

“MO-LIVE” & “ぼっちセッション” — 身体動作とスキンシップでサウンドを生成するシステム

松末千佳¹ 水野慎士¹

概要：本研究では身体動作とスキンシップに基づいてサウンドを生成システム“MO-LIVE”と“ぼっちセッション”の開発を行う。“MO-LIVE”は手で各関節に触れたり、足踏みすることにより音とオブジェクトを生成するシステムである。本システムでは複数人で体験することも可能であり、複数人でスキンシップやシンクロ動作を行うと、1人のときとは違った音が生成される。“ぼっちセッション”は身体動作によって楽器演奏を仮想的に体験出来るシステムである。ユーザがポーズをとると楽器が選択でき、仮想の楽器演奏が可能となる。また、後で記録したデータを合成し、仮想のセッションをすることも可能である。

“MO-LIVE” & “One-person Session” — Sound Generation Systems based on Body Motion and Contact

Chika Matsusue¹ Shinji Mizuno¹

1. はじめに

近年 Kinect や LeapMotion を始めとする安価で高性能なモーションキャプチャデバイスが発達、普及したことにより、マーカーレスで身体動作を取得することが容易になってきた。それにより、身体動作を伴ったデジタルコンテンツが増加している。

例えば谷本らは、Kinect を用いたブレイクダンスの支援システムを開発している [1]。このシステムはブレイクダンスの模範動画を再生後に自分の動きを録画し、その動きを模範のダンスと比較することでダンスの上達を支援するシステムである。録画をする際に kinect で骨格の動きを保存しており、その動きを模範と比較してアドバイスを表示している。ダンス関連のその他の研究例として、千田らは、ダンサーの動きに応じて背景画像をリアルタイムで生成するシステムを開発している [2]。萩原らは、空中での手書きジェスチャに基づく日本語入力システムを提案している [3]。システムではユーザーはアルファベットの形を空中に描き、kinect によって認識をさせ、かな漢字変換を経て、その結果をユーザーの肩に円形状に表示する。ユー

ザーは、表示された検索結果に触れることで文書を作成することができる。

このように多種多様な身体動作を伴ったデジタルコンテンツが存在する中でも我々は音を用いたデジタルコンテンツに注目した。

音楽を聴いて楽しむだけでなく、演奏して楽しむためには楽器を準備する必要がある、様々な種類の楽器の演奏を行うことは困難である。打楽器などであれば楽譜等が読めなくても音を出して楽しむことは可能であるが、様々な音色を楽しむためにはそれに応じた楽器の種類が必要となる。さらに複数人での合奏はより難易度の高いものとなる。そこで、仮想空間内の楽器と身体動作による演奏を実現することで、楽器演奏の敷居が下がり、多くの人にとってより身近なものになることが考えられる。同様のコンセプトを持つ研究として、加速度センサーや振動センサーを搭載したスティックを用いて仮想的にドラム演奏を体験するシステム (図 1) [4] や、Leap Motion, Kinect, AR マーカなどの各種センサを用いたエア楽器演奏システム [5] が提案されている。また、実際の指揮者に近い状況で仮想オーケストラの指揮を実現する研究が提案されている [6]。

そして、本研究では楽器や道具を用いずに身体動作のみで音楽の演奏を体験出来る 2 種類のシステムを開発した。

¹ 愛知工業大学情報科学部
Faculty of Information Science, Aichi Institute of Technology

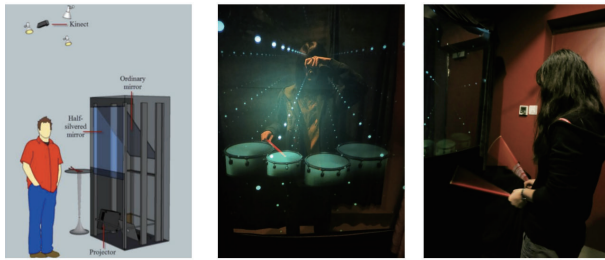


図 1 Augmented Reflection of Reality

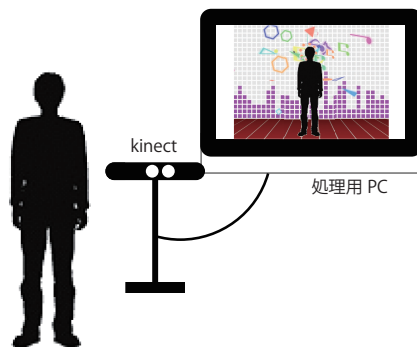


図 2 システム構成

これらシステムでは自分の身体を動かしたり触ったりすることで、様々な打楽器音を生成することができる。また、複数人でハイタッチなどのスキンシップをとることで様々なサウンドを生成する。さらに、バンド演奏で用いられる楽器を身体動作で仮想的に演奏することも可能であり、一人で仮想的にバンドセッションを体験することも可能である。

2. システムの概要

本研究で開発したシステムは、身体動作とスキンシップによって音とエフェクトを生成するものである。本システムは処理用 PC、ユーザの全身の三次元的な関節座標を取得するために用いる Microsoft Kinect、ユーザやエフェクト、楽器を表示するディスプレイによって構成される。図 2 にシステムの構成を示す。

開発システムは大きく 2 種類に分けることができる。1 つは、手で身体のような部位に触れたり足踏みをしたりすることで、様々な打楽器音を生成する“MO-LIVE”システムである。シンバルやトムなど 22 種類が用意されており、触れる部位や動作によって生成される音に変化する。そして、システムのディスプレイには様々な映像エフェクトが生成される。また、本システムは複数人の体験を可能としており、互いの関節に触れたり、シンクロした動作を行うと 1 人のときと違った音出力される。

もう 1 つは、身体動作によって楽器演奏を仮想的に体験出来る“ぼっちセッション”システムである。始めにユーザがポーズを取ることで、使用できる楽器を選択する。このとき、システムの画面には選択した楽器の画像がユーザ

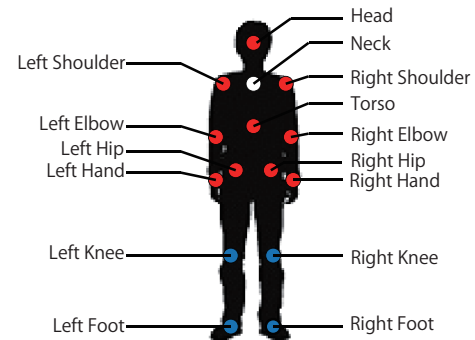


図 3 kinect によって取得する関節

を映したビデオ映像に合成される。次に、選択した楽器に応じた身体動作を行うことで、選択した楽器を仮想的に演奏することが可能となる。演奏結果は記録することができるため、各楽器を演奏したあとにシステムで同時再生することで仮想的なバンドセッション演奏も可能である。

3. 実現方法

本研究で開発したシステムは、Microsoft Kinect を用いてユーザ 1 人に対して図 3 に示す 15 カ所の関節点の三次元座標を取得して、ユーザの身体動作を認識する。

3.1 MO-LIVE

“MO-LIVE”では手による関節への接触と足踏みによって音とエフェクト画像を出力する。手による関節の接触で音を出力可能な箇所は図 3 の赤色で記した箇所である。また足踏みは青色の関節の座標を元に足踏みの判定を行っている。また、システムではユーザの身長に基づいて性別を子供、女性、男性の 3 種類に分類しており、ユーザの頭上にそれぞれの性別を表すマークを表示している。複数人モードでは性別の組み合わせによって生成する音やエフェクトが変化する。

3.1.1 身体動作の認識

手の接触判定は手と各関節との距離によって行っている。Kinect で取得した各関節点座標には、図 4 のようにサウンド開始距離とサウンド再開可能距離を設定している。システムは関節点と左右の手との距離を逐次計算しており、フラグが ON の状態で距離が十分に大きい状態からサウンド再開可能距離を下回ったとき、システムは打楽器音の生成を一度だけ行って、フラグを OFF にする。そのあとで引き続き手と関節の距離がサウンド再開可能距離よりも小さい場合でも、フラグが OFF の場合には新たな打楽器音生成は行わない。そして、距離がサウンド再開可能距離を上回ったときにフラグが ON となり、再びサウンド開始距離を下回ったとき打楽器音の生成を行うことができる状態になる。以上の処理を逐次行うことで、まるで身体が関節点が様々な打楽器であるような感覚で、手で身体を触りながらサウンドを生成することができる。

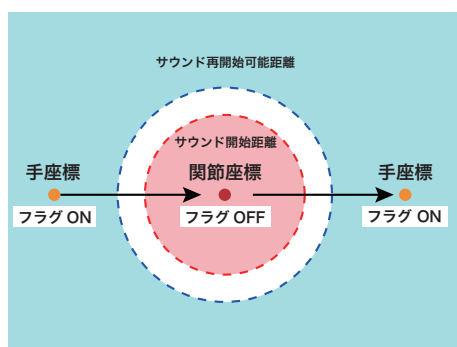


図4 接触判定における手座標としきい値の設定

足踏みの判定は、膝関節の座標に基づき足の上げと踏みの判定を組み合わせることで行う。足の上げ判定は、一方の足の膝関節が他方の足の膝関節に比べて閾値以上高い位置にあった場合とする。そして足上げの状態から、両膝座標間の距離が閾値以下になった場合に足踏みがされたと判定して、打楽器音の生成を一度だけ行う。この処理を逐次行うことで足踏みをしながら打楽器音を連続的に生成することができる。

ジャンプの判定は、両足関節点と床面の距離に基づいて行う。両足の高さの平均が閾値以上の場合に跳躍判定を行い、跳躍音を生成する。その後、両足の高さが閾値以下の場合着地判定を行い、着地音を生成する。この処理を逐次行うことでジャンプ音を生成することができる。

3.1.2 音とエフェクトの出力

3.1.1 節で述べた手法で手の関節点への接触、足踏み、ジャンプなどの身体動作が行われたと判定されると、あらかじめ用意された音源ファイルを再生することでサウンドの生成を行う。音源は、クラッシュシンバル、スネア、トムなど主にドラムで用いられる打楽器音やトランペットなどの管楽器音である。関節ごとに異なる音源が設定されており、体の異なる部位に触れたり足踏みしたりジャンプすることで、様々な打楽器音を生成することができる。

システムの画面には音の発生と同時に CG オブジェクトも生成される。オブジェクトの種類と色と大きさはランダムで決定しており、発生後には平行移動と回転移動をしながら徐々に消えていく。足踏みの場合には踏んだ場所から同様のエフェクトと円筒形のオブジェクトが出現する。図5にオブジェクト発生の様子を示す。

3.1.3 複数人モード

本システムは1人の身体動作による打楽器音の生成に加えて、複数人での身体動作によるサウンドを生成することができる。複数人の場合にそれぞれのユーザーが自身の関節に触れる場合には、1人は打楽器音、それ以外はトランペットのメロディの生成を行う。

一方、お互いの関節に触れ合う場合には自身の関節に触れた場合とは異なるピアノ和音を生成する。このとき、ユーザー同士が触れた高さを取得して、高さに応じて音高を



(a) 手によるオブジェクト発生



(b) 足踏みによるオブジェクト発生

図5 オブジェクト発生の様子

変化させることにより、メロディを作ることができる。また、全員で同時に両手を挙げたときにはシンクロモードとなり、ギターメロディが流れ始め、その曲に合わせて体でのドラム演奏をすることができる。

また、複数人モードではユーザーの組み合わせによってエフェクトが変化する。例えば男女の組み合わせで体験した場合には、エフェクトが花に変化する。ユーザーの中に子供が含まれている場合には背景が森のイラストに変化し、動物の観客が出現する。図6に複数人モードでのエフェクトの変化の様子を示す。

3.2 ぼっちセッション

3.2.1 概要

バンド演奏仮想体験システム“ぼっちセッション”では、バンド楽器のドラム、ギター、ベース、キーボードの各演奏を身体動作で仮想的に体験できる。

初めに、ユーザーはシステムの前でポーズを取ると、関節点の位置に基づいてポーズを認識して楽器が選択され、システムで構築した三次元仮想空間に表示される。そして、各楽器に応じた身体動作によって楽器を仮想的に演奏することができる。各楽器の演奏結果は記録することができ、すべての演奏記録を同時再生することで、仮想的なバンドセッション演奏も体験することができる。

3.2.2 楽器の選択

楽器の選択は、Kinectで取得した関節点の座標を用いてユーザーのポーズを認識する。楽器選択におけるポーズと楽



(a) 男女の場合



(b) 子供が含まれる場合

図 6 複数人モードにおけるエフェクトの変化

器の関係を以下に示す.

- ドラム：両手を頭部より上に挙げたポーズ
- ギター：左手のみを肩の高さに挙げたポーズ
- ベース：左手のみを頭部より上に挙げたポーズ
- キーボード：両手を胸の高さに挙げたポーズ

楽器が選択されると、仮想空間内にユーザと選択された楽器が表示される。楽器の位置はユーザの位置やポーズに応じて変化する。例えばギターとベースは左手でネック部を持っている状態のため、ユーザの左手の位置の応じて楽器の角度も変化する。

ユーザが両手を胴体に沿わせて直立姿勢をとると選択された楽器が解除され、その後新たに楽器を選択し直すことが可能である。

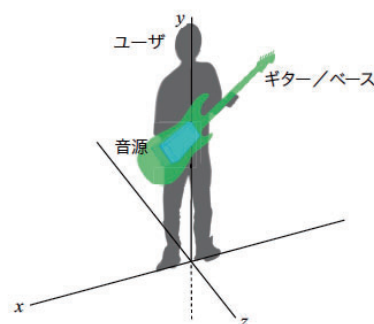
3.2.3 楽器の演奏

各楽器にそれぞれの音源ファイルが用意されている。用いる音源の種類は以下の通りである。

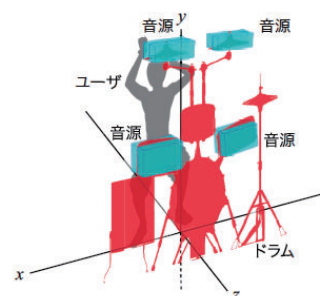
- ドラムセット：シンバル 3 種，ドラム 5 種
- ギター：C を基音とする和音
- ベース：低い単音 C
- キーボード：単音 C，C を基音とする和音

音源は各楽器に応じた場所に配置されている。各楽器の音源位置を図 7 に示す。

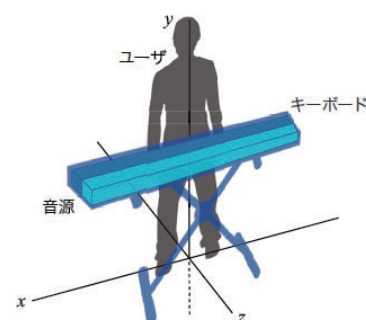
演奏方法は各楽器によって異なるが、基本的には Kinect で取得した右手や左手の位置が音源に達したときに音源ファイルを再生してサウンドを出力する。ドラムセットの場合、右手と左手がそれぞれ音源に達したときにサウンドを出力する。ギターとベースの場合、左手と腹部との距離によって音高を変化しながら、右手が音源に達したときに



(a) ギター/ベース



(b) ドラムセット



(c) キーボード

図 7 仮想バンド演奏における各楽器の仮想空間での音源配置

サウンドを出力する。キーボードの場合、右手と左手が音源に達したとき、その位置によって音高を変化しながらサウンドを出力する。

3.2.4 バンドセッション演奏の実現

仮想的にバンドセッション演奏を行うために、各楽器の演奏中に取得した全身の関節座標を逐次保存している。保存した各楽器演奏中の関節座標ストリームを用いて 3.2.3 節の手法を用いて楽器を同時に演奏することで、セッションの様子を再現することができる。また、セッション画面時にユーザの手にはマイクが表示され、セッションにボーカルとして参加することが可能である。

4. 実験

4.1 実装環境

本システムを実装するにあたって、使用した機材は以下の通りである。

- MacBook Pro
 - OSX 10.10.3
 - 2.8GHz, Intel Core i7
 - 8GB MM
- Microsoft Kinect

4.2 MO-LIVE

1人で行った場合、提案手法によって身体動作に応じて様々な打楽器音が生成されることを確認した。またシステムの画面にはサウンド生成と同時に様々なオブジェクトが発生することを確認した。ただし、素早い動きには対応できない場合があった。

2人で行った場合でも、それぞれの身体動作に応じて様々な打楽器音とオブジェクト発生ができることを確認した。そして、お互いに様々な位置でハイタッチなどのスキンシップを行うことで、音高を変化させながら和音を生成できることを確認した。そしてお互いが両手を挙げるとシンクロモードとなり、ギターメロディに合わせて身体動作で打楽器音を生成できることを確認した。ただし、ユーザが移動しながらすれ違った場合には、ユーザの関節点追跡に失敗して、正しくサウンドが生成されない場合もあった。図8に2人で様々なサウンドを生成している様子を示す。

また、本システムは2015年7月に開催された愛知工業大学オープンキャンパス及び、11月に開催されたせとまるとデジタルまつりにてデモ展示を行った。システム体験者に任意でアンケートへの記入を行ってもらい、合計で30名からの回答を得た。システムは楽しかったかどうか尋ねたところ、97%が楽しかったと回答した。ところが、自分の思ったような演奏ができたか尋ねたところ、64%はできた、30%はどちらともいえないと回答した。自分の思った通りの演奏ができたという回答が伸びなかった理由として、周りに体験者でない人が多く立っていた場合に、周りの人を誤検知したことや、体験者が小さい子供だった場合に上手く関節点を取得できなかったことなどが挙げられる。また、日光の入る展示スペースでデモを行った場合に、日光によってKinectによるユーザ情報の取得が阻害された場合もあった。

4.3 ぼっちセッション

システムを用いてユーザが演奏している様子を図9に示す。ユーザ認識後、姿勢による楽器選択、楽器の切り替え



(a) 個別のサウンド生成



(b) 2人のスキンシップによるサウンド生成



(c) 2人でのシンクロ動作によるサウンド生成

図8 2人のユーザによる身体動作に基づく打楽器音生成の様子

が正しく動作することを確認した。ギターやベースはユーザの移動や左手の動きに応じて正しく対応している。そして、ユーザの身体動作によって選択した楽器を仮想的に演奏できることを確認した。複数の音源を配置しているドラムやキーボードでの両手による複数のサウンドの同時出力、ギター、ベース、キーボード演奏時における身体動作に応じた音高変更も適切に行われていることを確認した。図9ではユーザが仮想的にギターを演奏している。

そして、各楽器の演奏中の関節座標を逐次保存して、演奏終了後に関節座標ストリームを用いてすべての楽器の演奏を同時に再現することで、バンドセッション演奏を仮想的に行うことができることを確認した。図10に仮想的なバンドセッション演奏の様子を示す。

5. おわりに

本論文では、楽器や道具を用いずに身体動作のみで音楽



図 9 仮想的にギターを演奏している様子



図 10 仮想バンドセッションの様子

の演奏を体験出来るシステムを開発した。まず MO-LIVE では 1 人、もしくは複数人で関節に触れたり、足踏みをしたり、シンクロ動作を行うことで音やエフェクトを表示させるというインタラクティブコンテンツを実現した。ぼっちセッションでは仮想楽器での演奏と、最大 4 種類の楽器による個人でのセッション体験が可能なシステムを実現した。

実験では音とエフェクトの表示を確認できたが、素早い動作に反応できていない場合があり、音やエフェクトが生成できない場合や、人が交代するときに正しく人を認知することができない場合があった。また、アンケート結果により自分の思った通りの演奏をできたと感じられるようにする必要があると考えられる。今後の課題としてエフェクトを増やしたり、複数人ならではの効果を増やすことで、よりコンテンツとして楽しいものになるように改善を行う。また、大人数での体験を可能にし、大人数でも人の入れ替わりをスムーズにできるように改良したい。

謝辞 本研究を進めるにあたり、システムの基盤技術を開発した小畑翔太郎氏に感謝する。本研究の一部は科研費(26330420)による。

参考文献

- [1] 谷本祐一：Kinect を利用することによるダンス支援システムの提案，早稲田大学大学院基幹理工学研究科 2012 年度修士論文，2013.
- [2] 千田真弓，小原悠太郎，菅野駿哉，本村健太，千葉則茂，ダンスパフォーマンスを拡張表現する高さを活かしたインスタレーション，NICOGRAPH2013，pp.121-122，2013.
- [3] 萩原 正人，益子 宗：KooSHO - 空中での手書きジェスチャーに基づく日本語入力環境，言語処理学会 第 19 回年

次大会発表論文集，2013.

- [4] Hongdo Fu, Wing Ho Andy LI, Augmented Reflection of Reality, SIGGRAPH 2012, Emerging technology, Article No. 3, 2012.
- [5] 岩谷亮明，澤田秀之，VR エンタテインメントに向けたエア楽器演奏システム，情報処理学会 インタラクション 2014 論文集，C1-4 pp. 587-592，2014.
- [6] 洲崎美里，砂川宗一郎，森谷友昭，高橋時市郎，仮想オーケストラを指揮するための演奏記号を考慮した楽器の自動生成，情報科学技術フォーラム講義論文集，11(3)，pp. 355-356，2012.