

# 事例を利用したミックスダウン支援インタフェース

谷井章夫<sup>†,††</sup> 後藤真孝<sup>†††</sup> 片寄晴弘<sup>†,††</sup>

ミックスダウンにおけるエフェクタや音量、音像定位の設定情報が異なれば、同じ音素材を用いても、楽曲の印象は大きく異なる。的確なミックスダウンを行うには高度の技能と経験が必要であり、アマチュアにとって意図した結果を得ることは困難である。近年、ミックスダウンのためのソフトウェア環境が充実してきており、商用音楽の多くがその環境上で制作されるようになってきている。しかし、その環境はハードウェアのミキサやエフェクタをそのままソフトウェアに置き換えたものであり、経験の乏しいアマチュアの使用を前提とした支援については十分に考慮されていない。本論文では、既存事例からミックスダウンのデザインを転写することによって、そのデザインを支援する方法論を提案し、その方策、技術的な準備、基礎的評価、可能性について論じる。

## An Mix-Down Assistant Interface with Reuse of an Example

AKIO YATSUI,<sup>†,††</sup> MASATAKA GOTO<sup>†††</sup> and HARUHIRO KATAYOSE<sup>†,††</sup>

Mix-down (track down) occurs at the final stage of commercial music production. Selecting effectors for each sound track and setting the parameters of each effector balance the sound from each track in stereo (mix-down design). The mix-down process greatly influences the final sound quality. Recent professional mixing tasks have been done on digital audio workstations, i.e., software on PCs. In this sense, amateur musicians have entered the realm of professional production. However, it is difficult for mix-down beginners to obtain the design they want. One rational way of supporting mix-down is by using examples. In this paper we propose a mix-down supporting interface that copies an existing mix-down design to the given music, and describe its functions, evaluations and possibilities.

### 1. はじめに

近年の商用音楽制作は、ハードディスク・レコーディング・システムと呼ばれる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。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、最終的にステレオトラックにまとめ上げる作業である。商用音楽制作における中核的な作業であり、楽曲の完成度を大きく向上させることができる。また、ミックスダウンにおけるデザイ

ンが異なると、楽曲の印象が大きく変わる<sup>2)</sup>。

個々の楽曲に対するミックスダウンは、唯一ではなく、様々なミックスダウンを施すことが可能である。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に取り組む。イメージ通りの「聞こえ」に仕上げるためには、各音素材に適用するエフェクタの組み合わせと適用順序、各種パラメータの設定に至るまで、適切に行なう必要があり、ミキシングエンジニアが持つ高度な技能と経験を要する。したがって、アマチュアがミックスダウンに取り組んでもイメージ通りに仕上げることは難しく、良質なマスターデータが得られないという問題があった。ミキシングソフトウェアは、基本的に、ハードウェアミキサやエフェクタをソフトウェアに置き換えたものであり、技能と経験の乏しいアマチュアがミックスダウンを行うのは容易ではない。

<sup>†</sup> 関西学院大学

Kwansei Gakuin University

<sup>††</sup> 科学技術振興機構さきがけ研究 21「協調と制御」領域

”Intelligent Cooperation and Control,” PRESTO, Japan Science and Technology Corporation (JST)

<sup>†††</sup> 産業技術総合研究所

National Institute of Advanced Industrial Science and Technology (AIST)

現在、世界的に普及しているシステムの1つに、ProTools (Digidesign, Inc.)<sup>1)</sup>がある(図1)

ボリュームバランスやパンニング、エフェクタ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。現在、ProToolsには200種類以上のエフェクタがプラグインとして用意されている。



図 1 ProTools を利用したミックスダウン環境



図 2 イメージ伝達における事例利用の有効性

そこで我々は、経験豊富なミキシングエンジニアが仕上げた楽曲におけるミックスダウンデザインをユーザが持ち込んだ楽曲に転写するという方式を提案し、その具体化を進めてきた<sup>3)4)</sup>。以下、第2章でイメージ伝達における事例利用の有効性と、事例を利用することで実現可能なインタフェースについて議論し、第3章でミックスダウンデザイン転写時に参照される「A, B, サビなどの音楽構造」と「音色・奏法」に関する情報の解析手法について述べる。第4章では提案インタフェースの予備的なユーザビリティ評価を紹介し、第5章で本インタフェースの応用領域について検討する。

## 2. 事例を利用したミックスダウン支援インタフェース

### 2.1 イメージ伝達における事例利用の有効性

本章では、アマチュアのためのミックスダウン支援に求められる事項と、それぞれの用途に向けたインタフェースの形を整理する。

アマチュアの制作者（クライアント）が、自身の制作中の楽曲のミックスダウンをプロのミキシングエン

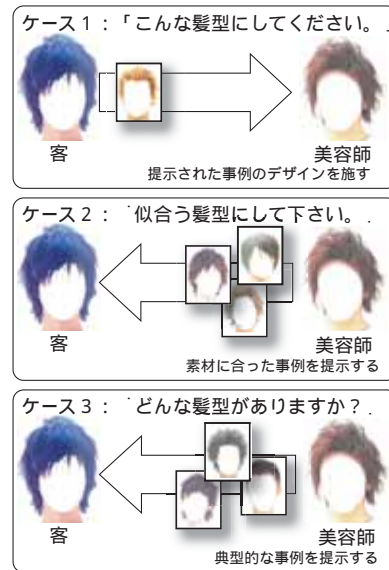


図 3 事例利用により実現する3つのケース

ジニアに依頼する状況を考える。ミキシングエンジニアに完全に任せてしまうという状況以外では、クライアントは楽曲の「仕上げ」のイメージをミキシングエンジニアに伝える必要がある。しかし、そのイメージを言葉で説明することは容易ではない。形容詞を並べてなんとかイメージを伝えるということも考えられるが、自分の持っているイメージが相手の心の中に再構築できたかどうかを確認すること自体そもそも困難である。このような状況に際し、デザイン現場では「事例」が参照されてきた。イメージ伝達における事例利用の有効性については、ヘアメイクのやりとりを思い浮かべればわかりやすい(図2)。

### 2.2 事例活用型のデザインインタフェース

クライアントとエンジニアのインタラクションに着目した事例活用型のデザインインタフェースとして3つのケースが考えられる。その3つのケースをヘアメイクの例に対応させたものを図3に示す。

ケース1) こんな「聞こえ」にして下さい。

クライアントが好みの事例を持ち込み、エンジニアに提示するケースである。この場合、エンジニアは、提示された事例のミックスダウンデザインをクライアントの楽曲に適用する。

ケース2) 素材を活かして下さい。

音楽スタイルや楽器構成などを考慮し、楽曲の特徴に合うようにミックスダウンを実施するというケースである。この場合、エンジニアは条件を満足する類

ヘアメイクの例では、顔型、髪質、毛量などが素材に相当する。

似事例を検索し、クライアントに提示する。クライアントは提示された中から好みの事例を選択し、エンジニアはその事例に基づいてミックスダウンを実施する。ケース3) どんなデザインがありますか？

クライアントが典型的なデザインをいくつか聴いてみたいというケースである。エンジニアは、典型的なミックスダウンデザインを施した楽曲を複数提示する必要がある。クライアントは提示された中から好みの事例を選択する。

以上3つのケースに対応したミックスダウン支援インタフェースを実現するためには、音楽構造の理解、音色・奏法の理解、アノテーション利用が中心的な課題となる。

### 2.3 アノテーションの利用

ミックスダウンのデザイン転写において利用が想定されるアノテーション情報(注釈)は、外的な付帯情報と、楽音に直結した情報に大別される。前者は、例えば「ホラームービーでのBGMに使われた」「ダンス音楽に定評のあるエンジニアによってデザインされた」などの情報である。後者は、楽器種名、奏法、音楽的意味(構造)などである。ミックスダウンデザインには過去の膨大な資産が存在するが、実際の制作過程において、後者の楽音に直結するアノテーションはほとんど実施されてこなかった。対象を分析することによってアノテーションを得る意義は大きい。

ここで、ミックスダウンが施されていないユーザの楽曲に対して、既存事例のミックスダウンデザインを転写する状況を考える。図4に示すように、ミックスダウンデザインを転写するには、転写元と転写先のトラックを対応付ける必要がある。トラックの対応付けは、音色や奏法に基づいて行なうのが自然である。音色や奏法は、楽曲全体を通して不変ではない。RWC研究用音楽データベース(ポピュラー音楽<sup>5)</sup>100曲の制作過程情報(Pro Tools データ)においても、次のような特徴がみられた。

- A メロ・B メロ、サビ等の音楽構造ごとに楽器種編成が変化したり、奏法が変わることが多い。
- 音楽構造ごとにミックスダウンを変化させるために、同楽器種による演奏を複数トラックに分割することが多い。

つまり、音楽構造と音色や奏法の特徴をとらえて、対応する音楽構造に対してミックスダウンデザインを転

ヘアメイクの例では、ショートレイヤー、ウルフモヒカン、ボウズ、アフロなどが典型的なデザインとして上げられる。この場合、クライアントの楽曲の楽器編成の類似性を考慮する必要はない。

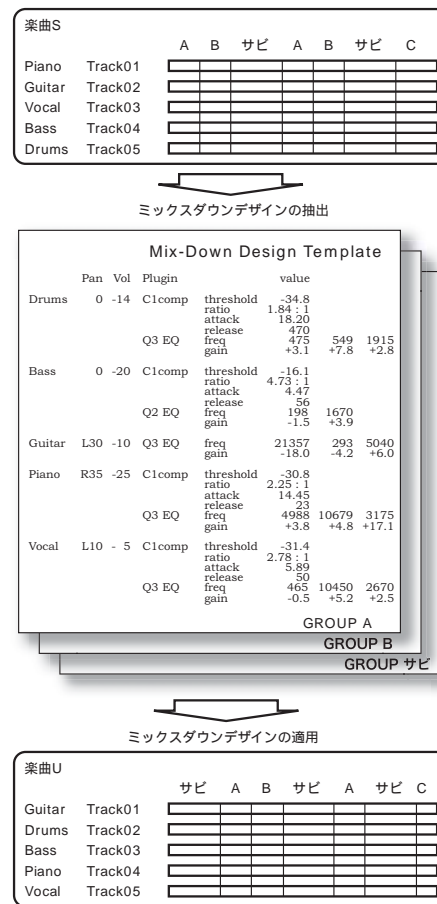


図4 ミックスダウンデザインの抽出と適用

写する必要がある。

本研究では、各トラックの楽音データに対して、音色・奏法、音楽構造の解析を行なう。ここで得られた情報は、アノテーションの他、最も似ている編成(構成)の曲を提示するという用途(ケース2)において利用される。以下、音楽構造の解析、音色・奏法の解析について述べる。

### 3. 音楽構造の解析、音色・奏法の解析

各トラックの音素材データの入力に対し、音楽構造の解析、音色・奏法の解析までの処理を以下のような手順で実施する。

- (1) 全トラックの音素材データの入力
- (2) ボーカル、ドラムス系音色、その他の音色の3つのカテゴリーへの分類
- (3) ドラムス系音色と判断されたトラックからのビート単位抽出
- (4) ボーカルトラックに対してのビート単位を利用した音楽構造解析

## (5) 残りのトラックに対する音色・奏法の解析

以下、音楽構造の解析、音色・奏法の解析について述べる。

### 3.1 音楽構造の解析

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)<sup>6)7)8)</sup> 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。

本研究では、主としてポップス楽曲を対象とする。ポップス楽曲には、ボーカル等、メインになるメロディラインが存在する。トラックの中からメロディライン候補を抽出し、その繰り返し構造から A メロ、B メロ、サビなどの音楽構造を判定することができる。本研究が扱う楽曲データは既にパートごとに分離されている。また、ドラムパートが抽出できれば、そこからビート情報を得ることも可能でありため、問題設定としては、後藤の RefraiD よりもシンプルなものとなる。そこで本研究では、まずドラム系音色のトラックを特定し、ビート単位を抽出する。その後、メロディパートを特定し、その基本周波数推移に対して相関分析を施し、繰り返し構造を抽出することによって音楽構造を抽出する。

#### 3.1.1 ビート単位の抽出

ドラム系音色(スネアドラム、バスドラム、タムタム、ハイハットシンバル、ライドシンバル、クラッシュシンバル)とそれ以外の楽音の比較を行ったところ、前者特有の特徴として、音の立ち上がりから 10ms 以内に 1300Hz 以上の周波数成分にパワーが集中することが確認された。そこで、アタックから 10ms 内の 1300Hz 以上のパワー値と周波数を掛け合わせた値の総和をドラム系音色判断の指標とし、その指標が閾値以上のものをドラム系音色と判断することにした。

ドラムスがマルチトラックで収録されていることも考慮し、スペクトル包絡の類似性、音の持続時間、楽曲全体における発音回数により、具体的な打楽器音名を判断し、そのトラックに対し、楽器名およびドラム系楽器という情報を付与する。ドラム系楽器の IOI (Inter Onset Intervals) を追跡することで、ビート単位を算定する。

#### 3.1.2 ボーカルパートの特定

ポップスのメインメロディパートは、一人のボーカリストの歌唱として収録されることがほとんどであり、基本的に音素材データに和音が含まれることはな

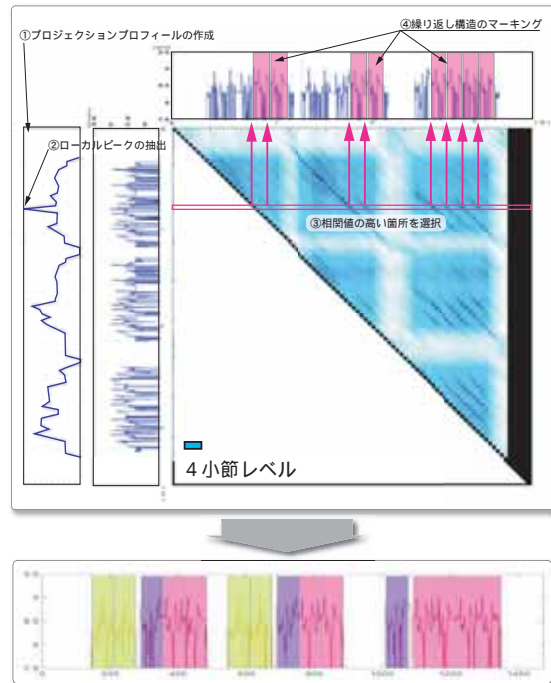


図 5 音楽構造の抽出

い。よって、音素材データの単音の含有率を、ボーカルパートを特定する指標の 1 つとして利用する。

自己相関関数に基づくピッチ抽出手法は、単音の音素材を対象とする場合には正しい結果を与える可能性が高いが、和音を対象とする場合、推定値があばれてしまうという性質がある。一方、パワースペクトルに対して、ある周波数に対し、その整数倍の周波数成分のパワー成分をその足し込んでいくという操作を行うことにより、音素材データに含まれる基本周波数のパワー成分を突出させることができる。

本研究では、上記の 2 つの手法によりピッチを抽出した結果を比較し、その一致率を見ることで単音か和音かを判定することにした。さらに、ベース音など単音で演奏されやすい音色との区別を行うため、音域や、ピッチ抽出結果の揺れに関しても考慮することにした。

#### 3.1.3 メロディラインの繰り返し構造の解析

上記、ボーカルパートの推定の結果、この時点ではボーカルパートのピッチ推定は完了している。ボーカル中のピッチ遷移に対して、ビート単位毎にずらして自己相関値を計算することによって、音楽の繰り返し構造を分析する。この様子を図 5 に示す。この図は、

ピッチ抽出結果が、40Hz~900Hz に入っているかを見る。

RWC 研究用音楽データベース<sup>5)</sup>RWC-MDB-P-2001 No.13 のボーカルパートの音素材データ<sup>2)</sup>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。得られた繰り返し構造のうち、最も繰り返し回数が多いものをサビと判断する<sup>6)7)8)</sup>。

### 3.2 音色・奏法の解析

#### 3.2.1 セグメンテーション

本研究で取り扱う音色・奏法の特徴量には、アタックポイントを基準としたものが多く、正確にアタックポイントを同定する必要がある。ポップス楽曲は、通常、ピアノ、オルガン、ディストーションギター等の楽器によって構成される。ギターのピッキング音やピアノの打鍵音などは、音の立ち上がりにおける非正常成分が問題となり、アタックポイントの同定は容易ではない。そこで、本研究では以下のような手法によりアタックポイントの同定を行なう。

- (1) 移動平均により平滑化したパワー包絡近似線において、直前のパワーとの差が閾値以上であるポイントをアタックポイント候補とする。
- (2) パワースペクトルから高調波成分を除去した非倍音成分のパワーを足し合わせ、その時系列データを作成する。閾値以上の値を持つピークをアタックポイント候補に追加する。なお、ディストーションギターなどのノイズ性の高い音色に対しては、過度に候補を採用してしまう可能性があるため、閾値を高めにとる。
- (3) リリース直前の候補をアタックポイント候補から外す(経験的処理)。

#### 3.2.2 音色の解析

楽器音の音源同定を目的とした研究に北原らの手法<sup>9)</sup>がある。北原らの研究では、ディストーションギターなどのノイズ性の高い音色は対象としていない。本研究では、ノイズ性の高い音色についても扱う必要があるため、楽器音響学の知見<sup>10)11)12)</sup>や楽器音の音

源同定に関する研究<sup>9)</sup>を参考に、以下に示す79個の特徴量を抽出することにした。

#### (1) スペクトルに関する特徴量

- 1: 周波数重心(各高調波成分のパワー値を重みとする周波数の重み付き平均を求める。)
- 2: 全高調波成分のパワー値の合計に対する基音成分のパワー値の割合
- 3 ~ 30: 全高調波成分のパワー値の合計に対する基音から*i*次までの高調波成分のパワー値の合計の割合( $i=2,3,\dots,29$ )
- 31: 奇数次の高調波成分(基音含む)と偶数次の高調波成分とのパワー値の合計の比
- 32 ~ 40: 音が鳴り続けている時間(周波数成分全体のパワーがしきい値を超えている時間)に対して、その高調波成分の鳴り続けている時間(パワー値が同じしきい値を超えている時間)が*p*%である高調波成分の個数( $p=10,20,\dots,90$ )

#### (2) パワーの時間変化に関する特徴量

- 41: パワー包絡線の線形最小二乗法による近似直線の傾き
- 42 ~ 58: 発音開始直後*t*秒間のパワー包絡線の微分係数の中央値( $t=0.15,0.20,\dots,0.95$ )
- 59 ~ 75: 最大パワー値と、発音開始から*t*秒後のときのパワー値の比( $t=0.15,0.20,\dots,0.95$ )

#### (3) ノイズ性に関する特徴量

- 76 ~ 79: アタックポイントから一定区間に含まれる非倍音成分のパワーの総和(平均値, 標準偏差, 最小値, 最大値)

#### 3.2.3 奏法の解析

本研究では、奏法の特徴量として、以下の17個の特徴量を抽出する。

- 1: 発音回数  
抽出されたアタックポイントの総数を発音回数と判断する。
- 2 ~ 5: 発音間隔  
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。各アタックポイントの間隔の平均値, 標準偏差, 最小値, 最大値を求める。
- 6 ~ 9: 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。各アタックポイントのパワー値の平均値, 標準偏差, 最小値, 最大値を求める。

- 10～13：メリハリの付け具合  
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。具体的には、アタックポイントからパワー包絡線が閾値以下に達するまでに要した時間を捉える。閾値以下に達する前に、次のアタックポイントが現れた場合は、線形最小二乗法による近似直線を求め、閾値以下に達するまでの時間を予測する。得られた時間の平均値、標準偏差、最小値、最大値を求める。

- 14～17：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。含まれる音高群の最も低い音高の基本周波数を解析し、その平均値、標準偏差、最小値、最大値を求める。

以上の特徴量を1小節ごとに算出し、その平均値をアノートしていく。

### 3.3 実験と検討

#### 3.3.1 ボーカルパートの推定

楽曲データ3曲に対して、ボーカルパートの特定に関する本手法の認識精度を調べた。ボーカルパートの抽出に対しては、94%の認識精度が得られた。この際の6%の誤認識は、コーラスパートに対しての誤りであった。メインボーカルパートの抽出という観点からは、100%の認識精度が得られている判断される。一方、非ボーカルパートの非ボーカルパートとしての除外率は92%であった。誤認識のほとんどは、ボーカルの音域内で演奏されていたギターのソロパートであった。これらは、ボーカルの歌唱のない前奏や間奏、後奏において演奏されたものである。誤認識ではあるが、構造解析処理において、不利な方向に働く誤りではない。なお本研究では、通常、ボーカルパートの歌唱のない区間を、前奏・間奏・後奏として判断している。

#### 3.3.2 音楽構造の解析

RWC 研究用音楽データベース (ポピュラー音楽)<sup>5)</sup>RWC-MDB-P-2001 No.13, No.18, No.45, No.64, No.82, No.88, No.98 の楽曲データに対して、提案手法により音楽構造の解析を行なったところ、100%の認識率が得られた。今後は、他ジャンルへの対応も行っていく予定である。

## 4. インタフェースのユーザビリティ評価

アマチュアバンドによるミックスダウンデザイン転写インタフェースの利用を想定し、予備的なユーザビ

リティ評価を行なった。被験者は、オリジナル作品を作り、ライブ活動を行なっているアマチュアバンドのメンバー4人である。各被験者のプロフィールを表1に示す。

表1 被験者プロフィール

|       | 演奏パート      | 作曲経験 | Mixdown 経験 |
|-------|------------|------|------------|
| 被験者 A | ギター (15 年) | 10 曲 | 10 曲       |
| 被験者 B | ベース (7 年)  | 3 曲  | なし         |
| 被験者 C | ドラム (1 年)  | 2 曲  | なし         |
| 被験者 D | ボーカル (5 年) | 10 曲 | なし         |

評価実験データとしては、転写元の楽曲の6パターンのミックスダウン事例と、各デザインを別の楽曲へ転写した複数のミックスダウン結果を用意した<sup>2)</sup>。自作の楽曲と他者の制作した楽曲とで、評価に違いが生じるかを判断するため、転写先の楽曲には、被験者らのバンドのオリジナル楽曲2曲とRWC 研究用音楽データベース<sup>5)</sup>RWC-MDB-P-2001 No.95 の計3曲を使用した。尚、転写元と転写先でのトラック数が大きく異なるため、そのマッピングの仕方に複数の可能性が存在する。また、エフェクタの中には、そのパラメータを楽曲のビートタイムや調性に合わせて調整する必要がある可能性を持ったものもある。その評価のため、一曲につき、平均25個のミックスダウンパターンを生成し、聴取実験およびインタビューを実施した。以下、実験内容と結果、考察について述べる。

### 実験 1

まず、転写元の楽曲の6パターンのミックスダウン結果を十分に聞いてもらい、それぞれの聞こえの特徴について、感じたことを自由形式で書き出してもらった。次に、転写先の楽曲の6パターンのミックスダウン結果をランダムに提示し、どのデザインを転写したミックスダウン結果であると思うかを尋ねた。

その結果、各被験者の正解率は、被験者 A は 77%、被験者 B は 83%、被験者 C は 38%、被験者 D は 61%となった。被験者 C 以外は、全パートに関してまんべんなく特徴を書き出したが、被験者 C は、ドラムパートの聞こえや配置に関することしか記述していなかった。ドラムにのみ注意が集中していたことが正解率の低下につながった可能性がある。

### 実験 2

転写元と転写先のミックスダウン結果の対応を被験者に示したうえで、転写先のミックスダウン結果に関

本研究では、1つの楽曲に対して聞こえの異なる6パターンのミックスダウンを実施してもらうよう、プロのミキシングエンジニアに依頼した。

して以下の項目に関してそれぞれ 10 段階で評価してもらった。

- ミックスダウン結果としての完成度
- 転写元の聞こえとの類似度
- ミックスダウン結果の聞こえに対する好感度

まず、ミックスダウンの完成度については、必ずしも満足のいくミックスダウン結果が得られていないことが分かった。その要因に関するインタビュー事項としては、「雑音が残っている」「音像定位と音量のバランスが悪い」「楽曲のイメージに合っていない」などのコメントがあった。満足のいく完成度を得るためには、単純なデザインの転写だけでは不十分であり、人手による最終的な作り込みが不可欠であるといえる。

転写元の聞こえとの類似度については、転写元と転写先の楽曲の楽器編成が類似している場合に評価が上がり、楽器編成の異なる場合に評価が下がる傾向があった。このことから、類似した楽曲の検索機能の重要性がわかる。また、楽器編成が異なっても、聞こえの変化が顕著にわかるエフェクト処理が施されている場合は、聞こえの類似度が上がることが確認された。

最後に、ミックスダウン結果の好みについては、どの楽曲に対しても 1 つのデザインを好む被験者、楽曲ごとに異なるデザインを好む被験者、いずれのデザインも気に入らなかった被験者など、個人差が大きい状況が確認された。

#### 実験 3

ある 1 つのデザインに関して、ギターパートのトラック対応付けの異なる 6 パターンのミックスダウン結果を用意し、どの対応付けが最適かについて尋ねた。

結果、被験者 B と C からは、「どの対応付けでも特に問題ない」というコメントが得られた。一方、被験者 A と D からは、「どの対応付けも一長一短であり、どれが最適であるかは一概にはいえない」というコメントが得られた。このことから、トラックの対応付けについては、必ずしも正解はなさそうである。インタフェースとしては、システムが数パターンのトラックの対応付け案を提示し、最終的にはユーザが決定するといったタイプのものが良さそうである。

#### 実験 4

エフェクタのパラメータの中には、楽曲のビートタイムや調性に合わせて設定する必要性のあるものがある。パラメータを単純転写すると、転写元の楽曲のビートタイムや調性に設定されてしまう。このことが問題となるか確かめるため、ビートタイムと調性に関して、転写元の楽曲に合わせたミックスダウン結果と転写先の楽曲に合わせたミックスダウン結果を被験者

に提示し評価を求めた。ビートタイムを転写元のパラメータに合わせた場合、ドラム演奏者である被験者 C は違和感を感じたが、残りの被験者には特に違和感を感じないという意見を述べた。調性を転写元に合わせた場合に関しては、音程のズレを不快に思う被験者が多かったが、逆にそのエフェクト効果を面白いとする指摘もあった。

#### 課題

今回の実験において使用したミックスダウン転写元のデザインはマルチ収録であったのに対し、転写先事例はいわゆる一発取りとして収録されたものであった。前者のトラック数が 50 を越えていたのに対し、後者が 8 であった。また、後者のトラックには、他の音源の音が混ざっていたり、収録時のレベルでベース音の音質にも問題があった。これらの状況は、ユーザビリティ評価を実施する上では、最も極端なケースであったと考えられる。今後、より精緻な検証を行うため、収録条件の近い対象を用いて実験を実施するとともに、アマチュアおよびプロの双方の視点でのユーザビリティ評価を進めていく必要があると考えている。

## 5. 応 用

「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大きな特徴である。

本研究で提案したミックスダウン支援インタフェースの利用形態としては、次のようなものを想定している。ひとつは、ミキシングエンジニアやアマチュアの制作者による利用である。本来膨大にかかるミックスダウン時間を短縮したり、発想の幅を広げる上で、ミックスダウンデザインテンプレートの利用は有用である。今回のユーザビリティ評価では、その効果を裏付けることはできなかったが、過去の自己のミックスダウン事例等、収録プロセスが似通ったものであれば、ここでの提案がより機能するものと期待している。

もうひとつは、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、という利用形態である。そのイメージとしては、グラフィックイコライザを高機能化させたようなものである。ミックスダウン経験のない一般の被験者 20 人を対象として、異なるミックスダウンデザインが適用された楽曲を聴取したときにデザインの違いを判断できるかを調査したところ、被験者全員がその差を識別することがで

きた。また、被験者ごとに、各自が聴きやすく好ましいと感じるミックスダウン結果も異なっていた。以上の結果は、ミックスダウンにおけるデザイン転写がアミューズメントの一つとして成立することを示唆している。ブロードバンド時代に突入し、楽曲データがトラック別に収録された形で流通することもそう遠くはないだろう。この利用形態の実現可能性は高いと考えている。

以上のような可能性の一方で、事例を利用することによって生じる課題もある。事例参照型のシステムにおいては、ユーザのイメージに適うミックスダウン事例が、データベース上に登録されている必要がある。このことは著作権にも抵触する。根本的に問題解決をはかるためには、音楽ビジネスモデル自体を変えていく必要があるだろう。着メロや着うたがこれほどまでに定着していることや、プリインストール型の mp3 プレイヤーが大ヒットしていることからすれば、ミックスダウンデザイン自体を商品（作品）として流通させるビジネスは十分に成立すると考えている。

## 6. おわりに

本論文では、既存事例からミックスダウンのデザインを転写することによって、そのデザインを支援する方法論を提案し、その方策、技術的な準備、基礎的評価、可能性について論じてきた。ミックスダウンのデザイ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音色と奏法について解析する必要がある。本稿では、インタフェースに対する考え方を示した上で、音楽構造、音色・奏法の情報の解析処理を示した。また、予備的なユーザビリティ評価について紹介し、ミックスダウンデザインにおける事例利用の可能性について論じた。近年、主成分分析、ICA、SOM などを用いて、音楽ジャンルの判別<sup>13)</sup><sup>14)</sup> や楽器音の判別<sup>15)</sup> に関する研究が注目を集めている。今後は、これらの手法を応用し、楽器種・奏法の自動クラスタリングを実現していきたい。また、さらなるユーザビリティ評価を実施していく予定である。

謝辞 本研究は、科学技術振興機構さきがけ研究 21 「協調と制御」領域の支援を受け実施されました。また、実験等に御協力いただいた川崎善之氏、中嶋啓人氏に感謝いたします。

## 参 考 文 献

- 1) Digidesign : <http://www.digidesign.com/>
- 2) <http://ist.ksc.kwansei.ac.jp/~katayose/members/yatsui/mixdown/>

- 3) 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).
- 4) 谷井章夫, 片寄晴弘: Mixdown 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, 2004 年度人工知能全国大会 (第 18 回) 論文集, 1E3-05 (2004).
- 5) 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).
- 6) 後藤 真孝: "リアルタイム音楽情景記述システム: サビ区間検出手法", 情報処理学会 音楽情報科学研究会 研究報告 2002-MUS-47-6, pp.27-34, (2002).
- 7) 後藤 真孝: "SmartMusicKIOSK: サビ出し機能付き音楽試聴機", 情報処理学会論文誌, Vol.44, No.11, pp.2737-2747 (2003).
- 8) Masataka Goto: A Chorus-Section Detecting Method for Musical Audio Signals, Proceedings of ICASSP 2003, pp.V-437-440 (2003).
- 9) 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).
- 10) 山口公典, 安藤繁雄: "短時間スペクトル分析法の自然楽器音への適用", 音響誌, 33,6,pp.291-300,1997 .
- 11) 早坂寿雄: "楽器の科学", 電子情報通信学会, 1992 .
- 12) 安藤由典: "楽器の音響学", 音楽之友社, 1996 .
- 13) Roberto Basili, Alfredo Serafuni, Armando Stellato: Classification of Musical Genre: A Machine Learning Approach, 5th International Conference on Music Information Retrieval-Proceedings of ISMR2004, pp.505-508, (2004).
- 14) Cory McKay, Ichiro Fujinaga: Automatic Genre Classification Using Large High-Level Musical Feature Sets, 5th International Conference on Music Information Retrieval-Proceedings of ISMR2004, pp.525-530, (2004).
- 15) Mauricio A. Loureiro, Hugo B. de Paula, Hani C. Yehia: Timbre Classification of a Single Musical Instrument, 5th International Conference on Music Information Retrieval-Proceedings of ISMR2004, pp.546-549, (2004).