

実世界の音楽音響信号に対するドラムスの音源同定手法を利用した ドラムイコライズシステム INTER:D の開発

Development of Drum Equalizer System INTER:D Using Drum Sound Identification Method for Real-World Music Signals

吉井 和佳[†]
Kazuyoshi Yoshii

後藤 真孝[‡]
Masataka Goto

奥乃 博[†]
Hiroshi G. Okuno

1. はじめに

我々は音楽を聴くとき、CD プレイヤーなどに搭載されているグラフィックイコライザを用いて楽曲の音量バランスを調節することがある。従来のグラフィック（パラメトリック）イコライザは、音楽音響信号に対して帯域抑制・通過型フィルタを適用し、該当帯域周波数のスペクトル成分を抑制あるいは増幅するという原理で動作する。この方式では、任意の周波数帯域のスペクトル成分を調節することができる一方、ある楽器の音量のみをピンポイントで調節することはできなかった。例えば、市販品のボーカルキャンセラでは、ボーカルの周波数帯域（400Hz 周辺）を抑制すると、ボーカル以外の周波数成分まで大きく減衰してしまう問題がある。

この問題を解決するため、音量調節したい楽器に由来するパワー成分のみを調節できるイコライズ方式 **INTER (Instrument Equalizer)** を提案する。INTER を用いることで、ユーザはその楽器パートの音量だけを他楽器に影響を与えることなく調節しながら音楽を楽しむことが可能になる。INTER を実現するには、音楽音響信号からイコライズ対象となる楽器個体のパワースペクトルを再構成し、発音時刻をすべて検出する必要がある。

本稿では、我々が提案したドラムスのパワースペクトルに関するテンプレート適応・マッチング手法 [1] を利用して、ドラムイコライズシステム **INTER:D (INTER for Drums)** を開発したので報告する。プロトタイプシステムでは、バスドラムとスネアドラムの音量調節が可能である。また、ドラムスの音量を完全にキャンセルし、全く別の音色のドラムスに置き換えてしまう機能も実装されている。ドラムスは現代の音楽において重要な役割を果たしており、音量や音色は楽曲の与える印象と密接な関係がある。そのため、さまざまにドラムスの設定を変えて試聴できる INTER:D は、初心者からプロユーザまで幅広いニーズがあると考えられる。

2. イコライズ方式 INTER

本章では、提案するイコライズ方式 INTER の基本原理について説明する。INTER を利用するには、イコライズ対象となる楽器個体 (X とする) のパワースペクトルを、テンプレート P_X として用意する必要がある。以下にイコライズ処理の流れを示す (図 1)：

1. 音楽音響信号中から楽器個体 X のすべての発音時刻を検出する。
2. 各発音時刻から始まるスペクトルを抽出し、音量調節の大きさに合わせて一定比率まで減少させた P_X

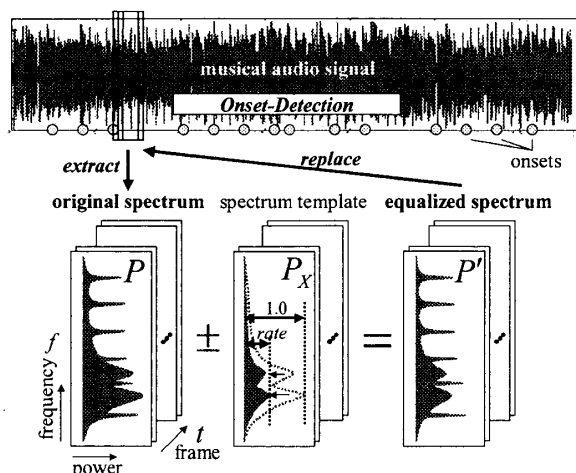


図 1: INTER におけるイコライズ処理の流れ

を加減算する。この処理によりイコライズされたスペクトルが得られ、もとのスペクトルと置換する。いま、抽出されたスペクトルを P 、イコライズ後 P' とすると、以下を満たすように P' を求める。

$$P'(t, f) = P(t, f) \pm (\text{rate} \cdot P_X(t, f)) \quad (1)$$

ここで rate とは、楽器個体 X の音量の変化率を表す定数 ($0 \leq \text{rate} \leq 1$) である。また、 (t, f) ($1 \leq t \leq T$ フレーム, $1 \leq f \leq F$ ビン) でパワースペクトルを構成する各パワー成分を示す (T, F は正の定数)。

このとき位相の設定方法が問題になる。パワーは単純な増減処理ができるが、 P' のスペクトル計算時に誤った位相を設定すると、逆フーリエ変換後に虚部が残り、正しい信号が得られない。この問題を解決するため、 P の位相をそのまま再利用する。

3. すべての発音時刻についてスペクトルを置換することで得られた楽曲全体のスペクトルを逆フーリエ変換すれば、イコライズされた音響信号が得られる。

INTER では、イコライズ対象楽器個体 X を他の音色の楽器個体 Y に置き換えることもできる。まず、各発音時刻について $\text{rate} = 1.0$ として P_X を減算し、逆フーリエ変換を行うことで、楽器個体 X の音量が完全にキャンセルされた音響信号を得る。次に、置き換えたい楽器個体 Y の音響信号を各発音時刻において加算する。

3. ドラムイコライザ INTER:D

INTER を実現する上での問題点は、実世界の楽曲で使われているドラムスの音色はバリエーションに富むため、あらゆる音色のドラム音テンプレートを事前に用意でき

[†] 京都大学大学院 情報学研究科 知能情報学専攻

[‡] 産業技術総合研究所 (AIST)

表1: バスドラム・スネアドラム音源同定結果

piece number	bass drum		
	recall rate	precision rate	F measure
No.11	100 %	100 %	1.00
No.30	93 %	94 %	0.93
No.50	97 %	88 %	0.92
No.52	94 %	90 %	0.92

piece number	snare drum		
	recall rate	precision rate	F measure
No.11	95 %	92 %	0.93
No.30	100 %	96 %	0.98
No.50	67 %	96 %	0.78
No.52	90 %	97 %	0.93

ないことである。この問題を解決するため、我々が提案したパワースペクトルに関するテンプレート適応・マッチングによるドラムスの音源同定手法 [1] を利用する。

3.1 ドラムスの音源同定手法の利用

本手法は、テンプレートとなるパワースペクトルをバスドラム・スネアドラムごとに1つしか必要とせず、それを楽曲に使用されているイコライズ対象となるドラム音に適応させていく。さらに、適応したテンプレートを用いてテンプレートマッチングを行い、音響信号中からすべての発音時刻を検出することができる。

我々は手法の有効性を検証するため、ポピュラー音楽データベース RWC-MDB-P-2001 [2] 中の4曲で音源同定実験を行った。評価は再現率、適合率、F値で行う。

実験結果を表1に示す。結果から、90%前後の高いF値、適合率が得られ、テンプレート適応手法が正しく動作することが示された。よって、我々の提案した音源同定手法が INTER:D に有効に利用可能であるといえる。

3.2 INTER:D のプロトタイプ実装

適応済みテンプレートと検出された発音時刻を用いて INTER:D のプロトタイプを実装する。スペクトル解析には、44.1kHz の音響信号に対して、シフト長 441 点、窓幅 4096 点の短時間フーリエ解析を利用する。また、テンプレートのサイズは $T = 15$, $F = 2048$ とする。

適応済みテンプレート P_X には、ドラム成分だけでなく、小さな調波成分や雑音的な成分が重畳している。そのため、これをそのまま楽曲のパワースペクトルから加減したのでは、ドラム以外の成分まで余計に調節してしまう問題がある。そこで、 P_X 中の小さなパワー成分を加減するときには変化率 $rate$ を下げ、ドラム以外の成分まで加減してしまう影響を小さくする。式(1)の変化率 $rate$ を次式(2)の関数 $R(t, f)$ に置き換える(図2)。

$$R(t, f) = \frac{P_X(t, f)}{P_M} \cdot rate \quad (2)$$

ここで P_M とは、 P_X 中での最大パワー成分である。

4. 評価実験

INTER:D の性能を評価するため、3.1 節の4曲に対してイコライズ実験を行った[§]。イコライズした信号を5人

[§] <http://winnie.kuis.kyoto-u.ac.jp/members/yoshii/FIT2004.html> にて実験に使用した音響信号が試聴可能である。

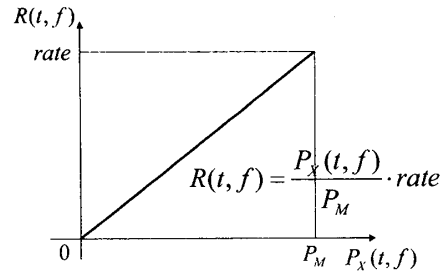


図2: パワー成分の大きさによる音量変化率の調整

表2: INTER:D 評価実験結果 (他楽器への悪影響の大きさを5人が評価した平均を示す。値が小さいほどよい。)

piece number	suppress	
	bass drum	snare drum
No.11	3.8	4.6
No.30	2.0	4.4
No.50	3.2	7.2
No.52	3.2	6.0

の被験者が試聴し、イコライズ対象以外の楽器への悪影響の大きさを相対評価した。従来のグラフィックイコライザで低域抑制した場合を10とし、INTER:Dによりバスドラム・スネアドラムそれぞれを抑制する場合の悪影響の大きさを数値で評価した。実験結果を表2に示す。INTER:Dの方が他の楽器への悪影響が小さく、優れていることが分かる。バスドラムに比べ、スネアドラムのイコライズが難しいことも示された。また、女性ボーカルの曲ではボーカルが消えやすいという意見があった。

次に、INTER:Dを用いて楽曲の使用ドラムスを他の音色のドラムスに置換した信号を試聴してもらい、アンケートを行った。音色が完全に置換されているにもかかわらず、自然な音響信号が生成されているという意見が多かった。これは、ドラムスを完全にキャンセルしただけではドラムス以外の成分まで消え、不自然さが目立つが、そこに他のドラム信号を足しこむことで、消えた部分が聴覚上マスキングされたためだと考えられる。これらのことは、本手法の有用性を示している。

5. おわりに

本稿では、音量調節したい楽器のスペクトル成分のみを調節し、他楽器のスペクトル成分への影響が小さいイコライズ方式 INTER を提案した。さらに、我々が提案したドラムスの音源同定手法を利用して、INTER方式で動作するドラムイコライズシステム INTER:D を開発した。実際のポピュラー音楽を用いたイコライズ実験により、本システムが有効に機能することを確認した。より自然な聴感を実現するためには、置き換えたスペクトルの前後を滑らかに接続する処理などが考えられる。今後、機能をさらに改善・追加していく予定である。

参考文献

- [1] 吉井和佳, 後藤真孝, 奥乃博: “テンプレート適応を利用した実世界の音楽音響信号に対するドラムスの音源同定”, 情報処理学会音楽情報科学研究会 研究報告 MUS-53-12, No.127, pp.55-60, 2003.
- [2] 後藤真孝, 橋口博樹, 西村拓一, 岡隆一: “RWC 研究用音楽データベース: 研究目的で利用可能な著作権処理済み楽曲・楽器音データベース”, 情報処理学会論文誌, Vol.45, No.3, pp.728-738, 2004.