

タイムラインの考え方を取り入れたサウンドインタラクションシステムの開発環境

齋藤 龍宏* 田村 隼* 真鍋 匠* 高橋 伸†

概要. 人の動きに対して音を出力するシステム（サウンドインタラクションシステム）は、ゲームやフィットネストレーニングなどの分野で利用されている。しかし、このようなシステムはプログラミングの専門的な知識を有する技術者が開発している場合が多い。そこで、我々はプログラミングの専門的な知識を有さない人が容易に開発できるようなサウンドインタラクションシステム開発環境の構築を目指している。本開発環境では、人体の骨格情報とサウンド出力情報を、動画編集におけるタイムラインの考え方をを用いて結びつけることで、サウンドインタラクションを実現できる。本論文では開発中のサウンドインタラクションシステム開発環境における人体の骨格情報とサウンド出力方法の結びつけ方法について概要を紹介する。

1 はじめに

ゲームやフィットネストレーニングなどの分野でサウンドユーザインタフェースが利用されている場面は増えつつある。人の動きに対して音を出力するシステムをサウンドインタラクションシステムと呼ぶ。サウンドインタラクションシステムに関する研究として、我々は音によるトレーニング支援の研究 [1] を行った。しかしプログラミングの専門的な知識を有さない人がサウンドインタラクションシステムを開発することは難しい。そこで GUI を用いたサウンドインタラクション開発支援環境の提案を行っている [2][3]。本稿では新たなアプローチである、タイムラインの考え方を取り入れたサウンドインタラクションシステムの開発環境について述べる。

2 サウンドユーザインタフェース開発環境

2.1 概要

本開発環境は骨格認識モジュール、サウンドインタラクション記述モジュール、サウンド出力モジュールの3つのモジュールで構成されている。図1にシステムの構成図を示す。骨格認識モジュールは、ユーザ動作の関節座標情報を取得するモジュールである。入力ソースとしては、カメラ入力と動画ファイル入力ができる。お手本となる動作の動画ファイルの認識や、ユーザの動作をカメラ等でリアルタ

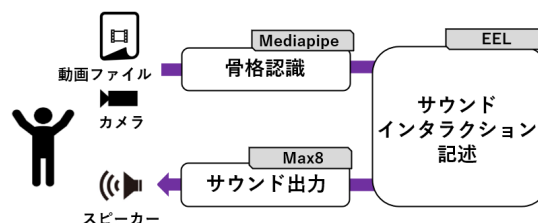


図 1. 提案する開発環境の概要

イムに認識することができる。サウンドインタラクション記述モジュールは、認識された関節座標情報を用いて、どのように音を出力するかを設定するモジュールである。注目する関節位置や関節間の距離などをスケルトンで指定すると、その値の時間変化がグラフとしてプロットされるので、それを利用して音を出力する条件を設定することができる。サウンド出力モジュールはサウンドインタラクション記述モジュールから送られてくる情報に従ってサウンド出力を行うモジュールである。

実装は、骨格認識モジュールに MediaPipe¹、サウンドインタラクション記述モジュールに EEL² (Python, JavaScript)、サウンド出力モジュールに Max8³ を用いた。

2.2 インタラクション記述例

サウンドインタラクションを記述する場合、特定の時間だけサウンドインタラクションを実行させたい場合や、複数の動作を連結させて順番に実行させたい場合、複数の動作を同時に実行させたい場合がある。そのような場合で容易にサウンドインタラクションを記述できるようにするために、本研究では

Copyright is held by the author(s). This paper is nonrefereed 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 システム情報系

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

²<https://github.com/python-eel/Eel>

³<https://cycling74.com/>

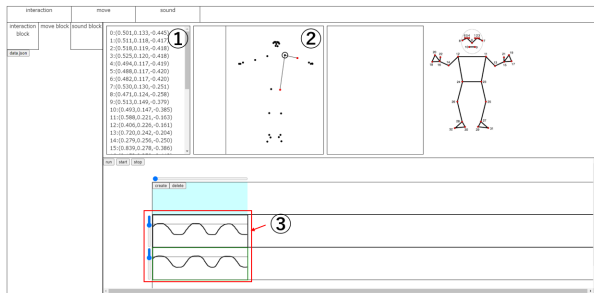


図 2. 開発環境のインターフェース

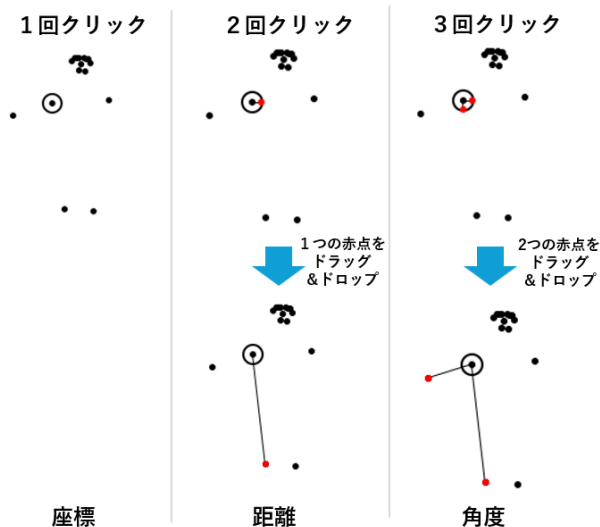


図 3. スケルトン上では座標, 距離, 角度を指定できる

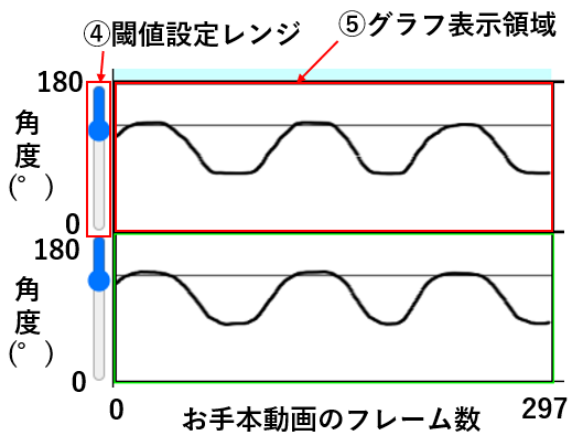


図 4. 図 2 の③を拡大し軸情報を加えたグラフ

実装する開発環境にタイムラインの考え方を取り入れた。ユーザ動作が時間軸上に表示され、それに対して音出力条件を設定することでサウンドインタラクションを記述することができる。

図 2 に本開発環境を用いて、お手本に近いサイドレイズができているかどうかを判断するサウンドインタラクションシステムの記述例を示す。ユーザの肩の角度が、お手本の肩の最高角度を超えると、ブザー音が鳴るというインタラクションである。お手

本の骨格から左右の肩の角度を抽出し、その情報をもとに音出力条件を設定し、その音出力条件をユーザの骨格に適用することで実現できる。

まず、お手本の関節情報からスケルトンを用いて、音出力条件を設定したい関節を抽出する。図 2 の①にお手本の関節座標の値が表示され、②にその座標をプロットしたスケルトンが表示される。スケルトンの中から目的の黒点をクリックすることでフレーム毎に座標、距離、角度のいずれかのデータを取得することができる。黒点を 1 回クリックで座標、2 回クリックで距離、3 回クリックで角度の情報を扱うことができる。図 3 にスケルトン上での座標、距離、角度の表現方法を示す。3 回クリックで出現する 2 つの赤点をドラッグアンドドロップで移動させ、目的の関節の位置を表す黒点と重ね合わせることで 3 点の角度を取得することができる。

次に、抽出した関節情報の時間変化をグラフとして表示させる。図 4 に図 2 の③を拡大し、軸情報を加えた図を示す。⑤をグラフ表示領域という。グラフ表示領域は、図 2 の②で指定したユーザ動作の関節座標、2 つの関節座標の距離、3 つの関節座標の角度などのグラフがプロットされる。今回の記述例では、横軸はフレーム数 (0~297)、縦軸は角度 (0° ~ 180°) である。

最後に、プロットされたグラフを参照しながら、音出力条件を設定する。図 4 の④を閾値設定レンジといい、つまみを上下に動かすことで閾値を設定することができる。今回は閾値以上であればサウンド出力モジュールに情報を送信する動作を行っている。閾値設定レンジの値を、グラフ表示領域にプロットされている y 軸の値の最大値と重ね合わせることで「最も大きいお手本の角度よりも、ユーザの角度が大きければ音を出力」を表現することができる。

また、グラフ表示領域はタイムライン上に設置されているため、縦に複数のグラフ表示領域を設置することができる。これにより、複数の関節情報に対して閾値を設定することができる。さらに、グラフ表示領域の横幅を狭めることで、音出力条件を適用する時間を設定することができる。

3 おわりに

本稿では、タイムラインの考え方を取り入れたインタラクション開発支援ツールについて紹介した。ユーザ動作のスケルトンをクリックすることで、グラフ表示領域に座標、距離、角度の情報がプロットされる。プロットされたグラフに対して閾値設定レンジを利用して閾値を設定することにより、インタラクションを記述することができる。

謝辞

本研究は科研費 23K24881 の助成を受けたものである。

参考文献

[1] Wang, R., Takahashi, S., Shizuki, B., Kawaguchi, I. (2020). Voice-Based Bodyweight Training Support System Using Smartphone. In Design, User Experience, and Usability. Case Studies in Public and Personal Interactive Systems. HCII 2020.

Lecture Notes in Computer Science, vol. 12202, pp. 370–379.

[2] 坂田 和輝, 田村 隼, 高橋 伸. サウンドインタラクション開発支援の環境設計. 第 30 回インタラクティブシステムとソフトウェアに関するワークショップ(WISS2022), pp. 1–3, 2022.

[3] 田村 隼, 齋藤 龍宏, 真鍋 匠, 高橋 伸. サウンドインタラクション記述用ビジュアルプログラミング環境の設計と実装. 第 31 回インタラクティブシステムとソフトウェアに関するワークショップ(WISS2023), pp. 1–3, 2023.