Audio Mixing Displays: The Influence of Overviews on Information Search and Critical Listening

Josh Mycroft, Tony Stockman, Joshua D. Reiss

The Centre for Digital Music, Queen Mary, University of London.

j.b.mycroft@qmul.ac.uk, t.stockman@qmul.ac.uk,
joshua.reiss@qmul.ac.uk

Abstract. Due to the compartmentalisation of the mix into separate channel strips, Digital Audio Workstation displays often result in global attributes and relationships between channels becoming hard to discern. This is despite the fact that overviews of mix information may help support direct visual comparison and make the visual data more coherent, thereby reducing the cognitive load involved in navigation and freeing up resources for critical listening (the focused listening to details of an audio mix, including dynamics, tone, blend and stereo placement). In this study participants were asked to locate visual information from a 24-channel mixer using interface designs with and without overviews. At the same time as locating the visual information, participants were required to identify changes in the panning (stereo positioning) of specified tracks in an audio mix. Results suggest that overviews not only allow more efficient search for visual mix information, but also significantly improve concurrent critical listening tasks.

Keywords: Audio Mixing, Digital Audio Workstations, Interaction Design, Critical Listening, Overviews.

1 Introduction

The Channel strip design, while being the most widely adopted for Digital Audio Workstations (DAWs) may not be the best suited to exploring the mix [1]. Due to the compartmentalisation into separate channel strips, global attributes and relationships between channels such as the relative levels of audio effects, gain levels, pan positions and processing become hard to discern [2]. Furthermore the need to navigate through several separate channels can inhibit the engagement and 'flow' of the mixing process and impede the user's ability to quickly respond to the programme material [3]. Previous work by the authors has also found evidence that increased interface navigation associated with the channel strip design may overload working memory and reduce aural acuity [4].

Alternative visualisations of the mix that support a better understanding of the global mix space have been proposed, though not widely adopted. For example, the stage metaphor has been suggested as a viable alternative to the channel strip [5,6,7,8]. In this metaphor the mix is broken down into three planes; width, depth and height. The enhanced global understanding of the mix that this provides may help users to quickly see patterns within the mix, avoid common errors such as masking or

bunching of elements within a certain stereo position [9] and allow any outliers (in terms of volume, pan etc.) to be easily attended and selected [10].

Using a stage metaphor, an overview of a channel strip mixer and a scrolling channel strip interface, this paper investigates the extent to which visual presentation affects the detection of global mix information. Furthermore the authors aim to assess whether by reducing navigation (and subsequent issues of disorientation and cognitive load) overviews may facilitate improved concurrent critical listening (the focused attention to details of an audio mix, including tone, blend, dynamics and stereo placement). It is anticipated that the results of this study may be of benefit in providing heuristics for interfaces that better support screen-based audio mixing workflow.

2. Study Design

2.1 Participants

Nine participants were recruited from Music Technology staff and students at City and Islington College, London (7 male, 2 female and aged 18-43) all with a minimum of one-year experience mixing audio on computers. All participants were required to give informed consent to participate in the study. The study was conducted in accordance with the guidelines of the University. The Ethics Committee of Queen Mary, University of London, approved the details of the study.

2.2 Visual Task

Three versions of a 24-channel interface showing volume and pan-position were designed using Max/MSP. The designs consisted of a channel strip with all 24 channels shown on a single page without the need to navigate (fig 1, top), a design using a stage metaphor presented on one page without the need to navigate (fig 1. middle) and a channel strip mixer requiring scrolling navigation to view all 24 channels (fig 1, bottom). The scrolling channel strip was included as it is the predominant metaphor used in DAWs, and provides a reference against which to measure the effectiveness of the overview designs.

For each of the interface designs participants were asked to answer four questions about the visual information and select the correct answer from a drop down menu above each interface (fig 1). The questions were designed to test quick visual referencing (i.e. which channels are panned to extremes, whether the mix has more channels above or below the centre volume) as well as more specific visual referencing questions (i.e. how many channels have volume or pan set between certain values; what is the panning/volume positions of specific named channels). Each of the four question types were asked in each interface design so that a direct comparison of the time to find particular visual information could be analysed for each interface. Participants were presented four occurrences of each interface type (with the order randomised for each participant) making twelve screens in total.

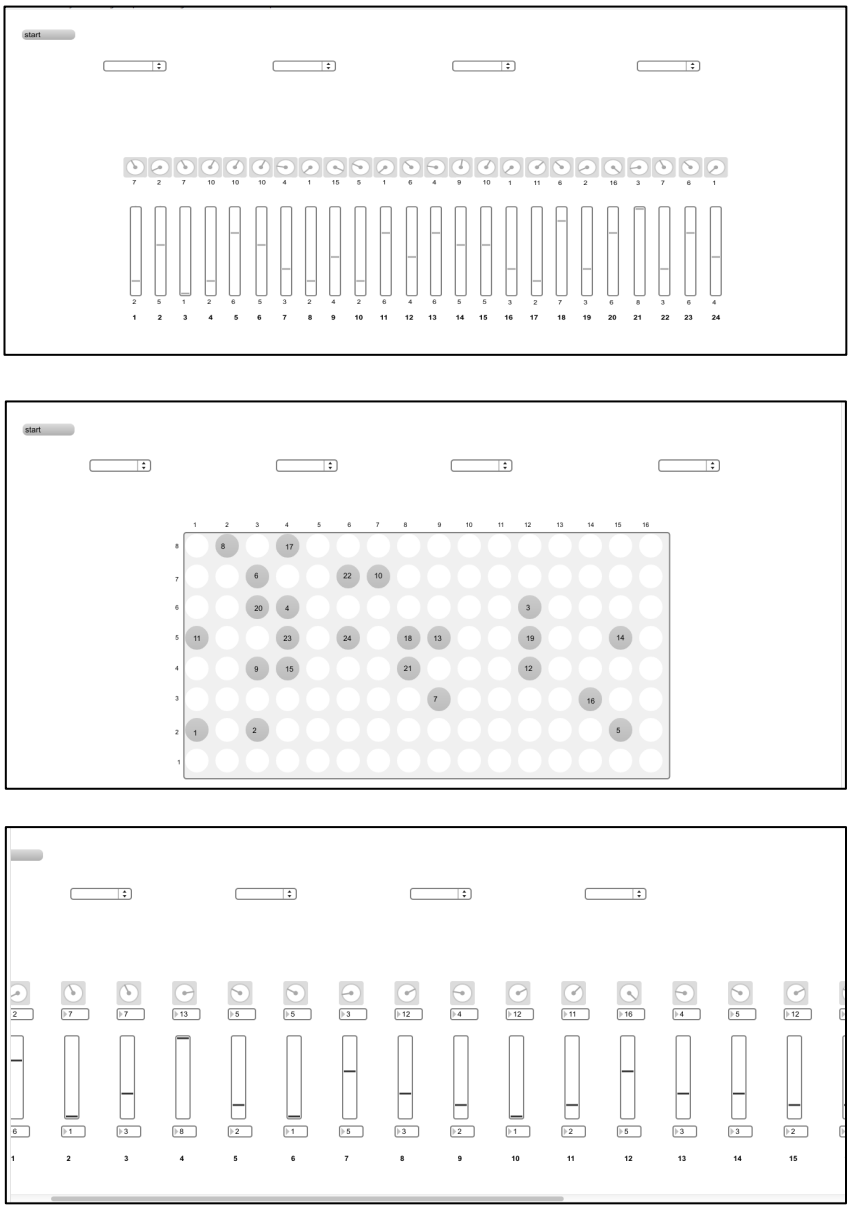

Fig 1. *Top;* the 24 channels as a mixer overview, all channels are on one page without the need for navigation. *Middle*; the 24 channels as a 'stage' overview, the left right position represents pan position while the up down position represents volume, channel numbers appear in the circles. All channels are on one page without the need for navigation. *Bottom;* The 24 channels as a traditional scrolling design; the faders and dials are larger in this design requiring navigation to view the channels (which do not all fit onto the one page).

2.3 Listening Task

While the participants were undertaking the visual task they were played a twelve-channel audio mix (duration 44 seconds). The mix was played twelve times for each participant (corresponding to the twelve visual interfaces). Each time the excerpt was played, four of the instruments within the mix (namely vocal, snare, flute and tambourine) were randomly panned either left, right or centre. As soon as the excerpt had finished the mixer screen was automatically closed and participants were asked to select the correct pan position of two of the instruments (chosen at random) from a drop down menu with the categories; left, centre, right or couldn't tell (this last option was included to avoid participants guessing the answer if they were unsure).

On completion of the study, each participant was asked to comment on how they perceived their performance in each interface type in terms of their success in locating the visual information and correctly detecting the audio panning.

3. Analysis and Results

The amount of correctly answered visual and auditory questions (and the time taken to answer) were recorded and analysed for each participant per interface type. From this the mean average and standard deviation was calculated for the three interface types (tables 1 and 2). The mean and standard deviation generated confidence levels (at 95%) showing the range of the true population per interface type (figures 2 and 3). A comparison of time and accuracy by question type was also analysed to quantify whether the various interface designs supported particular types of visual search better than others.

The analysis of the listening task per interface type reveals that the stage overview design provided significantly higher amounts of correctly identified audio panning than the scrolling interface (fig 2). While the mixer overview also provided higher amounts of correctly identified panning (relative to the scrolling interface) it was not significant. This finding is in line with the results from previous studies that suggest that simplifying visual search, specifically by reducing interface navigation, may result in improvements in concurrent auditory processing [4, 11, 12].

In terms of the visual search task, there were also significant differences between the three interface designs. The scrolling interface had the lowest amount of questions correctly answered, significantly less in fact than both the mixer and stage overviews (fig 3). While the amount of correctly answered visual search questions in the stage overview was not significantly greater than the mixer overview, it was an increase. This may be due to the fact that in the stage overview both the panning and volume of the channels were represented within the same User Interface object, making it better suited to perceptual limits that dictate that only a few items are attended to at any one time [13].

Table 1. The mean and standard deviation of amount of audio panning positions correctly identified per interface type.

Interface Type	Mean	Standard Deviation
Mixer Overview	5.11	1.76
Stage overview	5.77	1.09
Scrolling	3.88	1.16

Table 2. The mean and standard deviation for amount of visual search questions correctly identified per interface type.

Interface Type	Mean	Standard Deviation
Mixer Overview	7.66	0.86
Stage overview	8.66	1.73
Scrolling	5.6	1.22

Another finding from the analysis was the time taken for particular visual search tasks. The stage overview provided the quickest times when it came to recognizing patterns within the mix (e.g. whether more channels were panned left than right, whether more channels had volume below rather than above half way) which tallies with literature which suggests that overviews allow the user to effectively comprehend the relationships between data and discern global patterns more easily [14].

However, the slowest times in the stage overview occurred when participants were required to find information about specific channels. Due to the random distribution of the channels (they were not numerically ordered as in the other two designs) it took participants a longer time to locate the channels within the interface. Future studies aim to address this by finding ways to query the data using dynamic query filters so that the visual information can be displayed in a rapid, incremental and reversible manner [15].

Finally, when asked which of the interfaces had allowed the participants to successfully match both visual and listening task, the majority responded that the stage overview had been the least successful (though in fact this was the opposite). The explanation for this may be due to the unfamiliarity of the stage overview; interface designs which are common will often be accepted as the most natural by users, even though they may not represent the best possible interaction [16].

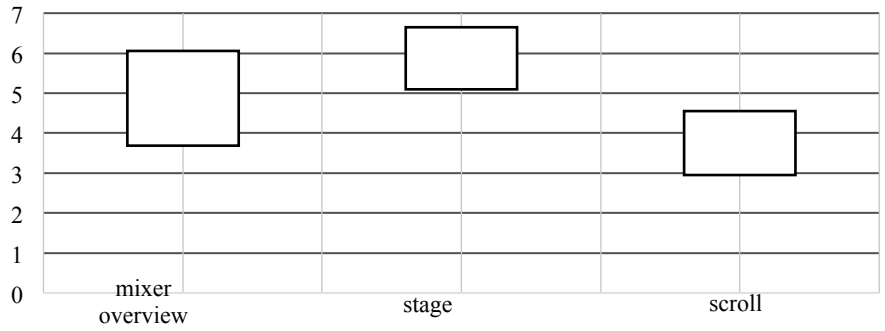

Fig. 2. Confidence intervals (95%) for the amount of correctly identified panning positions per interface type. While there is an overlap between the scrolling interface and the mixer overview, the stage overview shows a significant increase compared to the scrolling design.

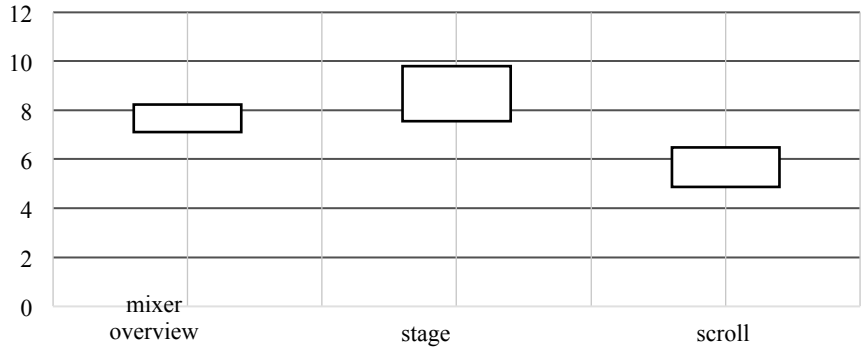

Fig. 3. Confidence intervals (95%) for the amount of correctly answered visual questions per interface type. The mixer and stage overviews show a significant increase compared to the scrolling design. The highest amount of correctly answered visual search questions occurred in the stage overview.

Conclusion.

This study suggests that compared to scrolling interface designs, overviews can be effective in improving participants' ability to discern visual information while undertaking concurrent critical listening tasks. Under complex navigation, as often found in large multi-track audio mixes, user orientation becomes a key issue and providing well-designed global views of the data is an important criterion for successfully navigating the information space [17]. Furthermore, the results of this study suggests that, in line with the authors previous work, reducing the need to navigate the interface to find visual information can significantly improve the users abilities to hear concurrent audio changes to the programme material [4].

While further work is needed to ascertain the amount and variety of information that can be effectively discerned in mixing overviews the results of this study provide some early indications of the benefit that overviews may have for mixing workflow, especially in light of the unlimited tracks offered by current DAWs and the growing use of smalls screen and tablet computers for audio mixing workflow.

References

1. Cartwright, M, Pardo, B, Reiss. J. Mixploration: Rethinking the audio mixer interface. In 19th Int. Conf. on Intelligent User Interfaces, pages 365–370, 2014. (2014)
2. Theberge, P. Any Sound You Can Imagine. Wesleyan University Press. (1997).
3. Szalva, W. Behind the Gear. Tape Op Magazine, No.73, pages 10-11. (2009).
4. Mycroft, J., Reiss, J.D., Stockman, T. The Influence of Graphical User Interface Design on Critical Listening Skills.Sound and Music Computing (SMC), Stockholm, (2013).
5. Moore, Allan, and Ruth Dockwray. "The Establishment of the Virtual Performance Space in Rock." Twentieth-Century Music 5 (2): 219–41.Twentieth-Century Music, 5/2, pp. 219–41. (2010).
6. Hodgson, J. Understanding Records: A Field Guide To Recording Practice. Continuum. (2010).
7. Gibson, D. The Art of Mixing. ArtistPro Press. (1997).
8. Moylan, William. Understanding and Crafting the Mix: The Art of Recording. 2nd Edition Oxford, Focal Press. (2007).
9. Case , A. Sound FX: Unlocking the Creative Potential of Recording Studio Effects. Elsvier, Oxford, UK. (2007).
10. Wolfe, J.M., Klempen, N., &Dahlen, K. Post-attentive Vision. The Journal of Experimental Psychology: Human Perception and Performance, 26(2): 693-716. (2000)
11. Dehais, F., Tessier, C., Chaudron, L. GHOST: experimenting conflicts countermeasures in the pilot's activity. Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI) 18, 163e168. (2003).
12. Sörqvist, P., Ljungberg, J. K., and Ljung, R. A sub-process view of working memory capacity: evidence from effects of speech on prose memory. Memory 18, 310–326. (2010).
13. Rensink, R. The Management of Human Attention in Visual Displays.In Human Attention in Digital Environments. Edited by Claudia Roda. Cambridge University Press. (2012).
14. Shneiderman Designing the User Interface: Strategies for Effective Human-Computer Interaction. Addsion-Wesley, Maryland. (1998)
15. Mendelzon, A. Visualizing the World Wide Web. In Proc. of the working conference on advanced visual interfaces (AVI). (1996)
16. Harrower, M. and B. Sheesley. 'Designing Better Map Interfaces: A Framework for Panning and Zooming', Transactions in GIS, 9, 77–89. (2005).
17. Jul, S., & Furnas, G. W. Critical Zones in Desert Fog: Aids to Multi-scale Navigation. In Proceedings of User Interface and Software Technology (UIST 98) ACM Press, 97-106.(1998).