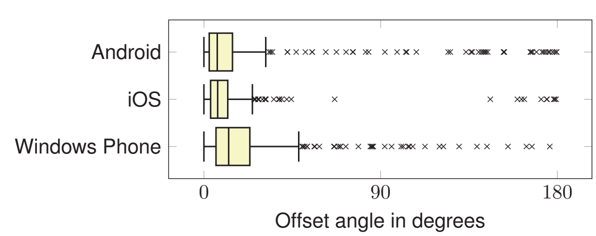



Figure 9: Measured Angle Offset

We present a more detailed perspective on the average time to take a picture as well as the average angle offset in

[5] Note that we used a different frequency spectrum for the animals in order to increase the value of the conducted study.

Table II. Note that the results revealed a significant difference between the detection of the two targets. Altogether, the analysis revealed comparable results with respect to the different mobile platforms. However, significant differences still exist for several aspects. For example, on iOS the average detection time for the two targets as well as the angle offset vary significantly. Currently, we are working on respective configuration features, to account for these differences.

| OS | Target | Time | | Offset | |
|---|---|---|---|---|---|
| | | Average | SD | Average | SD |
| All | Bird | 11.5 s | 9.1 s | 17.8° | 25.4° |
| | Frog | 14.2 s | 9.8 s | 19.1° | 22.6° |
| Android | Bird | 11.3 s | 11.4 s | 23.6° | 37.5° |
| | Frog | 13.4 s | 11.8 s | 16.1° | 21.5° |
| iOS | Bird | 8.5 s | 4.8 s | 11.3° | 16.6° |
| | Frog | 11.8 s | 7.3 s | 16.2° | 21.5° |
| Windows Phone | Bird | 14.8 s | 8.9 s | 18.6° | 14.0° |
| | Frog | 17.5 s | 9.0 s | 25.0° | 24.1° |

Table II: Average Playing Time and Angle Offset

## V. Related Work

Mobile serious games have already been considered in the context of personalized healthcare [8], [9]. Furthermore, they are used to determine vital parameters in various life situations [10]. Regarding our context, several mobile approaches exist for providing feedback to the general hearing ability. However, these approaches mainly focus on a person's hearing sensitivity at different frequencies [11]. Recently, a computer-based game was presented for sound localization training for tinnitus patients [5]. However, no mobile approach has been developed yet. As presented for directional hearing, new requirements must be considered compared to these approaches. For example, 3D audio features must be realized on different mobile operating systems.

## VI. Summary & Outlook

The overall goal of our research was to develop a generic framework for serious games allowing for clinical interventions in the auditory domain. The framework was evaluated in the context of an application for treating chronic tinnitus.

Using mobile serious games offers promising perspectives for supporting targeted rehabilitative training. First, patient motivation can be significantly increased by using serious games instead of conventional training procedures. Second, patients get immediate feedback to their individual performance, which, in turn, provides further motivation. High levels of motivation and enjoyment are critical for adherence to training and learning effects.

In the context of the developed mobile game, the conducted study revealed that users are able to localize sounds with appropriate accuracy. Furthermore, the results showed that platform-specific adjustments must be performed to enhance user experience as well as overall accuracy. Regarding overall technical implementation, the considered mobile platforms are generally capable of measuring the hearing performance of patients. For this purpose, several sensors were integrated. Additionally, complex calculations on the mobile devices became necessary. In general, a platform must carefully consider platform-specific differences. For example, the built-in support for required 3D audio features differs significantly. Based on these lessons learned, we are developing a generic platform that enables clinicians to configure required mobile serious games in the absence of IT experts to enable further individualization. In particular, they shall be enabled to specify required parameters of mobile games on their own.

## References


[1] A. Elgoyhen, B. Langguth, D. De Ridder, and S. Vanneste, "Tinnitus: perspectives from human neuroimaging," *Nature Reviews Neuroscience*, 2015.

[2] R. Pryss, M. Reichert, B. Langguth, and W. Schlee, "Mobile crowd sensing services for tinnitus assessment, therapy, and research," in *Int'l Conf on Mobile Services*. IEEE, 2015, pp. 352–359.

[3] R. Pryss, M. Reichert, J. Herrmann, B. Langguth, and W. Schlee, "Mobile crowd sensing in clinical and psychological trials–a case study," in *28th Int'l Symposium on Computer-Based Medical Systems*. IEEE, 2015.

[4] J. Rauschecker, A. Leaver, and M. Mühlau, "Tuning out the noise: limbic-auditory interactions in tinnitus," *Neuron*, vol. 66, no. 6, pp. 819–826, 2010.

[5] K. Wise, K. Kobayashi, and G. Searchfield, "Feasibility study of a game integrating assessment and therapy of tinnitus," *Journal of Neuroscience Methods*, vol. 249, pp. 1–7, 2015.

[6] E. M. Wenzel, M. Arruda, D. J. Kistler, and F. L. Wightman, "Localization using nonindividualized head-related transfer functions," *The Journal of the Acoustical Society of America*, vol. 94, no. 1, pp. 111–123, 1993.

[7] V. Algazi, R. Duda, D. Thompson, and C. Avendano, "The cipic hrtf database," in *Workshop on the Applications of Signal Processing to Audio and Acoustics*. IEEE, 2001, pp. 99–102.

[8] S. McCallum, "Gamification and serious games for personalized health," *Stud Health Technol Inform*, vol. 177, pp. 85–96, 2012.

[9] P. Klasnja and W. Pratt, "Healthcare in the pocket: mapping the space of mobile-phone health interventions," *Journal of biomedical informatics*, vol. 45, no. 1, pp. 184–198, 2012.

[10] S. Hardy, A. El Saddik, S. Göbel, and R. Steinmetz, "Context aware serious games framework for sport and health," in *Medical Measurements and Applications Proceedings*. IEEE, 2011, pp. 248–252.

[11] A. Kam, J. Sung, T. Lee, T. Wong, and A. van Hasselt, "Clinical evaluation of a computerized self-administered hearing test," *Int'l Journal of Audiology*, vol. 51, no. 8, pp. 606–610, 2012.