

Spatial Mixer

Cross-Device Interaction for Music Mixing

Kjeldskov, Jesper; Paay, Jeniece Michelle; Sørensen Nilsson, Aleksander; Plejdrup, Kasper; Thomsen Pedersen, Mette

Published in:

Proceedings of the 31st Australian Conference on Human-Computer-Interaction, OzCHI 2019

DOI (link to publication from Publisher):

[10.1145/3369457.3369465](https://doi.org/10.1145/3369457.3369465)

Publication date:

2019

Document Version

Accepted author manuscript, peer reviewed version

[Link to publication from Aalborg University](#)

Citation for published version (APA):

Kjeldskov, J., Paay, J. M., Sørensen Nilsson, A., Plejdrup, K., & Thomsen Pedersen, M. (2019). Spatial Mixer: Cross-Device Interaction for Music Mixing. In *Proceedings of the 31st Australian Conference on Human-Computer-Interaction, OzCHI 2019* (pp. 85-94). Association for Computing Machinery (ACM).
<https://doi.org/10.1145/3369457.3369465>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Spatial Mixer: Cross-Device Interaction for Music Mixing

Jesper Kjeldskov¹, Jeni Paay^{1,2}, Aleksander Nilsson¹, Kasper Plejdrup¹, Mette Pedersen¹

¹ Aalborg University, Aalborg, Denmark

² Swinburne University of Technology, Melbourne, Australia
jesper@cs.aau.dk, jpaay@swin.edu.au

ABSTRACT

When workspaces for commercial music production are digitized they are generally configured based on existing analogue music production environments, using the “channel strip” metaphor to organize the multiplicity of devices involved in the process. The channel strip metaphor is derived from the traditional design of mixing consoles with parallel sliders acting as separate channels. An alternative to this is the “stage” metaphor, which visually represents sound elements as discrete actors on a stage for manipulating settings. To study cross-device interactions in music mixing, we designed the *Spatial Mixer*, an app that uses the stage metaphor, to work in concert with *Ableton Live*, an existing music composition program. A study of *Spatial Mixer* in use showed that the stage metaphor and spatial visualizations of music tracks and effects was intuitive for users to interact with and helped them work across both pieces of software to effectively create their musical composition.

CCS CONCEPTS

• Human-centered computing → Interaction design process and methods

KEYWORDS

Spatial Mixing, Cross-Device Interaction, Interaction Design, Stage Metaphor, Music Production, Surround Sound

ACM Reference format:

FirstName Surname, FirstName Surname and FirstName Surname. 2018. Insert Your Title Here: Insert Subtitle Here. In *Proceedings of ACM Woodstock conference (WOODSTOCK'18)*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/1234567890>

1 Introduction

Within cross-device interaction, a lot of research has been conducted on how multiple devices can be used at the same time to make tasks more manageable or to understand which interaction technique best suits a given context.

OZCHI'19, December 2–5, 2019, Fremantle, WA, Australia
© 2019 Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-7696-9/19/12...\$15.00
<https://doi.org/10.1145/3369457.3369465>



Figure 1. *Spatial Mixer* as used in music mixing (bot. left).

Music production is an environment where several devices are used concurrently in a complex, dynamic and unpredictable way and so offers an interesting testing ground for exploring new cross-device interactions, and opportunities for enhancing user experience [3, 16, 17]. Music mixing happens on a computer with software designed to emulate analog mixing consoles [14]. The complexity in mixing music is around balancing the mix of the depth of the music, the spatial relation between the tracks, finding a balance between the volume of the instruments and adding specific effects to make the sounds unique. Commercial music is usually mixed for stereo because the application domain is almost exclusively stereo for example, listening through headsets or listening to the radio. The application domain for surround sound is more limited, primarily to movie theaters, and home theaters with surround sound speaker systems.

This study explores how new cross-device interactions can be designed to enhance an application for spatial music mixing with surround sound utilizing the spatiality of the mixing room. Our prototype music mixing app, *Spatial Mixer*, shown in figure 1, is designed and developed to work in concert with the music composition program, *Ableton Live*. *Spatial Mixer* is built upon the source code of a project created by Gelineck et al. [6]. They explore how the stage metaphor can be used in music production to provide a spatial, new and more intuitive way to control panning and volume while mixing music. Additionally, the *Spatial Mixer* app provides the music composer with a two-dimensional view of the tracks from the music mixing program, *Ableton Live*, placed in a box representing the stage. The listener is placed in the middle to facilitate surround sound composing. The composer can change parameters in either *Ableton Live* or *Spatial Mixer* and the other will change accordingly.

2 Related Work

This research sits within spatial mixing and cross-device interaction, within the context of music mixing. In this section, we present relevant research from these fields.

2.1 Spatial Mixing

Music Mixing Surface is an application designed to enable composition of music on a tablet [6, 7]. The interface design of the app is based on the stage metaphor. Most standard composing programs, such as *Ableton Live*, use the traditional channel strip metaphor. The stage metaphor uses a “stage” as a place where each of the channels are represented as widgets within the music mix. The placement of these widgets on the stage determines panning and volume. This gives the composer an overview of the different channels’ proximity to each other and thereby their level of volume in relation to each other. Gelineck et al. [8] compared the performance of the stage metaphor with the channel strip metaphor for adjusting volume and panning and concluded that there was no significant difference between the two. However, users preferred the stage metaphor, finding it more intuitive.

RITZ [13] is a system built to provide real-time spatialization of sound sources. Nixdorf and Gerhard [13] describe how the system is designed to be a practical and powerful interface and one of the few viable for spatializing sound in real-time, such as during a concert. They suggest the need for dynamic speakers that allow the system to adjust to speaker positions.

Carrascal and Jordà [2] developed a multitouch application that also uses the stage metaphor. This application gives users an alternative way to interact with the parameters available on a mixing console. They compared the user interface of their application to the user interface of a physical mixing console. Five participants with no previous experience mixing music were compared with a sound engineer using the application. The study found that overall participants performed better on the application than the physical console, spending less time composing a mix they were satisfied with. They also found that use of the stage metaphor in the design made it much more intuitive for novice users who could easily understand the user interface and started mixing the music more quickly than with the mixing console.

2.2 Cross-Device Interaction

Schmidt et al. [15] studied the interactions for mobiles as well as shared interactive displays. In their work, they present cross-device interactions between a mobile and shared interactive displays. They proposed the use of mobile interaction styles across all displays. They developed a series of interactions for both input and output between the two surfaces and described application concepts in which they could potentially be used. They concluded that further research would be beneficial and would include an evaluation of the techniques over a longer period, expanding interactions to work across multiple surfaces.

Conductor [9], is a prototype framework that sets a baseline on how to create cross-device applications. The framework allows devices to be connected in a variety of ways, essentially extending the functionality of one device to another, allowing for more advanced interactions than a single device can perform. In their study, participants make use of up to 10 tablet devices to solve a sense-making task. Hamilton and Wigor [8] found that each test participant made good use of cross-device functionality to make the task easier for themselves.

JuxtaPinch [12], is an application that explores the use of multiple devices in the setting of sharing photos. With three different pinch gestures, users can pair multiple devices to display a single picture across all the connected devices. The purpose was to create a flexible way of sharing pictures when co-located with others. Nielsen et al. [12] found that participants engaged in the new interaction for sharing pictures with each other but found simultaneous use of the application challenging.

Dong et al. [4] conducted an interview study with 29 designers and developers, to better understand the difficulties faced when designing cross-device experiences. They identify challenges that need to be addressed in order to help ease the designing of multi-device systems. They also suggest the need for additional research on simulating multi-device experiences, to lower the barrier for developing good multi-device experiences.

Wäljas et al. [19] proposed a list of three themes for cross-platform usability, based on qualitative data collected through a field study. The characteristics they found are: Composition, Continuity and Consistency. In the study, participants kept diaries about their interactions and experiences with a set of services. Looking at the data and correlating the services with user activities and context of use, the researchers found these three distinct characteristics of cross-device environments that have an impact on the user experience of the cross-platform services.

The 4C framework, by Sørensen et al. [18] explains the interactions that take place within digital ecosystems. There are four themes in the framework: Communitality, Collaboration, Continuity, and Complementarity. *Complementarity* covers the idea that in interactions with multiple devices, one device is complemented by another providing the same or additional functionality. The principles of *Remote Control* and *Extension* each cover a different type of interaction to achieve this. The principle *Remote Control* maps the possible interactions from one device to another, allowing the new device to control the other without necessarily adding new functionality. The principle *Extension* on the other hand is where a device can complement the activity on another device with new functionality or other interactions.

2.3 Music Mixing

Traditionally, the production of commercial music can be a time consuming, expensive and complicated process with multiple steps, involving musicians, producers and a lot of analog

equipment [10]. Mixing consoles had to be very large in order to support multiple channels and all their controls. For each channel, the composer applies a series of different effect components to the tracks, such as an equalizer or a compressor, which they use to adjust the sound of the track. If the composer wants to apply the same effect component to multiple tracks they can reroute the tracks to a “bus” and then add the effect to all sound tracks [11].

Digital Audio Workstation (DAW) is a digitized work surface modeled on multitrack tape recorders and mixing consoles to fit traditional work paradigms [5, 14]. This replaced the expensive physical equipment a composer would normally need to acquire, opening up new possibilities for importing or creating effects and easily add multiple effects to a track to get the sound they want. The interface and general feel of the DAW stayed true to the original physical mixing console by retaining the layout with switches, wheels and faders, and the idea of a mixing tape [5]. This mixing tape is represented in the channel stripe metaphor. The DAW makes it possible to connect to multiple devices and hence apply cross-device interactions.

In comparison with the DAW, the **Music Mixing Surface** app [6] using the stage metaphor for mixing music in stereo, can show the position of the tracks on the iPad indicating panning and volume. The composer can work spatially and literally see where a track is positioned rather than having to imagine where each specific track is placed with respect to its panning and volume. The app was designed to be stand-alone but could also remote control the DAW through a simplified interface to make use of the stage metaphor. However, it does lose some of the functionality that the DAW can provide.

3 Prototype

Based on related work, illustrating the benefits of the stage metaphor [7, 8], **Spatial Mixer** was designed and developed to mix music in surround sound. It was designed as a companion app to a DAW to benefiting from the reduced complexity, usability and improved user experience of interacting with the stage metaphor interface. **Spatial Mixer** was also designed to work as an Extension, as described in the 4C Framework [18], with a focus on surround sound mixing as a more complex case than stereo.

The system consists of the DAW program, **Ableton Live**, on a desktop computer coupled with **Spatial Mixer** on an iPad (see figure 2). **Spatial Mixer** extends the functionality of **Ableton Live** by providing another way to interact with it. Additionally, they are coupled in such a way that changes made in one application are mirrored to the other. The **Spatial Mixer**, built around the stage metaphor, provides a visual representation of the track’s position relative to a listener on the stage. From the user’s point of view, spatial mixing should be more comprehensible, providing a visual representation rather than the abstracted channel stripe view with data shown as values.



Figure 2. The connected system, with **Ableton Live** running on the desktop computer and **Spatial Mixer** on the iPad.

4.1 Ableton Live

DAWs are extensively used in music production today. **Ableton Live** is just one example of a DAW. A benefit of **Ableton Live** over competitor DAWs such as **Logic Pro X**, for this study, is its accessibility to a MIDI API. The MIDI API makes it much easier to setup communication between the system components to be able to explore interactions between different devices.

To understand the setup and how the prototype works, it is important to know a bit about how **Ableton Live** works and looks. **Ableton Live** is a DAW used in both live performance and music composition and production, offering functionality such as effect plugins, and controls for music and effects. **Ableton Live**’s interface is partitioned into three main parts: a) an area where the channels are displayed; b) an area where the effects for a channel are placed and can be controlled; and c) the menu where the user can access a variety of tools, for instance effects, plug-ins and audio samples (see figure 3).



Figure 3. Ableton Live: a) channels (top right area), b) effects (bottom right area), and c) menu (top left area).

In area a) a channel can contain a track or work as a bus where multiple tracks’ outputs are held. Each channel has its own set of controls. On each track, there is a menu for selecting the audio source and destination. A track can have senders which control

how much of the track should be sent to a return channel, which can then be used to decide which speaker the sound is sent to. In figure 3, there are four senders (circle with a radius line), one for each return channel which represent a speaker each. A channel also has a fader to control the volume, a sender for panning and a button for muting. Each channel can have a series of effects attached to it. A track's effects are displayed when the track is selected. The Music Mixing Surface (area b, left hand side), is a group of eight senders where seven of them are mapped to different parameters of different effects. The parameters in the Music Mixing Surface are the ones affected by the *Spatial Mixer*.

4.2 Spatial Mixer

The *Spatial Mixer*'s design is based on the stage metaphor, where tracks are positioned on a stage (see figure 4). The tracks' positions are visualized as bubbles, called "track bubbles", in the spatial room relative to a user, called the "stage person".



Figure 4. Spatial Mixer interface on an iPad

A track bubble has a variety of interactions available to the stage person, who is at the center of the display. Around the stage person are two rings. The inner ring illustrates the area of the volume and the outer ring illustrates the area of the fall-off effect on a track. In the top left of figure 4, are a set of mix buttons which are used to switch between different mixes of the same music piece. A mix can be copied from one mix button to another to facilitate making changes to mixes with the same point of origin. At the right-hand end of the set of buttons is an undo button, which when tapped, will undo the last change the user made. On the right-hand side of the stage is a list of the different tracks. Above that list are four buttons. The first button from the left locks the stage person in position, so it cannot be moved accidentally. The second button shows or hides the list of tracks. The third and fourth buttons are used to connect the *Spatial Mixer* to *Ableton Live*.

A track bubble can be moved relative to the stage person, to indicate which direction the track's sound should come from and how loud it should be. Figure 5 (A) shows the track bubbles on

the stage, together with the track's volume and senders. Figure 5 (B) shows track bubbles that have been dragged to a new location. Moving track bubbles results in changes to the track's sender and volume levels. By moving the track bubble away from the stage person, the volume of that track will decrease until the track bubble crosses the inner ring, after which the volume will remain constant. If the track bubble crosses the outer ring the volume turns off. In this example, the Vocal's track volume has been turned down, while the Kick's track volume has been turned off, because the track bubble is outside the outer ring. The sends (speakers) for both bubbles have changed as well. This interaction corresponds to adjusting the volume slider and four sender levels individually in *Ableton Live* interface.

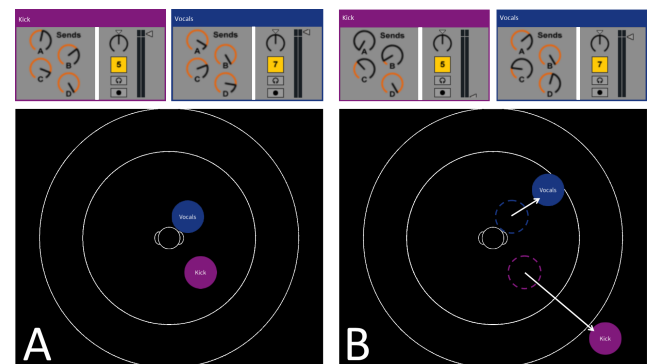


Figure 5. The Track Bubbles being moved showing corresponding changes to controls

The stage person can also be moved which adjusts all the tracks' sender levels and volumes (see figure 6).

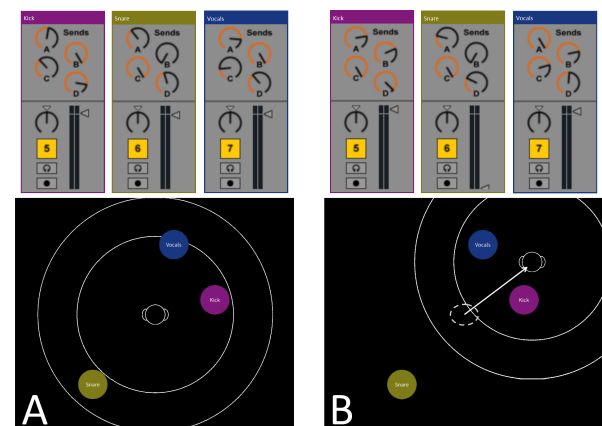


Figure 6. The Stage Person being moved showing corresponding changes to controls

In figure 6 (A) the stage person stands in the middle of the stage with track bubbles around them. In figure 6 (B) the stage person has been dragged while the track bubbles remain static. This movement of the stage person results in all the track bubble

values changing, reflected in the change in controls. If a track bubble ends up between the inner ring and the outer ring, then the fall-off effect is applied to the track. The closer the track bubble is to the outer ring the bigger the effect of the fall-off. The smaller the value, the more powerful the fall-off effect.

The inner and outer rings determine the volume and fall-off and they can also be resized. Resizing the inner ring changes the minimum value of the volume when crossing the inner ring. The smaller the inner ring the higher the minimum volume when crossing the inner ring. The fall-off effect is affected by the distance between the inner and outer rings. The bigger the ring the smaller the increase in fall-off effect when moving the track bubble toward the outer ring. The stage in Figure 7 has all the track bubbles together showing the track's volume and fall-off effect levels.

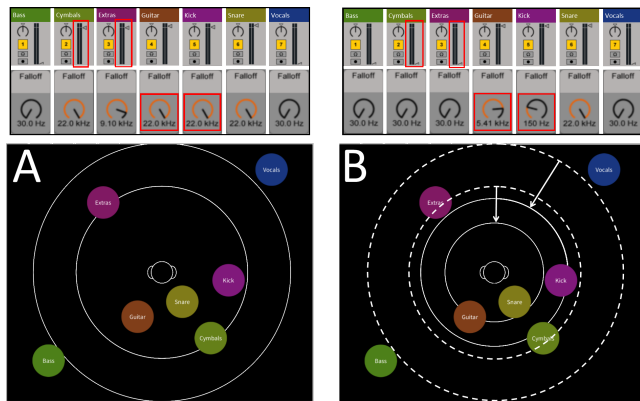


Figure 7. Effect of moving the inner and outer rings.

Figure 7 (B) shows the effect of moving both inner and outer rings

This movement results in Extras and Cymbals having zero volume since they are now outside the outer ring. Kick and Guitar now have fall-off because they are between the two rings.

A track can be muted by tapping the track bubble which corresponds to muting it using the controller in *Ableton Live*. This can also be done by tapping the track in the list of tracks. The tap gesture results in the track being muted in *Ableton Live*. A user can with the long press gesture, on the track bubble or the track in the list of tracks, mute all other tracks allowing that single track to play solo. Only one track can be soloed at a time. When a track is muted its color fades both on the track bubble and in the list of tracks (see figure 8). This effect carries over into *Ableton Live*.

By double tapping on a track bubble, a user can open the Effect Menu which displays the track's effects as slices (see figure 9 (A)). The user can then, by dragging the slices bigger or smaller, adjust the three preset filters on the equalizer and the compressor threshold in *Ableton Live*. In figure 9 (C) the user has dragged the reverb effect bigger, also changing the effect in *Ableton Live*.



Figure 8. A long pressing on a track bubble creates a solo track, and mutes all other tracks

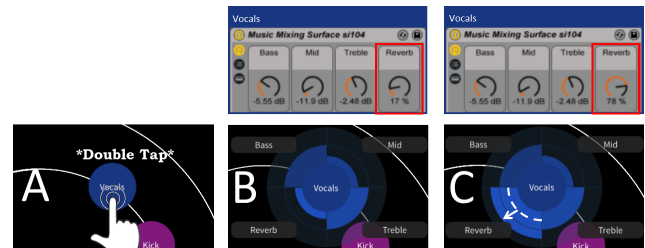


Figure 9. A double tap on a Track Bubble opens its Effect Menu, dragging to the Reverb corner increases that effect (C)

The user can also pinch a track bubble to make it bigger or smaller. This will adjust either the compressor or expander, respectively. As long as the track bubble is bigger than the standard size, the compressor is applied in degrees depending on its size. The bigger it is made, the greater the compressor effect. In figure 10, (A) to (B), as the bubble is made bigger, the compressor threshold moves from 6dB to -24.9dB. If the bubble is made smaller than standard size, the expander is applied.

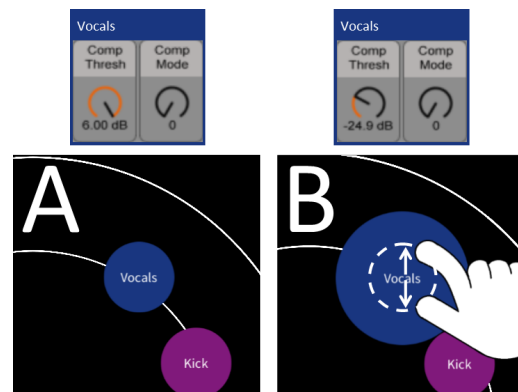


Figure 10. Pinching a Track Bubble to make it larger applies the compressor effect

4.3 Implementation

The prototype system was developed by modifying code provided by Steven Gelineck, from on the *Music Mixing*

Surface prototype [7]. The communication between *Spatial Mixer* and *Ableton Live* was changed to include surround sound. This included implementing surround sound in the app using new algorithms for calculating sender levels, volume and fall-off values for *Ableton Live*. However, there was a problem implementing how a track's position was updated based on the volume or fall-off without colliding with the position based on the senders' levels. Therefore, the volume and fall-off from *Ableton Live* are not used for positioning the tracks. The implementation of surround sound included the values necessary to facilitate it, as well as gestures and additional interaction techniques to operate it.

The system comprises *Ableton Live*, a plug-in (LiveOSC) for *Ableton Live*, a MIDI bridge and the *Spatial Mixer* app on the iPad. Development was focused on the *Spatial Mixer* app and the MIDI bridge, while *Ableton Live* and the plug-in were used as is. *Spatial Mixer* connects to the MIDI bridge on the computer, which ensures that it communicates *Ableton Live*, with the help of the plug-in. The MIDI bridge code base was modified to add more routes to facilitate the implementation of surround sound and add changes to existing functionality.

5 Use Study

We conducted a study of the *Spatial Mixer* in use to evaluate the cross-device interaction between the desktop computer, running *Ableton Live*, and the iPad, running the *Spatial Mixer* app. The focus of the study was on understand how people interacted between the two devices to support the music production activity.

5.1 Participants

We recruited five participants (P1-P5), aged 21 to 28, all male. All studied the University subject, "Popular Music and Sound Production", and were in the second or third year of the University degree. The participants were asked to rate their experience with different technologies and music mixing software programs and devices, on a scale of 1 to 10 where 1 was no experience and 10 was extensive experience. On experience with touchscreens, they rated on average 6.2, with the highest being 10 and the lowest 4. On experience with *Ableton Live* they rated an average 2.0, with 3 as the highest and 1 as the lowest. None of the participants had any experience with working with surround sound. Three of the 5 participants owned a smartphone, and 2 of 5 owned a tablet, though they commented, that they did not really use it. Only one of the participants did not own either a smartphone or a tablet.

5.2 Test Setup

The test was conducted at Musikkens Hus in Aalborg in a music practice room. The setup included an iMac with iOS X: El Capital (10.11.4) running *Ableton Live 9* (9.6.1), an iMac keyboard, an iMac mouse and an iPad Air 2 (64GB) with iOS 9.3.2, running *Spatial Mixer*. In addition, the room was fitted out with four

speakers, used to set up surround sound. A mobile phone was used as a timer for the tasks.

5.3 Procedure

The use study involved three hours with each of the five participants. Each participant was asked to complete three tasks using the system, and this was followed by a semi-structured post interview. In all three tasks the participants had to edit music tracks to improve their sound quality and use surround sound. Sound quality describes the style of the sound and is used to compare sounds with each other. All tasks had a time limit with a visible timer to give the participants an indication of how much time they had left. A test facilitator managed the test and helped the participants if they had any problems. A test logger logged the tasks and interview.

To have baseline data for comparison, the first task used only the *Ableton Live* program. This was also to give the participants a change to familiarize themselves with *Ableton Live*. The first task lasted 30 minutes and they worked with the music number "Heard it Through the Grapevine" by Marvin Gaye. After this base task the participants were given two tasks involving cross-device interactions between *Ableton Live* and *Spatial Mixer*, which lasted 55 minutes each.

Before the second task, participants were shown a demo of the interactions available in *Spatial Mixer*, on the iPad. During the second task, participants had to edit "Heard it Through the Grapevine" again, but they were given subtasks to make sure they used different aspects of the *Spatial Mixer*. The third task was an open task involving editing "Next To You" by The Police. This task did not have any subtasks, and participants were free to choose whether they wanted to use the *Spatial Mixer*, *Ableton Live* or both.

All use tasks and interviews were recorded on a video camera, focusing on the user interactions. The desktop computer and iPad were recorded using a screen recorder. The post use interview was fully transcribed and the tasks were logged to document the data and extract important information. The data was then sorted into different themes, using Thematic Analysis. These resulting themes were used to suggest design implications for designing cross-device interactions in the setting of spatial mixing in surround sound.

5.4 Limitations

Some activities from the study did not make it to the final analysis, so we worked with the data we had. The music files for the first two tests were out of sync, making it more difficult for participants to mix it. Some recordings were lost due to technical difficulties, including the screen recording for Task 1 for P1 and the camera recording of the last part of P3's Task 3. P4 stopped Task 2 before time claiming there was nothing left to do.

6 Results

After qualitative analysis of the data collected, several themes emerged. These are presented in this section.

6.1 Spatial Mixing

From the data collected, it was evident that all participants liked using the stage metaphor as a visual representation of the tracks, within a room context. As P1 stated,

"To be able to create a sound stage on this [the iPad]. That is brilliant, because it is actually a cool feature to be able to visualize it. And you have some programs in Logic [Logic Pro] which can do this a bit, which can get you to visualize it in some way, but not as concrete as this." - P1

Not only did it give them a clear visual of the music they were composing but it also allowed for a more hands on approach to the actual composing, as they could use the touch feature of the iPad to imitate some of the analog feel. As P3 said,

"It gives the possibility for your creativity/musicality to come through without you thinking too much about the boundaries/limitations there are, the ones you have acquired one way or the other. You skip all the boring numbers. It is something like the old school analog gear. There were not so many numbers either. There you had to turn it and feel one's way. It makes it more musical in a way, you do not sit thinking in numbers." - P3

Participants explained that using surround sound was only feasible if it was meant for something more than the music in itself. They explained that when they compose, they do it in stereo because the ordinary listener would not have access to a surround sound system or they would listen to the music through the radio which only broadcasts in stereo. As P4 explained,

"The idea about using surround sound is fine, but in practice the ordinary listener uses two speakers or headphones or listens in mono on their phone, then I do not know what I should use surround for." - P4

6.2 Cross-Device Interaction

In the study, P3 found a track in *Ableton Live* and then used the iPad to adjust that track's compressor. While doing this he shifted his attention between the desktop computer and the iPad. Other participants also tried to change a track's effect on the iPad but looked at a different track on the desktop computer. P5 was confused by the fact that the color of a track on iPad did not match the color of the same track on the desktop computer. He commented on the connection between the two devices,

*"The only problem was that the colors were f***ed [did not match] compared to the program [Ableton Live]." - P5*

P4 did not see a reason to both interfaces operating. He said that if *Spatial Mixer* app was on the tablet, it might as well be able to do everything *Ableton Live* can do. The notion of moving all

the functionality to the app was also indicated by P2 when he talked about the mobility of the app,

"The advantage is if you could have a mixing-program on the iPad, then you could sit anywhere and mix. You just need the iPad and then connect to a computer. Then you could make any place your mixing-studio if you just make sure the place is alright. The advantage is that it is portable and easy to use." - P2

Participants used the tablet as a way to get an overview and make general modifications, then make more advanced modifications on the desktop computer. P5 confirmed this,

"It gives another kind of overview than normally, because I use the computer to give an overview of where I am, while the app gives a greater overview because it is designed as it is." - P5

6.3 Change in Music Production Process

In doing the first task, all the participants would open the effects for the tracks one after the other and make adjustments to the compressor or equalizer before adjusting volume or panning, this way of working was confirmed by P2,

"Normally when I mix I usually do not make any panning at all until way into the process, usually everything is on top of each other in the center and then I make small adjustments on the equalizer and compressor." - P2

When presented with the *Spatial Mixer*, they immediately started to spread out the tracks on the iPad stage and experiment more with positioning than they did when working with *Ableton Live* alone. This was commented on by P5,

"I experimented more with the panning part because it is so easy compared to sitting there and then you have a track where you pan a bit that way [indicates he moves a track to his right], and then you find the other track and pan it a bit out there [indicates he moves a track to his left]. Here it is easier to quickly move things around and get them thrown everywhere." - P5

All participants made good use of the mute function on the iPad even though *Ableton Live* had this feature too. They also made good use of the solo function since *Ableton Live* did not provide this functionality when using return tracks. P1 expressed how the app changed his way of thinking and how he found it fun,

"In reality when it is about new technology is thinking out of the box. We start making the practice room in the mix, and that is a completely different way of thinking." - P1

6.4 Touch

During the test, some participants expressed really liking working with touch compared to the keyboard and mouse. They found the touch interactions familiar, as they often use them in other apps. They could just go ahead and interact with the system right from the start. What really made it intuitive to

work with was the fact that it felt similar to the analog equipment, as explained by P3,

"When you move over to the iPad, the information flow is drastically slowed and limited. You work more with static and simple forms, meaning they do not pulse or give you new numbers all the time. And I think it is there you see the intuitiveness and physical aspect." - P3

Some of the participants opened the Effect Menu of a track on the iPad and then just listened as they moved the slider back and forth. They generally preferred to make use of the touch interaction whenever possible to change an effect's parameter as mentioned by P4,

"[...] because it is easier to use your fingers, and when you have a screen where you can work with it graphically and use your fingers, it gives another kind of precision." - P4

6.5 Mapping

All the participants used the computer at some point to make adjustments to the effects of the tracks. Most were adjustments the iPad did not provide. Four of 5 participants returned to the iPad when they had made their adjustments. As P4 commented,

"I keep returning to the computer, because I feel I am missing some parameters to work with, and it makes it really hard to make a great mix on this one [the iPad]." - P4

When P4 turned away from the iPad to make adjustments on the desktop computer, he tended to stay there and make further adjustments that the iPad also could do. Participants also commented that the compressor effect on the iPad should adjust the ratio parameter instead of the threshold parameter,

"This means the more I expand, the more it reacts, but it is not the same, as it actually makes the sound bigger." - P2

Regarding the compressor, P1 mentioned that he did not like the compressor changing to an expander when you pinched it small enough. P1 said that the volume drop is too little near the stage person and too great further away from the stage person. On the other hand, P4 commented that he thought overall, the system was too sensitive.

Some participants adjusted volume or sender levels in *Ableton Live* but this was not represented on the iPad and when they moved the track on the iPad the adjustments were overridden.

6.6 Novice vs. Expert Users

Some participants thought the app was too simplistic. P5 said,

"I just do not think I am the target group. I think it might be more suited for novices." - P5

Other participants commented that they liked the metaphor behind the app as it made it easier to use. As P2 said,

"I really really like you have moved outside the metaphor typically used in music production programs which is about

a tape recorder or other studio gear. I really like this metaphor. It is straightforward and intuitive and I think it appeals to a lot who do not necessarily know how a traditional music program works. But have the idea that 'I just want the sound to come from over there or from over there or have this shape. I really like that.'" - P2

6.7 Usability Problems

All of the participants at one point had difficulty remembering which interaction did what, so they had a tendency to solo or mute a track when they wanted to open the Effect Menu. P1 commented that it was something he would get used to with time,

"There are of course things you have to get used to, how you push. It comes to you nice and easy. And about making it a fluent work process. It comes nice and easy too." - P1

None of the participants understood the principle behind the volume and fall-off circle. They either left them alone completely or placed them randomly, making the fall-off circle really big and the volume border really small. Alternatively, they placed a track somewhere and then adjusted the fall-off effect on the desktop computer. When asked about their understanding of the fall-off and volume circle P5 said,

"I never got the hang of how it worked, because I probably just ended up having a lot here in the middle [around the stage person]." - T5

Doing Task 3, P5 asked if it was possible to get a bigger screen for the iPad as he found the iPad screen too small and he mentioned, that it might become a problem if there are a lot of tracks, which modern music tends to have.

7 Discussion

In this section we discuss the implications of the findings with respect to the spatial mixing concept, the use of the stage metaphor and how this affects cross-device interaction.

7.1 Spatial Mixing

Although we added surround sound into our app because it is an interesting case in mixing, and for added complexity of tasks, it became apparent that surround sound was rarely used in general mixing practice. However, spatial mixing is still well suited for mixing in stereo and having the surround sound capability could lead to new inspirations when mixing music.

From the recordings made by our participants, it is possible to see that they rarely used the senders in *Ableton Live* and did not create any particular sound setup. One participant explicitly commented that it was hard to mix surround this way using the senders. This could indicate, that they were not really comfortable with the representation of the tracks and did not have an understanding of the sender's effect.

When participants worked with *Spatial Mixer*, they were much more encouraged to use the spatial room and created different kinds of setups of the tracks compared to when they only used *Ableton Live*. Participants would, for instance, move a track around the stage person while listening to the changes in the sound. This kind of interaction was not observed in Task 1, suggesting that they could better understand the spatial aspects of music using the stage metaphor than with the channel stripe metaphor. One participant said he did not find the *Spatial Mixer* useful, and yet, even though he reported not liking the app, he still used it for controlling the surround sound settings. This shows that for sender control, the stage metaphor is easier to use than the channel strip metaphor used in *Ableton Live*.

In using the stage metaphor, participants were more enticed to work with spatial mixing and were capable of setting up the tracks the way they wanted. The stage metaphor made it easier to do spatial mixing by providing users with an intuitive understand of how to apply different effects in their production. This is supported by statements from the participants, who commented that this new way of producing music is enjoyable, simpler and supports more creativity in the production. This indicates that for an extension to a DAW to be attractive to users, it needs to use a meaningful metaphor and facilitate user interactions that are easily understandable.

The way the stage metaphor is implemented in *Spatial Mixer*, means the volume and fall-off on a track depend on the distance to the stage person. Limits can be set using the two rings around the stage person. However, we found that none of the participants understood the rings and their purpose. One participant, after having their function explained to him, commented that in modern music production this would not make sense, since in most cases you do not aim for acoustic representation but a more electronic one, where you can adjust volume and fall-off effect independently. Therefore, it seems that users might not want what is a realistic case, but instead the functionality they are accustomed to. If an extension app is providing a new way to produce an effect that is not a standard parameter, such as volume or panning, then the user should be able to control whether they want to use it or not.

7.2 Cross-Device

In working with the cross-device system, participants would use the iPad to get the general overview of the tracks and to make use of the Effect Menu to modify pre-defined parameters in the *Music Mixing Surface* of *Ableton Live*. When there was something they could not do on the iPad they would turn to the computer and do it in *Ableton Live* instead. As some of the functionality was not mapped to reflect changes on devices, the system could end up in a state where the settings on both no longer corresponded. This caused frustration with participants, as their work got overwritten when they changed things on either device. Sometimes when modifying parameters on the iPad, participants would look at the computer monitor, while others would just focus on listening to the modifications they

were making. This indicates that even though they could hear the difference, some participants wanted to see the exact changes they were making to the parameter using the iPad.

During the study, there was sometimes confusion around making modifications to a track on the iPad and identifying which track it was on the computer. This indicates that participants were not able to easily link tracks on the iPad to the correct track in *Ableton Live*. Some participants compensated for this by selecting the track they wanted to work with on the computer before going to the iPad to continue working on it. This illustrates that participants had troubles identifying tracks across the two devices. This indicates that an app extension needs to make it very clear how an element on the app corresponds to that same element in the DAW. This can be achieved by concurrently highlighting the representation of any item selected on one device, on the other device. Another way to achieve this is to clearly indicate which tracks are linked on the computer and iPad, through for instance, clear color coding.

7.3 Change in the Music Mixing Process

One major change in the participants' normal mixing experience was the change from stereo to surround sound, which many of them had not tried before. Without the *Spatial Mixer* participants did minimal spatial mixing, but as soon as they got the app, they experimented with the panning and tried different setups of the surround sound. One participant told us that he would normally do panning as one of the last steps in his production, but with the *Spatial Mixer* he did it as one of the first things and afterwards began changing the sound effects.

In using the extension and the stage metaphor the mixing also became more physical for the participants, feeling closer to using a physical mixing surface. One participant said he liked that it became more about feeling how the music should sound instead of working with numbers. He liked the "*Hands-on*" feel of using touch, rather than a mouse, as it made it more natural. Most participants primarily used the *Spatial Mixer* and only moved to *Ableton Live* when there was something they could not do in the *Spatial Mixer*. However, one participant felt that the *Spatial Mixer* was inaccurate and he could not use it to mix in the exact way he wanted to. This indicates that mixing can be about how it feels to do it, but it also needs to be very accurate.

Cross-device interaction allowed participants to move between the computer and iPad and do work on both. Often participants used *Spatial Mixer* for an overview and *Ableton Live* for working with tuning effects. Participants did not express frustration with shifting between devices, indicating that well designed cross-device interaction can support music production without a big learning curve. The main problem with working across devices occurred when *Ableton Live* was receiving input, causing a lot of values to change. This was a critical problem as two participants lost work they had done and felt very annoyed at having to redo it. Therefore cross-device interaction can be used to support music mixing but the communication between

devices needs to be robust and accurate, so that users do not experience problems or information loss.

7.4 Mental Model

Participants quickly understood how the stage metaphor worked, indicating that it is a good way for users to get an overview of the tracks and their proximity relative to the listener and each other, as opposed to the channel strip metaphor. Despite the comments about the effects being too simplistic, many of participants returned to the iPad to adjust either the mid, bass, or treble effects or simply to begin working with another track after having made more advanced adjustments in *Ableton Live*. This shows that participants liked having the analog feel of adjusting sliders. The comments of simplicity also indicate that the participants actually understood what they were able to achieve in the app, and when they should use *Ableton Live* for more complex adjustments. So, given that they thought the extension was too simplistic at first, because it was missing a lot of the functionality they were accustomed to, they are still able to see the purpose of combining both devices. In the setting of mixing music, it is also important to create a similar "hands-on" feeling as the composer has with analog equipment.

8. Conclusions

In this research we focus on cross-device interaction in the context of spatial mixing. The use study showed that a cross-device system combining the stage metaphor with the channel strip metaphor successfully supported user understanding of how to mix music with surround sound.

Our findings suggest that when designing a cross-device system for music mixing, it is really important to understand the context in which the system will be integrated. It is important to evaluate which functionality should be extended while balancing the complexity of the extension. Additionally, any action on either device should have a corresponding change on the other device, to ensure consistency across devices. Also, users should be able to quickly identify the same element across the different platforms in the cross-device setup. This is to ensure that the users don't get confused when navigating between systems or make errors because they thought they were working on something else. Communication between devices has to be robust and accurate, to avoid confusion and loss of data, and maintain predictability in the system. We also found that the addition of a touch device helped achieve the "hands-on" feel of analog consoles but need to provide similar precision to that of the physical gear they represent. The stage metaphor supported spatial thinking and was successful in adding to the systems usability, as well as the user enjoyment of using *Spatial Mixer*.

Spatial Mixer provided a new way for the users to interact across multiple devices in the context of music production. Our study highlighted important considerations for designing extension music mixing apps that make use of the spatial aspects of music composition, especially when working with surround sound.

ACKNOWLEDGMENTS

We thank Steven Gelineck for fruitful discussions of the stage-metaphor, and all test subjects for their participation.

REFERENCES

- [1] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3 (2), 77-101.
- [2] Juan Pablo Carrascal and Sergi Jordà. 2011. Multitouch Interface for Audio Mixing. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 100-103.
- [3] Rodrigo de Oliveira and Heloísa Vieira da Rocha. 2005. Towards an Approach for Multi-device Interface Design. In *Proceedings of the 11th Brazilian Symposium on Multimedia and the Web*, 1-3.
- [4] Tao Dong, Elizabeth F. Churchill, and Jeffrey Nichols. 2016. Understanding the Challenges of Designing and Developing Multi-Device Experiences. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*. ACM, New York, NY, USA, 62-72. DOI: <https://doi.org/10.1145/2901790.2901851>
- [5] Matthew Duignan, James Noble, and Robert Biddle. 2010. Abstraction and activity in computer-mediated music production. *Comput. Music J.* 34, 4 (December 2010), 22-33. DOI: https://doi.org/10.1162/COMJ_a_00023
- [6] Steven Gelineck, Jesper Andersen, and Morten Büchert. 2013. Music mixing surface. In *Proceedings of the 2013 ACM international conference on Interactive tabletops and surfaces (ITS '13)*. ACM, New York, NY, USA, 433-436. DOI: <https://doi.org/10.1145/2512349.2517248>
- [7] Steven Gelineck and Dannie Korsgaard. 2014. Stage Metaphor Mixing on a Multi-touch Tablet Device. In *Audio Engineering Society Convention 137 Los Angeles: Audio Engineering Society*.
- [8] Steven Gelineck, Dannie Korsgaard, and Morten Büchert. 2015. Stage- vs. Channel-strip Metaphor - Comparing Performance when Adjusting Volume and Panning of a Single Channel in a Stereo Mix. In *Proceedings of the international conference on New Interfaces for Musical Expression (NIME 2015)*. The School of Music and the Center for Computation and Technology (CCT), Louisiana State University, Baton Rouge, Louisiana, USA, 343-346.
- [9] Peter Hamilton and Daniel J. Wigdor. 2014. Conductor: enabling and understanding cross-device interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 2773-2782. DOI: <https://doi.org/10.1145/2556288.2557170>
- [10] Brian J. Hraacs. 2012. A Creative Industry in Transition: The Rise of Digitally Driven Independent Music Production. *Growth and Change* 43 (3), 442-462.
- [11] Sean McLaughlin. 2014. Mixing with iZotope - Principles, tips and techniques. *iZotope*.
- [12] Heidi Selmer Nielsen, Marius Pallisgaard Olsen, Mikael B. Skov, and Jesper Kjeldskov. 2014. JuxtaPinch: exploring multi-device interaction in collocated photo sharing. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services (MobileHCI '14)*. ACM, New York, NY, USA, 183-192. DOI: <https://doi.org/10.1145/2628363.2628369>
- [13] JJ Nixdorf and David Gerhard. 2006. RITZ: A RealTime Interactive Tool for Spatialization. In *Proceedings of the 14th ACM International Conference on Multimedia*. ACM, 687-690.
- [14] Charles Roberts. 2008. Multi-Touch, Consumers and Developers. Technical report, Media Arts and Technology Program - University of California.
- [15] Dominik Schmidt, Julian Seifert, Enrico Rukzio, and Hans Gellersen. 2012. A cross-device interaction style for mobiles and surfaces. In *Proceedings of the Designing Interactive Systems Conference (DIS '12)*. ACM, New York, NY, USA, 318-327. DOI: <https://doi.org/10.1145/2317956.2318005>
- [16] A. Seffah and H. Javahery. 2005. *Multiple User Interfaces: Cross-Platform Applications and Context-Aware Interfaces*. Wiley.
- [17] Katarina Segerstahl. 2009. Crossmedia Systems Constructed around Human Activities: A Field Study and Implications for Design. In: Gross T. et al. (eds) *Human-Computer Interaction - INTERACT 2009*. INTERACT 2009. Lecture Notes in Computer Science, vol 5727. Springer, Berlin, Heidelberg.
- [18] Henrik Sørensen, Dimitrios Raptis, Jesper Kjeldskov, and Mikael B. Skov. 2014. The 4C framework: principles of interaction in digital ecosystems. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14)*. ACM, New York, NY, USA, 87-97. DOI: <https://doi.org/10.1145/2632048.2636089>
- [19] Minna Wäljas, Katarina Segerstahl, Kaisa Väänänen-Vainio-Mattila, and Harri Oinas-Kukkonen. 2010. Cross-platform service user experience: a field study and an initial framework. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services (MobileHCI '10)*. ACM, New York, NY, USA, 219-228. DOI: <https://doi.org/10.1145/1851600.1851637>