Haptic and Visual feedback in 3D Audio Mixing Interfaces

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

Dan Overholt
Aalborg University Copenhagen
Department of Architecture, Design and Media Technology
A. C. Meyers Vænge 15, 2450 Copenhagen, Denmark
dano@create.aau.dk

ABSTRACT
This paper describes the implementation and informal evaluation of a user interface that explores haptic feedback for 3D audio mixing. The implementation compares different approaches using either the LEAP Motion for mid-air hand gesture control, or the Novint Falcon for active haptic feedback in order to augment the perception of the 3D space. We compare different interaction paradigms implemented using these interfaces, aiming to increase speed and accuracy and reduce the need for constant visual feedback. While the LEAP Motion relies upon visual perception and proprioception, users can forego visual feedback with interfaces such as the Novint Falcon and rely primarily on haptic cues, allowing more focus on the spatial sound elements. Results of the evaluation support this claim, as users preferred the interaction paradigm using the Falcon with no visual feedback. Furthermore, users disliked active haptic feedback for augmented perception of 3D space or for snapping to objects.

Categories and Subject Descriptors
H.5.1 [Multimedia Information Systems]: Audio input/output.; H.5.2 [User Interfaces]: Input devices and strategies.

General Terms
Human Factors, Design, Experimentation

Keywords
User Interface, 3D Audio, Mixing, Haptic Feedback, Mid-air gestures, Leap Motion, Visualisation, Haptic User Interface, Novint Falcon

1. INTRODUCTION
Spatial 3D audio is becoming more and more common in cinema, games and other artistic contexts. Technology has matured and commercial standards for 3D audio are fairly well established. While research on user interfaces for mixing of 3D audio is still in its infancy, some interesting approaches are emerging. Commercially, user interfaces currently implement traditional faders and rotary knobs as well as 2D joysticks for positioning and manipulating audio sources in three dimensions. However, more progressive commercial interfaces include the EVO with 3DAW from Fairlight, which adds support for the LEAP motion controller enabling users to position audio sources using free hand movements.

A recent study investigated different approaches comparing the LEAP motion controller to a more traditional mixing console interface. This study suggested that while there was great potential in using an integral controller with 3 degrees of freedom, where movements are directly coupled to movements of audio sources in 3D space, there were challenges to do with especially precision, ergonomics and fine-tuning because of the lack of tangible or haptic feedback of the interface. Additionally, the results of the study suggested that engaging in interfaces that rely heavily on visual feedback decreases the listening experience (a phenomenon also observed by Lech & Kostek).

Here we extend the research by investigating the use of a haptic feedback device for assisting the user not only in augmenting the feel for the 3D space, but also potentially en-
Figure 2: The GUI displays a 3D cube representing the available positioning space. Each virtual sound source is represented by a coloured sphere and the cursor is represented by a grey sphere. Shadows help the user perceive depth positions of each sphere.

Figure 3: When the cursor is close enough to a sound source it can be selected and it turns white.

Several approaches to mid-air musical control, including the control of spatial audio, exist \cite{11, 10, 9, 1}. After having explored mid-air gesture controls \cite{5} we were interested in incorporating haptic feedback. The technology has matured and several devices and software libraries are available, including the Open Source Haptics library \cite{11} used in this study — see more in the Implementation section. Sinclair provides a good overview of related works on using haptic feedback devices for musical purposes \cite{17}. Further advances in this area include \cite{13, 2, 3}. Most relevant for the research presented here is the work by Melchior et. al \cite{11}, which compares different forms of haptic feedback with mouse and slider input for controlling different movements of audio sources in 3D space. We build upon their research by exploring methods for augmenting the perceptual feel of the 3D space (by haptically providing information about the current 3D position), for fast and precise selection and positioning of audio sources and for decreasing the dependency on visual feedback when mixing for 3D audio.

3. SYSTEM DESIGN

The haptic 3D audio mixer is built around a virtual 3D space, where one is able to move and position virtual sound sources in relation to a centred listening position (see \cite{5} for more details). For rendering of the 3D audio, we use a binaural audio setup built in Max/MSP based on the binaural spatial panning tool, which is part of the Spat software suite. Here it is possible to control the distance, azimuth angle and elevation angle for an arbitrary amount of audio channels. For the exploratory evaluation presented later we used 6 different sound sources each representing differences in timbral, rhythmical and frequency content.

The virtual space is visualised as a 3D cube in which audio sources can be positioned by the user — see Figure 2. Each sound source is visualised as a coloured sphere with a text in front of it communicating the number and name of the sound source (for instance “3. clicks”). An additional grey sphere in front of each sound source shows the current 3D position of each audio source.
sphere represents the position of the controller (the *cursor*), when not interacting directly on any of the sound sources. Once the user moves this cursor close to a sound source, the coloured sphere of that sound source turns white indicating that the sound source can be selected and thereafter manipulated — see Figure 3. Finally, a white cube represents the location of the listener in the center of the 3D space.

### 3.1 Control Schemes

The focus of this paper is on assessing different approaches to using haptic feedback while manipulating sound sources in 3D. Therefore we have implemented several different control schemes, which all provide different forms of haptic feedback. Each input scheme is implemented using the Novint Falcon, which is commonly used as a game controller but also heavily used as a low cost solution in academic contexts. Additionally, we were interested in contrasting this to mid-air non-haptic interaction, which is why we have implemented an input scheme based on the LEAP motion control. Thus, the following six control schemes were developed:

- **LEAP Motion controlled** - users control the position of the audio source by hovering their hand above the LEAP motion. They select/deselect sound sources by pressing/releasing the space bar on a computer keyboard.

- **Falcon with no feedback** - users control the position of the audio source with the position of the Falcon controller. No force feedback is induced, so the user only feels the passive mechanical haptic feedback of the robotic arm itself. The user presses/releases a button on the device using the same hand as is used to manipulate the translation of the controller in order to select/deselect the desired audio source.

- **Falcon with pull towards centre** - same control scheme as above, but with a slight gravitational pull towards the centre of the virtual space. This was implemented in order to provide haptic feedback as to where in the 3D space one is located. The feedback emulates that of having an elastic band pulling towards the center (the further away from the center, the more force is applied).

- **Falcon with pull towards centre and snapping to sound sources** - same control scheme as above, but with a slight snapping to sound source locations as one moves close enough. This emulates the sort of haptic feedback one feels in the centre position of most stereo panning knobs. Note that the snapping only occurs when no sound source is selected as it is only meant to help the user select a sound source (when actually moving a sound source around one wants to move freely without sudden movements each time one passes another source.).

- **Falcon with pull towards centre, snapping to sound sources and/or key press to target sound sources** - same control scheme as above. Additionally we provide the user with an option to press a key on the computer keyboard (1-6), which induces a force that pushes the haptic device to the position of the sound source associated with the number that was pressed. The idea here is to enable the users to take their eyes off the screen and concentrate solely on positioning and listening.

- **Falcon with pull towards centre, snapping to source and/or key press to target sound sources and no visual feedback** - same control scheme as above. However, here we remove any form of visual feedback, forcing the subjects to rely solely on the haptic and auditory feedback.

Several other haptic feedback schemes could be imagined including restricting the user to certain planes on demand, or to auto-aligning to other sound source’s x/y/z positions (as seen in graphics software such as Photoshop and Illustrator). An interesting approach is also to first define a trajectory path and secondly trace along the path paying attention to timing (as implemented in [11]).

Finally, we must note that when adding more forces to the system, one also increases the risks of instability leading to oscillation because the haptic feedback acts as a force feedback loop and because different forces are added together. With careful tweaking of the parameters, the goal for us has been to make the haptic feedback subtle enough to avoid oscillation while still being strong enough to be useful for the interaction. Future studies will include more structured testing of the optimal parameters in order to experimentally understand this balance.

### 4. IMPLEMENTATION

In this section we briefly provide an overview of the overall system implementation, after which we describe how force feedback for center pull, snapping and sound source targeting was implemented.

The graphical interface and LEAP motion input was implemented in Objective-C/Cocoa using OpenHaptics for displaying the graphical interface and the LEAP SDK for handling input. As mentioned earlier, the audio engine was implemented in Max/MSP using the binaural spatial audio objects that are part of the Spat software package from Ircam (we use the standard KEMAR HRTF). Finally, the position and button press data received from the Falcon and the force data sent to the Falcon is controlled in the same Max/MSP patch using Max objects distributed as part of the Open Source Haptics library. All the haptic functionality is thus programmed in Max/MSP, which acts as both

---

https://www.novintfalcon.com
https://www.leapmotion.com
http://openhaptics.org/
Regarding the force feedback, three different overall functionalities have been implemented: center pull, snapping and sound source targeting. The center pull is simply implemented as a spring/damping system constantly exciting a subtle force on the controller in the opposite direction of its position relative to the center. Here the spring stiffness was set to 35 N/m with a damping of 0.1 N/(m/s).

The snapping has been implemented by emulating a subtle stick/slip motion similar to that of plucking a string (using the contact-pluck abstraction in the Open Haptics library). Here we have used a stiffness of 130 N/m, damping of 0.5 N/(m/s) and a minimum displacement difference for contact of 6 mm. A simple thresholding makes sure that the stick/slip emulation is only active when the position of the cursor is actually near enough to the position of the sound source to select it.

Finally, the automatic targeting of sound sources has been implemented using the same approach as the center pull. When the user hits a number key (for instance number “3” on the computer keyboard), we induce a strong force (stiffness of 500N/m, damping of 8 N/(m/s)) pulling the controller towards the position of audio channel number 3. In order to avoid the Falcon overshooting, which would result in a strong force pulling it back in the opposite direction, which would then lead to oscillation, we do not set the center of gravity of the force to the desired position immediately. Instead we slowly interpolate between the position of the cursor and the position of the audio source over a timespan of 400 ms. This creates a smooth motion ending at the position of the selected audio source.

5. EVALUATION

An exploratory pilot study was conducted in order to gather initial feedback from users as to the potential of the proposed control schemes. The goal was to understand whether the introduction of haptic feedback would (1) augment the feeling of the 3D space, (2) improve the speed and precision of the selection and positioning of the sound sources and (3) lower the dependency on visual feedback. Three test subjects (both with more than 8 years of mixing experience) were asked to try all six different control schemes in an exploratory fashion. For each they were given the task to imagine that they had just acquired the interface and were trying it out for the first time exploring its capabilities. As part of the exploration they were also asked to create a rough mix of the 6 audio channels. After having explored each control scheme they were asked to comment on their experience. Finally, after having tried all schemes they were asked to compare them in terms of the three goals mentioned above. Detailed notes were taken during the test sessions. The notes were analysed using a critical incidents approach where critical events relevant to the overall purpose of the evaluation were identified. These were then categorised into different themes and used to compare the different interfaces.

6. RESULTS & DISCUSSION

Since there were only three participants in the study, the following results are only indicative. They serve to provide an initial idea of the potential of different haptic interaction schemes. A future study will have to be conducted in order to properly assess this potential.

In regards to the LEAP motion control scheme all participants found the interaction reasonably natural and easy to use with minor comments including it being a little too sensitive and not as ergonomically pleasant as the Falcon. However, when beginning the first interaction scheme implementing the Falcon all participants were surprised at how much easier, faster and more accurate they were able to select and position the sources. After having explored the control scheme with no active haptic feedback, one participant stated:

"Wow, this is 10 times better.. I am much faster and more secure"

The difference experienced was not due to any imprecisions of the LEAP motion controller. In fact the LEAP motion works at a frame rate of 200 fps and with an accuracy of below 0.2 mm [19], and it was actually perceived as having less latency than the Novint Falcon. The experience of the LEAP motion being less accurate in use was most likely due to the nature of the freehand gestures having no haptic feedback, thus making the hand less steady.

The control scheme only implementing passive haptic feedback was preferred over both the center pull and the snap-
The control scheme implementing key presses for automatic translation to a sound source worked well but was not preferred over the simple Falcon implementation with no active haptic feedback. However, all three participants stated that they would probably be able to work fast and accurate with it after some practice. It was observed that participants used a combination of relying on the visual feedback moving the cursor towards the sound source they wanted to target while also hitting the key for that sound source.

Participants all strongly preferred the last control scheme implementing the Falcon with key press to target sound sources with no visual feedback. This was surprisingly apparent even when just observing the participants. They would gaze into the distance or close their eyes to really listen to the audio content in a completely different way than with visual feedback. All participants were really surprised about the difference it made having no visuals and how intuitive and natural the controls felt. One stated that

"I was really worried at first, whether I would be able to locate the sounds or use the controls properly with no visual reference, but I was much better than before - it felt like I was much more forced to use my senses. It just felt totally different."

Another participant who seemed very surprised about how intuitive the controls were stated:

"Really really cool... it was almost scary how intuitive it felt. I have never tried anything like it. It felt like I was almost inside the sound."

Other comments compared the feeling to how a physical mixing console can perform compared to using a Digital Audio Workstation (DAW), where the lack of visual feedback for the console forces one to listen more. The same phenomenon was also observed in an earlier study by the authors when testing tangible tabletop mixing interfaces [3].

7. CONCLUSION
We have presented a new interface for 3D spatial audio mixing, which implements different forms of haptic feedback using a Novint Falcon device. Different haptic interaction schemes were compared to a mid-air freehand gesture control scheme using the LEAP Motion in an informal evaluation. Haptic control schemes included (1) center pull, for augmentation of the perception of 3D space, (2) snapping to objects, for faster sound source selection and (3) automatic translation to selected targets, in order to reduce dependency on visual feedback. Although the evaluation was very informal, we strongly believe that there is a great potential in using a haptic controller like the Falcon for 3D audio mixing, because it achieves many of the same benefits as found on a more traditional mixing console. These include tangible feedback for a sense of precision and especially the reduced need to look at a screen, which leads to the users focusing more on the task at hand—namely listening. Interestingly, participants did not appreciate the active haptic feedback involved with center pull and snapping, stating that it felt counter-intuitive. Future studies will include a performance test of speed and accuracy testing improved schemes as the ones tested above using a methodology similar to the one used in [13].

8. ACKNOWLEDGEMENTS
We would like to thank Dannie Korsgaard for helping interface with the LEAP Motion controller and Francesco Grani for initial thoughts on the research. Finally, we would like to thank the participants who helped evaluate the interface.

9. REFERENCES
[6] Steven Gelineck, Dan Overholt, Morten Büchert, and Jesper Andersen. Towards an interface for music mixing based on smart tangibles and multitouch. In


