

Mix-down 支援システムにおける非言語系インタフェースに関する検討

Consideration about a Non-Language Interface in Mix-Down Assistant System

谷井章夫
Akio Yatsui

片寄晴弘
Haruhiro Katayose

関西学院大学
Kwansei Gakuin University

さががけ研究 2 1
PRESTO, JST

If the setting information on mixing differs, even if it uses the same sound material, the impressions of a musical piece differ greatly. For this reason, in order to do exact mixing, high skill and experience are required. Therefore, it is difficult to obtain the result meant although amateur do mixing. In recent years, in commercial music work, mixing was performed more often using the software on a computer. The work which the mixing engineer was doing using a hardware mixer and an effector is enabled on a computer, and there was no support for amateur. In this research, know-how of mixing which a mixing engineer with abundant experience has is template-ized, and the amateur proposes a system reusable in his work process.

1. はじめに

ミックスダウンとは、レコーディングによって録音された各トラックの音素材に対し、音量や音像定位の調節とエフェクタによる音質加工処理を施し、ステレオトラックにまとめあげる作業である。商用音楽制作における中核的な作業であり、最近では、計算機上のソフトウェアを利用して実施されることが多い。プロのミキシングエンジニアは、完成後の「聞こえ」をイメージした上で、それを目標としてミックスダウン作業を行う。イメージした通りの「聞こえ」を得るためには、エフェクタの選択から各種パラメータの設定に至るまでの、きめ細やかな作業が必要である。高度な技能と経験を要し、アマチュアが意図した結果を得ることは難しい。

我々は、経験豊富なミキシングエンジニアの持つミックスダウンのノウハウをテンプレート化しておき、経験の乏しいアマチュアが自分の制作過程で再利用できるシステムを提案してきた[谷井 03]。デザイン転写を実施する際には、A, B, サビなどの音楽構造、楽器種や奏法の類似性がタグとして利用される。本論文では、システムの概要を説明し、特に、人間にとって直感的な認知的特徴をシステム内部で取り扱うための音響信号の解析処理について述べる。

2. システム概要

1つの楽曲に対して施すべきミックスダウンデザイン*1は唯一に決まらない。同じ音素材から構成される楽曲データに対して、異なるミックスダウンを施した場合の事例を用意した[Demo]。これらは、プロのミキシングエンジニアに依頼し制作したものである。ミックスダウンデザインが異なると、楽曲の印象が大きく変わることがわかる。

本研究では、ミキシングソフトウェアとして、世界的に普及しているレコーディング・エディティング・ミキシングシステムの1つである Pro Tools (Digidesign, Inc.) [Digidesign] を使用する。ここで、楽曲 S のミックスダウンデザインを楽曲

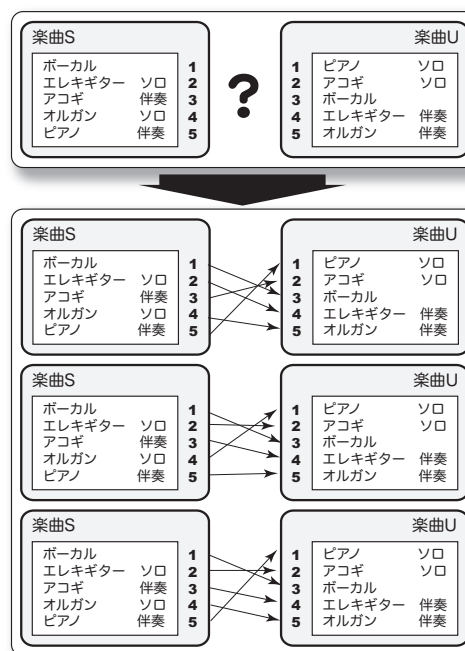


図 1: トラックの対応付け

Uへ転写する状況を考える。図1に示すように、ミックスダウンを転写するには、転写元と転写先のトラックの対応付けが必要である。一般的に、楽器種や奏法により対応付けされるが、その組み合わせパターンは複数考えられる。楽器種と奏法のどちらをどの程度重視するかによって、そのパターンが決まる。これら楽器種や奏法は、楽曲全体を通して不変ではない。Aメロ・Bメロ、サビ等の音楽構造ごとに楽器種編成が変わったり、奏法が変化することが多い。また、音楽構造ごとにミックスダウンを変化させるために、同楽器による演奏を複数トラックに分けることも多い。

そこで、本システムでは、ミックスダウン転写の前準備として、音楽構造を解析し、(音楽構造ごとの)各トラックの楽器種と奏法を解析し、これらの情報をアノテートするシステムを実現する。システムが保有している楽曲に関しては、事前に

連絡先: 谷井章夫, 関西学院大, 〒 669-1337 三田市学園
2-1 関西学院大学 理工学部 片寄研究室, tel.&fax.
0795-65-7861,

<http://ist.ksc.kwansei.ac.jp/~katayose/members/yatsui>

*1 ボリュームやパンニング, エフェクト処理の順序やパラメータなど, ミックスダウン作業で設定する全ての情報を指す。

アノテートしておく。ユーザの楽曲は、提示時に解析し、アノテートする。これらのアノテーションを楽曲データの信号処理によって実施する。

我々は、ユーザ主体のミックスダウン支援システムの実現を目指している。転写元楽曲の選択や、トラック対応付けの際の楽器種、奏法の優先度の指定など、随所にユーザに決定権を与える*2。システムはあくまで様々なミックスダウンデザインを提示する立場にある。システムは、選択された楽曲のミックスダウンデザインを、ユーザの指定した楽器種、奏法の重みをもとに、ユーザの楽曲にどのように当てはめるのが適当かを様々な視点から提示する。以下に、ユーザが自身の楽曲をシステムに提示してから、システムによりミックスダウンデザインが転写されるまでの一連の流れを示す。

1. ユーザ
自身の楽曲データをシステムへ提示する。
2. システム
提示された楽曲に対して、音楽構造、楽器種、奏法を解析し、アノテートする。
3. ユーザ
類似楽曲検索における重み（音楽構造、楽器種、奏法）と類似楽曲提示数を指定する。
4. システム
設定された重みをもとに類似度の高い楽曲をユーザの指定した曲数分提示する。
5. ユーザ
完成イメージに近い楽曲（好みの楽曲）を選択する。
6. ユーザ
音楽構造の対応付けの際に重視する項目を選択する（音楽構造名による対応付け、楽器種編成の類似性による対応付け）
7. ユーザ
トラックの対応付けにおける重み（楽器種、奏法）を指定する。
8. ユーザ
ミックスダウン転写パターン数を指定する。
9. システム
指定された重みを元にトラック間の類似度を計算し、音楽構造ごとに類似度の高いトラック同士を対応づける。トラックの対応付けパターンはユーザの指定数分だけ用意する。
10. システム
トラックの対応付けパターンをもとにミックスダウンデザインを転写する。
11. ユーザ
イメージした「聞こえ」に仕上がったものを選択する。

アマチュアにとっては、この段階でほぼ完成された結果を得ることができる。プロのエンジニアは、ここでの結果をベースにして修正作業を行っていくことで作業の効率化をはかることができる。なお、過程3と4は、完成イメージの浮かばないユーザに対して楽曲の選択肢を減らすために設けた類似楽曲検索機能に対するタスクであり、必要としないユーザは、無視できる。類似楽曲検索機能により、ユーザは、自身が用意した楽曲と、楽器編成や奏法の観点で類似する楽曲にどのようなミックスダウンが施されたかを視聴することができる。

以下、音楽構造の認識と楽器種・奏法の解析について述べる。

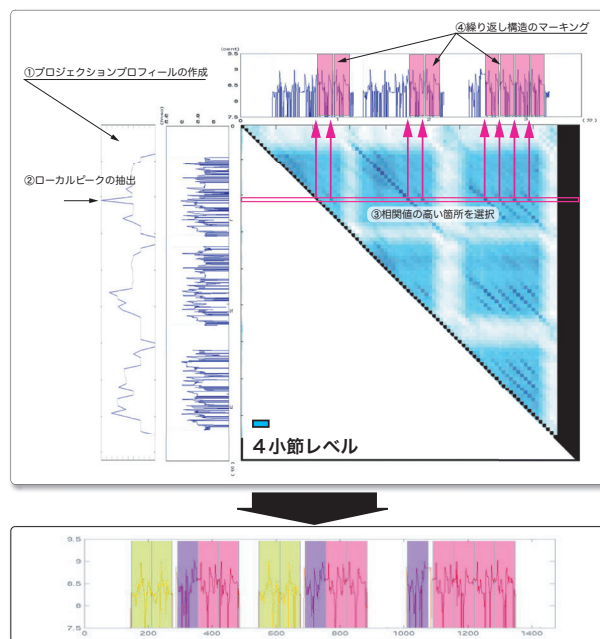


図 2: 音楽構造の抽出

3. 音楽構造、楽器種・奏法のアノテーション

3.1 音楽構造の認識

繰り返し構造に基づいて音楽 CD の音楽構造の自動抽出に取り組んだ研究として、後藤の RefraiD (Refrain Detecting Method)[Goto 03]がある。RefraiD は、音楽 CD 等による複雑な混合音を含む楽曲に対して様々な繰り返し区間の相互関係を調べることで、サビ区間を含む音楽構造を網羅的に検出する手法である。本研究が対象とする音響データは、既にパートごとに分かれており、またビートの利用もできることから、よりシンプルな手法で音楽構造の認識ができると考えている。人間は、主にボーカルの歌のメロディの繰り返し構造により音楽構造を判断している。そこで本研究では、ボーカルパートの基本周波数の時系列データの相関を利用し、繰り返し構造を抽出することにより、音楽構造を認識する。

まず、リズム系パートからビート単位を抽出する。続いて、ボーカルパートに対して自己相関関数を利用し、基本周波数の時系列データを得る。4 小節を窓幅として対応する基本周波数の音の並びとの相関値を求める。区間をビート単位分ずらし、同様に相関値を求めていく。図 2 は、RWC 研究用音楽データベース [後藤 01]RWC-MDB-P-2001 No.13 のボーカルパートの音素材データ [Demo] の基本周波数の時系列データにおける 4 小節レベルの相関マップである。相関の高い部分が、濃い点となって現れる。以下、相関マップをもとに音楽構造を得る手法を述べる。

1. 相関マップの行ごとに相関値の総和を算出し、プロジェクションプロフィールを作成する。
2. プロジェクションプロフィールのローカルピークを探索し、繰り返し構造の開始点を多く含む可能性の高い地点を得る。
3. 得られた行において相関値の高い箇所を探索し、繰り返し構造の開始点を抽出する。
4. 抽出された繰り返し部分にマーキングする。

その他のローカルピークに対しても、同様の処理を繰り返していくことで、複数の繰り返し構造が得られる。得られた繰

*2 例えば、ハードロック曲のミックスダウンデザインをクラシック曲に当てはめようが、ユーザの自由である。様々なパターンを当てはめながら、気に入ったミックスダウンデザインを最終的に利用すればいい。

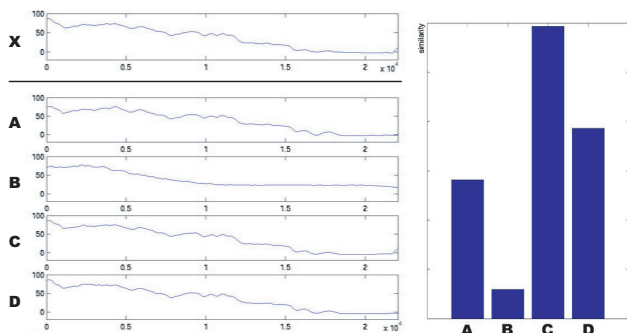


図 3: スペクトル包絡の類似度による楽器種の類似性の判断

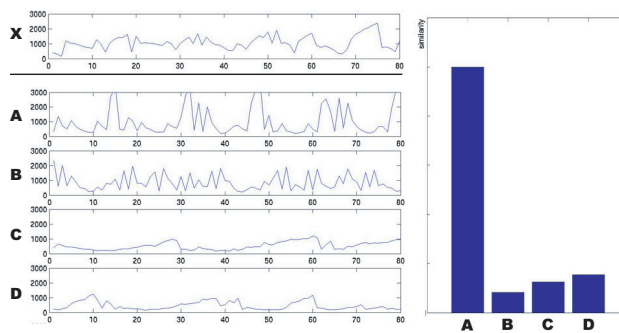


図 4: 発音回数による奏法の類似性の判断

り返し構造のうち、最も繰り返し回数の多いものをサビと判断する [後藤] .

3.2 楽器種・奏法の解析

楽器種の解析

本研究では、スペクトル包絡の時間的な変化の類似性により、楽器種の類似性を判断する。具体的には、発音開始直後から1秒間のデータを採取し、さらに複数区間に分割した上で、各区間ごとのスペクトル包絡を算出する。これを楽器種の特徴量としてアノテートする。本研究における楽器種の解析は、楽器種の特長を目的としていない。人間が認知的に近いと感じる音素材を結びつけることが目的である。図3に、スペクトル包絡の類似度により楽器種の分類を行なった例を示す。例では、音素材データA, B, C, Dを用意し、音素材データXのスペクトル包絡に対して最も類似度の高い音素材データを同類楽器種と判断する [Demo]。音素材データXは、音素材データCに対して最も類似度が高いという結果が得られている。

奏法の解析

本研究では、奏法の特徴量として、以下のものを解析する。

- 発音回数
ギターのピッキング時やピアノの打鍵時など、発音開始直後のスペクトル包絡は、乱れやすい。そこで、スペクトル包絡の乱れを捉えることにより、発音地点を抽出する。抽出された地点の個数を発音回数と判断する。図4に、音素材データXに対する音素材データA, B, C, Dの発音回数の類似度をみた結果を示す [Demo]。音素材データXは、音素材データAに対して最も類似度が高いという結果が得られている。
- 発音間隔の散らばり
発音回数が等しくても、発音の間隔が一定のものとはばらつきのあるものでは、奏法の印象が異なる。そこで、発音間隔の散らばりを捉える。
- 発音の歯切れ良さ
意図的に音を短時間で止めるような演奏と、止めない演奏とでは、奏法の印象が異なる。そこで、各発音の持続時間を抽出する。
- 音域と広がり
低音のみによる演奏、高音のみによる演奏、音域を広範囲に行き来する演奏とでは、印象が全く異なる。そこで、演奏内容の音域とその分布を抽出する。

- 強弱の付け具合

強い演奏、弱い演奏、強弱をつけた演奏とでは、印象が大きく異なる。そこで、各発音地点のパワーを抽出し、その分布を捉える。

以上の特徴量を1小節ごとに算出し、その平均値をアノテートする。

4. 実験と検討

「Aメロ, Bメロ, サビなどの構造記述」「楽器の音色」「奏法」など、人間にとって直感的な認知的特徴をシステム内部で扱おうとしている点が本システムの特徴である。現在のところ、提案手法により音楽構造の認識率が100%得られている。これは、転調を含んでいないポップス曲に限定したためと考えられる。今後は、転調を含んだ楽曲や他ジャンルへの対応も行っていく。奏法の認識においては、もともとノイズを多く含んでいる楽器音に対する認識率が下がった。これは、発音直後以外の箇所でも周波数の乱れが生じるためと考えている。今後は、アノテーションと類似判別の高度化における手法として北原の手法 [北原 04] 等を利用し、さらにシステムの高度化をはかり、精緻な評価を行っていきたい。

RWC研究用音楽データベース (ポピュラー音楽) [後藤 01] の制作過程情報 (Pro Tools データ) を使用し、RWC-MDB-P-2001 No.82 (ハードロック風) のミックスダウンデザインとRWC-MDB-P-2001 No.98 (ファンクロック風) のミックスダウンデザインをRWC-MDB-P-2001 No.88にそれぞれ手作業により適用する実験を行った [Demo]。

我々が提案したミックスダウン支援システムは、ミキシングエンジニアやアマチュアの制作者により利用されることで、本来膨大にかかるミックスダウン作業時間を短縮することに繋がり、発想の幅を広げる上でも有用である。一方、一般の聴取者が、楽曲素材に対して各自の好みのテンプレートを適用して楽しむ、という利用形態への発展も検討している。イメージとしては、高級なグラフィックイコライザである。実現には、ビジネスモデルから変えていく必要があるが、音楽流通の新たな形態が生まれる可能性もある。

その一方で、事例を利用することにより生じる問題もある。ひとつは、ユーザがイメージした「聞こえ」に近い楽曲が、システム内に存在する必要があるということである。この問題を根本的に解決するためには、音楽ビジネスモデル自体を変えていく必要がある。具体的には、アーティスト側が、ミックスダウンデザインを公開していくことで解決できる。着メロサービスが当初予想もしなかった程に普及していることから分かる

ように、ミックスダウンデザインの公開が、ビジネスとして成り立つ可能性は十分にありうる。また、著作権に関する問題も考えられる。音楽制作において、著作権として最も保護されているのは、作曲であり、ミックスダウンに関しては重要視されていない。また、一旦抽象化を行ったものについては合法である。本研究では、できる限り豊富なミックスダウンデザインを用意するために、現在、ミックスダウンデザインデータベースの制作に取り組んでいる。

5. おわりに

我々は、音楽制作におけるミックスダウンデザインのテンプレート化という新たな概念を提唱し、その楽曲からの抽出と再利用を可能にするシステムを提案してきた。ミックスダウン転写の前準備として、転写元と転写先の楽曲において、音楽構造ごとの各トラックの楽器種と奏法について解析する必要がある。本稿では、システム概要と音楽構造、楽器種、奏法の解析処理について述べた。現段階で、システムのデザインと音楽構造と楽器種・奏法の抽出処理の実装が終了している。今後は、各モジュールの高精度化とシステム統合を行う。また、実際に人間が認知したトラックの対応付けが実現しているかに関して評価していく予定である。

参考文献

[谷井 03] 谷井章夫, 後藤真孝, 片寄晴弘: ミックスダウンデザインの抽出と適用, FIT2003 (情報科学技術フォーラム) 情報技術レターズ Vol.2 2003, pp.109-110, (2003).

[Degidesign] <http://www.digidesign.com/>

[Goto 03] Masataka Goto: A Chorus-Section Detecting Method for Musical Audio Signals, Proceedings of ICASSP 2003, pp.V-437-440, (2003).

[後藤 01] 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: ポピュラー音楽データベースと著作権切れ音楽データベース, 情処研報音楽情報科学 2001-MUS-42-6, pp.35-42, (2001).

[北原 04] 北原鉄朗, 後藤真孝, 奥乃博: 音響的類似性を反映した楽器の階層表現の獲得とそれに基づく未知楽器のカテゴリレベルの音源同定, 情報処理学会論文誌, Vol.45, No.3, pp.680-689, (2004).

[Demo] http://ist.ksc.kwansei.ac.jp/~katayose/MXD/jsai_demo.html