AN AUDIO GAME APP USING INTERACTIVE MOVEMENT SONIFICATION FOR TARGETED POSTURE CONTROL

Daniel Avissar[1], Colby Leider[1], Christopher Bennett[1] and Robert Gailey[2]

[1] Frost School of Music, Music Engineering Group
[2] Miller School of Medicine, Department of Physical Therapy
University of Miami, Coral Gables, FL, 33124 USA
d.avissar@umiami.edu

ABSTRACT

Interactive movement sonification has been gaining validity as a technique for biofeedback and auditory data mining in research and development for gaming, sports, and physiotherapy. Naturally, the harvesting of kinematic data over recent years has been a function of an increased availability of more portable, high-precision sensory technologies, such as smart phones, and dynamic real time programming environments, such as *Max/MSP*. Whereas the overlap of motor skill coordination and acoustic events has been a staple to musical pedagogy, musicians and music engineers have been surprisingly less involved than biomechanical, electrical, and computer engineers in research efforts in these fields. Thus, this paper proposes a prototype for an accessible virtual gaming interface that uses music and pitch training as positive reinforcement in the accomplishment of target postures.

1. INTRODUCTION

For musicians, the notion that motor control development is aided immensely by audition is an apparent truism. Even in the absence of a musical instrument, a person is capable of whistling different pitches or tapping at different tempos by marking how subtle changes in their muscular configurations coincide with modulations in pitch, tempo, duration, velocity, and timbre. Here, an audio biofeedback (ABF) loop engages the senses of proprioception and audition, producing an acoustic reaction to biomechanical processes. For activities that are not inherently musical in nature, however, the inquiry into these processes may be augmented by *interactive movement sonification*, which refers to the mapping of proprioceptive information to synthesized acoustic events, and their subsequent auditory presentation in real time to the user [1].

Interactive movement sonification has already seen application in augmenting a wide range of physical activities, including electronic gaming, sports and exercise, and physiotherapy. In audio games such as Röber's Matrix Shot, users interact with virtual auditory environments in order to accomplish goal-oriented tasks through ABF [2]. Additionally, in a series of studies both indoor and outdoor rowing, users practicing indoor rowing were able to obtain motor precision more quickly when presented with real-time audiovisual biofeedback of their kinematic and dynamic motions than when presented with solely visual biofeedback [3] [4]. Recently, physiotherapists have begun to recognize movement sonification as an invaluable component in patient rehabilitation, as in the study by Chiari et al [5], which is examined in closer detail in this paper. In general, movement sonification increases one's proprioceptive awareness and accuracy in tasks of motor learning and coordination, and decreases the time spent in visual data mining [1].

Thus, the issue that remains is that which developers of earcons, auditory icons, spearcons, and musicons have been attempting to reconcile for over two decades: although digitized sound for use in sonification does not necessarily have to be *pleasant*, it most certainly cannot be *annoying* [6]–[9]. In this sense, target behavior may be intellectually motivated and reinforced by the successful completion of a musical task. Presented herein is VLimbo, an interactive audio game application, in which users tonally match a constant sine tone pulse with a pitch-shifted music file by moving their bodies about the trunk towards a target posture. It combines the accessibility of a smart phone's accelerometer with a two-part sonic user interface (SUI): the first part tests Chiari et al's study of posture correction as presented in [5]; the second part augments the Chiari algorithm in the VLimbo virtual gaming environment.

2. SYSTEM DESIGN

The interactive audio game prototype presented in this paper implements a sway-detection algorithm to produce ABF for the user. It comprises components similar to [5], with the exception that the data processing unit and audio output unit are consolidated into one unit by the *Max/MSP* programming environment. A signal flow of the project setup is illustrated in Fig. 1.

2.1. Sensory Unit

One of the design goals for the prototype was to utilize the data output of a common mobile phone's sensory hardware. The prototype accessed data output from the onboard Invensense MPU-6050 triaxial accelerometer, a hardware feature of the Samsung Galaxy SIII SGH-T999, with a range of up to $\pm 2\ g$, and 14-bit sensitivity. A sample of the device's measurements isolating each of the axes' gravitational acceleration (g) is provided in Table 1. The accelerometer was calibrated using the application native to the phone's Android 4.0.4 operating system. During its initial demonstration and testing, users held the phone to their chests with one hand, but the phone may also be adjusted into a light strap that can be placed either on the user's back [5], or chest.

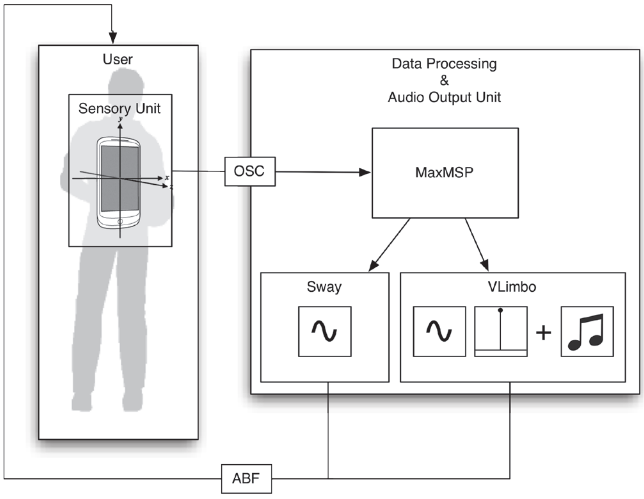


Figure 1: Signal flow diagram of the VLimbo interactive audio game environment.

Table 1: Invensense MPU-6050 g Measurements at 98.46 Hz Sample Rate (Ideal $g = 9.81m/s^2$)

Axis ($\pm$)	Average g (m/s^2)	Average Noise	Max Noise
$+x$	9.937	0.011	0.046
$-x$	9.818	0.011	0.042
$+y$	9.876	0.010	0.045
$-y$	9.646	0.010	0.059
$+z$	9.682	0.018	0.079
$-z$	10.23	0.017	0.069

2.2. Data Processing and Audio Output Unit

Data from the accelerometer was accessed and transmitted via an open-source application *Control*, which was available for free download on the Android Market and iTunes. *Control* allows users to create customizable, multilevel graphical user interfaces in JavaScript for communication with other electronic devices via the Open Sound Control (OSC) or MIDI protocols. Among the common widgets available in *Control*, the prototype used the accelerometer, multislider, and button widgets [10].

A central processing and audio output environment was programmed in *Max/MSP* [11] for rendering the data and human-computer interactions. When both the mobile device and the host computer were logged on to the same wireless network, data was sent as OSC messages and received by the *Max* patch's `udpreceive~` object, where it was parsed and routed respectively to the widget activated in the mobile interface. All accelerometer ranges were processed in the range of $[-90, 90]$ Cartesian degrees to correspond with the specifications outlined in [5].

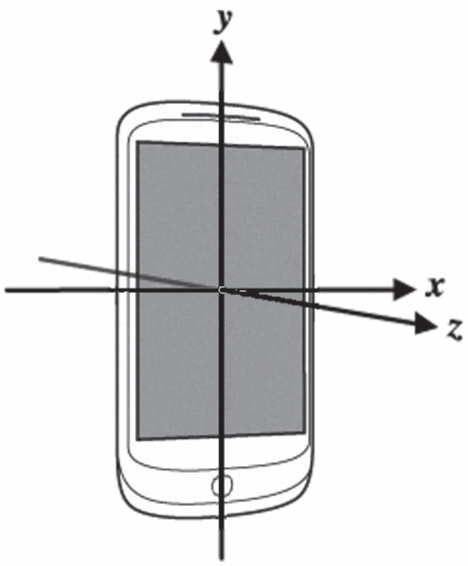


Figure 2: Accelerometer axis orientation relative to the mobile telephone used in the prototype.

2.3. Sonic User Interface

The *Max* patch provided two interfaces for audio rendering of posture and sway. The first SUI served as a training environment that reproduced the sway-measuring algorithm presented in [5]. When engaged, users could observe how slight changes in their posture produced changes in a continuous sine tone, which modulated by frequency when swaying in the anterior-posterior (AP) direction, and by stereo panning in the medial-lateral (ML) direction, aiming towards a target region of corrected posture. As in [5], this constant interaction with the ABF attempted to demonstrate the efficacy of sonification in controlling motor coordination.

The second SUI served as a virtual audio gaming environment similar to [1] and [2]. As in the first SUI, accelerometer data was used for the input, but the value of the target posture reference coefficient was varied, thereby requiring a different set of corrective interactions on behalf of the user. The rules of the VLimbo game were based on the classic party game, Limbo, and are outlined both in the *Max* patch help file, and in the following section of this paper. In general, perceived changes in frequency and stereo balance during the audition of a popular song, as compared to a reference (target) pitch played at the song's natural tempo, required users to adjust their posture accordingly in order to match the modulated song to the target pitch for a determined period of time.

3. IMPLEMENTATION

3.1. Posture Correction

The triaxial accelerometer degrees were input to the *Max* patch from the *Control* interface, and once parsed, were mapped similarly to [5]. Data from only the x and z accelerometer readings were sufficient in mapping to ML and AP, respectively; the y accelerometer data was omitted, as users were not required to perform any activities where they needed to rotate about this axis.

These parameters were used in adapting the posture correction algorithm presented in [5], with several adjustments. In particular, [5] defines a “reference region” (RR) at which the user will always exhibit a negligibly small amount of sway accelerations, which for their clinical purposes was relegated to the region within ± 1 degree from the user’s natural, upright stance. For the purposes of this study, however, in which the target balance point varied from trial to trial, and where clinical precision was not technically crucial to its success, the RR was given a modicum of flexibility. By default, RR was set to ± 5 degrees, but may be adjusted liberally.

Although the results from [5] were based on the triaxial accelerometer measurement technique as applied in [12], the sway-detecting algorithm required some technical adjustment for implementation in this study. For example, the equation for the resultant modulated frequency f of the pure sine tone as explicated in the former is:

$$f = m a_{AP} + f_0 \tag{1}$$

where $m = 0Hz$ inside RR, $m = 250Hz$ outside RR backward, $m = 600Hz$ outside RR forward, $f_0 = 400Hz$ at RR, and a_{AP} is the AP acceleration. The algorithm implemented in this study produced fewer errors after empirically adjusting the equation to:

$$\hat{f} = m\alpha_{AP} + f_0 \tag{2}$$

where α_{AP} corresponds to the z accelerometer reading scaled to a value within the range of $[-0.5, 0.5]$.

An equation is also provided for modulating the stereo panning effect as a function of the ML acceleration [5]. This study, however, implemented an equal power panning algorithm as a function of the x accelerometer reading, which is scaled to the coefficient α_{ML} within the range $[0, 1]$:

$$\begin{aligned} \hat{S}_L &= S_L \cos(\alpha_{ML}\frac{\pi}{2}) && (3)\\ \hat{S}_R &= S_R \sin(\alpha_{ML}\frac{\pi}{2}) && (4)\\ \hat{S} &= \hat{S}_L + \hat{S}_R && (5) \end{aligned}$$

where S_L and S_R are the left and right input signals, and $\hat{S}$ is the output panned stereo signal [13].

3.2. VLimbo

The VLimbo SUI used the same panning algorithm as in the posture correction SUI to modulate the panning of the audio playback, but AP movement was utilized for a different conceptualization of the game of Limbo. As mentioned previously, traditional Limbo requires players to adjust their posture according to their visual perception of the modulated “reference” height of the Limbo stick, held horizontally above the ground. A successful turn is completed once a user maintains a perceived target posture without losing balance, thereby clearing the stick while walking underneath it. Typically, music is played in the background to accompany gameplay, but is usually not necessary to the completion of the physical task.

In VLimbo, a WAV-format audio file of the quintessential “Limbo Rock” by Chubby Checker (a song in A-flat major) was loaded into the *Max/MSP* patch, and a reference pitch corresponding to the song’s tonic note of A-flat was set using the `kslider` GUI object. Some musical knowledge was required in order to determine this reference pitch. Next, the `qm.btrack~` object [14] calculated the song’s beat pulses in real time and output a bang at every pulse, simultaneously triggering a MIDI note at the reference pitch to play on every beat. Empirical trials using “Limbo Rock” yielded pleasant results for the beat tracker, accurately and consistently hitting either the down or the up beats.

In traditional Limbo, the degree of difficulty of a particular turn is determined by the vertical height of the Limbo stick as it is held parallel to the ground. The degree of difficulty in VLimbo was determined as the difference between the pitch of the sine tone pulse and the modulated pitch of the playback audio’s tonic note. The pitch modulation factor p as the user moved in the AP direction was calculated empirically as

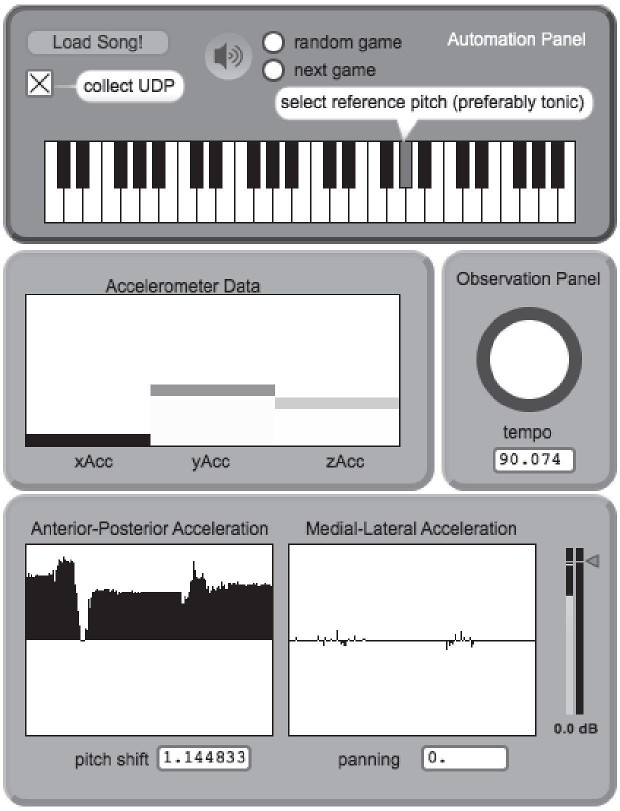
$$p = \frac{d}{50}\hat{\alpha}_{AP} \tag{6}$$

where $\hat{\alpha}_{AP}$ is the z accelerometer reading scaled to $[0.75, 1.25]$, and the degree of difficulty d is an integer within the range of $[41, 50]$. The difficulty d was reset at the beginning of each turn, and could be selected either linearly or randomly by triggering the corresponding button on the *Control* user interface. In the *Max/MSP* patch, p was passed as an input parameter to the `grainstretch~` object, along with the audio signal, thereby modulating the audio in pitch while maintaining its original tempo. With the mobile device either strapped onto the user, or slipped into a shirt pocket, the user compared the pulsing reference pitch to the tonic pitch of the audio playback, which was modulated by the pitch factor p. The user then moved in the AP direction until reaching the target posture, as in traditional Limbo, which occurred when p reached a value of 1.0 and the tonic pitch of the audio playback matched the reference pitch.

4. CONCLUSIONS AND FUTURE WORK

The purpose of VLimbo was to construct a framework that simultaneously demonstrated targeted motor control through the use of movement sonification as biofeedback, while testing its enjoyability as a motivating factor in the achievement of a physical task. It is important to emphasize that it has yet to be implemented in a study validating its effectiveness as a therapeutic tool, and its entertainment value as a game. While the quality of the sonic feedback as a motivating factor as compared to similar preceding studies is still under question, this paper does suggest that users may be better motivated to execute certain movements by feedback that does not *sound* like traditional feedback. Thus, the combination of the intellectually and physically challenging elements of altering musical elements of sound that is already musical in nature presents a basis for inquiry into musicianship as additional positive reinforcement in tasks of audition and sonification. When the user accomplished the task of matching the song to the reference pitch, in conjunction with the internal biochemical processes involved in successfully completing rather athletic feats of varying difficulty, a goal was positively reinforced without the aid of a coach or physiotherapist. The applications of this type of movement-targeted sonification in the realms of gaming, sports, and physiotherapy may be boundless, and have yet to demonstrate a significant scientific contribution.

As is depicted in Fig. 1, the individual components were written for separate pieces of software, each requiring a considerable amount of configuration. Thus, the first major improvement to the prototype will be porting the entire system to iOS and Android, which will be able to access the accelerometer data and process the audio directly on one software and hardware platform. Once ported into its own mobile application, the audio output may be



☒ collect UDP

tempo 90.074

pitch shift 1.144833

panning 0.

Figure 3: The VLimbo *Max/MSP* graphical user interface.

transmitted wirelessly, as in Airplay for Apple devices. More practically, future versions of the game will be accompanied by a timing algorithm that requires the user to maintain the target posture for a determined amount of time in order to pass a successful turn. In addition, the y accelerometer data may be implemented to increase the level of the game's acrobatic challenge. Looking further, the current demonstration assumes that the user possesses enough of a musical background to be able to identify the tonic pitch of the uploaded songs. Advances in pitch-detecting and cataloging algorithms should be able to compensate for this assumption, making the game accessible to a wider audience of users.

Finally, it should be noted that just as the end user's level of annoyance is taken into consideration in the design of other methods of sonification such as earcons and auditory icons, conscious efforts must be maintained in research on motion sonification to produce sounds that users will want to hear [1] [6] [7] [8]. Consideration of subjective sound sensitivity should also be taken when performing similar investigations in the field of physiotherapy, particularly when applied to motor control training of amputees and war veterans; members of this part of the population may be more likely to present signs of post-traumatic stress, and may perhaps respond differently to certain types of sonification [15]. In any case, it is a foreseeable certainty that the continued contribution of musicians and music engineers in addition to the current efforts procured mostly by biomechanical, electrical, and computer engineers for this type of research will be vital.

5. ACKNOWLEDGMENT

The author would like to thank Prof. Colby Leider, Prof. Robert Gailey, and Dr. Chris Bennett for their contributions and direction in this study. The author would also like to extend his gratitude to Einat and Gabi Kalinin for their boundless patience, support, and love.

6. REFERENCES

[1] T. Hermann, A. Hunt, and J. Neuhoff. *The sonification handbook*. Logos Verlag, 2011.

[2] N. Röber and M. Masuch. Leaving the screen: New perspectives in audio-only gaming. In *11th International Conference on Auditory Display*. Citeseer, 2005.

[3] A. Effenberg, U. Fehse, and A. Weber. Movement sonification: Audiovisual benefits on motor learning. *BIO Web of Conferences*, 1:00022, 2011.

[4] N. Schaffert, K. Mattes, and A. O. Effenberg. An investigation of online acoustic information for elite rowers in on-water training conditions. J. Hum. Sport Exerc., 6 (2):392-405, 2011.

[5] L. Chiari, M. Dozza, A. Cappello, F. Horak, V. Macellari, and D. Giansanti. Audio-biofeedback for balance improvement: an accelerometry-based system. *Biomedical Engineering, IEEE Transactions on*, 52(12):2108 –2111, dec. 2005.

[6] S. Brewster, P. Wright, and A. Edwards. A detailed investigation into the effectiveness of earcons. In *Santa Fe Institute Studies in the Sciences of Complexity-Proceedings*, volume 18, 2011.

[7] M. Bussemakers and M. Ab. When it sounds like a duck and it looks like a dog... auditory icons vs. earcons in multimedia environments. In *Proceedings of the International Conference on Auditory Display*, 2000.

[8] D. Palladino and B. Walker. Learning rates for auditory menus enhanced with spearcons versus earcons. In *Proceedings of the 13th International Conference on Auditory Display*, pages 274–279, 2007.

[9] R. McLachlan, M. McGee-Lennon, and S. Brewster. The sound of musicons: investigating the design of musically derived audio cues. 2012.

[10] C. Roberts. Control v 1.0, July 2011. `http://charlie-roberts.com/Control`.

[11] M. Puckette et al. Max 6 v 6.0.7, 2011.

[12] R. Mayagoitia, J. Lötters, P. Veltink, and H. Hermens. Standing balance evaluation using a triaxial accelerometer. *Gait & posture*, 16(1):55–59, 2002.

[13] V. Pulkki. Virtual sound source positioning using vector base amplitude panning. *J. Audio Eng. Soc*, 45(6):456–466, 1997.

[14] A. Robertson, A. Stark, and M. Plumbley. Realtime visual beat tracking using a comb filter matrix. In *Proc. International Computer Music Conference*, 2011.

[15] S. Orr, Z. Solomon, T. Peri, R. Pitman, et al. Physiologic responses to loud tones in Israeli veterans of the 1973 Yom Kippur War. *Biological Psychiatry*, 1997.