Technical Note

Note: Technical notes are aimed at promoting discussion. The views expressed are not necessarily those of the editors or the Australian Acoustical Society. Contributions are not formally peer-reviewed.

MANIPULATION OF INTERAURAL LEVEL COUPLED WITH A PERFORMER'S HEAD MOTION FOR HEADPHONE REPRODUCTION OF PIANO SOUND

Sungyoung Kim[1], Anandhi Ramesh[1] and William L. Martens[2]
[1]Sound & IT Development Division, Yamaha Corporation, Hamamatsu, Japan
[2]Faculty of Architecture, Design and Planning, University of Sydney, NSW, Australia

INTRODUCTION

Modern electronic instruments not only provide high quality musical sound reproduction, but also allow performers to select public or private monitoring (via loudspeakers or headphones). While the tonal quality of the musical sound can be matched relatively well between these two monitoring choices, headphone-reproduced sound has a potentially objectionable spatial aspect in that no change in the sound typically results when a performer’s head moves. While a straightforward solution would be to filter the sampled instrument sounds using dynamic update of Head Related Transfer Functions (HRTFs), coupled with a motion-tracking sensor that enables the headphone reproduction of spatially stabilized virtual sound source [1], such a system faces a significant challenge to performer acceptance since the HRTFs induce an apparent change in the instrument timbre. Since maintaining authentic character of the timbre is considered to be more important than enhancement of spatial attributes for most electronic instruments, a relatively simple cue, such as Interaural Level Difference (ILD), might provide a better solution for dynamic spatial sound processing in response to the performer’s head movement (as proposed in [2]). In order to find such effective yet simplified means for both spatially and timbrally stable headphone piano sound, an empirical study of head-related responses was executed in which measures were made of the level variation in broad frequency bands that occurred at a performer’s average head position as a function of two variables: one variable was the angle of rotation of the performer’s head in the horizontal plane (i.e. yaw angle) and the other was the pitch of the note played on the piano (i.e., the variation in the piano sound as each of the 88 keys were depressed).

ACOUSTICAL MEASUREMENT

In order to provide an empirical basis for a parametric manipulation of headphone-reproduced interaural level at a piano performer’s ears, a set of acoustical measurements were made using a sphere microphone (SCHOEPS KFM360) that was designed to capture signals in a manner that resembled human interaural differences in level and delay. The sphere microphone was placed at a location approximating the average location of a performer's head in front of the piano and aligned such that the center of the microphone system faced the piano's center, above which an additional cardioid microphone was positioned (this microphone was used to provide a reference for the signal captured by the sphere microphone). The sphere microphone was rotated ±40° variation, with higher angular resolution for yaw angles at which the performer’s nose was pointed nearer to the center of the piano. To capture consistent performance of the 88 piano notes for the 17 yaw angles tested, we utilized a MIDI-controlled acoustic piano (the YAMAHA DISKLAVIER) with notes played at a constant MIDI key velocity (100) and constant duration (3 seconds). The recorded data, comprising 1496 notes in total (17 yaw angles by 88 notes), was collected in a studio space that was not overly dry (i.e., not anechoic, as might be preferred for isolating the direct sound from the piano).

PARAMETERISATION

While the variation in level of the signal received at ear position that occurred with changes in the yaw angle of the receiver (the sphere microphone) was captured separately for all 88 notes of the piano, there was a practical consideration that required a parametric representation of these variations. Therefore, a simplified gain function was developed using a quadratic polynomial fit to data level obtained for a ±40° variation in yaw angle. The gain values were calculated relative to the RMS level of the same note captured by the reference microphone. This data was reduced to a representation of ILD that remained constant for subsets of adjacent notes to be reproduced by an electronic piano. It was decided that all frequencies below A4 (with fundamental frequency 440Hz) would have the same gain function applied, since the observed ILD variation was in fact quite similar at these frequencies. Next, five more subsets of notes were chosen, with note ranges that were initially determined through visual inspection of the ILD functions. Then the boundary MIDI-notes at which changes in the gain function would occur was iteratively adjusted to minimize the difference between the original level changes over all 88 notes and the parameterized level changes.

Table 1 shows for the resulting six ranges of note numbers and the two polynomial equations (for +yaw and -yaw, respectively) that parametrically represent the gain variation within each subset.

Table 1. Equations representing the gain variation for each ear

Note Number	Ipsilateral Equation (+yaw Θ)	Contralateral Equation (-yaw Θ)	Frequency Region (Hz)
1 to 48	$0.001\Theta^2+0.1213\Theta$	$-0.0008\Theta^2+0.1298\Theta$	< 440
49 to 56	$0.0025\Theta^2+0.2367\Theta$	$-0.0014\Theta^2+0.2144\Theta$	440 to 700
57 to 62	$0.0059\Theta^2+0.4222\Theta$	$-0.0021\Theta^2+0.2924\Theta$	700 to 1000
63 to 72	$-0.0044\Theta^2+0.1502\Theta$	$-0.0014\Theta^2+0.2348\Theta$	1000 to 1760
73 to 78	$0.005\Theta^2+0.5041\Theta$	$-0.0046\Theta^2+0.3513\Theta$	1760 to 2500
79 to 88	$0.0027\Theta^2+0.4284\Theta$	$-0.0026\Theta^2+0.2882\Theta$	> 2500

CONCLUSION

This paper derived a simplified means for controlling the level variation in a head-tracking, headphone-based reproduction of piano sound. The resulting gain functions fit well empirical data observed at a piano performer's ears for 1496 recorded notes. System users experience a satisfying externalized image of the piano sound that is stable both in spatial position and in timbre.

REFERENCES

[1] K. Inanaga, Y. Yamada and H. Koizumi, "Headphone System with Out-of-Head Localization Applying Dynamic HRTF (Head Related Transfer Function)", *Proceedings of the 98th Convention of the Audio Engineering Society*, Paris, France, 25-28 February 1995

[2] W.L. Martens and S. Kim, "Dominance of head-motion-coupled directional cues over other cues during active localization using a binaural hearing instrument", *Proceedings of the 10th Western Pacific Acoustic Conference (WESPAC10)*, Beijing, China, 23-25 September 2009

COMPETITION

Frame Your Physics

The Australian Institute of Physics is running a competition to promote the communication of physics to a general audience. Physics is an important science that helps us learn about the world around us, and if we are lucky, allows us to build useful technologies.

The aim of the competition is to present physics in a new light by making a video that presents a topic in physics to a general audience in an entertaining and informative way in under 3 minutes. The top videos will get an invite to a special presentation night at Questacon and will go in the draw for a share of $3000!

This competition is open to everyone. Entries must be submitted by 15th October 2012. So get filming, be creative and have some fun with it! And if you do submit an acoustics related video please let the AAS General Secretary know so we can use it for promotion.

Read all about it at:
www.act.aip.org.au/Frame_Your_Physics/Home_-_Frame_Your_Physics.html
[or just do a google search on Frame your Physics]