

歌声・伴奏音・打楽器音分離に基づく音楽演奏支援システム

土橋 彩香

池宮 由楽

糸山 克寿

吉井 和佳

京都大学 大学院情報学研究所 知能情報学専攻

1. はじめに

音楽の楽しみ方には、単純に音楽を聴くだけではなく、音楽に合わせて歌ったり、楽器を演奏したりすることも含まれる。この際、市販CDに録音されている音楽音響信号と同等の品質の伴奏音を利用できれば、音楽演奏が一層楽しめるようになると考えられる。ただし、歌唱を含まないカラオケトラックが収録されている音楽CDや、楽器演奏のためにある特定の楽器パートのみを除いたマイナスワンCDが販売されているものの、利用可能な楽曲はごく一部である。また、カラオケ産業では、新しいCDが販売されるたびに、専門家が耳で聴きながら手で楽譜（MIDI ファイル）に書き起こすことが行われている。この方式では莫大な労力が必要であるだけでなく、MIDIを用いて合成した伴奏音の品質はもとのCDには遠く及ばなかった。

これまで、MIDIで合成した伴奏音を用いた音楽演奏支援システムがいくつか提案されている。例えば、楽譜情報が既知であるという前提のもとで、自動伴奏システム [1] が提案されている。このシステムでは、ユーザの演奏をリアルタイムで解析することで楽譜中の演奏位置を推定し、テンポ変動や引き直しに追従することができる。一方、任意の楽曲に対して、楽譜情報を用いずに伴奏音を生成する音楽演奏支援システム [2] も提案されている。具体的には、音楽音響信号中のビート位置やベース音高の自動推定結果からドラムとベースからなる伴奏音をMIDIで合成し、与えられたコード付き歌詞を音響信号に同期させることで、音楽の進行に合わせて歌詞とコードをスクロール表示していた。

本研究では、市販CDに収録されている複雑な音楽音響信号を歌声・伴奏音・打楽器音に分離することで、演奏したいパートを除去した音響信号を自動生成し、音楽の進行に合わせて演奏に役立つ様々な音楽内容（ビート時刻やコード進行など）をスクロール提示するシステムを提案する。本システムでは、MIDIを用いずに元の音楽音響信号から直接伴奏音を生成するため、もとの豊かな音色が保持できる利点がある。また、音楽内容の自動推定には誤りが不可避であるという前提にたち、能動的音楽鑑賞サービス Songle [3] からそのつど音楽内容を取得することにした。本サービスでは、多くのユーザの協力で誤りが少しずつ訂正される仕組みを持つため、音楽内容を逐次アップデートすることができる。

2. ユーザーインターフェース

本システムはAndroidタブレット（HTC Nexus 9）上に実装されており、歌声・伴奏音・打楽器音の各パートの音量調整や、音楽の再生に合わせたコード・ビート情報のスクロール表示が可能である（図1）。

2.1 楽器パートの音量調節

図1右上部のスライダにより歌声・伴奏・打楽器音の音量調節を行う。ユーザが歌唱・演奏するパートの音量

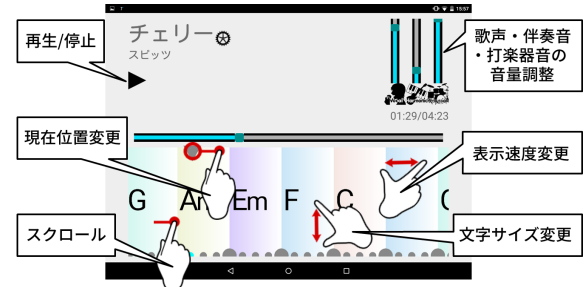


図1: ユーザーインターフェース

を小さく設定することで、カラオケやマイナスワン音源を生成できる。例えば、典型的なスリーピースのロックバンドの楽曲であれば、歌を歌いたい場合は歌声の音量を小さくしたり、歌を歌いながらギターも同時に演奏したい場合は、打楽器音のみを再生することができる。手軽に持ち運びが可能のため、ベースのメンバーの都合がつかずに三名全員が練習に参加できなくても、代わりに伴奏音のみを再生することも可能である。

2.2 演奏支援情報の提示

図1下部にはコード・ビート情報を表示されているので、ギター奏者はコードを見ながら演奏することができる。また、この範囲を左右にスワイプすることで再生位置の調節ができる（再生位置は中央部のシークバーでも調節可能）。横方向のピンチイン・ピンチアウトによりコード進行の表示範囲を変更できる。表示範囲を狭めると各コード名が確認しやすくなり、広げると後続のコードを確認しやすくなる。なお、この操作によって再生速度は変化しない。一方、縦方向のピンチイン・ピンチアウトにより文字サイズを変更できる。

コード、ビート表示部の左上部の三角形は現在の再生位置を示している。これを長押ししてそのまま左右に移動することで現在の再生位置を調整できる。これを右寄りにすると現在のコードが確認しやすくなり、左寄りにすると後続のコードを確認しやすくなる。

3. システム実装

本システムでは、音楽音響信号から歌声・伴奏音・打楽器音を得る。コード、ビート情報はクラウドソーシング型音楽鑑賞 Web サービス Songle [3] から取得する。

3.1 音源分離

歌声・伴奏音・打楽器音を分離するには、まず音源を歌声とそれ以外に分離し、後者をさらに伴奏音と打楽器音に分離することで実現する。図2に分離の全体図を示す。

楽曲中の歌声と伴奏の分離には、国際的な音楽認識コンテスト MIREX2014 の歌声分離トラックで最高性能を達成した手法 [4] を用いる（図3）。まず、入力のスเปクトログラムに対してRPCAを適用し、歌声成分を分離するためのRPCAマスクを得る。次に、分離した歌声からSHSに基づくビタビ探索を行い歌声F0軌跡を推定し、F0と倍音周波数のみを通過させる調波マスクを作成する。最後に、RPCAマスクと調波マスクを統合したマスクを入力に適用して分離歌声を得る。

A Music Performance Assistance System based on Source Separation of Music Audio Signals: A. Dobashi, Y. Ikemiya, K. Itoyama, K. Yoshii (Kyoto Univ.)

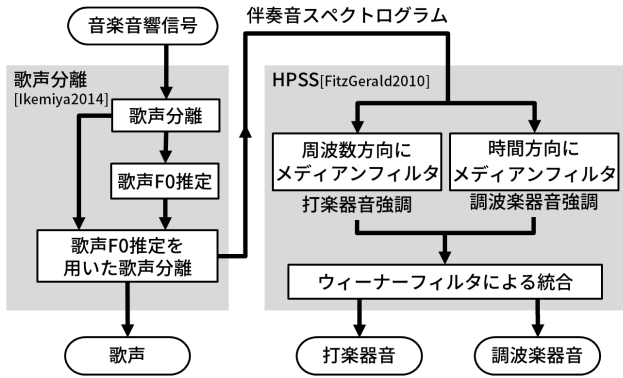


図 2: 歌声・伴奏音・打楽器音分離

伴奏音中の調波楽器音・打楽器音分離には、低い計算負荷と優れた性能からメディアンフィルタに基づく手法 [5] を用いる (図 4)。スペクトログラム中で調波楽器音は時間軸方向に、打楽器音は周波数軸方向に滑らかであることから、メディアンフィルタを用いてそれぞれ急峻な部分を除去することができる。最終的に、ウィナーフィルタを適用することで、調波楽器音あるいは打楽器音を強調したスペクトログラムが得られる。

3.2 音楽解析

演奏のための情報取得には能動的音楽鑑賞サービス Songle [3] を用いる。Songle では、Web 上に存在する任意の楽曲に対して、繰り返し構造・ビート時刻・コード・ボーカル音高といった各種音楽内容を自動解析した結果をブラウザ上に表示することができる。ユーザは認識誤りを適宜訂正ことができ、アノテーション精度が少しずつ向上していく。この機構を利用すると、タブレット端末上のデータの逐次更新が可能になる。

4. 実験

ユーザ評価によりこのシステムの有効性を確認した。被験者が以下の手順でシステムを利用し、ギターを演奏する様子を観察、実施後に聞き取りを行うことで、情報の表示の有無、情報の訂正の有無、各音源音量の変更の有無が演奏しやすさに与える影響を調査した。被験者はギター演奏歴 8 年の 23 歳男性である。

4.1 実験条件

1. 楽曲を聴取する
2. 演奏支援なしで演奏する
3. 演奏支援ありで演奏する

これを条件ごとに行う。条件は次の 3 つである。

- a コードおよびビート情報の提示あり・なし
使用曲:チェリー/スピッツ
- b 自動認識結果をそのまま提示・自動認識結果を手で修正したものを提示
使用曲:ポリリズム/Perfume
- c オリジナルの音源を使用・伴奏を抑圧した音源を使用
使用曲:あたしの向こう/aiko

4.2 実験結果

条件 a の結果から、コード情報の提示は演奏を容易にすることが確認された。条件 b の結果から、自動コード認識結果の提示はある程度の支援効果はあるものの、認識誤りは演奏を難しくする場合があります。正しいコード情報の提示は演奏を容易にすることが確認された。伴奏音

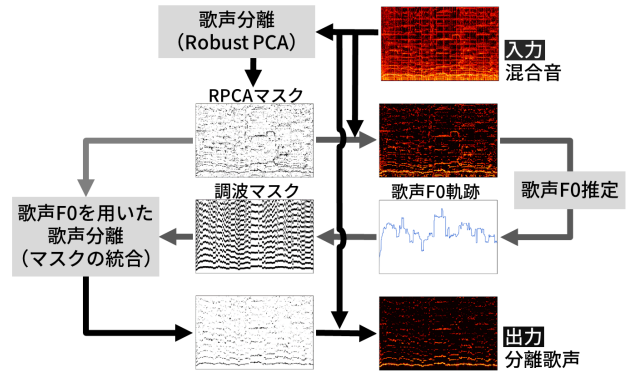


図 3: 歌声分離

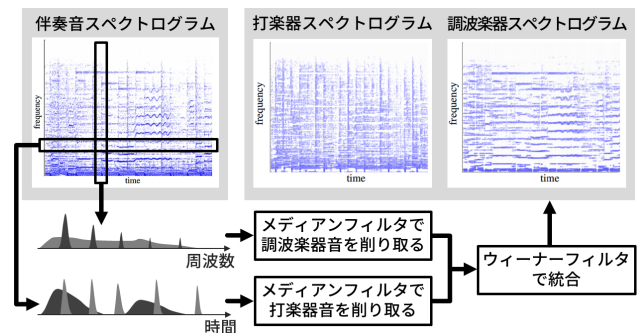


図 4: 調波音・打楽器音分離

の音量を下げることで、「音楽中の演奏と自分の演奏がぶつからずじっくりくる」、「ストロークを変えるようなアドリブがきく」という感想が得られた。システムの機能について、「わからないコードがあったのでコードダイアグラムを表示してほしい」、「盛り上がる場所がわかるように曲の流れがわかるといい」、「カポタストが使えるようにキーの変更がしたい」といった意見が得られた。

5. まとめ

本稿では歌声・伴奏音・打楽器音分離に基づく音楽演奏支援システムについて述べた。被験者実験において、演奏支援機能は実際に演奏を容易にし、演奏感を高めるといった感想が得られ、提案システムが演奏支援に有効であることが示唆された。今後は、楽曲構造など表示する情報の追加や、認識機能を付与し、ユーザの演奏に追従する再生機能の追加を行う予定である。

謝辞 本研究の一部は、科研費 24220006, 26700020, 24700168 および OngaCREST プロジェクトの支援を受けた。

参考文献

- [1] 中村栄太, et al. 任意箇所への弾き直し・弾き飛ばしを含む演奏に追従可能な楽譜追跡と自動伴奏. 情報処理学会論文誌, 54(4):1338-1349, 2013.
- [2] M.Mauch, et al. Song Prompter: An accompaniment system based on the automatic alignment of lyrics and chords to audio. In *ISMIR 2010*, 2010.
- [3] 後藤真孝, et al. Songle: ユーザが誤り訂正により貢献可能な能動的音楽鑑賞サービス. *Interaction*, 2012.
- [4] 池宮由楽, et al. 混合音中の歌声 F0 軌跡に対する歌唱表現転写システム. 音楽情報科学研究会 (SIGMUS 104).
- [5] D.Fitzgerald. Harmonic/percussive separation using median filtering. In *DAFx-10*, 2010.