

# サウンドインタラクション開発支援の環境設計

坂田 和輝\* 田村 隼† 高橋 伸‡

**概要.** ユーザの状況や動作に対して数秒程度の遅れで応答を実行するような、リアルタイム性のあるサウンド出力は細かな指示や情報伝達を可能にする。しかし、リアルタイムなサウンド出力を行うシステムを開発するためには、骨格認識や音声出力のための幅広い知識と技術が必要になるため、特に開発未経験者が日常的な作業やインストラクション目的で使用する事が難しいという課題がある。そこで、我々はGUIを用いたサウンドインタラクション開発支援システムを設計している。本システムは、ノードエディタを用いたビジュアルプログラミングモジュールで姿勢やジェスチャとサウンド出力を結びつけ、リアルタイムなサウンドインタラクションを実行する。本システムを使用することで、リアルタイム性のあるサウンド出力の実装に要求される専門的な知識やプログラミングを減らし、実装を容易にできる。

## 1 背景

近年、スマートスピーカーが普及し音声ベースのインタフェースが注目されている。画面を必要とせず日常生活への親和性が高いが、ユーザの言葉に返答するこのインタフェースは、現在は対話がメインでありリアルタイム性のある応答はしない。このリアルタイム性とは、ユーザの状況や動作に数秒以下の遅れで応答が実行されることであり、例えばトレーニングにおいて体勢を正すために言葉を用いた音声による指示を行うのではなく、身体がどの程度曲がっているのか、どう動かせば良いかを機械音や音楽の音程・リズムをなめらかに変化させた音を鳴らし続け適切な姿勢を指示することである。このようなリアルタイムなサウンド出力は音声出力以上に細かな指示や情報伝達を可能にする。

しかし、現在このような人の動きに合わせたリアルタイム性のあるサウンド出力を開発・利用するには、骨格認識やジェスチャ認識のための機械学習、midiや合成音声出力の知識を用いたプログラミングが必要である。こういった現状では、日常的な作業やインストラクション目的で開発未経験者が使用することは困難である。よって、リアルタイム性のあるサウンド出力の実装に要求される専門的な知識やプログラミングを減らし、実装を容易にするための開発支援システムを作成する。

## 2 関連研究

サウンドの開発支援として、ビジュアルプログラミングによるインタラクティブなサウンドの開発[1]ツールが存在する。ノードをつなぎ合わせることで音源再生やエフェクトの適用が可能である。人の動きに合わせたサウンド出力のためには骨格認識から認識データの受信を行えるノードや、目的の姿勢を指定できるノードが必要であるため、本研究ではそれらを実装した新たなビジュアルプログラミングツールを作成する。

人の動きに合わせたサウンド出力に関する研究として、音による作業支援の研究[2]が存在し、トレーニング時に正しい姿勢を保つための音声指示を行うシステムを作成した。しかし限られた姿勢に対して音声のみの指示であったため、本研究ではユーザの目的に応じたサウンド出力の開発を可能とする。

## 3 サウンドインタラクション開発支援システム

### 3.1 システム概要

本節ではサウンドインタラクションプログラミングを容易にするためのシステム設計について述べる。

システム構成は図1のようなになる。ユーザの姿勢を骨格認識モジュールが認識し、骨格情報などがプログラミングモジュールへと送られる。ユーザの設定をもとに、プログラミングモジュールは認識情報からサウンド出力命令を生成する。サウンド出力モジュールが受け取った出力命令通りに音を鳴らし、ユーザにリアルタイムなサウンド出力を与える。

### 3.2 骨格認識モジュール

本モジュールでは既存の画像認識システムを用いて骨格推定を行い、そこから得た値を入力として利用する。カメラで撮影した映像から画像認識システ

Copyright is held by the author(s). This paper is non-refereed and non-archival. Hence it may later appear in any journals, conferences, symposia, etc.

\* 筑波大学 情報理工学位プログラム

† 筑波大学 情報科学類

‡ 筑波大学 システム情報系

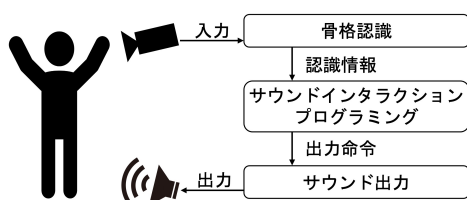


図 1. システム構成

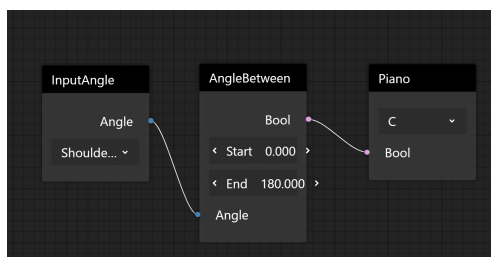


図 2. プログラミングモジュール

ムによって骨格情報や各関節座標を取得し、それらやそこから計算される関節間の相対座標、角度などの値を入力として利用する。さらに、ポーズや表情などのジェスチャを機械学習を用いて検出する。また、カメラ以外のセンサの値も同様に入力として利用可能にする。

### 3.3 サウンドインタラクションプログラミングモジュール

ビジュアルプログラミングツールにて用いられているノードエディタを参考に、プログラミングモジュールを作成する。モジュールの GUI は図 2 に示す。

本モジュールは骨格認識モジュールからの認識結果を受け取り、それに応じたサウンド出力命令をサウンド出力モジュールへ送信する。開発者はビジュアルプログラミングを用いてその認識データ入力とサウンド出力の紐付けを行う。例えば、肘の関節角度を入力として、その値の大きさによってピアノ音の音程を変化させることや、エディタ上で身体の姿勢を設定し、その姿勢を認識すると音楽の再生・停止を実行するなどの紐付けが可能である。

エディタ上へノードを配置、それらを線でつなぐことによってその紐付けを設定する。ノードには入力ノード、処理ノード、サウンド出力ノードの 3 種類があり、入力ノードは骨格認識モジュールからの骨格座標や関節角度、ジェスチャ認識結果を受け取り、繋がれた他のノードへその値を出力する。処理ノードは入力ノードから送られてくる認識情報をもとにサウンド出力を実行するための条件と実行タイミングを指定する。サウンド出力ノードは、処理ノードの結果を受け取り、ノードに設定されたサウンドを

サウンド出力モジュールへ送信する。

### 3.4 サウンド出力モジュール

サウンド出力命令を受けて音の出力を行う。プログラミングモジュールからの命令を受信し、ノードで設定された楽器の音や効果音、合成音声などを出力する。複数の出力命令を受けて同時に音を鳴らすこともできる。その他にも BGM やドラムなどをループ音源として常に出力し、条件を満たすことでの停止や再生をさせることも可能である。

### 3.5 システム試作

骨格認識モジュールには、mediapipe<sup>1</sup>とコンピュータに内蔵されている RGB カメラを用いる。mediapipe は Google 社が提供している ML ソリューションで 33 点の関節からなる骨格ポーズ認識や顔のメッシュ認識などが可能である。

プログラミングモジュールの実装にはデスクトップアプリケーションを開発するために Electron<sup>2</sup>を用いる。

サウンド出力モジュールは pygame.midi<sup>3</sup>モジュールを用いて midi 音源を出力する。出力する楽器音は General MIDI の 128 種類の中から用いる。

## 4 おわりに

本稿では、リアルタイムサウンドインタラクションの開発支援システムの設計を報告した。ノードエディタを用いたビジュアルプログラミングを提供し、身体の動きとサウンド出力の紐付けを容易に実現する。

今後は実装を進めた後、トレーニングや日常生活の中の実際の題材に応用して有用性を評価する。トレーニングやフィットネスに対して、例えばツールを使用し 3D アバタで指定した目標の姿勢から大きく異なると音を鳴らしたり、一定のリズムからずれた動きをすると BGM を大きくしたりなど、身体の動きや姿勢を客観的に認識するために使用することを検討している。また、スマートスピーカーのように使用することも想定し、音楽の再生・停止をジェスチャ操作できることに加えて、日常生活の中で変化する身体の姿勢や動きの速度に応じて流す曲や音量を変更するなどの使用環境にて、有用性の調査を実施する。

## 謝辞

本研究は科研費 22H03625 の助成を受けたものである。

<sup>1</sup> <https://google.github.io/mediapipe/>

<sup>2</sup> <https://www.electronjs.org/>

<sup>3</sup> <https://www.pygame.org/docs/ref/midi.html>

## 参考文献

- [1] <https://www.mi7.co.jp/products/cycling74/> (2022/11/23 確認) .
- [2] R. Wang, S. Takahashi, B. Shizuki, and I. Kawaguchi. Voice-Based Bodyweight Training Support System Using Smartphone. In A. Marcus and E. Rosenzweig eds., *Design, User Experience, and Usability. Case Studies in Public and Personal Interactive Systems*, pp. 370–379, Cham, 2020. Springer International Publishing.