



そこの君～！一生のお願いなんだけど 今から皆の前でこれ演奏してくんない!？ (圧)

Hey you! This is my last wish, but can you please play this in front of everyone right now?! (Pressure)

高宮駿¹⁾, 坂本峻平¹⁾, 瀬川湧斗¹⁾, 早水元気¹⁾

Shun TAKAMIYA, Shunpei SAKAMOTO, Yuto SEGAWA, Genki HAYAMIZU

1) 関西学院大学 工学部 知能・機械工学課程

(〒669-1330 兵庫県三田市学園上ヶ原 1, {hxp30390, iyh33531, gfg33527, ioe70007}@kwansei.ac.jp)

概要: 本企画では、唇の形状や呼気の強弱をリアルタイムにセンシングするマウスピース型デバイスを開発し、その情報を活用してトロンボーン演奏の VR 体験システムを構築する。このシステムは唇の形状、呼気の変化、スライドの正確な位置や動きといった繊細な身体感覚をセンシングし、インタラクティブな音響フィードバックと視覚フィードバックを行う。これにより、音楽経験の有無に関わらず、誰もが自然にトロンボーンの演奏感覚を獲得できる。単に楽器の模倣に留まらず、気軽に音楽表現の楽しさを体験できるエンタテインメントコンテンツの実現を目指す。

キーワード: マウスピース型デバイス VR 金管楽器 体験型音楽シミュレーション 画像情報処理

1. 企画概要

1.1 はじめに

あなたは学生時代、吹奏楽部の演奏を見て、「あの金管楽器を自分も吹いてみたい」と感じたことはないだろうか。特にトロンボーンは、スライドを前後させるという直感的な操作から、「なんとなく音が出せそうだ」と思わせる楽器かもしれないが、見た目のシンプルさとは裏腹に、音を出すために繊細な身体制御が求められる楽器である [1]。正確なスライド位置を感覚的に把握しながら、唇の形 (アンブシュア) や息の強さを繊細にコントロールする必要がある。そのため、未経験者にとっては最初の一步が極めて高いハードルになる。

近年、VR 技術を活用した楽器コンテンツは、ドラムやギターなどの視覚的・触覚的な操作が主体の楽器に偏重しており、口の動きや呼吸を伴う「管楽器」を題材としたものはほとんど存在しない。そこで本企画では、「口の動きをセンシングするデバイス」という新たなインターフェイスを開発し、金管楽器の演奏体験を VR 空間に構築する。

1.2 目的

本企画は、体験者のアンブシュアの形や呼気の強弱をリアルタイムにセンシングするマウスピース型デバイス、および手の動きを検知する物理スライド装置を開発し、これらの情報を統合してトロンボーン演奏の VR 体験を構

築することを目的とする。体験概要を図 1 に示す。

トロンボーンは鍵盤を押せば音が出るピアノや、穴をふさげば音程が決まるリコーダーとは異なり、直感が求められる楽器である。特に、呼気や唇の微妙な制御、そしてスライドの位置把握といった要素は、楽譜から直接読み取ることが難しく、演奏者の身体感覚に深く依存する。こうした不可視の動作をセンシング技術により可視化・可聴化することで、体験者が身体感覚を通じて自然に演奏感覚を獲得し、気軽に楽しめるエンタテインメントコンテンツの実現を目指す。

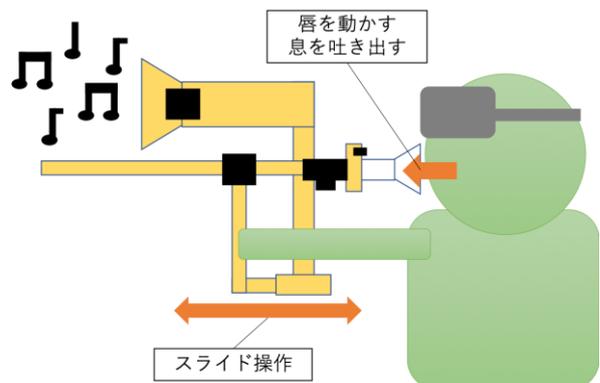


図 1 体験の概要図

2. システム構成

本企画のVRトロンボーン演奏体験システムは、主に三つの要素で構成されている。「口の動きをセンシングするデバイス」、「手の動きを検知するデバイス（トロンボーンのスライド動作）」、そして「視覚提示デバイス（HMD）」である。これらの要素はリアルタイムで連携し、体験者の動作を即座に視覚・音響情報としてフィードバックする。

2.1 口の動きをセンシングするデバイス

2.1.1 マウスピースデバイス

唇の形状を計測し、吐く空気の量をセンシングする機能を持つ。デバイスの概要を図2に示す。人間の口の平均的な大きさが縦20mm、横50mmであるため、デバイスはその口元に自然に収まるような形状とサイズで設計されている。このデバイスは、口に装着され、以下の二つの主要なセンシング要素を統合している。

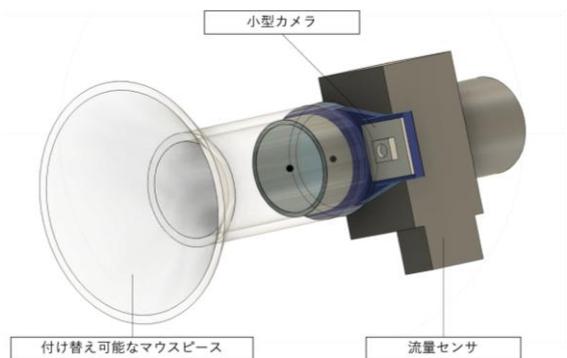


図2 口の動きをセンシングするデバイス

2.1.2 唇の形状検出

体験者の唇の動き（アンブシュア）をリアルタイムに捉える小型カメラと、それに最適化されたマウスピース設計および画像処理アルゴリズムによって実現する。マウスピース本体は、光造形方式3Dプリンタを用いて、透明度の高い樹脂で製作する。小型カメラをマウスピース型デバイスの外部に配置する。カメラより、マウスピースに適用された唇の形状、開口部の寸法、マウスピースに対する相対的な位置関係といったアンブシュアの変化をリアルタイムで撮影する。撮影された画像データは、画像処理システムへと送信される。

カメラをマウスピース外部に配置する設計において、唇の形状計測の際、視野の確保は重要な課題となる。特に、マウスピースの気道部分の構造自体がカメラの視野を遮り、唇の計測に影響を与える可能性がある。この課題に対し、以下の対策を講じる。

カメラは、体験者がマウスピースを構えた際に、唇の正面ではなく、左右いずれか一方（例：唇の右側70%の視野範囲）から唇を撮影する配置とする。この片側からの画像情報に基づき、画像処理によって以下の手順でアンブシュア形状を推定する（図3）。

1. 上唇と下唇の中心線検出

取得された画像から、唇の主要な構造である上唇と下唇の輪郭を抽出し、それぞれの中心線を特定する。

2. 唇の片側の形状解析

検出された中心線と取得された唇の画像情報を基に、開口部の形状、マウスピースとの接触面積などを計測する。

3. 左右対称性の仮定と全体形状の推定

金管楽器の演奏においては、安定した音を出すために一定以上の空気量を吐出する必要がある、その際のアンブシュアは、左右が大きくずれないという前提に立つ。この仮定に基づき、得られた唇の形状情報を利用して、唇全体の形状や開口部の大きさを推定する。これにより、限られた視野情報からでも、演奏に必要なアンブシュアの特徴量を効果的に把握することが可能となる。

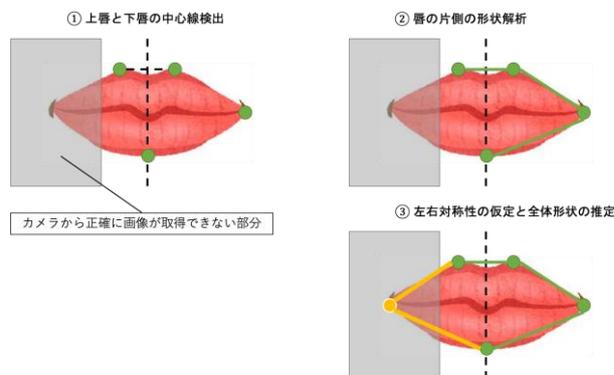


図3 部分画像からの唇の形状推定の流れ

2.2 手の動きをセンシングするデバイス

体験者のスライド動作を正確に検出するため、本システムでは以下のような専用の物理スライド装置を構築する。

1. リニアスライダ

摩擦抵抗を低減した設計のリニアスライダを用いる。これにより、継ぎ目のない滑らかな直線運動を再現する。体験者は抵抗なくスライダを前後させることができ、VR空間内の位置と物理的な手の感覚とのずれを最小限に抑える。

2. 赤外線距離センサ

スライダの現在位置を正確に計測し、VR空間におけるトロンボーンの高さを決定するためのデータとする。取得した位置データの時系列変化から、スライドの速度や加速度といった情報も算出する。これにより、スライドの動きの速さや勢いから生まれる音の変化も、VR空間上の音響表現に反映させることが可能となる。

2.3 視覚提示システム

視覚提示システムは、体験者に演奏状況とフィードバ

ックを提示する中心的な要素であり、HMDを用いてVR空間に構築される。HMD内には、以下の情報がリアルタイムで表示される。

1. 演奏ガイドとフィードバック

体験者が選択した楽曲の正しい音階やリズム情報が、視覚的なガイドとして提示される。これと同期して、トロンボーン型デバイスから取得される体験者の擬似的な演奏情報が、視覚的に重ねて表示される。これにより、体験者は自身の演奏と目標とする演奏との乖離をリアルタイムで把握し、直感的に修正することが可能となる。これは、インタラクティブな演奏支援として機能する。

2. 演奏スコア／達成度表示

演奏の正確さやリズム感をリアルタイムで評価し、スコアやグラフとして表示する。

2.4 システム構成図

本企画のシステム構成を図4に示す。

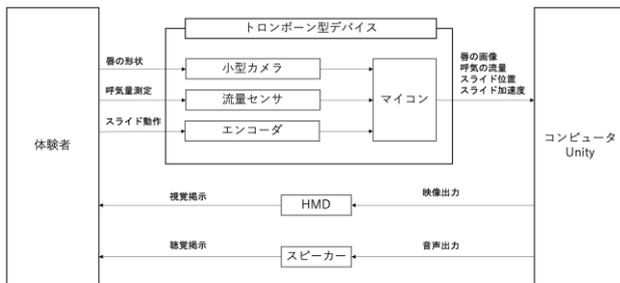


図4 システム構成

3. 体験の流れ

本システムのVRトロンボーン演奏体験は、以下のステップで構成されている。

1. 体験の準備

体験者はHMDを装着する。続いて、トロンボーン型デバイスのマウスピースを口につける。この段階で、デバイスが正しく装着され、システムに認識されているかを確認する簡易なキャリブレーションを行う。

2. 楽曲選択

VR空間には、様々なジャンルや難易度の楽曲がリストアップされた、インタラクティブな楽曲選択インターフェースが提示される。体験者は、自身のレベルや好みに応じて自由に楽曲を選択できる。

3. 演奏開始

楽曲選択後、システムは自動的に演奏モードに移行する。VR空間には、選択された楽曲の演奏ガイド（正しいスライド位置のインジケータ）がリアルタイムで表示される。体験者は、このガイドと同期しながら、マウスピース型デバイスへの呼気や唇の操作、そして物理スライド装置による手の動きを連動させ、バーチャルトロンボーンの演奏を開始する。システムは、体験者の演奏動作を即座に音響・

視覚情報としてフィードバックし、インタラクティブな学習体験を提供する。

4. 演奏終了

楽曲の終了、または体験者が任意のタイミングで演奏を停止すると、システムは演奏モードを終了し、結果表示画面に移行する。この画面では、演奏ガイドとの一致度に基づいた演奏スコアや達成度がグラフや数値で示される。

4. 音の高さの評価

バーチャルトロンボーンの音の高さを決定する際、本システムでは、スライド位置データとアンブシュア特徴量という二つの主要な入力情報を統合して処理する。これにより、トロンボーン演奏における音高決定の根幹を再現し、体験者の身体動作を正確な音として表現する。

4.1 スライド位置データによる基本振動数決定

物理スライド装置の距離センサから得られるスライダの絶対位置情報 $P(t)[\text{mm}]$ は、バーチャルトロンボーンの基本となる音の振動数 $F_{base}(t)[\text{Hz}]$ を決定する。トロンボーンの音高はスライドの位置に厳密に依存するため、スライド位置 $P(t)$ を基本振動数 $F_{base}(t)[\text{Hz}]$ にマッピングする関数 f_P を定義する。

$$F_{base}(t) = f_P(P(t)) \quad (1)$$

f_P は、各スライドポジションに対応する正確な音高に基づいて定義され、線形補間を用いて実装される。これにより、体験者がスライドを動かせば、それに対応する基本振動数がリアルタイムで生成される。

4.2 アンブシュア特徴量による振動数調整

2.1.1 節で示した唇の形状検出で得られる唇の開口部面積 $A(t)$ や縦横比 $R(t)$ 、唇とマウスピースの接触面積 $C(t)$ といったアンブシュア特徴量から、擬似的な唇の振動数 $F_{lip}(t)$ を算出する。これは唇の物理的な振動を直接検出する代わりに、その形状変化から間接的に振動を推測する本システム独自の試みである。

$$F_{lip}(t) = k_A \cdot A(t) + k_R \cdot R(t) + k_C \cdot C(t) + C_{base} \quad (2)$$

ここで、 k_A 、 k_R 、 k_C は面積、縦横比、接触面積に対する係数であり、 C_{base} は基本となる振動数オフセットである。これらの係数は、実験的に収集したアンブシュアデータに基づいて、音高の変化を再現するよう最適化された定数である。

4.3 最終的な音の振動数決定

最終的なバーチャルトロンボーンの音の振動数 $F_{final}(t)$ は、スライド位置から決定される基本振動数 $F_{base}(t)$ と、アンブシュア特徴量から算出される擬似的な唇の振動数 $F_{lip}(t)$ を統合して決定される。

$$F_{final}(t) = F_{base}(t) + \delta_F(t) F_{lip}(t) \quad (3)$$

ここで、 δ_F はアンブシュアによる振動数変化が最終的な音の振動数に与える影響度を調整する係数である。この統合により、スライドによる大まかな音高決定に加え、アンブシュアによる繊細なピッチコントロールの表現を可能にし、より豊かな演奏体験を実現する。

5. おわりに

本企画は、VR 技術を用いて楽器演奏の物理的ハードル

を低減し、未経験者でも気軽にトロンボーン演奏の奥深さに触れることを可能にする。この体験を通じて、単なるエンタテインメントに留まらず、音楽表現の新たな可能性を探求する場を提供する。

参考文献

- [1] ヤマハ株式会社. 楽器解体全書. 取得元:
https://www.yamaha.com/ja/musical_instrument_guide/trombone/play/, (参照 2025-05-23)