

反射音のリアルタイム畳込み機能を搭載した 3次元聴覚ディスプレイの開発*

☆三橋茂一(千葉工大院・工学研), △堀越健也, 飯田一博(千葉工大・工)

1 はじめに

PC 上で動作する聴覚ディスプレイが開発されている。その中には一次反射音を畳込む機能を有するものもある [1,2]。しかしながら、高次の反射音をリアルタイムに畳込むことは困難である。

本稿では、本研究室で開発された聴覚ディスプレイ SIRIUS [3]に、反射音をリアルタイムに畳込み、コンサートホールなどの音場をシミュレートする機能を搭載したので、システム構成および高速畳込みアルゴリズムについて報告する。

2 聴覚ディスプレイ SIRIUS の概要

2.1 機能

SIRIUS は Windows Vista/7 を搭載した PC 上で動作する。音源信号は WAVE 形式に対応する。HDD 上に頭部インパルス応答 (Head-Related Impulse Response: HRIR) データベース(応答長: 512 サンプル, 約 11 ms)を持つ。音源信号と HRIR を畳込み、リアルタイムな音像の制御が可能である。さらに、ヘッドトラック、3D 位置センサ (Kinect for Windows)を用いることで、受聴者の頭部運動、位置情報を取り込み、それらの変化を反映した音像制御が可能である。GUI 上でマウスを用いて音像位置の設定、畳込む HRIR の変更を行う。開発は Visual C++で行った。

2.2 処理の流れ

Fig. 1 に SIRIUS の処理フローを示す。プログラムの起動後、初期化処理を行いメインループに入る。メインループ中では、以下の処理を行う。

- ① ヘッドトラック、3D 位置センサから受聴者の頭部運動、位置情報を取り込む。
- ② GUI 上の音像位置の設定と頭部運動、位置情報から受聴者と音像位置の相対的

な角度、距離を算出する。

- ③ ②で算出した角度に対応する HRIR を呼び出す。すべての方向を含む HRIR データベースを用意することは難しいため、呼び出す HRIR は上昇角方向についてはパニング処理を施し、方位角方向については相当する ITD[4]を加えたものを生成し使用する。
- ④ 音源信号と生成した HRIR を畳込み、1 フレーム(512 サンプル)の両耳再生信号を得る。畳込みには、overlap-add 法を用いている。
- ⑤ 両耳再生信号をバッファへ送り、ヘッドホンから再生する。

以上の処理を繰り返し、リアルタイムな音像の制御を行う。

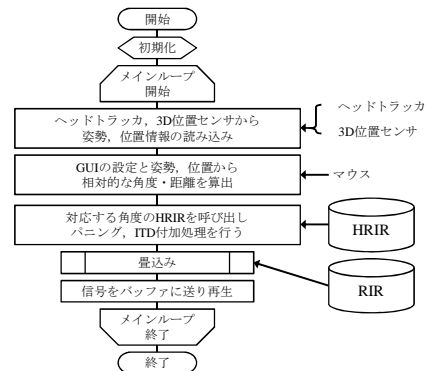


Fig. 1 SIRIUS の処理フロー

2.3 反射音のリアルタイム畳込み

HDD 上に、予め幾何音響シミュレーションによって計算された空間インパルス応答 (Room Impulse Response: RIR) を用意している。RIR と両耳の HRIR を畳込み、両耳空間インパルス応答 (Binaural Room Impulse Response: BRIR) を生成する。これを音源信号とリアルタイムに畳込み再生信号を得る。

生成した BRIR は応答長が長いため、単純な overlap-add 法ではリアルタイムに畳込むことは困難である。そこで、フレーム分割を用いた高速畳込みアルゴリズムを開発した。

*Development of three-dimensional auditory display with real-time convolution of reflections, by MITSUHASHI, Shigekazu, HORIKOSHI, Kenya and IIDA, Kazuhiro (Chiba Institute of Technology).

Fig. 2 に、その概念図を、Fig. 3 にフローチャートを示す。処理の過程は以下の通りである。

- ① Fig. 2 のように音源信号をフレーム長(512 サンプル)ごとに切り出す。切り出したブロックを S_1, S_2, \dots, S_n とする。
- ② ①と同様に BRIR(L,R)をフレーム長ごとに切り出す。切り出したブロックを $BRIR(L,R)_1, BRIR(L,R)_2, \dots, BRIR(L,R)_m$ とする。
- ③ j 番目のフレームの出力信号 $output(L,R)_j$ を得るために、以下の式に従って計算する。

$$output(L,R)_j = \sum_{i=1}^j S_i * BRIR(L,R)_{j-i+1} \quad (1)$$

- ④ 式(1)の畳込みには overlap-add 法を用いる。overlap-add 法のアルゴリズムに従い j 番目のフレームの出力信号 $output(L,R)_j$ を得るために、以下の処理を行う。

a) 式(1)に従い、各ブロック ($S_i, BRIR(L,R)_{j-i+1}, BRIR(R)_{j-i+1}$) をフレーム長の 2 倍の長さになるように 0 を詰める。

b) 各ブロックに対し FFT を行う

c) S_i と $BRIR(L)_{j-i+1}$ 及び、 S_i と $BRIR(R)_{j-i+1}$ それぞれの複素数積をとる。

d) c)で得た信号の L,R それぞれについて逆 FFT を行う。

e) d)で得た信号を L,R それぞれについて加算用の一時配列(1024 サンプル)に
加算する。

f) a)から e)を $i=j$ となるまで i を増加させながら処理を繰り返す。

g) f)で得た信号の前半(512 サンプル分)と前のフレーム($output(L,R)_{j-1}$)の計算で得た信号の後半(512 サンプル分)を加算する。

h) g)で得た信号の前半(512 サンプル分)を再生信号として再生用バッファに送る。

以上の処理を繰り返し計算することで、再生用の出力信号を得る。

2.4 リアルタイムに再生可能な BRIR 応答長

Windows 7 PC に搭載された CPU core i7 2.7GHz(4core, 8thread CPU) にて、OpenMP を用いた 6thread 並列処理において、BRIR の応答長 92160 サンプル(1.92 s)までリアルタイムに再生可能であることを確認した。ただし、これは受聴者の頭部運動および受聴位置の追従しない条件での値である。

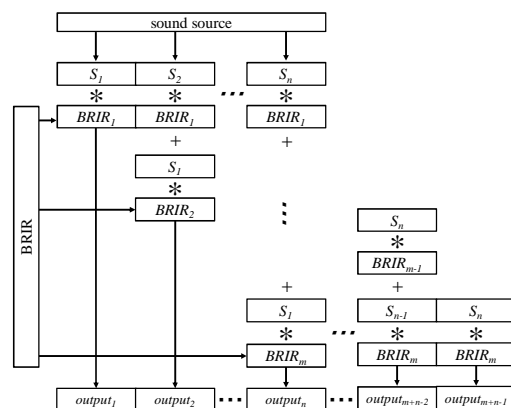


Fig. 2 リアルタイム畳込み処理概念図

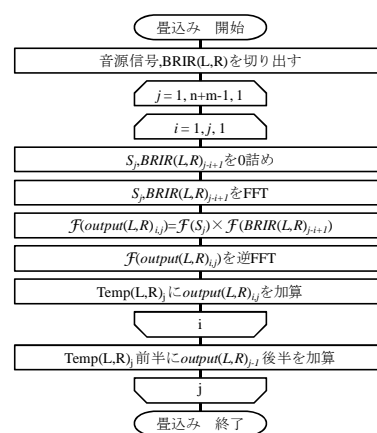


Fig. 3 リアルタイム畳込み処理フロー

3 おわりに

本稿では、我々が開発した反射音のリアルタイム畳込み機能を搭載した聴覚ディスプレイの概要について述べた。今後は、反射音のリアルタイム畳込みと頭部運動、受聴位置の変化への同時対応を考えている。

謝辞

本研究の一部は科研費(基盤研究(A)22241040)により実施した。

参考文献

- [1] J. D. Miller, E. M. Wenzel, *ICAD2002, 8th Intl. Conf. on Aud. Disp.*, Kyoto, Japan, 2002.
- [2] 大内他, 日本音響学会誌, 62(3), 224-232, 2006.
- [3] K. IIDA and Y. ISHII, *Proc. inter-noise 2011*, 428959, 8-pages(2011)
- [4] M. Morimoto *et al.*, *Acoustical Science and Technology*, 24(5), 267-275 (2003)