Towards a more Flexible and Creative Music Mixing Interface

Gelineck, Steven; Büchert, Morten; Andersen, Jesper

Published in:
ACM SIGCHI Conference on Human Factors in Computing Systems

DOI (link to publication from Publisher):
10.1145/2468356.2468487

Publication date:
2013

Document Version
Early version, also known as pre-print

Towards a more Flexible and Creative Music Mixing Interface

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

CHI 2013 Extended Abstracts, April 27–May 2, 2013, Paris, France.
ACM 978-1-4503-1952-2/13/04.

Abstract
This paper presents the ongoing work towards creating a novel interface for mixing music. It identifies 5 key design factors crucial for the development and evaluation of such an interface. It then shortly presents an initial prototype, which implements a stage metaphor control structure. Finally two initial evaluation activities are briefly presented and discussed, one of which compares multi touch to two different tangible user interface interaction schemes. Preliminary results suggest that tangible controls outperform and are preferred over multi touch.

Author Keywords
User interfaces; Music; Mixing; TUI; Tangible; Multi-touch; Stage Metaphor; Smart Tangibles

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

Introduction
This paper presents the early work towards designing a new interface for mixing music especially targeted towards expert sound technicians for especially studio sound production. Mixing is the activity of reducing numerous audio tracks to a stereo mix, manipulating volume,

Figure 1: The traditional mixing console control surface consists of sliders, knobs and buttons. Each vertical bar represents an audio channel with controls for volume, panning, filtering/limiting and aux sends. (Pictured is the Yamaha MG 166CX USB).

Figure 2: The digital mixer still implements the channel strip control structure where each audio channel is organized as fixed columns with sliders, knobs and buttons as main control devices. (Pictured is the Midas Pro1 IP).

panning and several other parameters of each track. The mixing console (or mixing desk), which is the primary interface for mixing music, has remained largely the same since its birth 40 years ago. It is still based on the notion of separate audio channel strips being manipulated using physical sliders, knobs, and buttons to control the various properties of each audio track.

The *analogue* mixing desk deploys a one-to-one mapping where every slider, knob and button has a dedicated function. See Figure 1. The major benefits of this design is that it is fast and intuitive. *Digital* consoles are getting increasingly popular based mostly on their reduced real estate, and extended functionality and flexibility. The interface is still based on the same channel strip control structure (see Figure 2) and from a usability point of view, most functionality is layered in sub-menus, breaking the one-to-one mapping, thus demanding more effort for the user. Thus, the research presented here attempts to balance the following two major requirements:

1) To challenge the interaction paradigm of the classic mixing console in order to free the musician/producer to explore more creative work flows. 2) To maintain the intuition and speed found in the physical interface of the traditional analogue mixing console.

The paper initially discusses what we see as key aspects in the design of novel audio mixing interfaces. We then go on to present an interface design that is evaluated in a musical context, as well as an initial general purpose usability test carried out to evaluate the *general* performance/preference differences between Mouse, Multi-touch, and in two versions of a Tangible User Interface (TUI).

Design Framework

When developing the framework for the intuitive mixing interface we have identified five crucial key factors:

Expert Users and the Experimental Context

The target group for the interface is professional music technicians and producers. It is important that the development takes into account the needs of professionals in all aspects. Besides establishing these needs by using qualitative interview based examinations of professionals, it is important to underline the importance of context when carrying out HCI usability tests of interfaces for musical expression (as suggested by [15, 5, 12]). The usability test presented later in this paper however, uses test subjects with no musical experience. We distinguish here between evaluations that have to do with musical interaction and general purpose interaction. The quantitative evaluation presented later is meant to establish a baseline for general purpose control of virtual widgets using different input technologies, while the qualitative evaluation is meant to evaluate these technologies in the context of music mixing.

Depth Mixing and the Stage Metaphor

Depth in music can be defined as the perceptual separation of sound sources in terms of distance from the listener. In order to create depth in the mix, an important goal while mixing is to emphasise this perception of sound sources positioned at different distances. Depth control is seldom explored in common music mixing interfaces (such as Ableton Live, Logic, Protools, etc.). The project will examine the importance of depth by exploring depth as a control unit, challenging the traditional channel strip metaphor.

A number of interesting systems have been proposed that challenge the tradition mixing desk metaphor. Peter

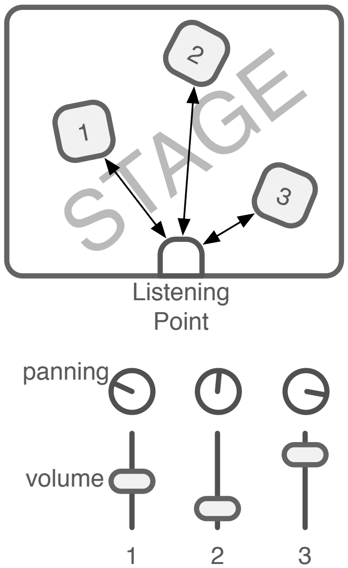

Figure 3: The stage metaphor for mixing audio works by placing virtual sound sources on a virtual stage with a virtual listening position. The volume of a track decreases the further away it is from the listening point. Likewise, the panning is determined by the angle between the listening point and the track. Here the relation between stage metaphor track positions (top) and tracks controlled using a traditional channel strip metaphor (bottom) is illustrated

Gibson proposed the "Virtual Mixer" [6], which is a virtual 3D mixing environment where each audio channel is controlled by manipulating the size and position of virtual 3D globes on a virtual stage. Moving the globes further away reduces the volume, moving the globe left or right pans the sound, and moving it up or down filters the sound. This approach was one of the first to re-define the control metaphor for music mixing and has inspired related systems where the so-called *stage metaphor* is used. The stage metaphor relates virtual objects to a virtual listening point in the same way as a sound source on a physical stage would relate to a listener. Here the panning and volume of each channel relates to its 2D position relative to the listening point (see Figure 3).

Diamante presents a GUI that not only implements the stage metaphor but also improves the visual feedback for the sound propagation of each channel [2]. Pachet and Delerue implemented the stage metaphor when developing a GUI for presenting music-end-listeners with capabilities of creating their own individual mix of a music track [13]. Carrascal and Jordà [1] explore the stage metaphor for controlling a mixing environment similar to the one presented here, focusing on a multi touch surface approach. They conclude with a preliminary usability test that tentatively suggests that the stage metaphor using this multi touch surface approach is preferred over a traditional mixing console.

Of course the stage metaphor also comes with a few obvious drawbacks. The most crucial of these is the organization and overview of the different channels. With the channel strip metaphor each audio channel is always located in the same place, which makes it easy to find and control. With the *stage metaphor* channel representations are scattered around the virtual stage area and can be difficult to overview—especially since professional systems demand at least 24 channels (often more). This is most likely why the channel strip metaphor is still predominant and why more flexible and experimental interfaces still keep this layout [11] for increased usability. We are examining various approaches to deal with this problem including, colouring, shaping, grouping, icons, and various localization functions.

Functionality Layering

Layering of functionality is often used for handling underlying complexity of an application. This can be because of limited real estate on a screen or in order not to overwhelm the user. However, layering is a big concern when dealing with music mixing. Because music is experienced in real time, there is a short window for making changes to sequences. This means that controls have to be fast and easy to get to, and layering may obstruct this speed. While acknowledging that in order to maximize the potential of the interface some layering must be implemented, we believe that layering must be reduced to the minimum - since it reduces fast intuitive interaction. A subgoal of the project is to examine how much layering is appropriate in order to maintain the speed and reliability of the interface. Exploring different layering methods for increasing the speed and intuitiveness will also be a part of the future development.

Creativity Support

Traditionally the post production phase of music mixing is regarded as a non-artistic activity of enhancing the musical track to fit the needs of the creative performer. However, more and more emphasis is put on the art of mixing as a creative activity demanding an increase in creativity support for the underlying mixing tools [8, 14].

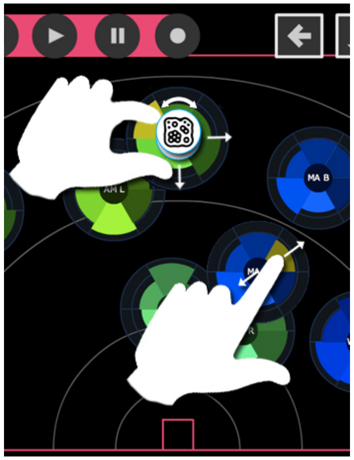

Figure 4: The initial mixing interface lets users control audio channels by manipulating widgets on an interactive surface using the stage metaphor. The position of each widget relative to the listening point determines the volume and panning of the associated channel. Widgets can be controlled using multi-touch, TUIs on all widgets, Few TUIs that are only deployed when needed, or using a smart tangible. See the interface in use at http://youtu.be/9WAI3FCzugE

We are very much interested in maximizing the creativity support of the interface, especially helping users explore new directions when mixing. Initial informal interviews with expert technicians suggest that technicians deploy various methods for exploring alternative workflows in order to explore different outcomes than those traditionally suggested by existing interfaces—by for instance setting dogmatic constraints for themselves, or using software plugins that were originally designed for other purposes. The flexibility of the proposed interface yields room for building these workflows into the underlying software.

Tangibility and Multi-touch
A major concern when developing the interface is how to physically interact with the underlying software. The stage metaphor described earlier demands that channels are somehow represented and manipulated by moving virtual sound sources on a 2D plane. Initially, a multitouch surface has was developed, but informal interviews with expert technicians underline the importance of tactile and haptic feedback while manipulating parameters. Therefore we have started the development of a Tangible User Interface (TUI) where virtual sound sources can be manipulated using tangible widgets—inspired by work by Jorda et al. [9]. This however, opens up several challenges in regards to the flexibility of the system as is also emphasized by Kirk et al. [10], which present an excellent overview of the challenges involved in designing hybrid TUI surfaces. The major concerns deal with maintaining the flexibility of the system when physical objects are deployed. For instance, there might be a need for virtual widgets to change appearance or position at different stages in the work flow. A solution, as proposed by Weiss et al. [17] could be to make the physical objects translucent in order to be able to display state changes while still maintaining the passive haptic feedback which is necessary here. Another solution is to use physical TUI objects not as representation of underlying data as is often seen in TUIs—but instead use them as manipulation tools that can be deployed when necessary as seen in [10].

Additional challenges lie in the often constrained functionality of TUI objects, where the only forms of manipulation are often reduced to positioning and rotation of the objects. This has lead to the development of so called *smart tangibles*, which are TUI objects embedded with micro controllers offering additional sensing for the control of additional functionality [3, 7] also helping with the problem of functionality layering discussed earlier.

Initial Design and Evaluation
In order to investigate the key factors discussed above, two research activities have been initialized. Firstly, an initial functional prototype has been developed, which is being informally evaluated. The interface lets sound technicians control up to 24 channels of audio by manipulating virtual widgets using different interaction schemes (see Figure 4). These include multi touch, tangible objects on all widgets, and so-called *smart tangibles* capable of sensing how they are grasped by the user (these can be used as manipulation tools to access and manipulate various parameters of each channel depending on how the user grasps the tangible). The interface makes it possible for the user to control 6 effect parameters by manipulating the length of 6 coloured effect sliders using one of the proposed control schemes (the design scales easily for control of additional effect parameters). Muting and soloing of tracks is also implemented by double tapping in the centre or at the perimeter of each widget. Finally, various randomization functions are implementing that lets the user experiment with mixes that lie outside their

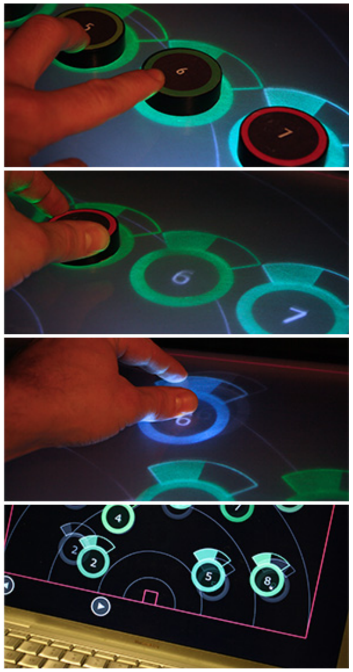

Figure 5: The four input technologies that were compared. From the top: TUI blocks (all), TUI blocks (few), Multi touch, GUI (mouse).

comfort zone. The evaluation of the interface involves an exploratory focus group session where *Tonemester* (sound engineer) students from the Classical Conservatory in Copenhagen explore the interface while mixing two 8-track music pieces using the different interaction schemes. The main purpose of the evaluation sessions is to validate the key design factors discussed above exploring solutions and further challenges. There is no room to go into detail with the evaluation here, but early results suggest that the conceptual model of the stage metaphor is much closer to the mental model of the user in how they conceive the mix. Moreover, the *smart tangibles* used only as tools (not representation) were preferred as they provided fast and intuitive access to underlying parameter manipulation.

Secondly, a preliminary test has been conducted to examine the performance differences in an acquisition task between *a) physical TUI blocks on all widgets; b) two TUI blocks each of which must be lifted and placed on a widget to manipulate it; c) multi touch; d) a traditional GUI using a mouse.* (See Figure 5).

Related work by Tuddenham et al. [16] based on [4] compares tangible controls; multi-touch; and mouse and puck, in a target acquisition and manipulation task. The actual widgets used in our experiment are accustomed to the design proposed above, and an additional control scheme where TUI objects used only as tools has been added. Other than that, the experimental method is the same as in [16].

In short 14 test subjects were given the task to manipulate the position and a variable of 8 virtual widgets. For each of the 4 conditions the goal was to keep the widgets as close to their corresponding moving ghost widgets as possible (see Figure 5). Each participant had one trial per condition and a trial lasted 90 seconds. A tracking error score was calculated as the roots-mean- squared Euclidean distance for all widgets to their corresponding ghost for each participant in each of the four conditions. Finally they were asked to rank the four approaches based on preference. Results indicate that TUI objects performed significantly better and where also preferred over the rest of the conditions. Figure 6 shows tracking error scores (in pixel distance) and ratings for each input technology (4 points for highest rank, 1 point for lowest).

Two major biases were identified that render most of the results non-applicable in a general sense. The TUI object tracking system was not fast enough to deal with the lifting and placing of fiducial objects in condition (b), which meant that widgets would sometimes detach from the physical block that was controlling them. Even with this frustration in mind participants still rated this input method higher than than Multi touch and GUI (mouse). As reported in [4, 16] we also had an issue with *exit errors*—as the user would remove his or her finger from a widget it would move slightly.

We are preparing a second experiment with a faster TUI object tracking, and re-positioning of the widget at end-touch to minimize exit errors. Finally, we are still analysing the results of the qualitative evaluation of the prototype while developing a second prototype that provides extended functionality to the user in terms of effect manipulation, grouping, edit/mix mode shifting, and an improved interaction scheme based on a new prototype of the *smart tangible*.

Conclusion

This paper has presented and discussed 5 key aspects to development and evaluation of a novel interface for mixing music targeted at professional music technicians

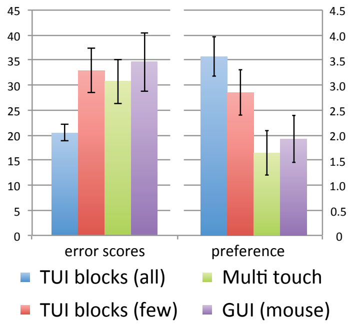

Figure 6: Left, tracking error scores calculated as roots-mean-squared Euclidean distance between widgets and their corresponding ghosts—lower scores indicate better performance. Right, preference rankings by input technology (highest ranking gave 4 points, while lowest ranking gave 1 point).

and producers. An initial prototype has been presented and an initial usability experiment has been conducted that shows potential in the use of a tangible user interface control scheme implementing the stage metaphor.

References

[1] Carrascal, J. P., and Jordà, S. Multitouch interface for audio mixing. In *Proceedings of New Interfaces for Musical Expression* (2011), 100–103.

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

[3] Feltham, F. Do the blocks rock: a tangible interface for play and exploration. In *Proceedings of the Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat*, ACM (2008), 188–194.

[4] Fitzmaurice, G., and Buxton, W. An empirical evaluation of graspable user interfaces: towards specialized, space-multiplexed input. In *Proceedings of Human Factors in Computing Systems*, ACM (1997), 43–50.

[5] Gelineck, S., and Serafin, S. A quantitative evaluation of the differences between knobs and sliders. In *Proceedings of New Interfaces for Musical Expression* (2009).

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

[7] Hochenbaum, J., and Kapur, A. Adding z-depth and pressure expressivity to tangible tabletop surfaces. In *Proceedings of New Interfaces for Musical Expression* (2011), 240–243.

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

[9] Jordà, S., Geiger, G., Alonso, M., and Kaltenbrunner, M. The reactable: exploring the synergy between live music performance and tabletop tangible interfaces. In *Proceedings of Tangible and Embedded Interaction*, ACM (2007), 139–146.

[10] Kirk, D., Sellen, A., Taylor, S., Villar, N., and Izadi, S. Putting the physical into the digital: issues in designing hybrid interactive surfaces. In *Proceedings of the British HCI Group Annual Conference on People and Computers: Celebrating People and Technology*, British Computer Society (2009), 35–44.

[11] Liebman, N., Nagara, M., Spiewla, J., and Zolkosky, E. Cuebert: A new mixing board concept for musical theatre. In *Proceedings of New Interfaces for Musical Expression* (2010).

[12] Marquez-Borbon, A., Gurevich, M., Fyans, A., and Stapleton, P. Designing digital musical interactions in experimental contexts. In *Proceedings of New Interfaces for Musical Expression* (2011).

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

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

[15] Stowell, D., Robertson, A., Bryan-Kinns, N., and Plumbley, M. D. Evaluation of live human-computer music-making: Quantitative and qualitative approaches. *International Journal for Human Computer Studies 67*, 11 (2009), 960–975.

[16] Tuddenham, P., Kirk, D., and Izadi, S. Graspables revisited: multi-touch vs. tangible input for tabletop displays in acquisition and manipulation tasks. In *Proceedings of Human Factors in Computing Systems*, ACM (2010), 2223–2232.

[17] Weiss, M., Hollan, J., and Borchers, J. Augmenting interactive tabletops with translucent tangible controls. *Tabletops-Horizontal Interactive Displays* (2010), 149–170.