

ドラム演奏ミキシングのための漏洩音抑制*

村松 知哉[†], 伊藤 克亘[†],

1 まえがき

ドラム録音で問題となる現象が Microphone Leakage(マイクの漏洩音)である。この音被りとは同じ空間に複数の楽器とマイクが存在するとき、意図していない楽器音がマイクに収音されてしまう現象のことである。図1は一般的なスネアマイクの短時間 RMS である。スネアドラムの音量は -12dB を推移し、バスドラムやシンバル類の漏洩音は -25dB を推移している。このマイク信号でスネアの音色を処理しようとしても漏洩音の影響が大きいため、スネア単体の処理には専門的な知識や経験が必要となる。そして初級者の演奏の録音では漏洩音が特に大きく、ミキシングで個々の楽器を処理することが難しくなってしまう。

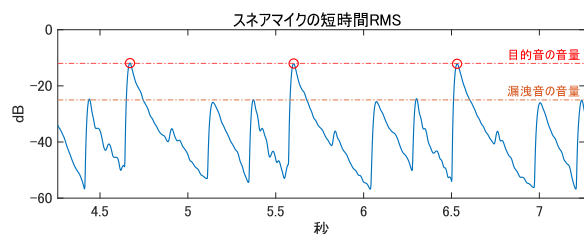


図1. スネアマイクの短時間 RMS

先行研究 [1] におけるウィーナフィルタは、漏洩音が目的音源に比べ小さく無視できることを前提としている。しかしシンバルの音を録音するためのマイク(ハイハットマイクやオーバーヘッドマイク等)については太鼓類の音量が大きく、その前提条件を達成できない。そのため一部のマイクにおける処理性能が悪くなる。

2 提案手法

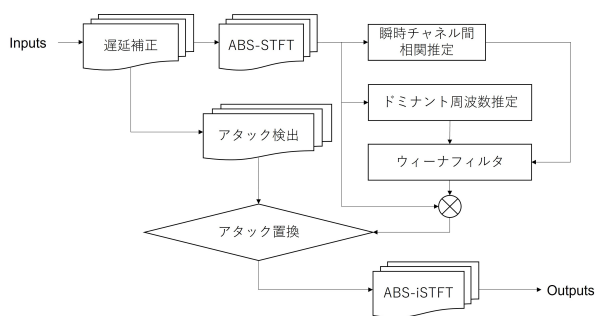


図2. 提案手法のフローチャート

本研究では、漏洩音が無視できない収音信号の処理を考案し、ミキシングの手助けとなるドラム録音に特化したマイクの漏洩音分離の手法を提案する。提案手法の処理の流れ(図2)は、まず入力された信号に近接効果とマイク同士の遅延の補正を行い、ABS-STFT(Adaptive Block Size-STFT)[6]で時間周波数領域に変換する。そしてウィーナフィルタの推定し信号へのフィルタ適用

を行う。またドラム録音のウィーナフィルタの精度を上げる補正項目として、瞬時チャンネル間相関推定、ドミナント周波数-PSD推定、共鳴音推定を考えている。この処理の流れは、先に波形ベースで処理できる補正を行いウィーナフィルタ推定の精度を上げるためである。最後にトランジェント検出を行い、検出されたフレームをウィーナフィルタに通過させないことによってアタック音が滲んでしまうことを防ぐ。

2.1 マイク間の遅延補正

マイク間の距離による遅延は音像を不明瞭にしてしまい、ミキシングやウィーナフィルタの推定において重要な問題となる。この遅延を補正するため、収音信号同士の相互相関を計算し一番楽器から離れているマイクの遅延に他のマイクを合わせるディレイ機能を作成する。

2.2 アタック検出と置換

ドラムのアタックに何かしらの音声処理を行うとアタックの歪みや滲みが起こる。これを回避するためにアタック検出 [3] を行い、アタックのみフィルタ処理を通過させない機能を作成。

アタックとは複数の周波数帯域で顕著なエネルギーピークが出現する短い領域の事である。収音信号にハイパスフィルタを掛け、各周波数帯域で複数の閾値を超える周波数帯域が検出された場合、その箇所をアタックポイントとする。またこのときの閾値の値は収音信号によって異なるが、図1では閾値を -12dB として検出を行っている。そして求めたアタックポイントをウィーナフィルタ処理前の収音信号に置き換える。このときスネアの音色やディケイの長さによって、矩形波窓やハン窓等の選択や置換する区間の長さ変更を行う。

2.3 ABS-STFT

音の立ち上がり早い楽器を STFT する場合、窓サイズを小さくすることで時間分解能が向上し音の立ち上がりを認識できるが、その代わり周波数分解能が低下する。そのため一つの周波数ビンで複数の倍音を表さなければならず、ミキシングを行うには音源が不明瞭になってしまう。逆に窓サイズを大きくすると周波数分解能は向上するが時間分解能が低下。そしてトランジェント区間の急峻な波形の変化では窓内の定常性が崩れてしまい Pre-echo[6] が発生する。

今回は Pre-echo の影響を少なくして周波数分解能を向上するために ABS-STFT(Adaptive Size Block-STFT)を使用する。これは Pre-echo が発生するトランジェントポイント(図3の赤い四角形部分)を検出し、その区間のみ STFT の窓長を短くするというものである。

このトランジェントを検出するために対象信号の短時間 RMS の変化量に着目し、閾値を超える地点をトランジェントポイントとする。

* : English Title Microphone-leakage suppression for drum performance mixing (Hosei Univ.) et al.

[†] 法政大学 情報科学部

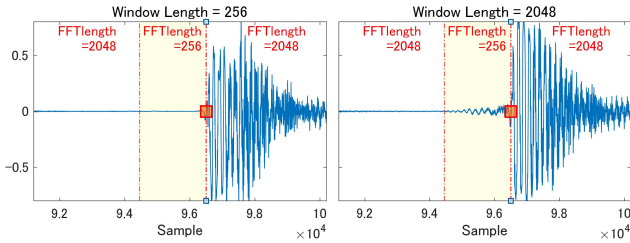


図 3. トランジェント検出ポイントと窓長を変更する区間

2.4 ドミナント周波数-PSD 推定

ウィーナフィルタ通過前後で変化のない周波数領域は音被りの影響を受けていない純粋な收音信号といえる。そしてその周波数領域に重みを付けたパワースペクトル密度を推定する。そのパワースペクトル密度を重みづけ関数 (1) としてフィルタに適用することで、ウィーナフィルタの処理精度を向上させる。

$$P_{s_m, s_m}(\kappa, \omega) = \gamma_m(\kappa) P_{s_m, s_m}(\kappa - 1, \omega) + (1 - \gamma_m(\kappa)) [P_m^D(\kappa, \omega) + P_m^R(\kappa, \omega)] \quad (1)$$

2.5 瞬時チャネル間相関推定

ある区間において、”目的音を奏でる楽器の收音信号” s 内の漏洩音が大きくなり無視できなくなるという状況は、結果的に”漏洩音となっている楽器の收音信号” u に s が相似しているということとなる。つまり漏洩音の影響が大きい区間の相関係数は大きくなり、影響が小さければ無相関になる。

3 実験と結果

本研究はどれだけ漏洩音が分離できているかを表す SIR(信号対干渉比)、どれだけ信号が歪んでいるかを表す SDR(信号対歪み比) の 2 つの評価方法を使用する。実験方法としては、ドラムフレーズを実際に叩きマルチトラックで録音し、その後各楽器単体音をそれぞれ收音。ドラムフレーズと各楽器単体音源を同じフィルタにかけ漏洩音を分離し、分離後ドラムフレーズは主観評価実験に使用する。分離後楽器単体音源は評価計算に使用し、SIR、SDR、分離前後の SIR Improvement を計算。そして計算結果を提案手法と先行研究手法で比較する。

今回は 2 種類の音色と 2 パターンの音量バランス、2 つのサスティン長の計 8 パターン (表 1) ドラム演奏を收音した。そしてスネア、バスドラム、ハイハットシンバルにマイクを設置し、その收音信号で分離実験を行う。

表 1. ドラムの演奏パターン

音色	音量バランス	サスティンの長さ
ハイピッチ	シンバルの音量 大	長い
ローピッチ	太鼓類の音量 大	短い

実験結果をそれぞれ評価したものが表 2 となっている。SIR Improvement は先行研究手法に比べて 9.15dB の向上がみられる。SDR については 1.5dB の低下となった。

表 2. 評価実験結果

	SDR	SIR Improvement
先行研究	9.25dB	15.27dB
提案手法	7.75dB	24.42dB
先行研究との差	-1.50dB	9.15dB

次に今回提案した補正項目の性能を評価する。アタック検出置換、ABS においては SIR、SDR どちらも変化は少なかったが、主観評価実験の評価から聴覚上での歪みが減った。瞬時チャネル間相関は SIR が 9.25dB と大幅な向上、SDR は -2.5dB の低下となった。

4 考察

SIR における 24.42dB の向上はミキシングのコンプレッション処理の平均的なリダクションレベルである -6dB を大幅に上回っており、過激なコンプレッション処理など積極的な音作りが十分可能な分離性能が得られた。

それぞれの補正項目の評価においてはアタック検出置換と ABS の計算上の性能が向上しなかったが、SDR の変化については歪み自体が信号にマスキングされている場合などが存在するため、数値だけで判断することは難しい。また瞬時チャネル間相関においては SIR の大幅な向上と SDR の低下が起こった。これは、目的音から離れているマイクと目的音のオンマイクで相関を求め重みづけを行ったことが原因と考えられる。太鼓類のマイクポジションの特性上リリースを收音しやすいマイキングになり、逆にオフマイクではリリースを收音しづらくなる。その結果リリース区間のスペクトル構造が変化することで相関が小さくなり、これによってリリース区間において目的音もフィルタで遮断してしまう。処理速度においては音質を担保するために長い窓長を必要とし、新たな補正項目を増やした結果、処理が重くなることが問題点に上がった。

5 あとがき

本研究では、チャネル間の相関関係とトランジェント-アタック区間への新たなアプローチから打楽器の漏洩音分離手法を提案した。評価実験から今回の提案手法でミキシングに十分な音質と分離性能で分離を行えることが分かった。この手法でミキシングにおける手間を減らすことが可能であると考え。リアルタイムでの処理が今後の課題だ。

参考文献

- [1] E. K. Kokkinis, et.al. "A Wiener Filter Approach to Microphone Leakage Reduction in Close-Microphone Applications," in IEEE TASLP, vol. 20, no. 3, pp. 767-779, March 2012.
- [2] E. Vincent, et.al. "Performance measurement in blind audio source separation," in IEEE TASLP, vol. 14, no. 4, pp. 1462-1469, July 2006.
- [3] X. Rodet and F. Jalliet, "Detection and modeling of fast attack transients" in ICMC, 2001, NA, France. pp.1-1.
- [4] H. Sawada, et.al. "New formulations and efficient algorithms for multichannel NMF," 2011 IEEE Workshop on WASPAA, 2011, pp. 153-156,
- [5] E. Kokkinis, et.al. "A New DSP Tool for Drum Leakage Suppression," Engineering Brief 108, (2013 October.).
- [6] Imen Samaali, et.al. "Attack restoration in low bit-rate audio coding, using an algebraic detector for attack localization," ISVC, Sep 2010, Rabat, Morocco.