

歌唱時に高音の発声ができるようになる VR 映像の提案と評価

坂名純太* 土田修平* 寺田 努* 塚本昌彦*

概要. 歌唱には、音程やリズムなど様々な要素が関わっており、発声もその一つである。良い発声を行うためには、音程に合った姿勢を保ちつつ、特定の筋肉に力を入れてその他の筋肉はリラックスさせることが必要となる。しかし、発声する音の種類によって適切な姿勢は変わるため、姿勢への意識が強くなると発声への意識が薄れてしまい効率が下がる恐れがある。そのため、専門知識をもたない人にとって歌唱時の発声練習は難しく、これまでの歌唱指導では「額から声を出すように」などのイメージで歌い方を伝えるような工夫が行われている。そこで本研究では、視覚刺激によって発声者の姿勢やメンタル状態を誘導すれば効果的な発声練習を実現できると考え、VR(Virtual Reality)映像を用いた視覚刺激を活用し、歌唱時の発声を暗黙的に支援するシステムの構築を試みる。発声者の高音の発声と発声時の姿勢に注目し、高音の発声がしやすくなると考えられる映像を作成し、映像の内容によってユーザの高音の発声のしやすさ及び発声時の姿勢に生じる変化について調査を行った。

1 はじめに

歌唱には、音程やリズムなど様々な要素が関わっており、発声の仕方その一つである。良い発声は、声に柔軟性を与え、声帯を効率的に動かすことで少ない力で大きな声を発することができ、歌唱中の喉の負担を軽減できる。フースラーら [1] は、声楽的な発声法について生理学的・解剖学的知見を取り入れ研究し、歌唱中の人体の構造を細部にわたって説明している。その中で、発声器官の能力を十分に発揮するためには、発声器官のすべてが的確な神経支配を受け、またその器官の筋肉が敏速で伸縮自在な張力を最大限に獲得する必要があると述べている。このことから、正しい発声には特定の筋肉に力を入れて、その他の筋肉はリラックスしていることが同時に求められるなどいくつかの発声に関わる器官を適切に動かすことが求められる。

しかし、発声の際に重要な筋肉は、普段の生活で使う機会が少ない箇所が多く、単に意識するだけでは正しく筋肉を使いこなすのは困難であり、専門的知識のない人には正しく使えているかの判断は難しい。そのため、発声に重要な筋肉を正しく使うことを補助する目的で、特定の姿勢や特定のイメージを持って練習する手法がよく用いられている。例えば高音を出すためには「額から声を出すイメージで」や「自分が階段を降りているようなイメージで」といった説明がされることがある。ただし、求められる姿勢やイメージは発声パターンごとに異なるため、歌唱時に姿勢やイメージをその都度切り変える必要があり、発声者への負荷が大きい。また、発声者自身の頭の中でリアルなイメージを想起しながらの練習では、発声への意識が薄れてしまい効率が下がる

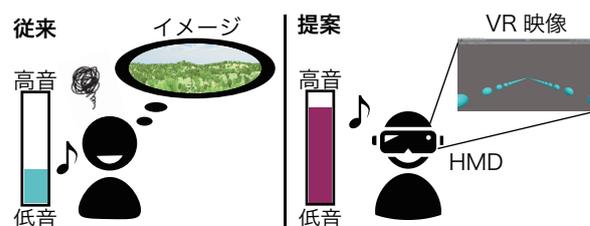


図 1: システム概念

恐れがある。

そこで本研究では、発声者が意識せずとも発声に重要な筋肉を正しく使うことを促せれば、より効果的な発声練習を実現できると考え、VR(Virtual Reality)映像によって視覚刺激を与えることで姿勢等に影響を与えられる [13] ことに着目し、VR映像を見せることで意識下でユーザの発声時における身体機能に影響を与え、良い発声を促す手法を提案する (図 1)。本稿では、発声練習時に見る VR映像によって重心位置および発声することのできる音程の上限に差異があるかについて調査した。このような映像効果と歌唱能力の関係が明らかになると、高音が出やすいカラオケ背景が作成できるなどさまざまな応用が可能になると考えられる。

2 関連研究

2.1 声楽的な観点からの発声研究

声楽的な観点から見て、歌唱時には発声器官の各役割を理解し、その多くをコントロールすることが求められる。石野 [2] によれば、地声や裏声といった声区は筋肉や呼吸法によって調整でき、音量、母音等によっても声区は変動すると述べている。歌唱

* 神戸大学大学院工学研究科

者は声区を意識しなくなることが理想であるが、そのためにも筋肉のメカニズムを知ることが必須であると述べている。また、理想的な発声は身体全体の筋肉の自然な連携を前提としており、その際に身体の様々な部位は無意識的に連動するとされ、意識した身体動作ではぎこちない発声となってしまふ恐れがあると石野は述べている。しかし、現在行われている発声のための姿勢や各器官を適切に動かす練習は手本を見たり聞いたりしてそれを真似るといった手法が用いられており、練習者は意識的に身体動作を行わなければならない。野口 [3] は意識と無意識という観点から、良い発声に大切なのは発声に関わる筋肉が良好に運動することであるが、発声者からすればそのような客観的な視点を意識すればするほど、筋肉は自然な運動性を失っていくと述べている。つまり正しい発声には、身体の発声に関わる諸器官が無意識のうちに適切に機能することが必要である。本研究では、無意識のうちに発声に関わる器官が正しく機能することを実現するための手法について提案する。

2.2 発声支援手法

歌唱時の発声支援システムとして、発声者の音声特徴量をフィードバックする方法が多く用いられている。斎藤ら [4] は、歌声に聴感的な響きを与える要因と考えられている 2.4~3.2kHz 付近の周波数帯域に着目し、歌唱指導前後でその周波数帯域の歌唱データがどう変化しているか比較することで習熟度を検討できることを示唆した。また、片岡ら [5] は強弱のある発声について、声楽経験の異なる被験者から音声データを取得し分析することで、習熟度をフィードバックするシステムを提案している。さらに、中野ら [6] は音声の音高を軌跡として可視化することで歌唱への理解を深めさせ、過去のデータやお手本となる歌唱データをゴースト表示することで練習効率を向上させる手法を提案している。

しかし、これらの手法はその時点での発声者の歌声を評価するものであり、直接的に発声練習の効率を上げるようなシステムではない。また、これらのデータをフィードバックしたとしても、有効に利用するためには専門的な知識が必要となり、練習者はこれらのデータから得たアドバイスを意識しながらの発声を行わなければならない。発声そのものへの意識が薄れてしまう恐れがある。本研究では、声楽的な知識を必要とせずに歌唱力の向上が可能となるシステムの構築を目的とする。

2.3 視覚情報が与える影響

VR 映像以外による視覚刺激

杉田ら [7] は拡張現実技術を用いた周辺視野が人間の重心知覚に与える影響を調べた。この実験では、被験者に実物のアタッシュケースを持たせ、その位

置に合わせて拡張現実を用いてバーチャルなアタッシュケースを重ねて表示し、形状・大きさを変化させることで重心がどの位置にあると感じられるか調査を行った。その結果、実際のアタッシュケースに比べ取っ手の位置がずれている表示をした場合、ユーザはずらした方向に重心が傾いていると感じる傾向が見られた。岡野ら [8] が行った実験では、LED を用いた周辺視野ディスプレイを被験者に装着させ、ルームランナー上でランニングをさせたところ、LED の光がランナーに向ってくるよう点滅させた際の体感速度は被験者の走行時の速度よりも速くなる傾向が見られた。また、吉川ら [9] は断面が半円状のレンズを無数に並べられたシートを用いた歩行誘導の実験を行った。レンズに少しずつ絵柄をずらした画像を格納することで、レンズに対して視線を任意方向に動かすアニメーションを呈示でき、被験者は絵柄をずらした方向に誘導されて歩行した。

VR 映像による視覚刺激が与える影響

VR 技術を用いて、身体機能の向上や特定の運動機能の促進を狙うシステム構築に関する研究は多数行われている。小貫ら [10] は生体センシングを用いた VR ゲームの利用が高齢者の運動能力にどう影響するか調査した。その結果、VR ゲームを用いることで被験者の歩行能力やバランス能力が有意に向上し、アンケート調査でも満足度が高かった。遠藤ら [11] は VR による歩行映像、モニターによる歩行映像、静止立位をランダムで被験者に提示した後に歩行させ、その際の重心動揺を計測したところ、前後方向の重心の移動について、VR 映像とモニター映像を提示した時の移動量は静止時の移動量に比べ有意に大きかった。石井ら [12] は HMD の情報処理を用いて人間の焦点に影響を与え、ユーザの歩行時の進行方向を誘導することを可能にした。このように VR 技術を用いて運動機能を向上させることや姿勢や重心に影響を与えることが可能であり、本研究ではこれを歌唱時の発声に応用することで練習効率上がるのではないかと考えた。

3 提案システム

我々は、ユーザが良い発声のために意識すべきことを無意識のうちに行わせることで、自然で良い発声が可能になると考えた。そこで本研究では、VR 映像による視覚刺激により身体や知覚を制御し、重心誘導を含めた意識すべきことを意識下で行わせ、ユーザの発声を支援する手法を提案する。提案手法では、被験者に VR 映像を提示可能な頭部装着型ディスプレイ (HMD: Head Mounted Display) を装着させ、作成した VR 映像を見せながら歌唱を行わせる。本論文で提示する映像は主に移動感覚を引き起こすようなものを用いた。没入感のある移動感覚提示により、高音の発声に適しているとされるつま先

重心(前方重心)の姿勢を無意識に促すことを意図している。つまり、本研究では良い発声の例として「高い音を発声できること」、意図する映像視聴効果を「移動感覚の付与による体勢制御」とした。VR 映像の作成には Unity を使用し、VR 提示可能な HMD には HTC 社の Vive(視野角: 110°, 重量: 640g, 以下 HMD) を用いた。

4 評価実験

本実験では、VR 映像を見ながら発声練習を行う際に、高音の発声と姿勢にあらわれる影響について調査する。

4.1 映像の選定

発声に適した姿勢は、発する声の音程や強弱などによってわずかに異なる。高音を発声する際はつま先に体重をかける姿勢が適しているとされる [1]。この知見をもとに、高音の発声に適した状態になるよう促す映像を調査するために 4 種類の映像を用意した。具体的には、高音の発声に適しているとされるつま先重心、前傾姿勢をとるよう促す効果が期待できる映像として、相対的にユーザが上または前方に移動しているかのような感覚を与える映像と、その逆の感覚を与える映像を計 4 種類用意した。具体的には、視界いっぱいに配置された球形オブジェクトが鉛直下向きにスクロールされ、相対的にユーザが上昇しているかのような感覚を覚えることを狙った映像(下方向)、球形オブジェクトが鉛直上向きにスクロールされ、相対的にユーザが下降しているかのような感覚を覚えることを狙った映像(上方向)、ユーザの両脇に球形オブジェクトが線上に配置されており、それらがユーザの前方から後方へと向かってくるように移動することで、相対的にユーザが前進している感覚を覚えることを狙った映像(後方向)、後方向の映像とは逆方向にオブジェクトが前方に移動し、ユーザが後退している感覚を覚えることを狙った映像(前方向)である。ここで、作成した映像の種類を表 1 に、上下方向のイメージ画像を図 2 に、前後方向のイメージ画像を図 3 に示す。なお、球形オブジェクトは半径 50cm の大きさに見えるように設定し、イメージ画像中の矢印は球形オブジェクトの移動方向を示している。また、上方向と下方向の映像では被験者とスクロールする球形オブジェクトとの距離は 11m となっており、後方向と前方向ではユーザの左右に配置された球形オブジェクトとの距離は 5m となっている。これはユーザがオブジェクトの接近により回避行動をとるような状況が起きないよう設定されている。

4.2 計測項目

映像が表示されてから各パートが終了するまでの間、Wii Balance Board[15] を用いて被験者の足圧

表 1: 作成した映像

映像名	説明と想定される効果
上方向	オブジェクトが上方向に移動 後傾姿勢となり高音が出なくなる
下方向	オブジェクトが下方向に移動 前傾姿勢となり高音が出せる
後方向	オブジェクトが向かってくる 前傾姿勢となり高音が出せる
前方向	オブジェクトが遠ざかってゆく 後傾姿勢となり高音が出なくなる

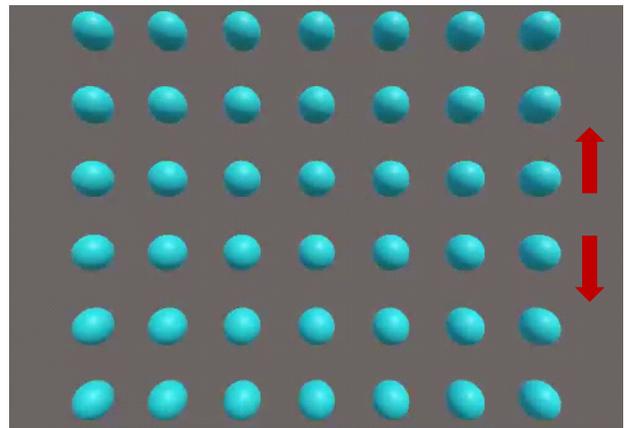


図 2: 上下方向のイメージ画像

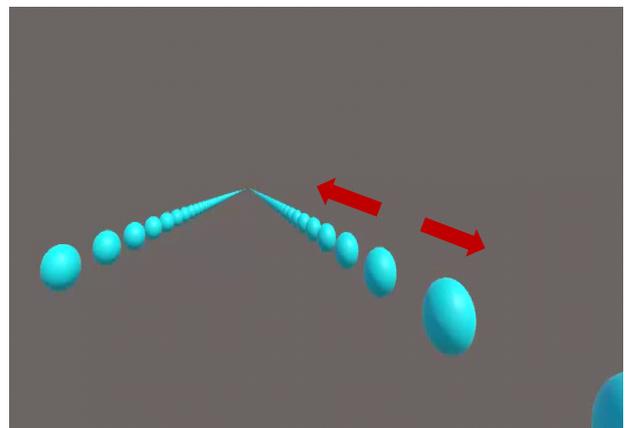


図 3: 前後方向のイメージ画像

中心(COP: Center of Pressure)を 100fps で取得する。COP は静止立位姿勢の重心動揺の測定において、重心位置として代替可能である [14]。Wii Balance Board を用いた足圧中心の測定には、公開されている Wii Balance Board Stabilometry System Ver 2.1[16] を使用する。

被験者に装着した無線ピンマイク (SYNCO-G1 2.4 GHz ワイヤレスオーディオ)[17] から、実験中

に被験者が発声した音声を WAV ファイル形式で保存する。その後、被験者の音声データから F0(基本周波数) を求め、被験者が地声で発声することのできた音程を音声データと F0 から判定する。また、実験中の被験者の様子をビデオカメラで撮影する。計測用ソフトウェアは Python を用いて作成した。

4.3 実験手順

計測時のシステム構成を図 4 に示す。被験者は実験環境に入室した後、発声練習に用いるピアノ音源を聞く。その後、Wii Balance Board の上に乗り、HMD とイヤホン、ピンマイクを装着する。HMD は PC に接続されており、被験者の動作のトラッキングのために HMD の位置を計測するためのベースステーションを 1 台被験者の正面、2.5m の距離に設置する。また Wii Balance Board 上では、被験者にとって自然な直立姿勢を取ってもらう。被験者に取り付けたデバイスと実験をしている様子を図 5 に示す。

被験者は HMD を額に載せて映像を見ない状態と、オブジェクトが移動する 4 種類の映像を見ながらの発声練習をランダムな順番で行う。なお、発声練習にはドミソミドの音を 1 セットとしたピアノの音源を使用し、音程は mid2C(262Hz) から始める。被験者はピアノの音源を 1 セット聞いた後に、同じ音程のドミソミドを発声する。ピアノの音程は段階的に高くなり、被験者が明らかに発声出来ない、もしくは裏声を用いなければ発声できない音階まで達した場合はそのパートの発声練習は終了とする。この判断は本論文の第一著者が行った。発声練習 1 回あたりにかかる時間は約 3 分であり、各発声練習の間には 5 分間の休憩を取る。被験者は 20 代男性 9 名(音楽経験、VR 経験なし)である。最後に、口頭もしくはチャットツールを用いてシステムに対するコメントを回答してもらった。

また、今回作成した映像を用いた発声練習に関して、高音の発声に与える影響がその後の本システムを用いない発声においても影響を与えるかどうかを調査するため、1 回目の実験から 1 週間後の同時刻に被験者 1 人当たり 3 回、同じ内容の発声練習をデバイス装着していない状態で行わせた。なお、この実験は神戸大学大学院工学研究科の研究倫理審査委員会の承認を得て行った(受付番号: 03 - 12)。

4.4 結果と考察

全被験者における、練習パターンごとの足圧中心の平均座標(左右方向、前後方向)を箱ひげ図で表したものをそれぞれ図 6, 7 に示す。これは Wii Balance Board の中心を 0 とした時の値であり、単位はいずれも mm である。また、左右方向では右が正、左が負の値であり、前後方向では前が正、後ろが負の値となっている。左右方向、前後方法におい

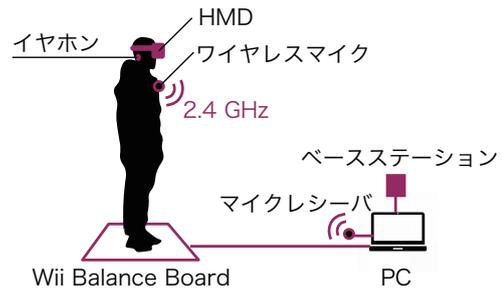


図 4: 計測時のシステム構成

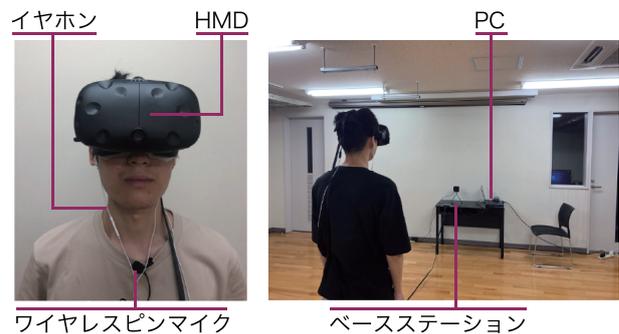


図 5: 実験環境

てそれぞれ一要因分散分析を行ったところ、提示する映像間に有意な差は見られなかった ($p > .10$)。左右方向について、映像なしと前方向での発声練習では被験者間で足圧中心のばらつきが小さく、下方向で特にばらつきが大きい結果となった。また、前後方向については、後方向の映像を用いた発声練習中の足圧中心のばらつきが小さく、残りの 4 条件ではばらつきがみられた。この値は、実験における体勢の変化を計測するために行ったものであるが、実験中は VR 空間を把握するため首を振る被験者が多く見られ、その際の重心変化が大きかったため今回の実験では姿勢の違いを表すような値が計測できなかった可能性がある。

また、地声で発声することのできた音声の最大周波数に対して、提示映像ごとに全被験者の平均値を求めた結果を図 8 に示す。なお、被験者の発声した地声と裏声の判定は、音楽経験者に音声データを聞いてもらい、主観的に判断した。提示映像について一要因分散分析を行った結果、有意差が見られた ($F(4, 32) = 5.02, p < .05$)。また Holm 法による多重比較を行なったところ、後方向と映像なし、後方向と下方向、後方向と前方向においてそれぞれ有意差が見られた ($p < .05$)。このことから、オブジェクトが前方からこちらへ向かってくる映像を見ながらの発声では、仮説の通り被験者が高音の発声に適した姿勢となり、より高い音の発声が出来るように

歌唱時に高音の発声ができるようになる VR 映像の提案と評価

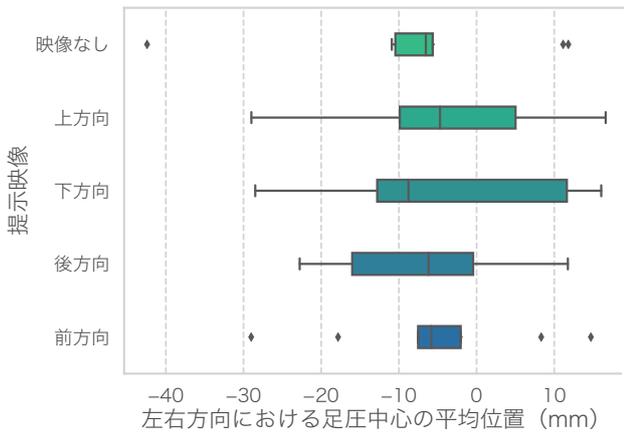


図 6: 各練習パターンの足圧中心の平均位置 (左右方向)

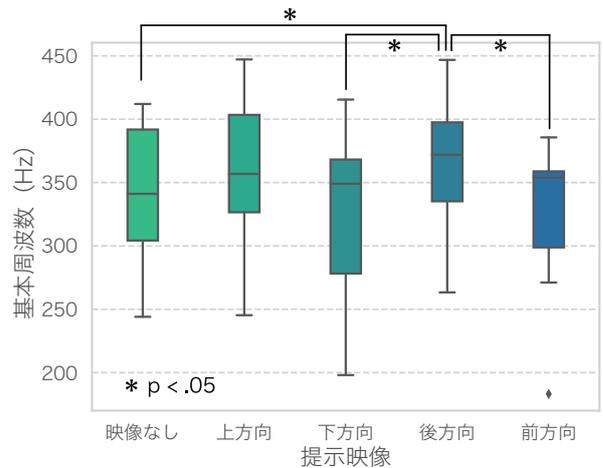


図 8: 各練習パターンにおける基本周波数の平均値

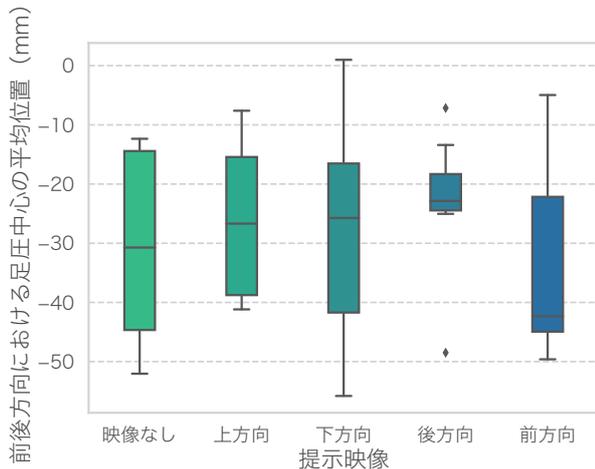


図 7: 各練習パターンの足圧中心の平均位置 (前後方向)

なった可能性が示唆された。また前方向の映像よりも後方向の映像で高音の発声できていたことから、高音の発声にはユーザが前方へ移動している感覚を覚える映像が好ましいと考えられる。以上の結果から、VR 映像を見せることにより高音の発声を促すことができるという可能性が示唆されたが、今回用いた手法では、正しい発声に関わる要因のうち重心移動以外に影響を与える効果が存在する可能性が高いと考えられる。今後は他の要因に焦点を当て、提案手法により発声が可能となる原理を解明することが課題である。

4.5 映像を用いた練習効果の持続性

VR 映像を用いた発声練習の効果の持続性を評価するための実験を追加で行った。VR 映像を見なが

らの発声練習を行ってから 1 週間後、同被験者には HMD を装着せずに、同じ内容の発声練習を 3 回行ってもらう。音声データから最も高い周波数を取得し VR 映像を見ながらの発声練習での最高周波数と比較した。各被験者ごとの、1 回目の実験と 2 回目の実験での最高基本周波数を図 9 に示す。2 回目の発声練習にて、1 回目の発声練習よりも高い音を発声できた、もしくは 1 回目の発声練習時と同程度の音程が発声できた被験者は 4 名 (被験者 E,F,G,I) であった。長期的な効果が期待できると判断できた被験者数が全被験者数の半数に満たないことから、本システムは繰り返し練習を行うことで長期的な効果が見込める、または映像を見ているときにのみ効果が現れるものであると考えられる。

5 まとめ

本稿では、高音の発声に適した映像見せることで、発声時に望ましい身体状態を無意識のうちに行わせるシステムの設計とその評価を行った。今回の実験では、高音に着目し低重心やつま先に重心を持ってゆく姿勢を誘導する効果が期待できる映像と、逆の効果が想定できる物を用意し、それぞれ映像なしの発声練習と比べどのような違いがあるのか調査した。その結果、オブジェクトが前方からこちらへ向かってくる映像を見ながらの発声では、より高音の発声ができるようになったという結果が得られた。また、本システムを用いた練習効果の持続性については、

今後の課題として、今回用いた映像よりも重心の移動に影響を与え発声を支援する映像の作成と、実際の歌唱に本システムを導入し、音程に合わせて適切な効果のある映像を切り替えて表示させることが挙げられる。そのためには、提案手法を用いた練習によって発声の質が向上した場合に、発声に関わるどの要因に影響を与えたのか解明することも重要で

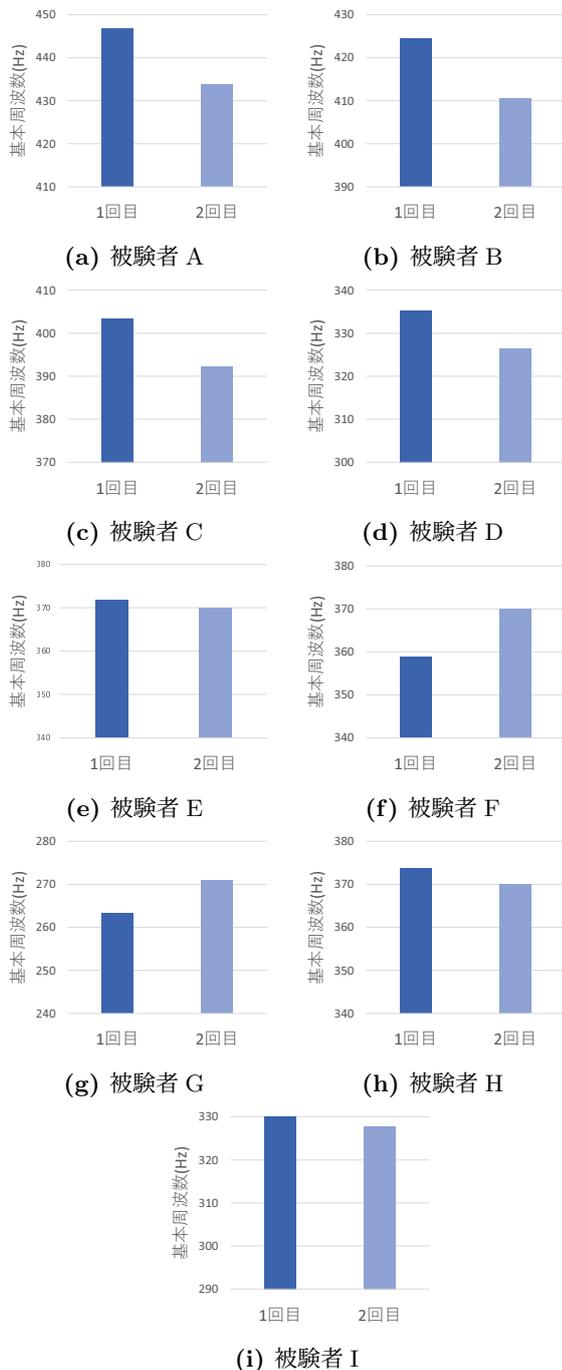


図 9: 各被験者が発声出来た音程上限の比較

ある。また、1人の被験者から上下方向のオブジェクト移動により浮遊感や下降感覚を覚え少し気分が悪くなったという感想があった。この被験者は軽度の高所恐怖症を患っており、これらの感覚に抵抗を感じたと考えられる。2章であげたLEDによる視覚刺激に比べ、VR映像による手法ではVR酔いや、ある状況や感覚に不快感や恐怖感を覚えるユーザーを考

慮した映像の作成も課題としてあげられる。さらに、実験後に被験者から得た感想としては、「オブジェクトに合わせて視線が移動してしまった」といったものがあり、作成した映像が狙っている重心移動の妨げになっていたということが考えられる。この対策としては、VR空間内にポインターを表示させ、被験者の視線を一点に固定するなどの処置を行って、自然な重心移動を促すことが有効であると考えられる。加えて、さらに多様な映像、条件を用いてさらに高音が発声しやすくなる映像や、高音以外の歌唱の要素を向上させる映像に必要な要因を調査する。また、複数の映像の順番と効果の関係を調べることや、各映像での長期的な練習による上達の程度に違いがあるのかも明らかにし、最終的にはカラオケにおける歌唱時の映像などにこれらの映像を組み込み、歌唱者にとって最も良い歌唱が可能となる環境の構築を目指す。

謝辞

本研究の一部は、JST CREST (JPMJCR18A3)の支援によるものである。ここに記して謝意を表す。

参考文献

- [1] フレデリック・フースラー, イブオンヌ・ロッド＝マーリング: うたうこと:発声器官の肉体的特質, 音楽之友社 (Jun. 1987).
- [2] 石野健二: 声楽発声のメカニズム, 宇都宮大学教育学部紀要. 第1部, No. 1, pp. 105–119 (Mar. 2015).
- [3] 野口雅史: 意識と無意識からみた歌唱の発声法, 新潟青陵大学短期大学部研究報告, Vol. 41, No. 41, pp. 151–161 (May 2011).
- [4] 齋藤 毅, 後藤真孝: 歌唱指導による歌声中の音響特徴の変化:歌唱ホルマントとF0動的変動に着目した音響分析, 日本音響学会講演論文集, pp. 457–458 (Sep. 2008).
- [5] 片岡靖景, 伊藤一典, 池田 操: 歌唱支援システム構築のための歌声の分析と評価, 情報処理学会研究報告音楽情報科学 (MUS), Vol. 1998, No. 74, pp. 23–30 (Aug. 1998).
- [6] 中野 倫靖, 後藤真孝, 平賀 讓: MiruSinger: 歌を「歌って/聴いて/描いて」見る歌唱力向上支援インタフェース, 情報処理学会インタラクティブ2007論文集, Vol. 2007, pp. 195–196 (Aug. 2007).
- [7] 杉田明弘, 溝口晃太, 木村朝子: 複合現実型視覚刺激が重心知覚に与える影響, 情報処理学会全国大会講演論文集, Vol. 71, No. 4, pp. 115–116 (Mar. 2009).
- [8] 岡野 裕, 雑賀慶彦, 橋本悠希: 速度感覚増強のための周辺視野への刺激提示手法の検討, 情報処理学会研究報告ヒューマンコンピュータインタラクション (HCI), Vol. 2008-HCI-127, No. 11, pp. 145–150 (Feb. 2008).

- [9] 吉川博美, 蜂須拓, 福島政期: ベクシオン場による歩行誘導手法の提案, 情報処理学会インタラクシオン, Vol. 2011, pp. 457-460 (Mar. 2011).
- [10] 小貫睦巳, 有田元英, 井上悦治: 生体センシング技術を使った仮想現実によるゲームが高齢者の運動機能に及ぼす影響について, 理学療法科学, Vol. 30, No. 6, pp. 811-815 (June 2015).
- [11] 遠藤佳章, 鈴木暁, 糸数昌史: ヘッドマウントディスプレイを用いた Virtual Reality 映像とモニター映像が立位重心動揺に及ぼす影響の違い, 理学療法科学, Vol. 33, No. 3, pp. 457-460 (Jan. 2018).
- [12] A. Ishii, I. Suzuki, S. Sakamoto and K. Kanai: Optical Marionette: Graphical Manipulation of Human's Walking Direction, Proc. of the 29th Annual Symposium on User Interface Software and Technology, pp. 705-716 (Oct. 2016).
- [13] 市川真澄, 渡邊 悟: 直立姿勢に対する視覚情報の影響 ((特集) 姿勢制御と感覚情報), バイオメカニズム学会誌. Vol. 15, No. 2, pp. 59-64 (Feb. 1991).
- [14] 藤原勝夫, 池上晴夫: 足圧中心位置と立位姿勢の安定性との関係について, 体育学研究. Vol. 26, No. 2, pp. 137-147 (Aug. 1981).
- [15] Wii Balance Board, <https://www.nintendo.co.jp/wii/rfnj/balance/index.html>.
- [16] バランス Wii ボードの重心動揺計ソフトウェア, https://researchmap.jp/blogs/blog_entries/view/112495/5a9c28ee606c50172546b83b5ea9939a?frame_id=513522.
- [17] SYNCO-G1(A1)-2.4 GHz ワイヤレスオーディオ, <https://www.syncoaudio.com/products/sy-g1-a1-wair>.