

合奏音楽音響信号からの1パート除去の検討*

松本恭輔, 小野順貴, 嵯峨山茂樹 (東大院・情報理工)

1 はじめに

本稿では, 合奏音楽音響信号から任意音を除去する問題を扱う. 指定したパートが除去できれば, 音楽音響信号のパート編集・自動伴奏用音響信号の作成・カラオケ作成などの需要に答えられるだけでなく, 音楽音響信号の加工技術の研究としても興味深い.

今回は, 本問題を一つの研究分野と捉えて問題を分類しその多様性を議論する. そして, この分野の一研究例として, 合奏音楽音響信号からの1パート除去を検討する.

2 音楽音響信号における除去問題の分類

まずは, 本分野の問題を概観する. 除去を行うには, 除去音が指定される必要があり, その指定方法によって問題の特徴が異なる. 除去音の指定方法には以下のような例が考えられる.

時間周波数 bin 時間周波数 bin ごとの適切な処理による除去を支援する

音響ストリーム 聴覚に関する知見等を利用し, 人間が知覚可能な一音単位の除去を行う

楽器 (パート) 指定された楽器・パートの音色等の特徴を持つ音を特定し除去する

音源位置 ステレオ音源, それ以上の多チャンネル音源に対して, 空間情報を利用した除去を行う

楽譜 音楽音響信号から, 入力された楽譜に対応する音を特定し, 除去を行う

参照音源 加工対象音源中の除去音に同期したハミング等に対する単ピッチ推定結果を楽譜の代わりに用い除去を行う

上記の指定方法を複数併用するなど多様な問題設定も考えられる.

上記情報のうち, 何を入力 (あるいはアルゴリズム中での推定対象) とするかに注目した問題分類の例を Table 1 に示す. 様々な従来研究が除去問題の一解法・部分的な解法として捉えられ, この表に含められる. ここからも, 本分野は多様な問題設定と課題を持ち, 広がりのある研究分野と考えられよう.

次節以降では, 本分野の研究の一例として, 全パートの楽譜情報既知である場合に1パートを選択除去する問題について検討を行う.

3 楽譜情報を用いた合奏音楽音響信号からの1パート除去問題

3.1 目的

合奏音楽音響信号と, 対応する楽譜 (MIDI 形式で与える) を入力とし, (1) 任意の除去パート音成分を0にし, ただし (2) 他パート音の発音部成分は残す, という問題を扱う. これは, 伴奏・カラオケ音源用に, (1) 除去パートが聞こえず, (2) 他パートが原音に近い音源を作成することを想定した問題である.

Table 1 音楽音響信号における除去問題の分類例:表中参照は各除去 (部分) 問題の手法として捉えられる従来研究例

入力 \ Ch 数	1	2	3~
音響ストリーム	[6]		
楽器	[1, 2, 10]		
音源位置		[4]	[7]
除去パート楽譜			
全楽譜	本稿後半		
参照音源	([5])		
全楽譜+音源位置			[8]
全楽譜+楽器	[9]		

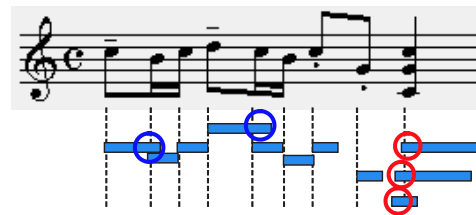


Fig. 1 テンポの変化以外の楽譜と実演奏のギャップ:発音時刻 (赤丸), 音長 (青丸) のずれを考慮せず, 求めた拍位置 (点線) に従った除去を行うと, 消し残しが生じる.

3.2 問題の特徴

本問題では, まず, 楽譜で与えた除去パートの各音符と, 音響信号中の音を対応づけて, 除去音を特定することが重要な課題である. 従来行われて来た多くの楽譜と音響信号の対応づけの研究は, 自動伴奏等への応用を想定し, 拍位置を推定していたが, 本問題を解く上ではこれは不十分である. 実演奏において, 各音の発音時刻は拍位置からずれることがある. また奏法や楽器の共鳴特性によって, 音長も, 楽譜上の音価とは一致しないことがある. それらのずれに起因する消し残し (Fig.1) をなくすために, 本問題では, 全体のテンポ変化の他に, 一音毎に発音時刻と, 消音時刻または他パートの同一周波数成分の発音時刻を調べ, 除去音区間を特定する必要が有る.

3.3 検討のための除去手法

上述の問題に対し, 特に音長と発音時刻のずれが顕著である, ピラフォン等, 響きが長時間残る楽器パートの除去を想定し, 以下の手法により問題の検討を行った.

1. テンポ変化への対処 実演奏におけるテンポの変化は, 従来のスコアアラインメントの手法で対応可能である. 今回は, [3] を参考に, 1次元 DP マッチングによる手法を用いた.

2. 一音毎のデヴィエーションへの対処 一音毎に除去区間を特定する目的で, 注目する音のノートナンバーに対応する矩形の楕円フィルタの出力の対数パワー

$$P := \log \left(\sum_i S_i F_i \right) \quad (1)$$

*"One part signal reduction from a performed ensemble music" by MATSUMOTO, Kyosuke, ONO, Nobutaka, and SAGAYAMA, Shigeki (The University of Tokyo).

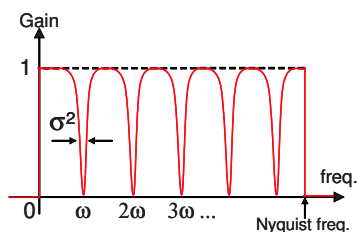


Fig. 2 除去音の音名に対応する周波数 ω とその整数倍を平均に持つガウシアンから成る楕円ノッチフィルタ

を考える．ここで S は短時間パワースペクトル， F は上述のフィルタ， i は周波数である．ピラフォン等打楽器系の楽器の振幅は，鋭く立上り滑らかに減衰すること，アラインメント結果から得られる発音時刻は真値と近いことを仮定すると，発音時刻は，もっとも近い P の極小点，消音時刻は，発音時刻後で P がある閾値を下回る最小の時刻として推定できる．除去パートの推定発音・消音時刻をそれぞれ t_{on}, t_{off} ，他パートの音のうち，フィルタ F を通過する周波数成分を持つ t_{on} の最近発音時刻の推定値を t'_{on} として，除去する区間は， $[t_{on}, \min\{t_{off}, t'_{on}\}]$ で求められる．ただし，この方法はバイオリン等，滑らかな振幅の減衰をしない楽器音への適用には限界があると考えられる．

3. パート除去 指定パートを除去し，他パートを原音に近いまま残すためには，本来は除去パート信号を原信号から減算するべきである．しかし信号中の指定パート信号の推定は容易でない．

聴覚に関する知見によると，調波信号中のある少数の倍音成分（1つおき等の特別な場合をのぞいて）を0にしても，知覚されるピッチは変化しない．そのため，フレーム毎のノッチフィルタにより，減算の代わりに乗算してパート除去を行ったとしても，同一時刻に複数の音が存在する合奏の音楽音響信号の中では，聴覚上，残りパートの音色への影響は大きくないと期待できる．よって今回は，各音毎に求めた除去区間を元に，短時間フーリエ変換の領域で，Fig.2に示すノッチフィルタを時間方向に連ねて構成した時間周波数マスクをかけることで除去を行った．

4 実験

前節の手法を用いた除去処理前後のパワースペクトル例を示し考察する．

4.1 実験条件

RWC 研究用音楽データベース（ジャズ）から RWC-MDB-J-2001 No.11（ピアノとピラフォンの二重奏）を用い，ピラフォンパートの除去を行った．サンプリング周波数 44.1kHz 16bit の音源に対して，フレーム長 4096 点（約 93ms），フレームシフト 1024（約 23ms）で分析し，楕円ノッチフィルタ（Fig.2）中のガウシアン分散を 15Hz とし，1～5 倍音までを除去を行った．

4.2 結果と今後の課題

処理前後のパワースペクトルと構成されたマスクを Fig.3 に示す．上下段赤い輪郭は，マスクと音源のパワーのおおよその位置関係を示すものである．ピラフォンパートの調波成分は，響きの残る音に対し，各音毎に適切な区間で除去されている．しかし，アタックの成分は周波数方向に広がり，除去がされていない．また，処理前のパワースペクトルより，ピラ

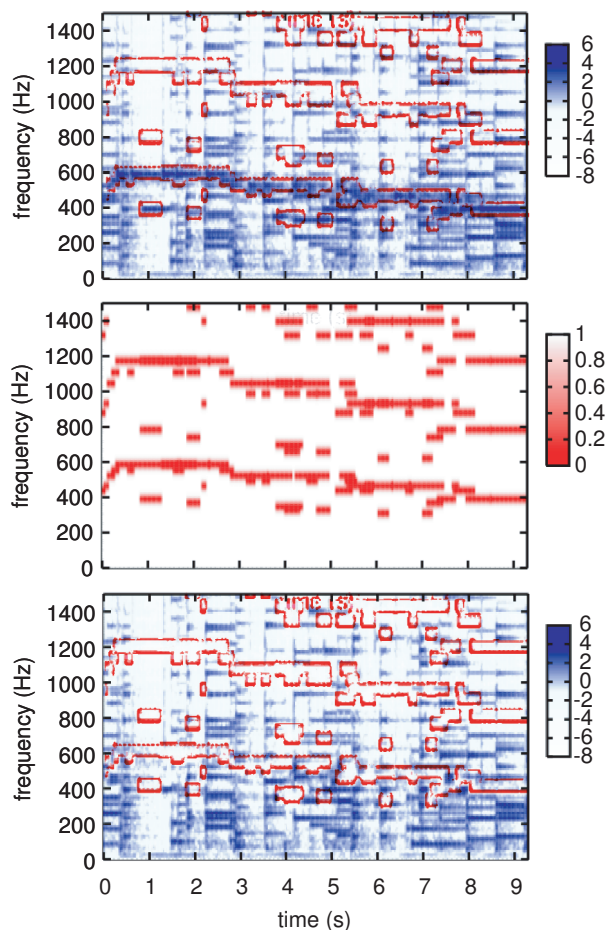


Fig. 3 処理前後のパワースペクトル（ログスケール表示）と時間周波数マスク：元の音源（最上段）に，時間周波数マスク（中段・各段赤線の輪郭）をかけて，各区間で調波成分を適切に除去した音源（最下段）を作成した

ラフォンの音は倍音成分が基本周波数成分と比較し，小さく，基本周波数成分のみの除去でも対応できたと予想される．対象音に依って除去する倍音成分を変化させること，または，倍音毎に除去区間をかえることなど，楕円ノッチフィルタによる除去にも検討の余地がある．

今後は，(1) アタック部分の除去，(2) 今回扱わなかった減算による除去を検討するとともに，第二節で論じた様々な問題を扱う．(1)(2) の課題解決に向け，特にステレオ音源に対し音源位置を利用する手法が有力な手法であろう．

謝辞 本研究の一部は，科学技術振興機構 CREST 研究課題「時系列メディアのデザイン転写技術の開発」として行われた．

参考文献

- [1] Kashino et al., SpeechCommun., Vol. 27, pp. 337-349 (1999).
- [2] Goto et al., SpeechCommun., Vol. 43, No. 4, pp. 311-329 (2004).
- [3] Orio et al., Proc. of ICMC 2001, pp. 129-132, (2001).
- [4] Pedersen et al., Proc. of ICA 2006, Vol. 3889, pp. 392-399 (2006).
- [5] Song et al., Proc. ISMIR 2002, pp. 133-139 (2002).
- [6] Virtanen et al., Proc. WASPAA 2001, pp. 83-86 (2001).
- [7] Viste et al., Proc. WASPAA 2003, pp. 25-28 (2003).
- [8] Woodruff et al., In Proc. of ISMIR 2006, pp. 314-319 (2006).
- [9] 糸山他, CrestMuse シンポジウム 2006, pp. 50-51 (2006)
- [10] 吉井他, FIT2004, LG-005 (2004).