ZOOZbeat – Mobile Music reCreation

Gil Weinberg
Georgia Tech Center for Music Technology
840 McMillan St.
Atlanta GA, 30332 USA
gilw@gatech.edu

Mark Godfrey
ZOOZ Mobile, Inc.
325 Trowbridge Walk.
Atlanta GA, 30350
mark@zoozmobile.com

Andrew Beck
ZOOZ Mobile, Inc.
325 Trowbridge Walk.
Atlanta GA, 30350
andrew@zoozmobile.com

CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.
ACM 978-1-60558-930-5/10/04.

Abstract
ZOOZbeat is a gesture-based Music reCreation studio. It is designed to provide users with expressive and creative access to music making on the go. ZOOZbeat users can compose user-generated songs based on generic beats in different styles or remix and modify commercially licensed songs. To play notes or trigger musical loops, players can shake the phone or tap the screen. Users can also record voice or other audio input into their songs and utelize tilt and shake movements to manipulate and share the music in a group. Design goals of the project focused on creating intuitive metaphors for mobile music making and maintaining a balance between control and ease-of-use.

Keywords
Mobile, music, gesture, creation, recreation

ACM Classification Keywords
J5. Arts and Humanities: Music

General Terms
Design, Human Factors, Performance

Introduction
ZOOZbeat is a gesture-based mobile musical studio designed to unlock musical expression and creative potential for novices and experts alike. It currently runs on iPhone, iPod Touch, and the Nokia N95 phones.

Unlike musical rhythm games such as Guitar Hero and Rock Band (where traditional eye-hand coordination and speed challenges are mapped to musical outcome), ZOOZbeat focuses on providing intuitive self-expression for non-musicians. Through a set of easily learned and natural gestures, ZOOZbeat users can enter musical material that is processed to fit different musical contexts in a loop-based sequencer. The loop-based interaction is forgiving of mistakes while providing structure. Users can perform additional gestures to manipulate and share their creations in a group. A key goal in designing ZOOZbeat was to provide immediate engagement and self-expression for novices, while providing a wide room for improvement and virtuosity that would engage experienced musicians as well.

Related Work

Advances in consumer mobile phones have facilitated massively distributed applications for music generation. Researchers in the field of New Interfaces for Musical Expression (NIME) have developed mobile music applications that can support the immediate, self-contained, and mobile nature provided by most acoustic musical instruments. Embedded accelerometers in mobile devices allowed researchers to experiment with gestural control of streaming audio [1] or self-contained audio synthesis [2]. Other input devices that have been used for musical and audio applications include touch screens [3,4], embedded cameras [5], and GPS [6, 7]. Recently, several music-making applications for mobile devices have emerged, from a breath-controlled wind instrument [8] to full-featured drum machines [9, 10] and synthesizers [11]. Unlike ZOOZbeat, these projects do not offer an integrated system that allows users to generate, edit, sequence, and share their compositions.

Application Structure – Songs and Gestures

At the core of ZOOZbeat is the song – a collection of MIDI instrument tracks, audio loops, and the organization of loops and related scales into a song structure. The format of the song allows full length, multi-sectional pieces to be condensed into short and manageable periods of time. Sections that have nearly identical backing tracks are reduced into a single background loop and attached to multiple parts of the song. Setting the act of music making in the context of a pre-organized song makes the process less intimidating to inexperienced users, allowing for uninhibited immediate engagement. For musicians, a set of complex muscle memory mappings defines the sensation of playing music. Subtle variations in finger and arm movements affect the feeling of playing a guitar, piano, or drum. For the non-musician, musical gestures are based on the experience of watching musicians play rather than their own muscle memory [12]. Our main design goal therefore was to find intuitive mappings between gestures and a loop-based music sequencer, so that musicians and non-musicians would immediately recognize and interact with.

Shaking

The most basic note entry mechanism in ZOOZbeat is shaking the phone. This gesture is similar to hitting a drum or striking a piano note. Input from the built-in accelerometer is analyzed for onset detection. The energy level of the onset is determined and mapped to the new note's pitch: the harder the energy of a shake, the higher-pitched the sound. In this way, melodic contours are created by adjusting the strength level of shaking. The shaking gesture was chosen as a main input form due to its familiarity and intuitiveness for both novices and professional musicians. A variation of

the shaking gesture is used in multi-player games to pass the music to other phones in the group. This feature simulates jam session scenarios such as in drum circles or in jazz improvisation, where the musical lead is passed among the musicians.

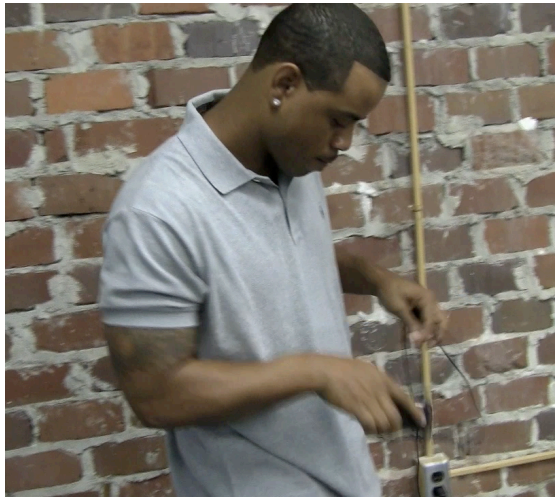

figure 1. Atlanta based hip-hop musician Cato shakes his phone to trigger and modify his recorded voice in ZOOZbeat.

Tilting

One gesture that is commonly associated with musical expression is tilting an instrument while playing. Classical musicians often move their violin or flute around during expressive and emotional sections. Rock guitarists often tilt their guitars up during intense solos, communicating an expressive and emotional impression to viewers [13]. In ZOOZbeat, users can experiment with such gestures by tilting their phone up and down and rolling it left and right. The tilting gesture transforms previously recorded material, subdividing note durations in the stored sequence. The sequenced notes are subdivided into 2 (by tilting upward) or 3 (by tilting downward) equal parts. The extra notes are considered ornamentations of the original melodic line, and their pitches are resolved by a set of deterministic rules governed by surrounding pitches.

Tapping

Another method of note entry in ZOOZbeat is tapping notes on a keypad. This involves pressing physical keys, holding them for the duration of the note, and releasing when finished. On touch-screen devices users can tap the screen instead. Lower notes are obtained by tapping lower keys (or tapping lower on the screen), and higher notes by pressing higher keys (or tapping higher on the screen). This method allows users to enter notes with a finer grain of detail than by shaking the phone.

Modes of interaction

Free Individual Play

In free play mode, after selecting a beat, a background loop is playing, and a selection of instruments or loops is displayed on-screen. In addition to percussive and melodic instruments users can also choose the microphone track, where they can record their voice or any other sound input. Each background loop is tied to a quantization level and scale so that the player can add new tracks in the appropriate tonal and rhythmic context. The user selects an instrument by choosing it from a menu and can add notes using shaking, tilting, or tapping gestures.

Group play

Two modes of group play are available –"Hot Potato" and "Free Play". In Hot Potato mode players pass

control of the song using a tossing gesture. Each player is assigned a track, which can be "tossed" via Bluetooth to the next player. The receiving player can add a new track, while its original creator can manipulate rhythmic aspects of the track by sending tilt information wirelessly. This allows participants to have a sense of ownership of their tracks while continuously engaged in the composition. In Free Play mode, players can play simultaneously while each controls one track. When players chose a different track to play, their previous track is made available for other players to modify, delete, or create their own sequences.

Software

A shared code base was written in C to maximize support of next generation phone platforms. The only requirements to run the code base are file input/output, an audio output stream, and the ability to run compiled C code. The advantages of a shared code base include the ability to synchronize among several platforms and the ability for multiple platforms to share common media files. The code base, inspired by popular music sequencers, organizes a song into samples, instruments, background audio loops and sequences.

Audio Engine

A sample is a piece of audio tagged with information such as sample rate, loop positions, base frequency (in MIDI note number) and current state. Each sample is constantly passed to the output stream. If the sample is marked as playing, then the appropriate audio bytes are fed to the output stream. This method allows for thread safety in platforms such as the iPhone, where samples are triggered by event threads and audio bytes are outputted by an audio thread. An instrument is a collection of samples, each given a MIDI-note value. Sequences are represented by a collection of instrument tracks, MIDI note numbers recorded, a scale, and a collection of audio loops. Notes are entered asynchronously into the sequence and instruments are triggered when a note is encountered. The built-in allows for sampling of live audio, typically the user's voice. One instrument track is dedicated to playing back these recordings. After recording, the sequencer divides the new sample into several sections and considers each note, with higher "pitch" mapped to later occurring sections of the sample. Users can then produce scratching or stuttering effects on the recording by entering and manipulating.

Wireless Communication

Different smart phones support different wireless protocols for data transport. On the Nokia N95, for example, we use the Bluetooth RFCOMM protocol to create a master-slave relationship between up to eight phones. On the iPhone, we use Apple's Bluetooth model to synchronize between two devices. The exact protocol used for synchronizing tempo and data is transport agnostic and is designed to be robust in the presence of dropped packets. The most challenging data to synchronize is the beat alignment. Even when two phones start at exactly the same time, slight changes in clock speed between the phones can put them out of phase by as much as a 16th note after a few bars. To synchronize the timing of two phones, we estimate the time it takes for a packet to move from one phone to the next. We send a number of messages back and forth and measure the amount of time from sending each message to receiving the response. Once the system is confident that there is low variance between most of the estimated timings, it sends a final sync message, signaling to the other phone that the next

beat is arriving in a specified amount of time. If the internal timing on the receiving phone is drastically different from the timing sent by the master phone, the receiving phone will adjust itself only slightly to prevent large jumps in sequencer timing.

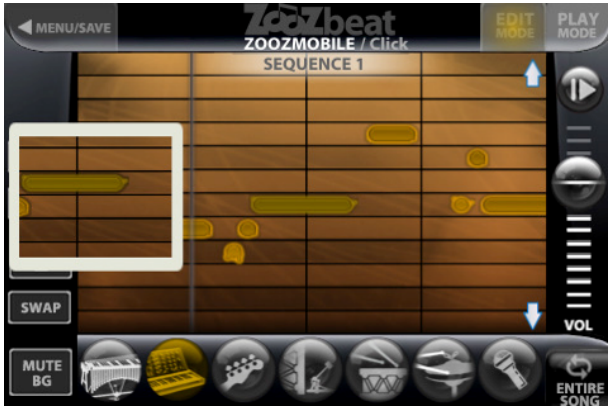

figure 2. ZOOZbeat Landscape Mode User Interface

User Interface

In line with our goal to allow natural gestural control over musical events and multi-user interaction, ZOOZbeat's user interface was designed to welcome users inexperienced with technology-aided music production while streamlined and powerful enough so as not to hinder musical expression. The less detailed Portrait Mode lands itself more to gestural interaction, while Landscape Mode (depicted in figure 2) allows for more detailed touch screen based editing including accurate note entry, note selection and movement, notes length change, track volume change, instruments swap the background beat mute. To avoid confusion and clutter, only the current track's notes are displayed. Users switch tracks (and thus instruments) by selecting the appropriate note icon. The notes animate as the curser moves over them which gives the application a whimsical sense of livelihood.

Preliminary User Feedback

The response to ZOOZbeat has been quite varied. Novice testers have been impressed by the ease in which sophisticated musical material is generated. More experienced users found the resulting music trite and the quantization algorithms heavy-handed. Musicians and non-musicians alike found the application enjoyable and at times surprisingly inspiring. Typically, musicians experienced with computer-based music production tend to look for more advanced features that were intentionally omitted for simplicity sake, such as tempo control or chromatic scales. Reviews of the gestural control have been generally positive. TechCrunch considers them a "perfect use of the built-in accelerometers… you can pick an instrument and simply tap, shake or tilt to create your own masterpiece" [14]. Wired.com's reviewer – Eliot Van Buskirk – was "impressed" and "thoroughly enjoying" creating loops, but also "really wanted to upload [his] creations" [15]. To address this need, we have developed the website zoozbeat.com, which allows users to upload, listen, download, and share their creations with other users. Interestingly, another common complaint was that the application seemed "pointless", as there is no goal or specific challenge to address. Familiar with games like Guitar Hero, typical users seem to tend toward more structured play, where their actions are guided by more than a desire to produce satisfying music. To allow for such challenges we are currently developing a contest engine that will provide prizes to collaboratively filtered "best songs".

Future Work

Currently under development are social interaction features that will enable users to play with ZOOZbeat in groups. This involves adapting our ad-hoc local area wireless communication protocol to client-server architecture, allowing players across continents to jam together. We are also exploring more play modes where users will be driven by competition without loosing the creative and expressive value of the interaction. Enhancements to the audio engine are also under development, which will allow for optimized audio effects that can be controlled real-time. These will not only satiate musicians eager for more production power but will also allow non-musicians the opportunity to intuitively explore other dimension of musical production such as timbre.

Citations

[1] Tanaka, A. Mobile Music Making. in Proc. of the International Conference on New Interfaces for Musical Expression *NIME 2004*, ACM Press (2004),154–156.

[2] Essl, G. and Rohs. M. Mobile STK for Symbian OS. In Proc. International Computer Music Conference *ICMC 2006,* New Orleans, USA, pp. 278–281.

[3] Geiger, G. PDA: Real Time Signal Processing and Sound Generation on Handheld Devices. In Proc. of the International Computer Music Conference. *ICMC 2003* Singapore, 2003.

[4] Geiger, G. Using the Touch Screen as a Controller for Portable Computer Music Instruments. In *NIME 2006*, 61 – 64.

[5] Rohs, M., Essl, G. and Roth M. CaMus: Live Music Performance using Camera Phones and Visual Grid Tracking. In Proc. *NIME 2006*, 31–36.

[6] Strachan, S., Eslambolchilar, P., Murray-Smith, R., Hughes, S., and O'Modhrain S.. GpsTunes: Controlling Navigation via Audio Feedback. In Proc. *CHI 2005*, 275-278.

[7] Tanaka, A., Valadon, G. and Berger, C. Social Mobile Music Navigation using the Compass, in Proc. of the Int'l Mobile Music Workshop, Amsterdam, 2007.

[8] Smule: Ocarina. http://ocarina.smule.com

[9] Intua: BeatMaker. http://www.intua.net/products.html

[10] Izotope: iDrum. http://www.izotope.com/products/audio/idrum/iphone

[11] Noise.io. Available: http://noise.io/

[12] Rodger, M. Issartel, J., and O'Modhrain, S. Performer as perceiver: perceiver as performer. Proc. of the Int'l Conf. on Enactive Interfaces *ENACTIVE 2007*, Grenoble, France, 2007.

[13] Cadoz C. and Wanderley, M. Gesture – Music. In *Trends in Gestural Control of Music*, M. M. Wanderley and M. Battier, Eds. Paris: IRCAM (2000), 71-94.

[14] ZOOZbeat Turns Your Phone Into a Music Studio. http://www.techcrunch.com/2008/11/14/ZOOZbeat-turns-your-phone-into-a-music-studio

[15] ZOOZbeat Turns iPhone into Beat Factory for Three Bucks. http://blog.wired.com/business/2008/12/ZOOZbeat-turns.html