

Windows上で動作する動的聴覚ディスプレイ*

大谷真, 平原達也 (富山県立大学・工学部)

1 はじめに

両耳間レベル差/時間差や周波数特性といった情報を含んだ頭部伝達関数 (Head-Related Transfer Function: HRTF)[1] を利用した聴覚ディスプレイが開発されてきた。このようなシステムには、頭部運動に対応できず頭部運動による両耳信号の変化という音源定位の手がかり [2] を提示できない、という問題がある。

そこで、近年、角度/位置センサで受聴者の頭部運動の情報を検出して出力を変化させる動的聴覚ディスプレイが開発されている。最近では、計算機性能の向上により、DSPハードウェアを用いずにソフトウェア上でも実現されている [3, 4, 5, 6]。

このようなシステムは音響空間の認知メカニズムを解明するためのツールとして有用であり、我々も角度センサを利用した動的聴覚ディスプレイを開発した。本稿では、本システムの構成について述べ、既存のシステムとの比較を行う。

2 システム構成

2.1 概要

本システムは、PC(OS: Windows XP)、ヘッドフォン、オーディオデバイス、角度センサ、により構成される (Fig. 1)。ヘッドフォン上部 (頭頂部) に取付けられた角度センサ (NEC/TOKIN MDP-A3U9S) から取得される受聴者頭部の方向情報に基づき、PC上でHRTFが選択される。音源信号を選択されたHRTFに畳み込んで得られる両耳信号がオーディオデバイス (Roland UA-101) を経由してヘッドフォンから出力される。角度センサから取得される方向情報に基づいてHRTFが更新され、受聴者が頭部を回転させた場合でも常に同じ絶対位置に仮想的な音源が提示される。開発は Visual C#で行い、角度センサの出力を取り込むインタフェース部分は C++ で開発した。

2.2 ヘッドトラッキング

本システムで用いた角度センサは、セラミックジャイロ、加速度センサ、地磁気センサ、の組み合わせにより、設置対象の姿勢角を検出する。USBによりPCに接続され、角度情報は8ms毎に更新

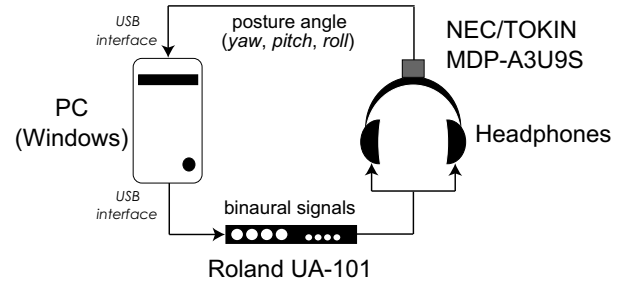


Fig. 1 System configuration

される。なお、角度センサからは yaw/pitch/roll の3つの角度が取得されるが、今回のシステムでは yaw 方向の頭部回転のみに対応している。

2.3 音響信号処理

HRTFと音源信号から両耳信号が合成される。この合成は畳み込みによって実現され、本システムでは overlap-add 法を適用した [7]。DA変換機に音響信号を出力するオーディオインタフェースとして、Microsoftのマルチメディア拡張API DirectSoundを用いた。DirectSoundでは、出力されるべき音響信号は、まずセカンダリバッファに書き込まれ、その後、プライマリバッファを経由してサウンドデバイスへと出力される。

Fig. 2に overlap-add 法 [7]を用いた音響信号処理の過程を示す。音源信号がフレーム毎に切り出され、受聴者の姿勢角と音源の絶対位置に基づいて更新されるHRTFに畳み込まれる。各フレーム毎に得られる信号は、重複加算された後にセカンダリバッファに書き込まれる。本システムではフレーム長は513ポイント (約12ms)とした。

2.4 音源信号とHRTF

様々な音源信号とHRTFデータベースに容易に対応できるように、両者のデータはプログラムの実行開始後にHDD上のファイルからメモリに読み込む仕様とした。音源信号としてWAVE形式 (標本化周波数44.1kHz, 量子化ビット数16bit, モノラル) に対応する。

HRTFデータベースにはHRTFの時間領域表現であるHRIR(Head-Related Impulse Response)が含まれている。HRIRのタップ数は512ポイントとした。データベースにはHRIRが一

*Dynamic auditory display operating on Windows. by OTANI, Makoto, HIRAHARA, Tatsuya (Toyama Pref. Univ.)

Table 1 Comparison among representative dynamic auditory displays

	OS	sensor	delay [ms]
DIVA[3]	Unix, Linux	Ascension Motion Star	110-160
RTVAS[4]	Linux	Intersense IS-900VWT	≤ 7
SLAB[5]	Windows	—	6.1 (excluding sensor delay)
Yairi <i>et al</i> 's system[6]	Linux	Polhemus FASTRAK	12.01
Our system	Windows	NEC/TOKIN MDP-A3U9S	50-60

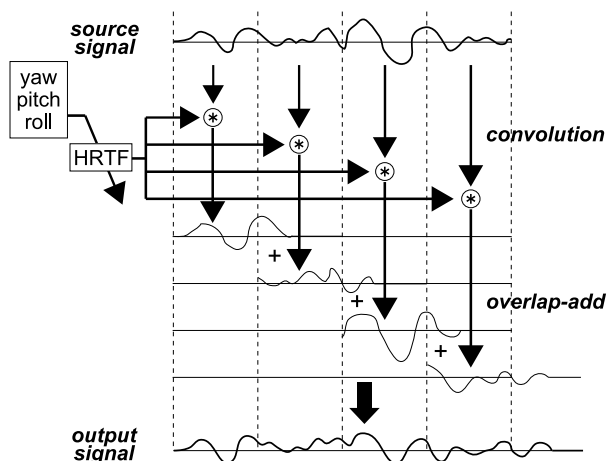


Fig. 2 Schematic diagram of signal processing

定の水平角度毎に左右1セットで保存されている。本システムでは、用意されていない音源位置に対するHRIRが必要になった際に、ソフトウェア内での補間を行わず、事前に補間を施すことによって空間分解能を高くしたデータベース(1度毎)を用いる事で対応している。この方式ではyaw方向だけでなくroll/pitch方向の頭部回転に対応する場合や水平面外に仮想音源を配置する際にはデータ量が大きくなってしまふ。したがって、そのような場合にはプログラム内で補間を行う仕様に変更する必要がある。現在のシステムでは水平方向のHRIRのみを用意すればよいため、HRTFデータベースのデータ容量は約1.4 Mbyteと小さくてすむ。

また、HRTFデータベースではなく、数値計算と頭部コンピュータモデルに基づいたHRTF高速計算システム[8]により計算されるHRTFを、TCP/IP接続により取得し利用する事も可能である。

3 既存のシステムとの比較

Table 1に、既存の代表的な動的聴覚ディスプレイと本システムの主な特徴を示す。既存のシステムの中には10 ms以下の遅延を実現しているも

のもある。一方、本システムにおける、センサの角度情報が取り込まれてから音響信号に反映されるまでの処理時間の計測結果は40-50 msであった。センサ更新間隔8 ms及びセンサ通信遅延を加えたシステム全体の遅延は約50-60 msである。これらの遅延に関しては、DirectSoundよりもレイテンシの少ないASIOの使用、フレーム長の短縮、などにより小さくしていく事が可能である。また、本システムは、それほどCPUパワーを必要とせず、IBM ThinkPad (Intel Pentium M 1.50-GHz)等の一般的なPCで問題なく動作する。

4 まとめ

本稿では、我々が開発したWindows上で動作する動的聴覚ディスプレイの概要について述べた。今後は、遅延の短縮などを行いつつ、本システムを利用して頭部運動を考慮した音響空間認知に関する検討を行っていく予定である。

参考文献

- [1] J. Blauert, *Spatial Hearing* (The MIT Press, London, 1997).
- [2] H. Wallach, *J. Acoust. Soc. Am.*, **10**, 270-274 (1939).
- [3] T. Lokki, Doctoral dissertation in Helsinki University of Technology, (2002).
- [4] J.W. Scarpaci *et al*, in *Proc. 11th Intl. Conf. on Auditory Display*, July (2005).
- [5] E.M. Wenzel *et al*, *Audio Engineering Society 108th Convention*, Pre-print, 5140 (2002).
- [6] 矢入 他, 日本バーチャルリアリティ学会論文誌, **11**(3), 437-446 (2006).
- [7] A.V. Oppenheim *et al*, *Discrete-time Signal Processing* (Prentice-Hall International, London, 1989).
- [8] M. Otani *et al*, *J. Acoust. Soc. Am.*, **119**(5), 2589-2598 (2006).