

# ミックスダウンの視覚化による支援システムの開発

北谷光浩<sup>†1</sup> 上松大輝<sup>†2</sup> 根上生也<sup>†1</sup>

**概要:** 視覚的に捉えながらミックスダウン作業を行うための、Mix-Down Visual Controller の提案と実装を行った。ミックスダウンは音楽制作において、聴きやすい音源を作るために非常に重要な工程であるが、適切に行うには高い技術や熟練した耳が必要であるため、従来は経験を積んだエンジニアが行う専門性の高い作業であった。近年、機材・ソフトウェアの低価格化や、作品発表の場が増えたことで、アマチュアの制作者がミックスダウンを行うケースが非常に多くなっている。しかし、既存のミックスダウン環境は熟練者が行うことを想定して作られているため、経験の浅いミックスダウン初心者にとっては敷居が高く、簡単に行えるものではない。本論文では、聴覚に依存した作業であるミックスダウンを、視覚化することでより分かりやすく全体像を把握しながら行うシステムについて、背景と構成を述べる。

## Mix-Down System with Visualization

MITSUHIRO KITADANI<sup>†1</sup> HIROKI UEMATSU<sup>†2</sup>  
SEIYA NEGAMI<sup>†1</sup>

**Abstract:** In this paper, we propose a system "Mix-Down Visual Controller". In music creation, the mix-down is an important part for it affects whether the song will be popular or not. However, it is hard for beginners to do mix-down. So our system makes mix-down easy by visualization in VR. Our system consists of parameter converter, visual interface, NUI(Natural User Interface)and OSC(Open Sound Control). We show what "Mix-Down Visual Controller" is and why it is needed.

### 1. はじめに

近年、音楽機材・ソフトウェアの低価格化や、インターネットで手軽に作品を公開できるようになったことで、多くの初心者が音楽を始めるきっかけになるほど、気軽にDTMを楽しめるようになってきている。特に最近ではYouTube やニコニコ動画のように気軽に投稿できるサービスが普及したことや、「初音ミク」に代表される VOCALOID がブームになったこともあり、自作の楽曲を公に発表する機会が非常に増えた。その結果、数多くの音楽作品がインターネット上に溢れることになったが、例えばメロディやアレンジは素晴らしいものの音のバランスが悪いため聴きづらくなってしまっているような作品も少なくない。

音楽制作の工程は大きく分けて、メロディラインの構築や、楽器や音色の選択を行う「作曲・編曲」と、複数の楽器同士のバランスを取る「ミックスダウン」に分かれる。作曲・編曲に比べて、ミックスダウンは初心者にとって抵抗がある。そのため、各楽器の定位がセンターに集中し過ぎて広がりなくなってしまう、音のバランスが悪くメインのパート（例えばボーカル）が埋もれてしまっている、といったようなケースに初心者は陥りやすい。ミックスダウンは本来、プロのエンジニアが熟練された耳を頼りに行う作業であるため、経験の浅い初心者が耳だけを頼りに行うと、どのパラメータを操作すればどのような変

化をもたらすのかが分かりづらく、適切な調整が行えない。また、ミックスダウンで扱うパラメータの多さも、初心者にとってはミックスダウンを分かりづらくする要因である。

本研究では、このような問題を解決するために Mix-Down Visual Controller システムを提案する。Mix-Down Visual Controller は、音のバランスを聴覚だけでなく視覚的にも表現することで、より全体像を把握しながらのミックスダウンを可能にすることを目的としている。また、複数あるパラメータを厳選することでパラメータの多さによる煩雑さを軽減することも目的としている。本稿では Mix-Down Visual Controller のインターフェースとパラメータの割り当て、システム構成について述べる。

### 2. 要件定義

初心者のミックスダウン支援に関連する研究としては谷井章夫らの「事例を利用したミックスダウン支援インターフェース」が挙げられる。前章でも述べたように、ミックスダウンの作業は初心者には難しいため、初心者がプロのエンジニアにミックスダウンを依頼する際に、事例を用いる事でより正確なイメージを伝えるための試みも行われている。谷井らによると、プロのエンジニアがミックスダウンした事例を初心者とエンジニアの間での意思疎通に用いる事で、より正確に楽曲の仕上げるイメージを伝達することが可能だと述べられている。事例を用いることで、次の3つのケースに対応したミックスダウン支援インターフェースを実現できるとしている。

i) クライアントが持ち込んだ好みの事例を楽曲に適用する

<sup>†1</sup> 横浜国立大学  
Yokohama National University  
<sup>†2</sup> 株式会社木星ラヂヲ  
Jupiter Radio Inc.

ii)数ある事例の中から類似度で検索を行いクライアントに提案を行う

iii)典型的な事例を提案しクライアントの好みのものを探す

谷井らの研究のように事例を用いる必要があるのは、初心者がミックスダウンのイメージを具体的なパラメータに変換する技術を持ち合わせていないからだと推測される。そのため、初心者がパラメータの操作を行ってイメージ通りに音のバランスを作り上げることは難しく、システムによってミックスダウンのイメージをパラメータに変換する補助を行う必要がある。

そこで、本研究ではパラメータを視覚的に表現することで問題を解決する。ミックスダウンの状態を視覚的に把握するためには、次の要件を考慮する必要がある。

i)視覚化した結果と音のイメージが一致する

ii)全パートの状態が、一つの画面内に表現されている

i は視覚化において非常に重要である。例えば、ギターパートが画面上で右寄りにあったとき、スピーカーから発せられるギターの音が左から聴こえてきたら、視覚情報による正しい把握を行う事ができない。また、ミックスダウンは個々のパートに注目して行うものではなく、全体を見通して行うものであるため、ii のようなインターフェースである必要がある。

本要件を満たすシステムである、Mix-Down Visual Controller について以下の章で述べる。

### 3. Mix-Down Visual Controller の構成

#### 3.1 インターフェース

Mix-Down Visual Controller は、パート毎に録音された歌や楽器などの音声に対応したオブジェクトを、VR 空間上で操作してミックスダウンを行うシステムである。

ミックスダウンで用いられる一般的なインターフェースは図 1 のようなものである。①の領域ではパート毎の基本的なパラメータを、②の領域ではトラックにインサートされている個々のエフェクトのパラメータをコントロールする。実際の PA 卓を元にデザインされたこのようなインターフェースでは、耳でバランスを聴きながらツマミやフェーダーを操作してパラメータを調整することでミックスダウンを行う。プロのエンジニアのように熟練した耳の持ち主であれば、視覚的な情報がなくとも耳で聴くだけで適切なミックスダウンが行えているかどうかの判断が可能であるが、初心者はその判断を行うことが難しい。そのため、DTM 関連書籍の初心者向けに書かれたミックスダウン講座等では、各パートのバランスを分かりやすく示すために図 2 のような配置図が使われることが多い。

Mix-Down Visual Controller では、図 3 のようなミックスダウンインターフェースを作成した。3次元空間上に各パート間のバランスを立体イメージで表現し、その立体イメ

ージの位置や大きさを操作することでミックスダウンを行う。どのパートをどの位置に配置するかを図 2 のようにイメージすると、音質や音量の変化を考える際の助けとなり、例えば「バランスが悪く聴こえるのはここに隙間があるからだ」といったようなことを視覚的に把握しやすくなる。そのため、「位置を右の方に配置すると定位が右寄りになる」「体積を大きくするほど Volume も大きくなる」といった具合に各パートの配置を動かすことで音の調整を行えば、ミックスダウンの状態を常に視覚的に把握できるため、非常に分かりやすくミックスダウンが行えると考えた。



図 1

Figure 1

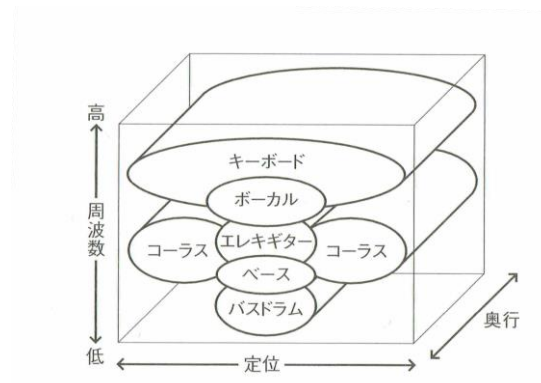


図 2

Figure 2

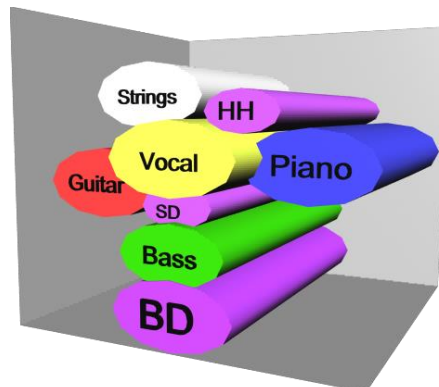


図 3

Figure 3

### 3.2 パラメータの割り当て

本節では、VR 空間とミックスダウンのパラメータとの紐付けについて述べる。パラメータはよく使われるものだけでも 1 パートにつき十数項目、場合によっては数十項目ものパラメータがあるため、全てを割り当てようとする VR 空間上の数値が足りなくなる。そこで、図 2 を参考に必要最低限のパラメータに絞り、VR 空間上の数値に割り当てる。パラメータを絞って簡略化することで、ミックスダウンに不慣れな初心者が混乱しにくくなることが期待できる。

今回作成した Mix-Down Visual Controller では、Pan と Equalizer の周波数帯域、Reverb、Volume をパラメータとして使用する。以下に、各々の項目について割り当てられる VR 空間上の数値と合わせて述べる。また項目 i, ii, iii におけるパラメータの割り当てを図 4 に示す。

#### i) Pan

Pan は音が聴こえる位置を意味し、定位と訳される。今回は 2ch のステレオミックスを対象とし、Right と Left の音量バランスを調整するパラメータとして扱う。オブジェクトの x 座標が+方向に移動すると Pan は R 方向に、-方向に移動すると L 方向に増加するように割り当てた。また、Pan は標準的な MIDI の定義に基づき 0~127 の値をとるものとする (0:最左翼,64:中心,127:最右翼)。

#### ii) Equalizer で増幅される周波数

Equalizer は音声の特定の周波数の増幅や、減衰を行う音響装置である。今回は操作の簡略化のために増幅のみを扱う。オブジェクトの y 座標が+方向に移動すると増幅する周波数が高く変化し、-方向に移動すると低く変化するように割り当てた。また、増幅する周波数の範囲はミックスダウンで処理される事の多い 50hz~12khz とする。

#### iii) Reverb

Reverb は入力された原音に対して残響音を付加する音響装置であり、奥行き感を演出することができる。一般的な Reverb は複数のパラメータを持つことが多いが、本システムでは簡略化のために原音と残響音の音量比のみを調整対象とし、その他のパラメータは一定とする。z 座標が+方向に移動すると原音の割合が増加し、-方向に移動すると残響音の割合が増加するように割り当てた。

#### iv) Volume

Volume は各パートの音量である。オブジェクトの大きさに比例して音量も大小するように割り当てた。物体が大きいほど音も大きくなるというのは感覚的に結びつき易いため、感覚的な調整を行う事ができる。また Volume は 0~100 の値を取るものとする。

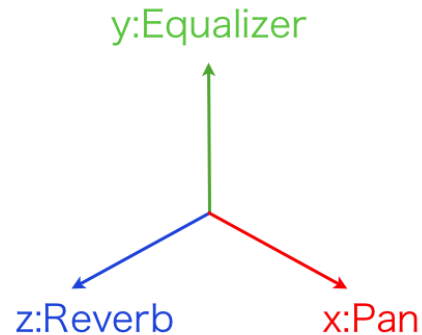


図 4  
Figure 4

## 4. システムの実装

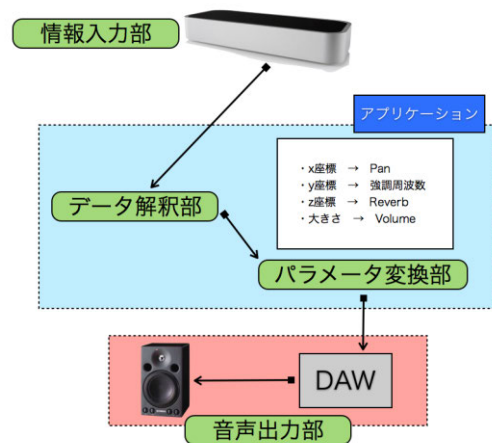


図 5  
Figure 5

Mix-Down Visual Controller のシステム構成を図 5 に示す。本システムは、データ入力部・データ解釈部・パラメータ変換部・音声出力部から構成される。

データ入力部では、VR 空間に表現された 3 次元空間上のオブジェクトを直感的に操作するために、Leap Motion を用いて手や指の動きを取得する。Leap Motion は認識範囲にある 2 本の手と、手に付随する 5 本の指を認識し、それらの位置・角度・ジェスチャ等のデータを取得する。本システムでは、Leap Motion で取得可能な以下の 3 つのデータを使用する。

i) 検出されている手と指の数。

ii) 1 番目に検出された手とその手の指の Position (Position は手や指の座標情報を値として返し、xyz の値を持つ)。

iii) 2 番目に検出された手の sphereRadius (sphereRadius は手の屈曲に適当な球体の半径を値として返す)。

なお、Leap Motion では、左右の手、及び各指が実際の手や指と連動して認識されず、デバイス内で認識された順に

ID が割り当てられるため、本稿では便宜的に 1 番目に検出された手を「手 A」、2 番目に検出された手を「手 B」として述べる。

データ解釈部では、データ入力部から送られた i, ii, iii のデータを解釈して、VR 空間上でのオブジェクトの位置・大きさに変換する。オブジェクトの位置は手 A の Position に、オブジェクトの大きさは手 B の sphereRadius に対応し、以下の 3 つの操作により変換が行われる。

1 つ目は、手 A の親指と人差し指の Position がほぼ重なり、かつその手において検出されている指の数が 1 である時 (Leap Motion は重なった 2 本の指を 1 本として認識する)、つまり図 6 のように親指と人差し指の先でつまむようなポーズを取った時、選択を行う。選択を行った時、親指と人差し指が重なる座標にオブジェクトがある場合、そのオブジェクトは選択された状態になる。検出されている指の数が 2 本になった時、つまり図 7 のように親指と人差し指の先が離れた時、オブジェクトは選択された状態から解除される。

2 つ目は、オブジェクトが選択されている状態で、図 8 のように手 A の Position を変化させると、オブジェクトの位置は 3 次元直交座標系の x,y,z の座標軸で -64~63 に収まる範囲内を移動する。

3 つ目は、同様にオブジェクトが選択されている状態で図 9 のように手 B の sphereRadius が変化すると、オブジェクトの大きさは 0~100 の値で変化する。

上記の 2 つ目、3 つ目の操作が行われた時、位置・大きさは画面に反映されると共にパラメータ変換部に送られる。

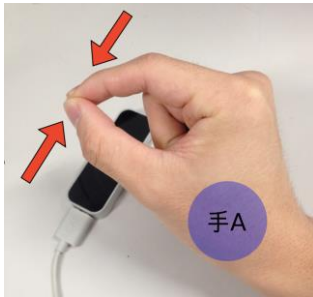


図 6  
Figure 6



図 7  
Figure 7

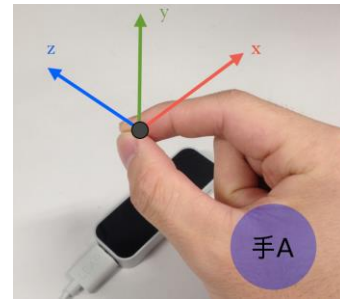


図 8  
Figure 8

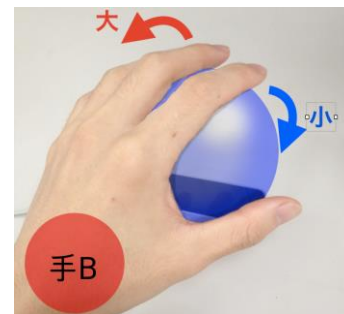


図 9  
Figure 9

パラメータ変換部では、データ解析部から送られたオブジェクトの位置・大きさ情報を 3.2. で述べたパラメータに変換する。3 章で述べたように、今回作成したシステムでは使用するパラメータを固定しているが、パラメータ単位で設定ファイルの記述を変更することで、変換処理の増減や、変換される閾値を変更可能としている。

音声出力部では、パラメータ変換部から送られてきたパラメータ情報を DAW を通して音に反映させる。音声出力部との通信方法として今回は実装の容易さから OSC を用いた。また DAW へ情報を送るには、電子楽器間の通信方法として用いられる MIDI を使用する必要があるため、OSC を MIDI に変換するソフトウェアを使用した。

## 5. おわりに

本稿では、視覚情報を用いて音のバランスを把握しながらミックスダウンを実現するシステムとして、Mix-Down Visual Controller を開発した。本システムは、ミックスダウン関連書籍で頻繁に用いられる配置図を参考に視覚化を行うことで、ミックスダウンにおける音のバランスを把握する難しさを軽減し、さらに直感的にパラメータを操作することを可能とした。

今後、様々なユーザーに Mix-Down Visual Controller と既存のミックスダウン環境の両方でミックスダウンを行ってもらい、音のバランスの把握のしやすさ等を比較するアン

ケートを実施し、ミックスダウンにおける視覚化の有用性を評価する。さらに、より分かりやすい視覚化の方法や、より直感的なパラメータの操作についても模索していく。

また、本システムでは視覚化を行うにあたってパラメータを絞ったために、簡易なミックスダウンしか行えず細かい調整を行うことができない。パラメータを位置や大きさだけではなく、色・輪郭の明瞭さ等の他の視覚的要素に割り当てることで、高度なミックスダウンにおいても視覚的な分かりやすさや操作の直感性を実現し、初心者から熟練者まで幅広いユーザーの使用に耐えうるツールとなることが可能であると考えられる。

## 参考文献

- 1) 谷井章夫, 後藤真孝, 片寄晴弘: 事例を利用したミックスダウン支援インターフェース. 情報処理学会インタラクション 2005. 2005
- 2) 杉山勇司: ”レコーディング／ミキシングの全知識”, リットーミュージック, 2004.
- 3) Leap Motion Developer : <<https://developer.leapmotion.com/>>, アクセス日 : 2013/11/24