

身体運動と相互作用する三次元音場再生システムの構築 -システムの基本構成-

☆小笠原圭祐[†], 宮寺真之介[‡], 池田雄介[‡], 渡邊祐子[‡], 伊勢史郎[‡]
([†]京大/JST・CREST, [‡]東京電機大/JST・CREST)

1 はじめに

我々は、物理的に忠実な三次元音場の再現を目指し、境界音場制御の原理に基づく三次元音場再現システム(以下、BoSC システム)の開発を行っている[1]。遠隔地に位置するヒトが、同時にある音響空間を感じながらコミュニケーションをとることができる聴空間共有システムの開発を最終的な目的として、BoSC システム“音響樽”を提案してきた。音響樽には1)インターネットによる音場の共有, 2)精度の高い音像定位, 3)内部空間での頭部の移動および身体動作が可能であるという特徴があげられる[2]。

本稿では、以上で述べた音響樽の特徴を活かした応用技術の一つとしてスポーツに着目したシステムを提案する。音響樽に加え、受聴者の骨格位置を座標情報として認識できるセンサーを新たに用いることで、直感的に身体運動と三次元音場が相互作用するシステム構築を行った。

2 システム概要

2.1 システムの基本構成

受聴者が体験できるコンテンツとして、視覚障害者の卓球競技であるサウンドテーブルテニスに着目した。サウンドボールと呼ばれ

る鉛玉が入った球を使用して、テーブル上で転がる球の音を手掛かりに打ち合うスポーツである。音情報のみを手掛かりに行うこと、および必要最低限な身体動作が打球動作(腕を振る)という比較的簡単な動きであることから、本競技が相互システムのコンテンツにふさわしいと考えた。以下、本システムをバーチャル卓球システムと呼ぶ。

バーチャル卓球システムの構成をFig.1に示す。バーチャル卓球システムでは、原音場の制御点上で得られた音源信号に、再現音場のスピーカから制御点までの伝達関数を打ち消す逆フィルタをあらかじめ畳み込んだ状態でデータベースに保存しておくことを考える。この音源をスピーカの入力信号として用いることで、境界音場制御の原理に基づき、再現音場において原音場を物理的に再現した三次元音場を得ることが可能になる。

以上のようにデータベースに用意された、相手側から自分側へ球が近づく音源と自分側から相手側へ遠ざかる音源を、音響樽内における受聴者の身体動作(打球行為)に対応させてリアルタイムに読み込みと切り替えを順次行うことで、受聴者は音響樽内で再現される仮想音空間上で卓球のラリーを体験できる。現在のシステムでは、受聴者の打つ動作を検

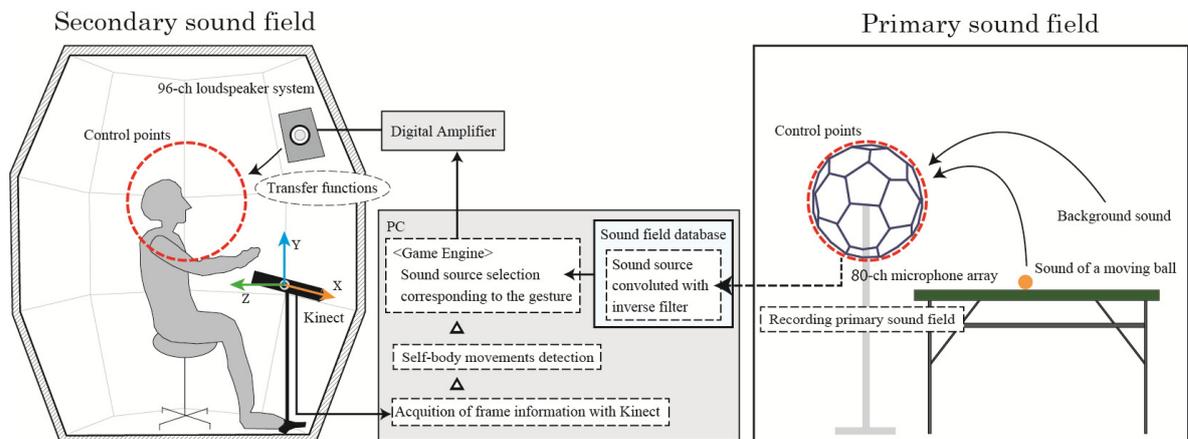


Fig.1 システムの構成

Development of the 3-D sound field reproduction system interacting with body action - Fundamental design of the system. by OGASAWARA, Keisuke[†], MIYADERA, Shinnosuke[‡], IKEDA, Yusuke[‡], WATANABE, Yuko[‡], ISE, Shiro[‡]. ([†]Kyoto univ./JST・CREST, [‡]Tokyo Denki univ./JST・CREST)

知して（方向・速さなどどのように打ったかは考えず），正面方向のみで球が往復する音場が再生される段階までを構築した。

2.2 卓球音の収録

システム構築にあたり，サウンドボールの転がる音を東京電機大学の防音室で収録した。収録には80個の無指向性マイクロホンを頂点に配置したC₈₀フラーレン構造のマイクロホンアレイを使用した。地面からマイクロホンアレイ中心までの高さは160 cmであった。収録した音のパターンは，卓球台の端3ヶ所から反対側3ヶ所に向けて打つ場合を双方向で行う18パターンに，球を打つスピードに緩急をつけた2種類を掛け合わせた計36パターンである。机の端50cmを始点とし，ラケットで打った後，反対側に落球するまでの音を収録した。台の中心を通る場合の，マイクロホン1番で収録した音源の時間波形をFig.2に示す。

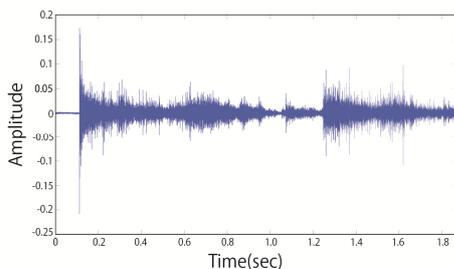


Fig. 2 収録音源の時間波形

2.3 身体動作入力インターフェース

Microsoft Kinectを身体動作の入力インターフェースとして使用した。音響樽壁面に沿って，カメラレンズが床から高さ70cmになるよう設置した。Fig.3にKinectによる人体の骨格位置検出の様子を示す。ソフトウェアの仕様により画像は左右反転して表示されている。画像認識処理により得られた人体の骨格

(Fig.3の赤丸)位置座標までの，カメラ位置を原点とする互いに垂直なX,Y,Z軸上での距離情報（mm単位）を本システムでは使用した。座標情報は1秒間に30フレーム所得され，時間的変化量の検出が可能である。適度な視

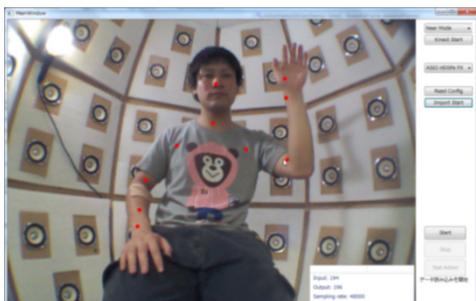


Fig. 3 Kinectによる骨格位置検出の様子

野角を確保するためKinectにズームカメラを搭載した。ズームカメラの画像のひずみによって実際の距離よりも大きい値が所得された。

3 身体動作の検出手法の検討

3.1 打球動作の分析

骨格座標情報から打球動作のジェスチャーを認識するプログラムの構成を検討するため，卓球の打球動作時の骨格座標情報の時間変化量を Kinect を使用して測定した。成人男女 8 名（うち男性 6 名）を被験者とした。構えた状態から打球動作を始め，球が当たると思われるポイントで手を止めるよう教示した。解析の結果，頭-右手間の奥行距離の差の値が，打球前と打球後のポーズを比べた際に最も顕著な変動（平均 225 mm）を見せたため，この値をジェスチャー認識のパラメータとして使用し，ある一定の値を超えると打球動作を行ったと認識するプログラム構築を行った。

3.2 打球動作認識に関する心理実験

打球動作と再現される音場のタイミングがもっとも適合する，頭-右手間の奥行距離の差のパラメータを心理実験により検討した。東京電機大学音響樽（スピーカ:FOSTEX FX120）を使用し，聴覚健全な 7 人の成人男女（うち男性 4 名）を被験者とした。ジェスチャー認識のパラメータを変えた条件下で，音響樽内において打球音のみを伴った打球動作を 20 秒間行ってもらい，各条件下で感じられる打球の実感を一対比較法により回答させた。頭-右手間の奥行距離の差のパラメータは 200 mm, 300 mm, 400 mm の 3 条件を用意した。結果，200 mm のパラメータでプログラム構成を行った場合，打球動作に再現音が上手くタイミングが合うように感じるという回答を全員から得た。

4 まとめ

身体運動と相互作用する三次元音場再生システムの構築を目指し，音響樽で再現された仮想音空間上で卓球を行うことができるバーチャル卓球システムを開発した。打球動作のジェスチャーを認識するためのプログラムおよびパラメータは実験を通じて検討した。

参考文献

- [1] 伊勢，日本音響学会誌，706-713，1997
- [2] 伊勢，音講論(秋)，3-5-13，2011.