Stage- vs. Channel-strip Metaphor

Gelineck, Steven; Korsgaard, Dannie Michael; Büchert, Morten

Published in:
Proceedings of the International Conference on New Interfaces for Musical Expression (NIME 2015)

Publication date:
2015

Document Version
Peer reviewed version

Stage- vs. Channel-strip Metaphor - Comparing Performance when Adjusting Volume and Panning of a Single Channel in a Stereo Mix

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

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

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

ABSTRACT

This study compares the stage metaphor and the channel strip metaphor in terms of performance. Traditionally, music mixing consoles employ a channels strip control metaphor for adjusting parameters such as volume and panning of each track. An alternative control metaphor, the so-called stage metaphor lets the user adjust volume and panning by positioning tracks relative to a virtual listening position. In this study test participants are given the task to adjust volume and panning of one channel (in mixes consisting of three channels) in order to replicate a series of simple pre-rendered mixes. They do this using (1) a small physical mixing controller and (2) using an iPad app, which implements a simple stage metaphor interface. We measure how accurately they are able to replicate mixes in terms of volume and panning and how fast they are at doing so. Results reveal that performance is surprisingly similar and thus we are not able to detect any significant difference in performance between the two interfaces. Qualitative data however, suggests that the stage metaphor is largely favoured for its intuitive interaction - confirming earlier studies.

Author Keywords

User Interfaces, Usability Testing, Stage Metaphor, Mixing, Interface, Evaluation, Channel Strip, Multitouch, Music Production

ACM Classification

H.5.5 [Information Interfaces and Presentation] Sound and Music Computing, H.5.2 [Information Interfaces and Presentation] User Interfaces—Input devices and strategies.

1. INTRODUCTION

Traditionally, user interfaces for mixing music (mixing consoles) have employed the channel strip metaphor for controlling volume, panning and other audio effects. An alternative to this channel strip metaphor is the *stage metaphor* - see Figure 1. Here the different channels in a mix are represented by virtual widgets, which can be positioned relative to a virtual listener [1, 3, 8, 14]. David Gibson was the first to introduce a serious mixing tool implementing the stage metaphor in a virtual 3D interface called the Virtual Mixer [7]. Gibson and others [1, 5] argue that this way of mixing is potentially more intuitive since it corresponds to the way one would listen to sound (perceiving distance and directionality) in the real world [11]. While some evaluations [1, 4] indicate that the stage metaphor might be more intuitive than the channel strip metaphor, there has never been a quantitative performance comparison between the two interface types. This paper takes a step back attempting to evaluate whether there is a difference between how fast and how accurately one is able to position a sound in a stereo mix using the stage metaphor compared to using the channel strip metaphor. In other words, here we focus on the usability in terms of performance, as opposed to the user experience.

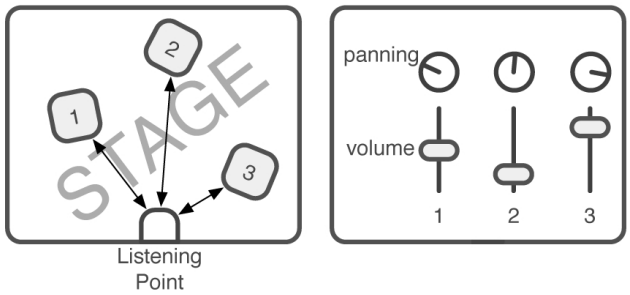

Figure 1: When employing the stage metaphor, volume and panning of an audio channel is determined by distance and angle in relation to a virtual listening point. Here the differences between stage metaphor channel positions (left) and channels controlled using a traditional channel strip metaphor (right) are illustrated.

The paper is organised as follows. Section 2 discusses related works within evaluation of control interfaces in general leading to the chosen methodology. Section 3 presents the actual experimental setup followed by results in Section 4. Finally, in Section 5 the results are discussed and concluded upon.

2. EVALUATIONS WITHIN NIME

Evaluation of musical devices, controllers, interfaces or instruments has been a growing concern within NIME. Wanderley and Orio [15] suggested several different methods for evaluation of input devices inspired by research carried

NIME'15, May 31-June 3, 2015, Louisiana State Univ., Baton Rouge, LA.

out within the field of Human-Computer Interaction (HCI). They proposed using simplified musical tasks for more objective evaluation of musical devices. Good overviews of different evaluation approaches are provided in [10] and [13]. Examples of where simple musical tasks have been used for evaluation of musical input devices include Marshall et al. [12], who compared input gestures for pitch modulation or Collicutt et al. [2], who compared four different input devices in order to estimate differences in player actions.

Especially the discussion of how simple a musical task can be when evaluating an interface for musical expression, is interesting to us. The answer to this question really depends on what is evaluated and for which intended context. For our purpose, we wish to measure how well users are able to place a sound in a mix using the two control metaphors, as described above. Although placing a sound in a mix seems like a simple task to do, it is more complex when carried out in the context of mixing music in the studio or live. The challenge for us has been to develop a musical task that represents a real musical goal in a certain context. In other words, our evaluation method must try to abstract the situation in which a producer, sound designer or audio technician has a musical idea (or thought) and tries to carry out that idea swiftly and accurately. At the same time we must simplify the experiment as much as possible in order to limit biases.

These considerations led to the following overall experimental setup. Test participants listen to a stereo track consisting of three separate channels of audio. Each channel is gained and panned differently. One of the channels is the test-channel that we ask the test-participants to replicate. After listening carefully, the test participant is asked to replicate the mix he or she heard.

We knew that the task should include more than the single channel in order to provide a reference for the channel that was to be manipulated. The number of additional channels (three) was chosen because it represented the minimum amount of channels that would still give a feeling of actually mixing a piece of music as opposed to just positioning a sound (it turns out that we may have simplified the task too much - more about this in Section 5).

3. EXPERIMENT

As explained earlier the goal of the experiment was to compare the channel strip metaphor to the stage metaphor when it comes to performance. The research questions were: are users (1) more/less accurate and/or (2) faster/slower in placing a single sound in a stereo mix when using the stage metaphor opposed to using the channel strip metaphor.

3.1 Methodological concerns

Each test-mix consisted of three channels, one of which had to be manipulated by the test participant. For each trial the test participant was asked to listen to a pre-rendered mix of the three channels (for instance drums, keyboards and bass), each with a different gain and panning. Participants were then asked to mix a piece of music with the same audio content as they had just heard, only focusing on one of the three channels. The other two channels were fixed in the correct positions, and the third channel (the one participants were asked to manipulate) started out being centred and gained to zero.

We chose to conduct the test using two different overall tracks, each containing different channels of audio—one contained bass, keyboards and **drums**, the other contained drums, congas and **guitar**—tracks in bold were the ones being manipulated. This was done to reduce any influence that a certain music type might have on the performance outcome. Each test participant went through several trials with different tracks, using different interfaces in a semi-randomised order.

An issue that often arises when comparing traditional interfaces with novel counterparts is that participants are often used to working with traditional interfaces. This gives the traditional interface an obvious advantage over the novel one. In order to explore this issue it was decided to test on both novice users (users with a musical background but with no mixing experience) and expert users (users with mixing experience of 3 years or more). This would let us estimate the effect of having experience with one interface over the other.

Figure 2: Interface used to represent channel strip metaphor mixing (left) and interface used to represent stage metaphor mixing (right).

3.2 Apparatus

Two tracks were set up in Ableton Live[1] each containing three mono channels. A Max[2] patch was built to receive input data from the control interfaces and to push that to Ableton Live for controlling volume and panning of each channel as well as for controlling simple playback. Additionally, the Max patch was used to store performance parameters for each trial. The two control interfaces used were (1) the Launchpad XL by Novation and (2) a custom built iPad app implementing the stage metaphor (see Figure 2). The iPad app was a simplified version of the Music Mixing Surface presented earlier by the authors [6]. The layout is kept as simple as possible with a virtual listener in the bottom center, three circles representing each channel, and buttons for controlling playback and for starting and stopping each trial. Circle distance to the listening position was mapped to MIDI Volume value in Ableton Live and circle angle to the listening position was mapped to MIDI Pan value in Ableton Live (an angle of -90 degrees corresponded to MIDI value 0 and +90 degrees corresponded to MIDI value 127). Max communicated with the Launchpad XL through standard MIDI and with the iPad app through Open Sound Control (OSC). See Figure 3 for an overview of the data-flow. Beyerdynamic DT 770 headphones were used for audio output from Live (and from Quicktime, in the case of the pre-rendered mix).

The reason we separated the audio engine from the controller was to make sure the experiment was not biased because of differences in audio quality. The reason for choosing the Launchpad XL and the iPad was to have a compact setup where both controllers were of similar size. We considered exchanging the Launchpad XL with an iPad app with virtual sliders and knobs (as used in [6]) in order to keep everything but the control metaphor equal between the two interfaces. However, we decided to use an interface with tangible sliders and knobs because it was a better representation of a *traditional* interface. Especially the rotary knob used for panning on traditional mixing consoles is not

[1] `https://www.ableton.com`
[2] `http://www.cycling74.com`

well replicated in GUI versions, where one normally grabs the knob by touching down and performs either a vertical slider motion or a directional angular motion to manipulate the associated value up or down.

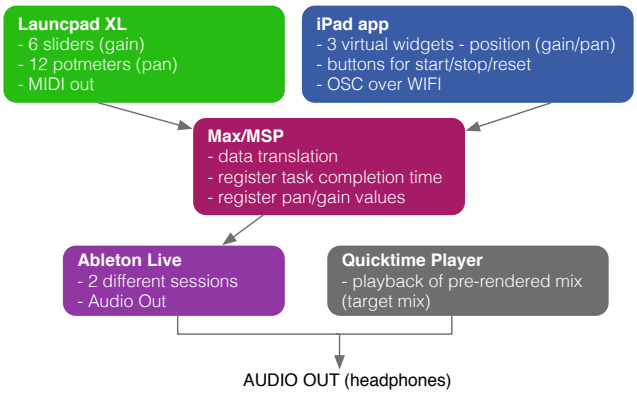

Figure 3: Overview of the overall system.

3.3 Test procedure

Experiments were carried out over two days. 15 participants (2 female, 13 male) took part in the evaluation (7 experienced and 8 novices). Each test took approximately 15-20 minutes. Initially, test participants filled in a questionnaire regarding demographics, including their musical and mixing experience. They were then presented with the overall test procedure. Each participant went through three setups - one practice setup and then one setup for each interface. Each setup was a combination of an interface (Launchpad or iPad) and a piece of music (Drums or Guitar). For each setup the participant went through 5 trials, each with a different mix of the three channels[3]. For each trial they went through the following steps:

- Participants practiced 1-3 times until comfortable with how the interface worked and what to listen for. After that the actual trial could begin.
- First they were asked to listen to a prerecorded mix. They were told to listen carefully to especially the channel for which they were supposed to replicate the volume and panning.
- After listening for 8 seconds, they turned to the interface. They were asked to hit a button starting the audio. They proceeded to adjust volume and panning of the target audio channel. When satisfied, they were asked to hit a stop button, which also registered their task completion time and the values for volume and panning. (They were not told that they were timed).
- They then went on to the next trial as explained above.

Finally, the participants were asked to comment on their experience, especially noting the difference between using the two interfaces and guessing which interface they were most accurate with.

4. RESULTS

Each of the 15 participants provided 5 trials for each of the two interfaces. This gives a sample size of 75 for each interface. For each sample, scores were calculated as the absolute difference between the MIDI value of the correct placement (the vol/pan setting used for the reference mix) and the MIDI value of the registered placement (vol/pan set by the participant). If the participant for instance listened to a mix where the drums were panned to the right (MIDI value: 93) and afterwards adjusted the pan of the drums to the center (MIDI value 64), he or she would obtain a score of $|64\text{-}93| = 29$. The lower the score, the better the performance.

Figure 4 shows mean scores and 95% confidence intervals for volume and panning for both iPad and Launchpad. No statistically significant differences were found.

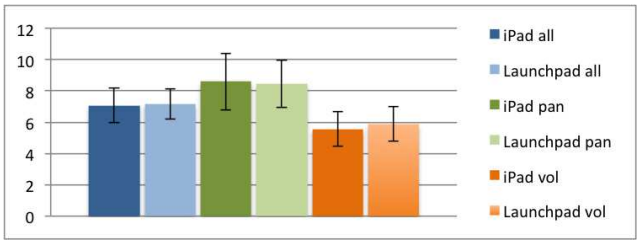

Figure 4: Means and 95% confidence intervals for iPad and Launchpad scores. The lower the score, the better the performance.

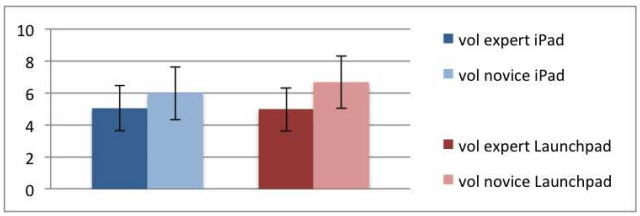

Figure 5: Means and 95% confidence intervals for expert and novice volume scores. Notice that volume scores differ more for Launchpad than for iPad.

The similarities between the scores for iPad and Launchpad were clear also when looking at expert and novice users separately. For both groups there was no significant difference between using iPad and Launchpad. When comparing the groups to each other, there was a significant difference between their ability to adjust panning in general (expert: 6.0 / novice: 10.7). There was also a difference between their ability to adjust volume (expert: 5.0 / novice: 6.3), but not statistically significant. Comparing task completion times also showed no significant differences, both when regarding novice and expert users separately and as a whole.

Interestingly, the volume scores for novices and experts were similar for iPad (see *vol expert iPad* and *vol novice iPad* in Figure 5), while the difference between novice and expert volume scores for the Launchpad was more explicit (see *vol expert Launchpad* and *vol novice Launchpad* in Figure 5). This may suggest that expert users are better at adjusting volume than novice users when using the Launchpad but only marginally better when using the iPad. This makes a lot of sense, since both experts and novices have no practice using the iPad interface (resulting in only marginally different scores). However, this is merely suggested, as the difference measured between expert and novice Launchpad volume scores falls short of statistical significance (p=0.12).

5. DISCUSSION

We were not able to detect any significant difference between using the two interfaces for this simple mixing task.

[3]go to `http://media.aau.dk/~stg/nime15_data.html` to listen to each mix and see target MIDI values for each trial

The only statistically significant difference was found between novice and expert users' ability to adjust panning in general.

The fact that no difference was found does not mean that no difference exists. While we were very careful to avoid biases in this experiment there might have been biases in regards to learning. It seemed that users would get more and more comfortable with the task the more trials they performed – this was also stated by some of the test participants. We did try to regard only the last 5 trials (out of 15) for each participant to see if this would yield any statistical significance, but the tendencies were the same.

After each evaluation session the test participants were asked which interface they thought they performed best with. Also here participants were mostly in doubt. Only a few dared make a decision and these revealed no tendency. While most of the participants had the feeling that they performed equally well with the two interfaces, most said that they found the stage metaphor interface more intuitive, enjoyable or inspirational to work with. Two of the expert users stated that they were surprised about how easy the interface was to learn. One directly stated: *".. I want to start mixing this way."*

There were discussions about the difference between being provided with a strong visual representation of the mix (stage metaphor) and just having to listen to what was going on (channel strip metaphor). There were several comments about how the iPad app made the user more aware of the "spaciousness" of the mix and in connection to this, the stage metaphor was suggested as a teaching tool.

Some stated that the Launchpad forced them to listen more and many closed their eyes while adjusting the slider and knob. However, the few who also closed their eyes while using the iPad app said that it was easier to connect the changes they listened to with the one-finger movement associated with moving the virtual widget about on the stage of the iPad. As one participant stated *"on the iPad I only have to move in one direction to get to where I want... with the mixer* [Launchpad] *I have to move two things at once"*. Another stated: *"with the mixer I had to think 'loud' and 'to the right', but with the iPad I just go directly... "*. In other words the integrality [9] of the stage metaphor might be more natural for adjusting volume and panning than the separable controls of the channel strip metaphor.

The measured performance similarity between the two interface types may also have been due to the simplicity of the task. One participant directly stated that he did not experience a difference in performance between the two, but if he would have been asked to manipulate multiple channels he would probably choose the iPad for better performance. We find this very likely, since every time a user is asked to manipulate a new channel, the amount of control points are doubled when using the channel strip metaphor (as opposed to using the stage metaphor). Note that this ease might decay at some point as channels increase leading to clutter and lack of overview. A future experiment is planned to explore this hypothesis.

6. CONCLUSION

We have presented an evaluation comparing the channel strip metaphor and the stage metaphor in terms of performance speed and accuracy. The task was to replicate a pre-rendered mix, but only focusing on positioning one of three channels in the mix. The channel strip metaphor was represented by a traditional mixing interface with knobs for panning and sliders for setting volume. The stage metaphor was represented by a custom built iPad app.

The study showed that there was no significant difference in terms of performance between the two interfaces. The simplicity of the task might have lead to this result and a future study will explore this. Qualitative feedback suggests that the stage metaphor is preferred for its intuitiveness, enjoyability and its ability to reveal the "spaciousness" of the mix. However, there is a risk that the visual feedback can take away focus from listening. Future studies will look into this.

7. ACKNOWLEDGMENTS

We would like to thank all the test participants.

8. REFERENCES

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

[2] M. Collicutt, C. Casciato, and M. M. Wanderley. From real to virtual: A comparison of input devices for percussion tasks. In *Proceedings of NIME*, pages 4–6, 2009.

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

[4] S. Gelineck, J. Andersen, and M. Büchert. Music mixing surface. In *Proceedings of the 2013 ACM international conference on Interactive tabletops and surfaces*, pages 433–436. ACM, 2013.

[5] S. Gelineck, M. Büchert, and J. Andersen. Towards a more flexible and creative music mixing interface. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pages 733–738. ACM, 2013.

[6] S. Gelineck and D. Korsgaard. Stage metaphor mixing on a multi-touch tablet device. In *Audio Engineering Society Convention 137*. Audio Engineering Society, 2014.

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

[8] A. Holladay and B. Holladay. Audio dementia: A next generation audio mixing software application. In *Audio Engineering Society Convention 118*. Audio Engineering Society, 2005.

[9] R. J. K. Jacob, L. E. Sibert, D. C. McFarlane, and M. P. Mullen, Jr. Integrality and separability of input devices. *ACM Trans. Comput.-Hum. Interact.*, 1(1):3–26, 1994.

[10] C. Kiefer, N. Collins, and G. Fitzpatrick. Hci methodology for evaluating musical controllers: A case study. In *Proceedings of NIME*, volume 8, 2008.

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

[12] M. T. Marshall, M. Hartshorn, M. M. Wanderley, and D. J. Levitin. Sensor choice for parameter modulations in digital musical instruments: Empirical evidence from pitch modulation. *Journal of New Music Research*, 38(3):241–253, 2009.

[13] S. O'Modhrain. A framework for the evaluation of digital musical instruments. *Computer Music Journal*, 35(1):28–42, 2011.

[14] J. Ratcliffe. Hand motion-controlled audio mixing interface. In *Proceedings of NIME*, 2014.

[15] M. M. Wanderley and N. Orio. Evaluation of input devices for musical expression: Borrowing tools from hci. *Computer Music Journal*, 26(3):62–76, 2002.