Music Mixing Surface

Gelineck, Steven; Büchert, Morten; Andersen, Jesper

Published in:
Proceedings of the 2013 ACM international conference on Interactive tabletops and surfaces

DOI (link to publication from Publisher):
10.1145/2512349.2517248

Publication date:
2013

Document Version
Early version, also known as pre-print

Music Mixing Surface

Steven Gelineck
Aalborg University CPH - Department of Architecture, Design and Media Technology
A.C. Meyers Vænge 15
2450 KBH SV
stg@create.aau.dk

Morten Büchert
Rhythmic Music Conservatory, Copenhagen
Leo Mathisens Vej 1, Holmen
1437 Copenhagen K
morten@mortenbuchert.com

Jesper Andersen
The Royal Danish Academy of Music
Rosenørns Alle 22
1970 Frederiksberg C
jesper.andersen@dkdm.dk

ITS '13, Oct 06-09 2013, St Andrews, United Kingdom
ACM 978-1-4503-2271-3/13/10
http://dx.doi.org/10.1145/2512349.2514592

Abstract
This paper presents a multi-touch based interface for mixing music. The goal of the interface is to provide users with a more intuitive control of the music mix by implementing the so-called *stage metaphor* control scheme, which is especially suitable for multi-touch surfaces. Specifically, we discuss functionality important for the professional music technician (main target user)—functionality, which is especially challenging to integrate when implementing the stage metaphor. Finally we propose and evaluate solutions to these challenges.

Author Keywords
User interfaces; Music; Mixing; Multi-touch; Stage Metaphor

ACM Classification Keywords
H.5.1 [Multimedia Information Systems]: Audio input/output.; H.5.2 [User Interfaces]: Input devices and strategies.

Introduction
Music mixing is the art of adjusting various audio parameters such as volume and panning (left/right positioning of the sound) of individual channels of audio in order to achieve a well-balanced collective mix [6, 9]. The traditional interface for doing so is based on the so

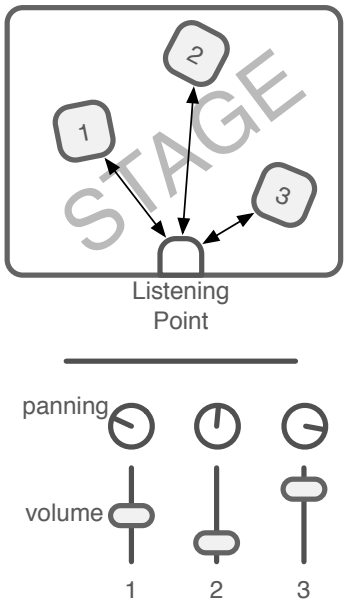

Figure 1: The stage metaphor for mixing audio works by placing virtual sound sources on a virtual stage with a virtual listening position. The volume of a track decreases the further away it is from the listening point. Likewise, the panning is determined by the angle between the listening point and the track. Here the relation between stage metaphor track positions (top) and tracks controlled using a traditional channel strip metaphor (bottom) is illustrated.

called *channel strip metaphor* where each channel of audio (for instance drum, vocal, guitar, etc.) is divided into a single column on the interface. Each column implements controls for adjusting the various parameters of the corresponding channel using physical buttons, knobs, and sliders. This control metaphor has not changed over the last 50 years—while the newest digital mixing consoles do provide a significant amount of additional functionality and flexibility, they are still based on the channel strip metaphor.

An alternative control scheme is the *stage metaphor*. Here audio channels are represented by virtual graphical widgets that can be positioned on a virtual stage. Their position relative to a virtual listening position determines the volume and the panning of the corresponding audio channel—see Figure 1. The metaphor simulates how sound sources placed at different positions are perceived in the real world, which (according to informal interviews with expert music technicians) is closer to how many users mentally picture an overall mix.

Several others have explored the stage metaphor including Peter Gibson's "Virtual Mixer" [5], Diamante's AWOL [3] and Pachet & Delerue's MusicSpace [8]. Most recently Carrascal and Jordà [1] implemented the stage metaphor similar to how it is implemented here—on a multi-touch surface. They tentatively concluded that the stage metaphor was preferred over the channel strip metaphor in a music mixing task. However, their implementation was rather limited including simple volume, panning and equalisation of few audio channels. We expand on their research by exploring how extended functionality (crucial for a successful real world implementation of the stage metaphor) can be implemented, dealing with challenges that arise while doing so. Previous studies by the authors [4] have exposed these challenges. Here we explore and evaluate several solutions.

System Functionality

The Music Mixing Surface is a control interface for controlling an underlying Digital Audio Workstation (DAW)[1]. In other words all of the audio processing is taken care of by the underlying DAW. It is implemented on a 27" LCD screen with a G4 multi-touch overlay[2][3].The main interaction area is the *stage* with the *listening position* located at the bottom centre—see Figure 2. Each audio channel is displayed as a graphical widget with a position on the stage relative to the listening position. The user is able to position the widget by dragging in the centre of the widget with one finger. When moving the widget left or right, the sound is panned to the left or right speaker, moving the widget closer to the listening position adjusts the volume upwards and vice versa.

Additionally, the user is able to adjust an optional amount of different parameters associated with the audio channel[4]. The effect sliders are constructed by dividing a circle into equally sized areas. The areas can be filled or empty depending on the value that they represent. These values are adjusted by dragging with one finger outwards or inwards within the area. This "pie" layout was chosen for its scalability and because it gives the user an impression of the values of each parameter by interpreting the overall shape of the virtual widget. Additional features have been implemented in order to account for previously encountered issues with the interface:

[1] In this case Ableton Live - http://ableton.com

[2] from PQ Labs - http://multi-touch-screen.com

[3] For more details about the technical implementation of the system go to http://media.aau.dk/~stg/tangibleMixSurface

[4] For the evaluation presented later these have been reduced to 6 effect parameters: bass, treble, compression, trim, reverb and delay

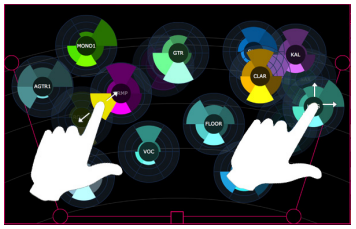

Figure 2: Overview of the main stage. Widgets represent audio channels, which are positioned in relation to the listening position represented by the pink square in the bottom centre. The distance from the listening position determines the volume of the corresponding audio channel. The angle determines the panning.

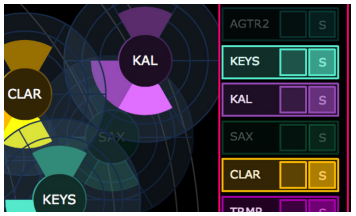

Figure 3: The interface implements a layering functionality, where widgets can be turned on or off by selecting the corresponding layer (in the right column). Here the Kalimba, Clarinet and Keys are turned on. Note that turning widgets on or off does not affect the sound in any way.

Clutter - Because the audio channels are not placed at fixed positions as in the channel strip metaphor, but are rather scattered around the stage area, widgets are likely to occlude each other making it difficult to locate specific channels. A visual layering mechanism was implemented, which makes it possible to turn the visual representations (widgets) of specific audio channels on or off. It is implemented as a layers menu similar to turning layers on or off in software such as Adobe's Photoshop or Illustrator[5]—see Figure 3.

Visual dominance - A *limited visuals* mode where only the centre of the widgets are visible has been implemented (see Figure 4) in order to provide a better overview while positioning widgets and in order to limit the visual burden on the user. A simple button toggles between *limited* or *full visuals* mode.

Stage layout - The stage metaphor inherently offers more limited space for positioning widgets the closer they get to the listening position. To compensate for this users now have the freedom to expand the mapping space by altering the stage layout. By dragging on one of the four corners of the stage the user is now able to widen the area closest to the listening position thereby giving room for a larger panning space. The user is free to change the layout at any time while mixing in order to gain the most intuitive overview[6].

Zoom - A simple pinch gesture performed anywhere on the main stage for zooming in and out has been implemented in order to work with more fine grained relationships between widgets.

[5] http://www.adobe.com

[6] Widgets automatically relocate to represent the volume/panning settings of their corresponding channels

Falloff coupled with Distance from Listening Position It has previously been suggested [4] that not only should a widget's increased distance from the listening position result in a decrease in volume of the associated audio channel. It should also result in an increase in reverberation together with an attenuation of high frequency content, which corresponds to the way we as humans perceive sound sources coming from further away [7]. This feature has been implemented as a so called *falloff effect* that can be set using a graphical slider. The higher the falloff effect the more each audio channel will be affected in terms of reverberation and filtering.

Randomisation - A randomisation function had been implemented in order to extend the explorability of the interface. It is now possible to set randomisation for either all or single widgets. Also, the amount of randomisation can be set by the user.

Exploration of different Mixes - Another feature for increasing exploration has been implemented, which lets users explore different versions of the same mix. They now have the opportunity to shift quickly back and forth between two or more mixes in order to explore and compare them instantly.

Increased visual information for the currently selected widget - A detailed overview of what the different parameter control sliders represent has been added. It is only shown as the widget is selected in order to reduce visual clutter (see Figure 5).

Additional features - Smaller improvements to the interface include *undo* functionality, *copy-pasting* of settings between channels, separate *volume trimming* for each channel, *master volume* control, *solo/mute*, and *transport controls*.

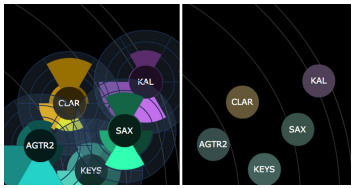

Figure 4: There are two visual modes: full visuals (left), and limited visuals (right).

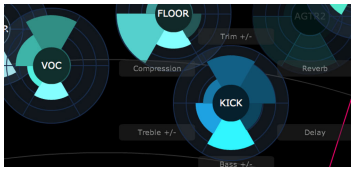

Figure 5: When selected, the names of the parameters are displayed. Here the 'KICK' is selected.

Figure 6: The Music Mixing Surface in use during one of the evaluation sessions.

Evaluation and Future Directions

An informal evaluation was conducted where two Audio Technician students and a Composer student from the Rhythmic Music Conservatory in Copenhagen were asked to mix a prepared piece of music using: (a) the Music Mixing Surface (see Figure 6); and (b) a traditional analogue mixing setup. After having tried both setups in a randomised order the participants were interviewed about the major differences between using the two setups. This, together with observations was analysed using a bottom up grounded theory [2] approach.

Results - Two of the participants emphasised how impressed they were with the speed at which one was able to achieve a good mix using the Music Mixing Surface—especially since they had never tried the interface before. Additionally, it was reported that using the analogue setup felt more like a remote control used to control each audio channel, while on the Music Mixing Surface the participants felt in actual contact with the channel through their fingers. The functionality for exploration of different mixes was really appreciated by all test subjects as having a powerful contribution to the creative potential of the Music Mixing Surface, while the randomisation functions were hardly used at all. Switching between *full visuals* and *limited visuals* was used a lot. The participants would fluctuate between the two modes as they shifted back and forth between editing single channels and balancing channels in relation to each other. They really appreciated the possibility of being able to reduce the amount of visual feedback in order to focus more on the task of listening. The layering feature where users were able to turn widgets on or off was also heavily used for focussing on specific audio channels or relationships between specific audio channels.

Future directions include implementing features for accessing additional underlying effect parameters (reported as being crucial during the evaluation session), features for grouping audio channels, as well as monitoring sound levels and controlling automation of effect changes over time. Additionally, we are exploring expanding the interface for mixing multi channel surround sound. Finally, we wish to explore how important the use of tangibles for various functions is compared to multi-touch. While tangibility seems important for fine control we seem to experience a shift in how well users are able to interact using multi-touch. However, this is yet to be evaluated.

References

[1] Carrascal, J. P., and Jordà, S. Multitouch interface for audio mixing. In *Proc. of NIME* (2011), 100–103.

[2] Corbin, J., and Strauss, A. Basics of qualitative research: Techniques and procedures for developing grounded theory. *UK: sage publishing* (2008).

[3] Diamante, V. Awol: Control surfaces and visualization for surround creation. Technical report, University of Southern California, Interactive Media Division, 2007.

[4] Gelineck, S., Overholt, D., Büchert, M., and Andersen, J. Towards an interface for music mixing based on smart tangibles and multitouch. In *Proc. of NIME* (2013).

[5] Gibson, P. *The Art Of Mixing: A Visual Guide To Recording, Engineering, And Production*. ArtistPro Press, 1997.

[6] Izhaki, R. *Mixing Audio: Concepts, Practices and Tools*. Focal Press, 2011.

[7] Malham, D. Tutorial article: Approaches to spatialisation. *Organised sound 3*, 2 (1998), 167–177.

[8] Pachet, F., and Delerue, O. On-the-fly multi track mixing. In *AES Convention*, Los Angeles (2000).

[9] Senior, M. *Mixing Secrets*. Taylor & Francis, 2012.