

# 身体動作による旋律線入力に基づく 音楽未経験者でも即興合奏可能な演奏インタフェースの試作

一ノ瀬 修吾\*, 白松 俊(名古屋工業大学)

Implementing an interface enabling music beginners to play improvised ensemble based on inputting pitch contour by body movement

Shugo Ichinose, Shun Shiramatsu (Nagoya Institute of Technology)

## 1. はじめに

旋律歌唱や旋律聴取における 3 つの処理側面として、リズム、旋律線、調性がある[1]。このうち、音楽未経験者の即興合奏を難しくするのは調性であり、リズムや旋律線については、比較的容易に直感的動作として表出できる。なお旋律線は旋律概形とも呼ばれ、大まかな音高の上下動を指す。本研究では、ユーザが旋律線とリズムを身体動作で入力すれば、システムの音高補正により調性の制約を満たす合奏ができる演奏インタフェースの実現を目指す。調性の制約については我々が過去の研究で提案した PFG Tonnetz (Prime Factor-based Generalized Tonnetz) [2]を用い、本稿では特に身体動作による演奏インタフェースに焦点を当てる。

身体動作による演奏インタフェースでは、(1)直感的な身体動作による旋律線の入力手法と、(2)減衰音と持続音の区別やオフセットタイミングを誤認識無く指定できる入力手法が課題となる、(1)の直感的な入力手法としては、Intel RealSense 3D Camera (以下、RealSense) を用い、手の動きを認識して音高の上下動 (すなわち旋律線) を入力する。RealSense は手指のジェスチャー認識が容易であり、(2)の減衰音と持続音の区別やオフセットタイミングを手指のジェスチャーで入力するのに適している。

## 2. 実装方法

Fig.1 にシステム構成図を表す。本システムの実装には、RealSense に対応し、音声処理にも適した Processing を用いる。まず、RealSense がユーザの手指の情報とジェスチャーを検出する。手の高さの変化が旋律線入力となり、ジェスチャーが持続音／減衰音の区別やオフセットタイミングの入力となる。あらかじめ与えられた背景楽曲に対し、時刻ごとに取りうる音高の調性制約を PFG Tonnetz [2] に基づき算出しておく。これを用い、入力された手の高さで近く制約を満たす音高を決定する。これにより、背景楽曲に対して不協和とならないサイン波を生成する。

持続音／減衰音の区別については、タップ動作 (手をカメラに向けながら前後に動かす) が減衰音の指定となり、サムアップ動作 (親指を上立てる) が持続音の指定となる。また、フィスト動作 (手指をすべて閉じる) で持続音を止めることができる。これらのジェスチャーは、どれも手の形状が異なり誤認識が起こりにくいため、自在に減衰音と持続音を区別し出力できる。左右の手で出力される演奏音は独立して

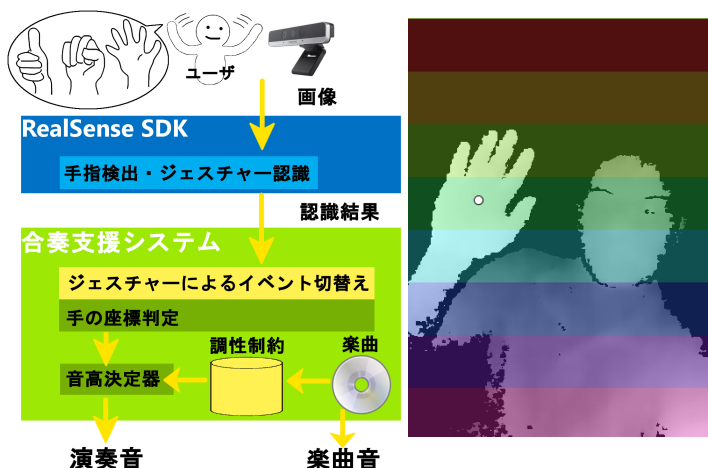


Fig. 1 System Architecture and User Interface

おり、2 和音まで出力可能である。

調性制約の計算に用いる PFG Tonnetz[2]は、音の周波数比を素数指数表現で表した上で指数を平面上にプロットし、その整数格子点で協和音程を表すモデルである。これにより算出された時刻 $t$ の調性制約を  $c(t)$ 、手の高さを  $x(t)$  とすると、Fig. 1 の音高決定器は調性制約を満たす周波数  $f(t)$  を以下のように決定する。

$$f(t) = \arg \min_{f \in c(t)} (f - f'(t))$$
$$f'(t) = f_{\text{tonic}} \cdot \exp(\alpha(x(t) - x_{\text{tonic}}))$$

なお、 $f_{\text{tonic}}$  は調の主音の周波数、 $x_{\text{tonic}}$  は主音に対応する手の高さ、 $\alpha$  は手の高さで音高の変化率の調性パラメータである。

## 3. おわりに

本稿では、音楽未経験者でも不協和音を出さず合奏できる演奏インタフェースを試作した。今後は、被験者実験によってユーザの意図どおりの合奏が可能か検証する必要がある。また、タップ動作による減衰音指定時に、ジェスチャーの速度による音量変更機能を追加する予定である。

## 文 献

- [1] 波多野 諄余夫: 音楽と認知, 認知科学選書, 第 12 巻, 東京大学出版会 (1987)
- [2] Shiramatsu, S., Ozono, T., and Shintani S.: A Computational Model of Tonality Cognition Based on Prime Factor Representation of Frequency Ratios and Its Application. *Proc. of SMC 2015* (2015) (to appear)