

マルチモーダルな拡張現実による音感の拡張

2022 年度 栗原一貴研究室
津田塾大学 理学研究科 情報科学専攻
MM2181 岡 あゆみ

目次

概要	3
1. はじめに	3
2. 絶対音感保持者に対する主観的な音の聞こえ方の調査	5
2.1 調査方法.....	5
2.2 音名の認識方法に関する調査結果	6
2.3 音名を認識できる数に関する調査結果.....	8
3. 絶対音感保持者への定性的調査.....	9
3.1 絶対音感保持者の欲しい音感や機能.....	9
4. 提案システム	10
4.1 <i>Otonona</i>	10
4.1.1 Otonona 設計指針.....	10
4.1.2 Otonona 概要	11
4.1.3 ユーザーインターフェース	11
4.1.4 Otonona システム構成.....	12
4.1.5 音の音名変換.....	13
4.1.6 ソーシャルメディアシェア機能.....	13
4.1.7 骨伝導ヘッドホンと密閉型ヘッドホン	14
4.2 <i>Otonona</i> の予備実験.....	15
4.2.1 実験方法	15
4.2.2 実験結果	16
4.2.3 予備実験から得られた知見	16
4.3 <i>Otonosa</i>	17
4.3.1 Otonosa 概要	17
4.3.2 音程の提示.....	17
5. 課題と展望	18
5.1 音感拡張の展望.....	18
5.2 <i>Otonona</i> の課題と展望	19

5.3 Otonosa の課題と展望.....	20
謝辞	20
参考文献	21

概要

絶対音感や相対音感に代表される音感は、音楽の認識や理解に影響を与える。本研究は、コンピュータを用いて人間の音感をマルチモーダルに拡張することを目的とし、音に対する新たな発見や体験を支援することを目指す。本論文では特に絶対音感に着目し、相対音感にも言及する。絶対音感は一部の人が6歳頃までに身につけることができると言われている能力である。程度に差はあるものの、この能力を持つ人は一般に、基準音を与えられなくとも音の名前を判断することができると言われている。これは音楽の理解に役に立つ可能性のある能力だが、大人になってからの訓練で身に付けることは難しい。一方、相対音感の有用性についても以前から注目されており、絶対音感保持者の中にも相対音感の必要性を語る者もいる。本論文では、絶対音感保持者に対して行った主観的な音の聞こえ方の調査結果と定性的見解を紹介し、マルチモーダルな拡張現実を用いて絶対音感と相対音感の有用性にアクセスするためのシステムの設計と実装、今後の展望を示す。

1. はじめに

絶対音感や相対音感に代表される音感は、音楽の認識や理解に影響を与える。絶対音感は聴き取った音の高さの認識に関するもので、個人差はあるものの、この能力を持つ人は一般的に基準音を与えられずとも聞いた音の音名を判定することができる。音の高さを詳細に把握することは音楽の体験や理解を深める可能性があるため、この能力が音楽活動において有利に働くと考える音楽家もいる。例えば最相氏の著書である“絶対音感”[1]で“作曲家の場合は、絶対音感があると便利だと誰もが口を揃える”と述べられている。また同じ本の中で尺八奏者である中村明一氏も“複雑化する現代音楽に対応するためには、邦楽奏者でも将来的には絶対音感を持っていた方がいいのではないかと感じている。”と述べる。現在、絶対音感に関する先行研究の多くはこの能力の獲得過程と、この能力を持つ人の認知に焦点が当たれている[2][3][4][5][6]。しかし榎原によると“絶対音感の発達には臨界期が存在し、6歳を超えると絶対音感習得が困難であることが指摘されている。”[7]とも言われている。

一方、相対音感についても以前からその有用性が語られてきた。曲の移調や移動ドへの対応には相対音感が有利に働くと考えられ、相対音感を習得するための書籍も存在する[8]。また、詳しくは第2章にて紹介するが、絶対音感保持者が絶対音感に加えて保持したい音感として相対音感を挙げることもある。

本稿では特別な訓練なしに人間の聴覚をマルチモーダルに拡張するためのシステムを提案する。絶対音感の有用性にアクセスするためのシステムをOtonona、相対音感の有用性にアクセスするためのシステムをOtonosaと名付けた。そしてこれらのフィードバックを

切り替え可能にすることで、複数の有用性にアクセスすることを目指す。

既存技術を紹介する。音を視覚で提示するものとして従来からの一般的なビジュアライザ、音名を視覚で提示するものとして一般的なチューナーが挙げられる。ビジュアライザは音や音楽を可視化するシステムであり、聴覚的な要素を視覚的な要素に変換して提示する。Apple 社の提供する iTunes アプリ(現在のミュージック.app)にもその機能が採用されていることから一般にも広く認知されている。図 1 はミュージック.app のビジュアライザ機能により音楽が視覚的に変換され提示される際の一例である。チューナーは主に楽器の音のピッチを基準に合わせる際に用いられるデバイスである。マイクから入力された音の音名と、基準となる Hz からどれくらいずれているかが可視化され提示される。図 2 はチューナーの一例である。音程を視覚で提示するものとしてはカラオケの画面に表示される音程バーが挙げられる。また、本来の音に編集を加えた後の音をユーザに聴覚的にフィードバックするシステムは Watanabe らのシステム[9]で実装されている。提案システムは音感に着目し音名をフィードバックするなどの機能面と、マルチモーダルな拡張現実を扱う点が既存システムと異なる。

既存技術を絶対音感的-相対音感的、視覚的-聴覚的という 2 軸を用いて 4 象限に分類すると図 3 のようになる。絶対音感的で視覚的な既存技術であるチューナーは主に象限 A に含まれるが、選んだ音を聴覚的に聞くことができるので象限 C にも関連する。相対音感的で視覚的な既存技術であるカラオケの音程表示は象限 B に含まれる。このように、聴覚でのフィードバックによってユーザの音感を拡張しようとしたものは既存技術がほぼない。現在の提案システムの実現部分は図 3 の A, C, D に該当する。

本論文の構成は、まず絶対音感保持者に対して行った主観的な音の聞こえ方の調査とその結果を紹介し、彼らが理想とする聴覚についての定性的見解にも触れる。それからマルチモーダルな拡張現実を用いてユーザの音感を拡張し絶対音感と相対音感の有用性にアクセスするためのシステム Otonona と Otonosa の設計と実装と予備実験とその結果を紹介し、最後に今後の展望を示す。

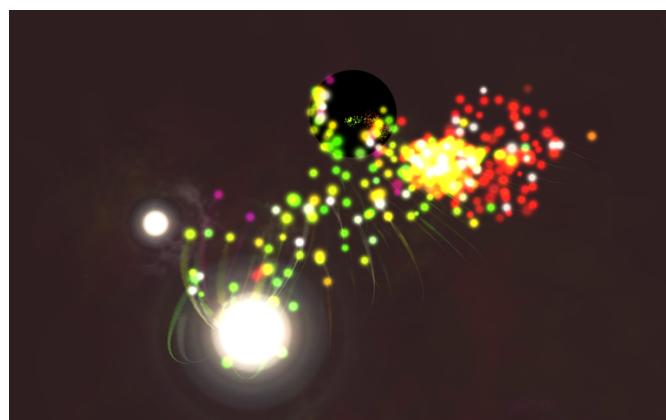


図 1 ビジュアライザによる視覚的提示の一例



図2 チューナーの一例

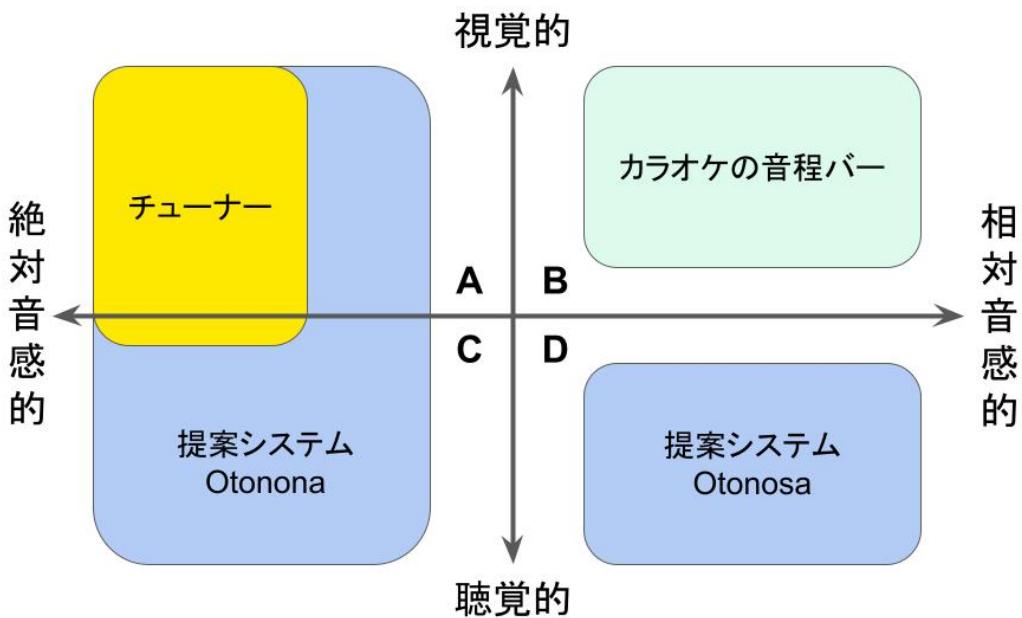


図3 既存システムと提案システムの実現部分

2. 絶対音感保持者に対する主観的な音の聞こえ方の調査

絶対音感保持者に対して主観的な音の聞こえ方に関する調査を行った。ここでは絶対音感保持者がどのように音名を把握しているかの調査結果と、音名を同時にいくつ判断できるかに関する調査結果を示す。

2.1 調査方法

音楽大学出身者を中心に絶対音感保持者を募集し、協力に応じた17名に対し、アンケート調査を行った。アンケートはGoogleフォームで作られ、回答者にはAmazonギフト券500円分が贈呈された。質問は以下の表1の通りである。

表1 質問項目と選択肢

質問	選択肢
問1. あなたは絶対音感を持っていますか？「その他」を選択した方は内容を教えてください	選択肢：はい・いいえ・その他
問2. あなたは楽器の音を聞いた時にどのようにその音名を感じることができますか？「その他」を選択した方は内容を教えてください	選択肢A：その音名の文字（「ド」や「C」など）が、その音の高さで発音されて、それが聞こえるように感じる。 選択肢B：その音名の文字（「ド」や「C」など）が、文字として見えるように感じる。 選択肢C：その他
問3. この質問は問2.でAと答えた方への質問です。音名はどのように聞こえますか？「その他」を選択した方は内容を教えてください	選択肢1：音名が、その楽器の音に重なって聞こえる。 選択肢2：音名が、その楽器の音に置き換わって聞こえる。 選択肢3：その他
問4. この質問は問2.でBと答えた方への質問です。音名はどのように見えますか？「その他」を選択した方は内容を教えてください	選択肢1：音名の文字を、視覚的イメージとして感じることができます。 選択肢2：音名の文字が、自分の実際の視界の中に現れて視認できる。 選択肢3：その他
問5. ピアノで和音を鳴らしたとき、平均的に何音まで音名を判断できますか？数値でお答え下さい	
問6. ピアノでランダムに音を重ねて鳴らしたとき、平均的に何音まで音名を判断できますか？数値でお答え下さい	

2.2 音名の認識方法に関する調査結果

最初に問1で調査協力者の全員が絶対音感を持つと自己認識していることが確認された。問2では選択肢A(その音名の文字（「ド」や「C」など）が、その音の高さで発音されて、それが聞こえるように感じる。)と回答した人が15人で全体の88.2%, 選択肢B(その音名の文字（「ド」や「C」など）が、文字として見えるように感じる。)や、その他と回答した人がそれぞれ1人ずつであり、音名を聴覚的に把握している人が大部分を占めることが分かった。

問3は問2で選択肢A(その音名の文字（「ド」や「C」など）が、その音の高さで発音さ

れて、それが聞こえるように感じる。)と回答した人に対して音名が聴覚的にどのように聞こえるか調べるために行った。結果は「音名が、その楽器の音に重なって聞こえる」を選んだ人が最多の9名となり全体の60%であった。ついで「音名が、その楽器の音に置き換わって聞こえる」を選んだ人が4名であり全体の26.7%であった。その他を選んだ人は2人であった。

問4は、問2で選択肢B(その音名の文字(「ド」や「C」など)が、文字として見えるようになる)と回答した人に対して音名が視覚的にどのように把握されるか調べるために行った。1名から聴こえた音が文字というより形になっているという内容の回答が得られた。この問4の調査は回答数が不足したため、今後調査を続けていく必要があると考えられる。問2と問3の結果を図4、図5にそれぞれ示す。

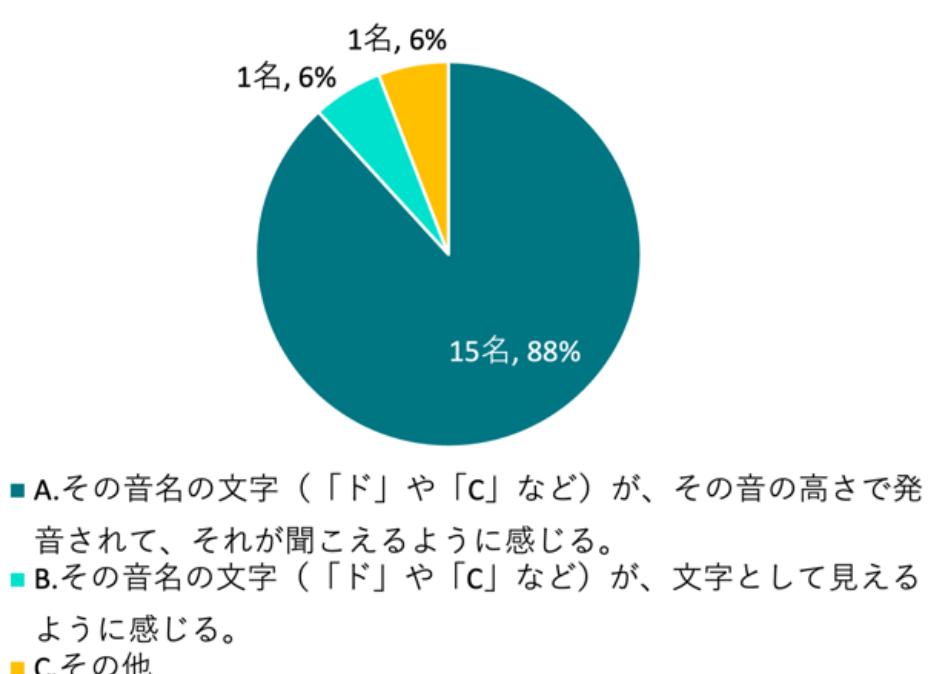
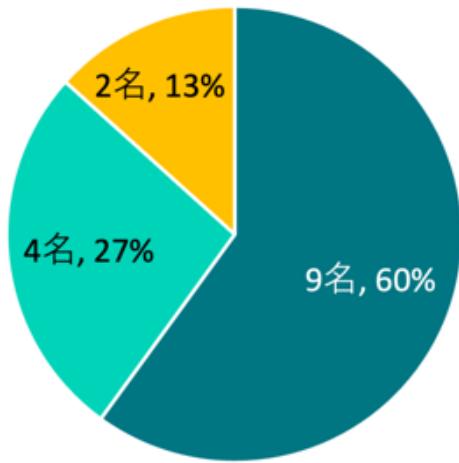


図4 音名の認識方法に関する調査結果(問2)



- 1. 音名が、その楽器の音に重なって聞こえる。
- 2. 音名が、その楽器の音に置き換わって聞こえる。
- 3. その他

図 5 音名の聞こえ方に関する調査結果（問 3）

2.3 音名を認識できる数に関する調査結果

次に、絶対音感保持者が音名を同時にいくつ把握可能なのかに関する調査の結果を示す。問 5 および問 6 の回答で、和音では平均 4.69 音（標準偏差 1.80）、ランダムでは平均 4.25 音（標準偏差 1.90）の音名を把握できることが分かった。図 6 は和音とランダムで音を提示した場合に音名を把握可能な数の平均値±標準偏差を示す。図 7 は協力者 17 名の回答である。なお、6 番の協力者からは、「和音に関しては上限なしで聞きとりが可能」と回答を頂いたため平均の計算から除外した。12 番の協力者はランダムでの音の提示に関しては無回答であった。

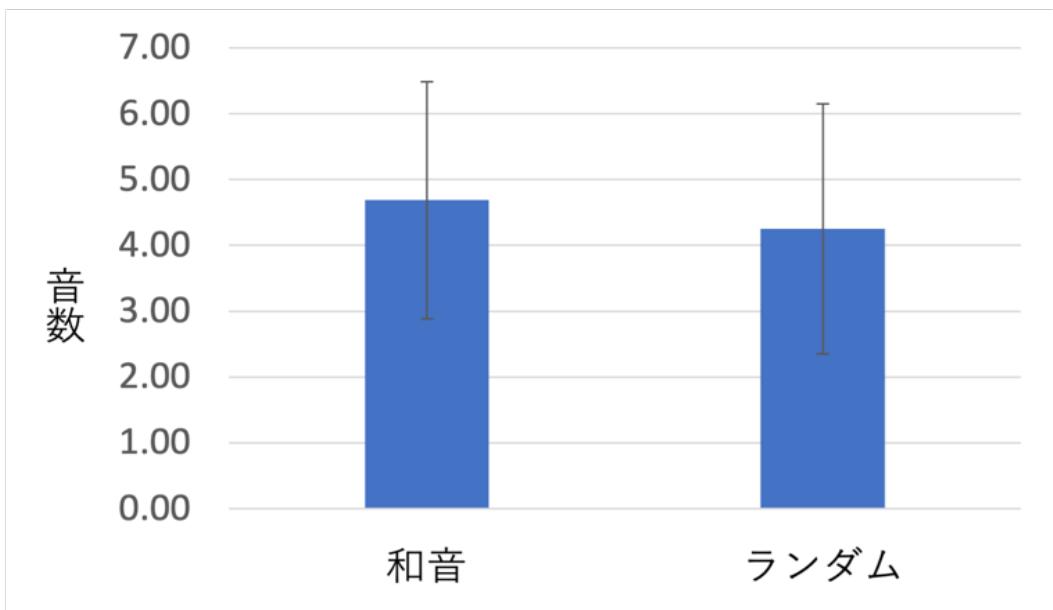


図 6 同時に把握できる音名の数の平均値

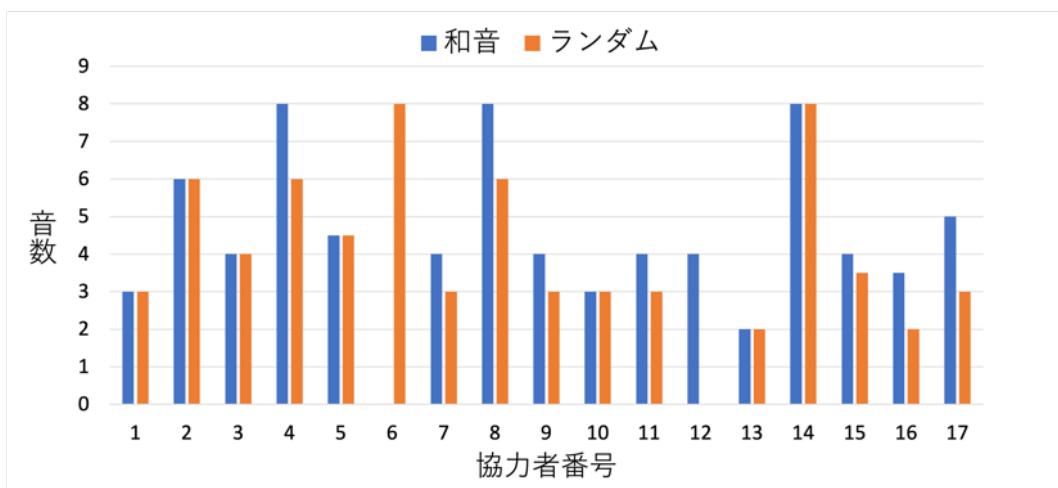


図 7 音名をいくつ把握できるかに関する調査結果

3. 絶対音感保持者への定性的調査

第2章での調査協力者17名に対して引き続き定性的調査を行った。Googleフォームで自由記述式の質問を用意し、任意で回答を求めた。次の節で詳細を紹介する。

3.1 絶対音感保持者の欲しい音感や機能

「こんな音感が欲しい、耳にこんな機能がほしいなどがあれば教えてください。」という自由記述式の質問に対し複数の回答が集まった。集まった回答を具体的に紹介すると「絶対

音感を持つ身からすると、相対音感の人がうらやましいのでどちらも欲しい。また、絶対音感を持っていない人が普段の生活で聞く音をどのように聞いているのかを体験してみたい。」「絶対音感のオンオフ機能が欲しい。聞こえてくる音の情報量が多くストレスを感じたりする時にオフにしたい。あとは絶対音感のない聴覚を体験してみたい。」「歌になるとピアノ程は正確に音が取れないので、その感覚が欲しい。」「もう少し定量的に音の評価ができるようになりたい。人間チューナーのような音感が欲しい。」「絶対音感を感じる周波数の上限が、約 4000Hz を超えて可聴域全体まで広がって欲しい。」「聴いた音をすぐに正確にアウトプットできる音感が欲しい。」「相対音感」「チューナーのように、具体的に何セント高いくらいまで分かれば良いと思う。」と多様な回答が集まった。ピッチをより定量的に把握することや、オンオフ機能、音色に関わらず音高を詳細に把握すること、音名を把握できる限界の音域の拡張、絶対音感と同時に相対音感の有用性にアクセスすることへの要望があることが明らかになった。

4. 提案システム

提案システムは人間の音感をマルチモーダルに拡張することを目的とする。提案システムはスマートフォンアプリとして実装され、第 2 章、第 3 章での調査を踏まえ現時点では特に絶対音感と相対音感の有用性にアクセスすることを意図した機能を備える。その上ユーザーがそれらの機能を自由に切り替えることができる。本システムは絶対音感に対応した機能として音名の聴覚的、視覚的なフィードバックを有し、相対音感に対応した機能としては聴覚的な音程のフィードバックを有する。本システムが有する前者の機能を Otonona、後者を Otonosa と命名した。

4.1 Otonona

4.1.1 Otonona 設計指針

本研究で実装した Otonona は絶対音感の有用性にマルチモーダルにアクセスすることを目的とする。第 2 章での調査を踏まえ、Otonona の設計指針をまとめる。調査の問 2 で選択肢 A(その音名の文字(「ド」や「C」など)が、その音の高さで発音されて、それが聞こえるように感じる。)と回答した人が全体の 88.2%、選択肢 B(その音名の文字(「ド」や「C」など)が、文字として見えるように感じる。)と回答した人が 6% であったことから、音名を聴覚と視覚でフィードバックすることは絶対音感の疑似体験として妥当であると判断した。

また、問 3 の調査で「音名が、その楽器の音に重なって聞こえる」を選んだ人が 60% であり、ついで「音名が、その楽器の音に置き換わって聞こえる」を選んだ人が 26.7% であったことを踏まえ、聴覚フィードバックの提示方法を 2 種類提案する。外界の音に音名を重畳する方法と、外界の音と音名が置き換わる方法である。この 2 種類のフィードバックは骨伝導ヘッドホンを用いるか、もしくは通常の密閉型ヘッドホンを用いることで切り替え

可能である。本システムを骨伝導ヘッドホンと共に使用する場合は外界の音に音名が重畠され、密閉型ヘッドホンと共に使用する場合は外界の音が音名に置き換わって聞こえる。

また、Otonona は屋内外での使用を想定しているため、持ち運びしやすいようスマートフォンのアプリケーションとして実装され、視覚的に提示された音名を画像として保存、シェアできる機能を持つ。

4.1.2 Otonona 概要

Otonona はユーザの音感をマルチモーダルに拡張し絶対音感の有用性にアクセスすることを目的としたシステムである。聴覚フィードバックとして、スマートフォンのマイクから読み取った音をリアルタイムで音名に変換し、ボーカロイドの初音ミク[10]がその音名をその音の高さで歌う。視覚でのフィードバックとしてはスマートフォンの画面に音名が表示される。その音名は風景と一緒に画像として保存することができ、ソーシャルメディアにシェアすることができる。図 8 に Otonona の概要を示す。

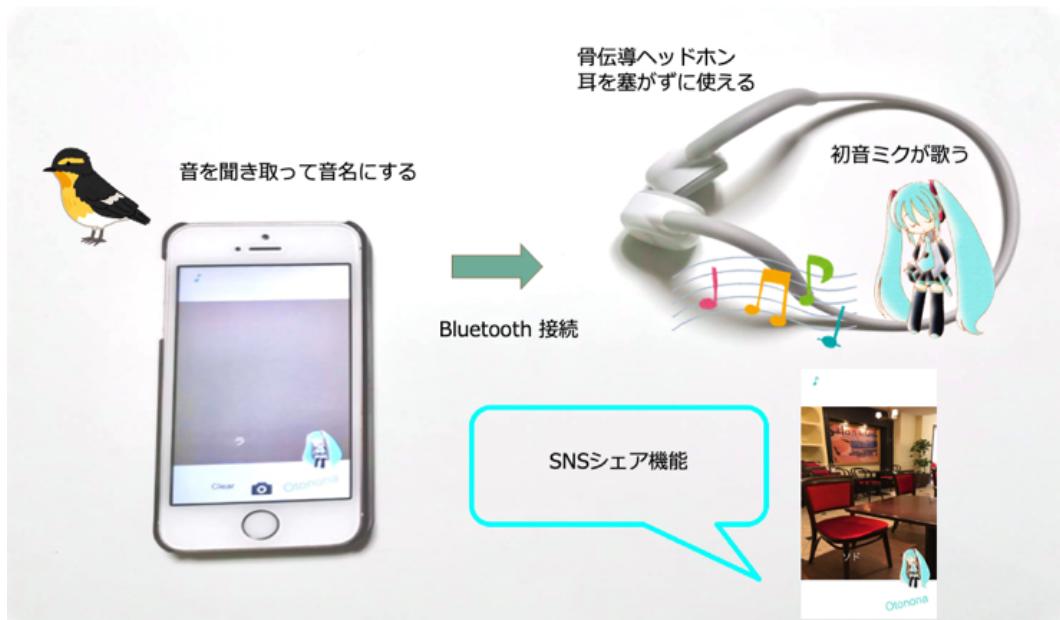


図 8 Otonona の概要

4.1.3 ユーザーインターフェース

本システムのユーザーインターフェース(UI)を図 9 に示す。本システムは Home 画面と Setup 画面の 2 画面で構成される。Home 画面の上部に音名表示部分があり、マイクから聞き取られた音がリアルタイムで音名に変換されこの部分に文字として表示される。同画面の中央部にはスマートフォンのカメラから読み込まれた風景がリアルタイムで表示されている。画面の下部には、シェアまたは保存するための画像を生成するボタン、現在表示されている音名をクリアするためのボタン、それから Setup 画面への遷移ボタンがある。

Setup 画面の上部には Home 画面への遷移ボタンがある。中央部にマイクの敏感さを調整するためのスライドバー、絶対音感機能 Otonona と相対音感機能 Otonosa を選択するためのトグル 2つがある。

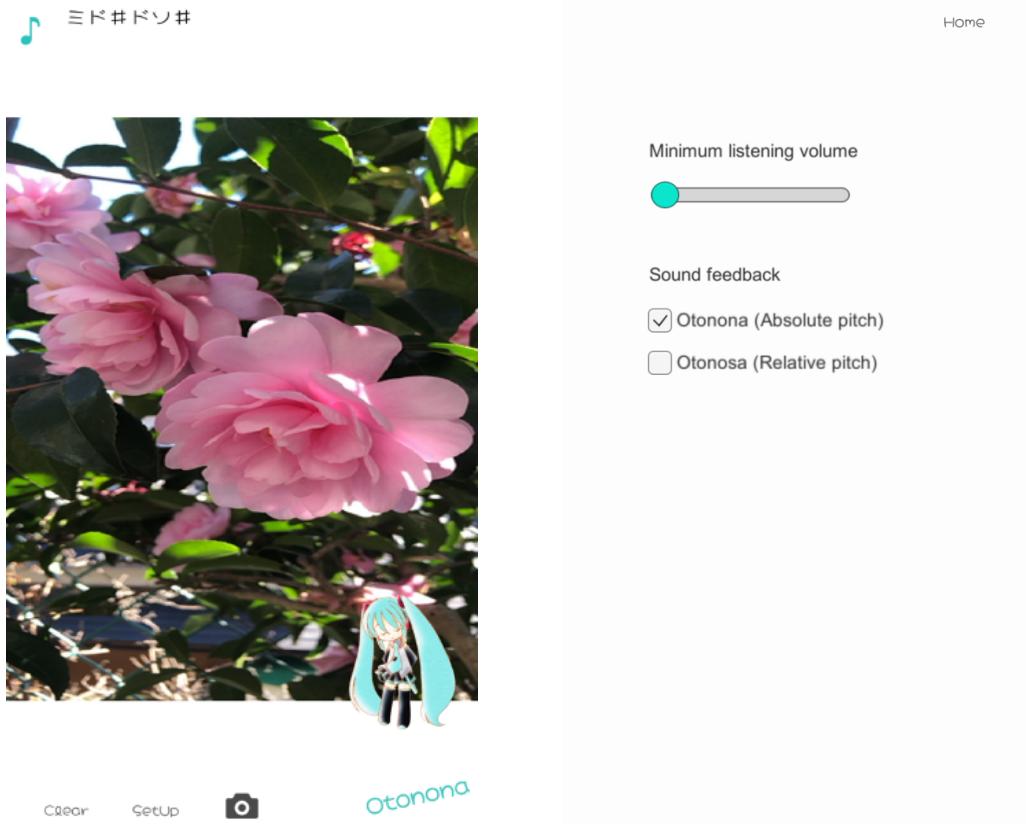


図 9 Home 画面(左), Setup 画面(右)

4.1.4 Otonona システム構成

Otonona のシステム構成を図 10 に示す。図 10 のように、Otonona は外界の音をマイクから聞き取る。それからフーリエ変換によりその音のスペクトラムを取得し、一番大きな値をとっている周波数を楽音に割り当て音名を判定する。聴覚での提示として、ボーカロイドにその楽音の高さで音名を発音させる。視覚での提示として、聞き取られた音の音名をカメラから読み取られた風景に重なるように表示する。音名は聴覚的には骨伝導ヘッドホン、または密閉型ヘッドホンを通してユーザに提示される。音名は視覚的には音名を画面に表示することでユーザに提示される。

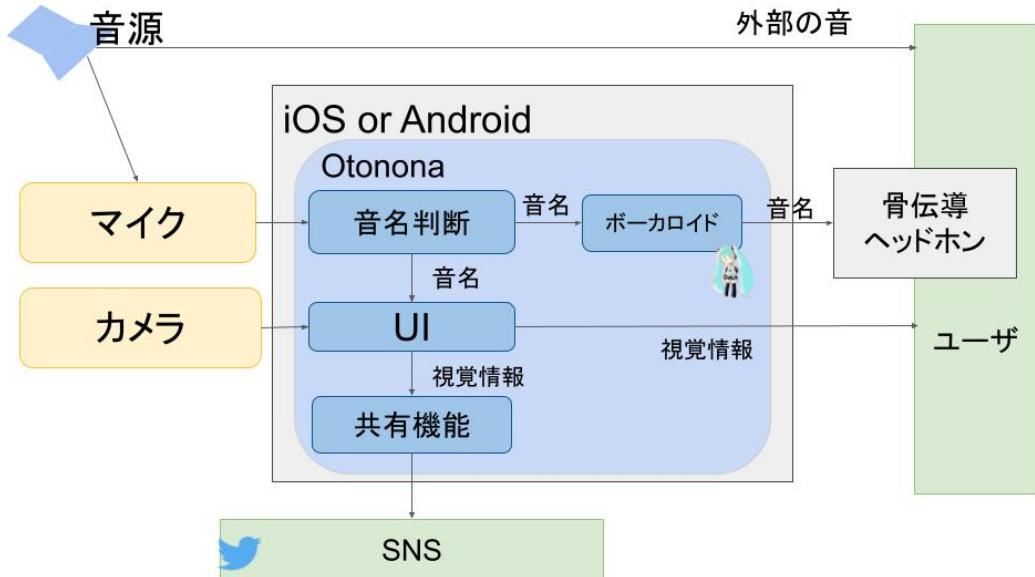


図 10 骨伝導ヘッドホンを用いる際のシステム構成図

4.1.5 音の音名変換

ここでは鳴っている音を音名に変換する仕組みを説明する。まず、マイクから入力された音をフーリエ変換によりスペクトラムに分割し、その中で一番大きい値をとっている周波数を取得する。その周波数を楽音に割り当て、音名に変換する。最大値が一定の値以下であるときは無音であるとし、変換を行わないようにすることでノイズの拾い方を調整した。

4.1.6 ソーシャルメディアシェア機能

Otonona はカメラから読み込んだ風景と音名を Home 画面の UI 上に表示している。ユーザは任意の瞬間に UI 上のカメラのアイコンをタップして風景と音名の入った画像を生成することができる。この画像は保存したりソーシャルメディアにシェアしたりすることができる。図 11 は生成される画像の一例である。

♪ ドソファ♯レ♯



Otonona

図 11 生成される画像の一例

4.1.7 骨伝導ヘッドホンと密閉型ヘッドホン

Otonona の使用方法として、骨伝導ヘッドホン、または密閉型ヘッドホンと組み合わせることを推奨する。これによりボーカロイドが発音した音をマイクが再取り込みすることを防ぐ。加えて、本システムと骨伝導ヘッドホンを用いた場合は外界の音に重畠する形で音名を同時に聞くことができ、密閉型ヘッドホンを用いた場合は音名を外界の音に置き換えて聞くことができる。図 12 は骨伝導ヘッドホンの使用イメージである。



図 12 骨伝導ヘッドホンの装着イメージ

4.2 Otonona の予備実験

絶対音感の有用性にアクセスするためのシステムである Otonona の性能調査のための実験を行った。この実験は Otonona の聴覚フィードバックのみに焦点を当て、ユーザが基準音を与えられずにどれくらい正確にピアノの音名を判定できるか調査するためのものである。調査協力者は 2 名でそれぞれ音楽経験者と音楽未経験者である。Otonona の聴覚フィードバックがある状態でピアノの音を聞いて音名を書いてもらうよう依頼した。協力者 1 名につき、音列に#を含む場合と含まない場合とに分け 2 度実行した。

4.2.1 実験方法

実験に用いる音は C3～B3 の 1 オクターブ内の音から選んだ 10 音の音列とし、その 10 音は毎回 Java のプログラムを用いてランダムに決定した。10 回の内、同じ音が連続した部分がある場合はもう一度音列を生成し、同じ音が連続している部分のない音列を使用した。協力者は Otonona がインストールされた iPhone に Bluetooth 接続した骨伝導イヤホンを装着し、電子ピアノの鍵盤に背を向けノートとペンが置かれた机に向かって立った。iPhone 本体は実験施行者が持ち、電子ピアノに接続されたスピーカーから約 1m の距離に保った。実験施行者が 1 音ずつピアノを弾き、協力者は Otonona の聴覚フィードバックを受け、その音名を書き取る。協力者が「OK」と言うのを聞いた後、施行者は次の音を弾く。

4.2.2 実験結果

Otonona を用いた際の音名聞き取り実験の結果を表 3~6 で示す。音楽経験者の結果は#を含む音列と#を含まない音列で差がなく、どちらも正答率 90%であった。音楽未経験者の結果は#を含む場合で正答率 40%であり、#を含まない場合で 90%だった。

表 3 音楽経験者の音名聞き取り実験結果(#を含む)

回	1	2	3	4	5	6	7	8	9	10
正解音名	ファ	ラ	シ	ファ	ミ	ラ#	レ	レ#	ミ	ラ
回答	ファ	ラ	?	ファ	ミ	ラ#	レ	レ#	ミ	ラ
結果	○	○	×	○	○	○	○	○	○	○
正答率	90%									

表 4 音楽経験者の音名聞き取り実験結果(#を含まない)

回	1	2	3	4	5	6	7	8	9	10
正解音名	ラ	レ	ラ	レ	ファ	ラ	ソ	ド	レ	ラ
回答	ラ	レ	ラ	レ	ファ	ラ	ド	ド	レ	ラ
結果	○	○	○	○	○	○	×	○	○	○
正答率	90%									

表 5 音楽未経験者の音名聞き取り実験結果(#を含む)

回	1	2	3	4	5	6	7	8	9	10
正解音名	ファ#	ラ#	ミ	シ	ファ#	シ	ラ#	シ	ラ	レ#
回答	?	?	ミ	シ	?	シ	?	ミ	ラ	?
結果	×	×	○	○	×	○	×	×	○	×
正答率	40%									

表 6 音楽未経験者の音名聞き取り実験結果(#を含まない)

回	1	2	3	4	5	6	7	8	9	10
正解音名	ミ	ラ	ファ	ド	シ	ラ	ド	シ	ド	ファ
回答	ミ	ラ	ファ	ド	シ	ラ	ド	シ	ソ	ファ
結果	○	○	○	○	○	○	○	○	×	○
正答率	90%									

4.2.3 予備実験から得られた知見

音楽未経験者が#を含む音列の音名を答える際、正答率が 40%と落ち込んだ。この理由として、音楽未経験者は音名そのものに慣れておらず、どのような聴覚フィードバックを受けたのか予想するのが難しいということが挙げられる。実験終了後に実験協力者に所感の聞き取りを行ったところ、「ファシャープ」や「ソシャープ」など、#の付いた音名のフィードバックを受けた際に、何を言われたのか聞き取れず混乱が生じたという意見が上がった。また、音楽経験者と未経験者が共通して失点しているのが、母音が共通する音名の聞き取りである。例えばドとソ、ミとシなどの聞き間違えで正答率が下がっている。

現在の Otonona の聴覚フィードバックは急速な音列のパッセージに対応するため、音名がボーカロイドにより素早く発音される。音源の速さや発音の聞き取りやすさを調整する

ことにより、正答率が向上する可能性がある。また、実験のデザインとして現在は音名を書き取ってもらっているが、音名を選択式で答えてもらうようにすることで結果が変わる可能性がある。また、提示される音名の例を予め確認してもらうことで混乱が防げるであろうと考えられる。

4.3 Otonosa

4.3.1 Otonosa 概要

Otonosa はユーザの音感を聴覚フィードバックにより拡張し、相対音感の有用性にアクセスすることを目的としたシステムである。本システムは音と次の音の音程(interval)を聴覚的にフィードバックする。この機能は Setup 画面で選択することで使用することができる。基本的なアルゴリズム、システム構成、視覚フィードバックは Otonona と共通である。

4.3.2 音程の提示

聴覚フィードバックにより音程を提示する。音程とは 2 音間の隔たりを指し、それを度数で表したものである。楽音からシャープやフラットを取り除き幹音とした時に、音の間がいくつ離れているかを数えたものを基本とし、その後「長」「短」「完全」「減」「増」などを使ってより詳細に表すことができる。例えば、ドとドは同じ音なので 1 度、ドとレは 2 度、ドとミは 3 度である。Otonosa は現時点では幹音の音程を測り、ボーカロイド「初音ミク」の発音により聴覚で提示する。例えば前に聞き取った音から現在聞き取った音の音程が一度の時は「iti」、二度の時は「ni」と初音ミクが発音する。

また 1 オクターブを超える音程を複音程と呼ぶが、これは単音程に直して表すことができる。本システムでも複音程を単音程として表示することを選択した。例えば 1 オクターブと 4 度離れた音程を「11 度」と表すことができるので、Otonosa では初音ミクが「jyuiti」と発音する。一例として基準音を C3(通称：真ん中のド)とした時の各音との音程と、Otonosa での初音ミクの発音の対応を表 7 で紹介する。

表7 音程とフィードバックの例

基準音	比べる音	音程	Otonosa
C3 (真ん中のド)	C3	完全1度	iti
	C#3	増1度	iti
	D3	長2度	ni
	D#3	増2度	ni
	E3	長3度	san
	F3	完全4度	yon
	F#3	増4度	yon
	G3	完全5度	go
	G#3	増5度	go
	A3	長6度	roku
	A#3	増6度	roku
	B3	長7度	nana
	C4	完全8度	hati
	C#4	増8度	hati
	D4	長9度	kyu
	D#4	増9度	kyu
	E4	長10度	jyu
	F4	完全11度	jyu-iti
	F#4	増11度	jyu-iti
	G4	完全12度	jyu-ni
	G#4	増12度	jyu-ni
	A4	長13度	jyu-san
	A#4	増13度	jyu-san
	B4	長14度	jyu-yon
	以下省略		

5. 課題と展望

5.1 音感拡張の展望

図13に示すように提案システムは現時点では相対音感の視覚的フィードバックに対応していない。現時点では相対音感の視覚的フィードバックとは、カラオケの画面に表示されるよう

な音程のバーを想定しているが、これを実装して 4 つの象限を網羅することは今後の展望といえる。その上、その 4 象限のフィードバックをユーザが自由に切り替えたり、重ねたり、自由な比率で混ぜ合わせたりできるようにすることを検討している。フィードバックを重ねるとは数種類のフィードバックを同時に受け取ることを指す。たとえば、絶対音感のフィードバックと相対音感のフィードバックを同時に使用すれば、音名と音程を一度に把握することが可能になる。また、フィードバックを自由な比率で混ぜるとは、4 種類のフィードバックの強弱をそれぞれ独立に調整することで聴覚や視覚のマルチモーダルフィードバックを任意に決定することを指す。たとえば、絶対音感の視覚フィードバックを強めれば音楽の音名を意識し、相対音感の視覚フィードバックを強めれば音程が作り出す形に意識が向くと考えられる。聴覚フィードバックは音量や変換する音の範囲の選択にて強弱を決定し、視覚フィードバックは UI に表示されるグラフィックの濃さや文字の大きさで強弱を決定できると考えられる。

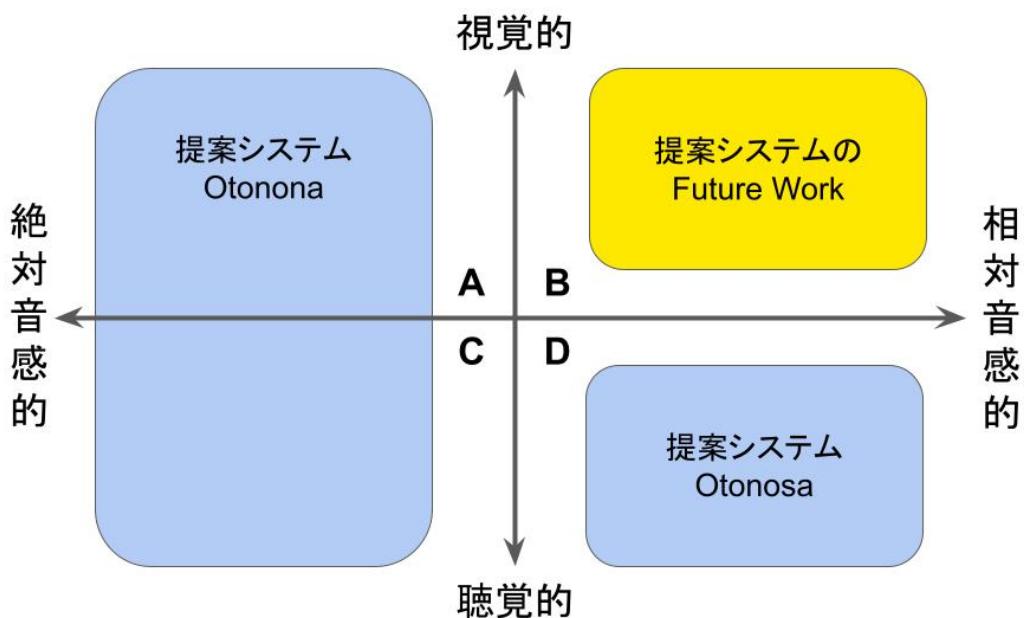


図 13 提案システムの実現部分と Future work

5.2 Otonona の課題と展望

本稿で提案した Otonona は、外界の音をスペクトラムに分割し、そのスペクトラムの内、一番大きな値を取っている周波数の音名を提示する。すなわち、同時に複数の音が鳴っている場合に対応できていない。本稿の第 2 章で示した調査により、絶対音感保持者は平均して和音なら 4.69 音、ランダムに音を鳴らした場合には 4.25 音の音名を判断できることが分かった。このことから Otonona の今後の展望として、複数の音名を同時に提示するということが挙げられる。これは楽音から成る音楽の和音の移り変わりや音楽の構造を理解する

上で有効であると予想される。

複数の音名を同時に提示する場合、聞こえている音すべての音名を提示するのか、あるいはその一部の音名を提示するのかといった「範囲選択」の問題が発生する。今回の調査で絶対音感保持者に対し音の聞こえ方を自由記述してもらった結果によると、この「範囲選択」ができるかどうか、およびそれが自分の意思で可能かどうかについては個人差が大きいことが示唆されている。今後調査を進めることにより、どのような機能が絶対音感の疑似体験あるいは有用性へのアクセス方法としてよりふさわしいかを考察する必要がある。

また、現在はたとえば 450Hz の音が検出されたとき、それに最も近い音名は 440Hz の「ラ」であり、ボーカロイドは 440Hz で「ラ」と歌唱するアルゴリズムをとっている。つまり実世界の音の周波数と、歌唱される音の周波数は多くの場合一致しない。この不一致により、骨伝導ヘッドホンを用いた場合に音名の標準的な周波数と外界の音の周波数との差異を知覚可能にすることが可能であるのが提案手法の特徴である。一方でボーカロイドが 440Hz ではなく 450Hz で歌唱するアルゴリズムも考えられ、この手法には外界の音の情報をより多く保存しているという別の利便性がある。どちらのほうが絶対音感の疑似体験として、あるいは絶対音感の有用性へのアクセス方法としてよりふさわしいかも検討していくたい。

また、第 5 章の予備実験では初音ミクにより発音された音名がユーザによっては聞き取りにくい可能性が示唆されたため、聴覚的フィードバックの最適化の後、より大きな規模で性能評価実験を行いたい。

5.3 Otonosa の課題と展望

本システムの音程の聴覚的フィードバックは、現在は幹音の差のみを考慮し反映している。例えば「完全 5 度」「増 5 度」「減 5 度」はすべて 5 度とみなされ、「go」と発音されることでフィードバックされる。これは発音時間を短くすることで音の速い変化に対応するためだが、より詳細に音程を把握可能にするためには検討の余地がある。このシステムの視覚的フィードバックは現在 Otonona と共に音名を表示するのみであるが、これを音程の提示に特化させれば、音程をより詳細に把握することも可能だと考えられる。また、Otonosa に関しての性能調査やユーザスタディを行いたい。

謝辞

本研究を遂行するにあたり丁寧なご指導を賜りました栗原一貴先生に感謝いたします。また栗原研究室の皆様、ご意見や技術的助言を下さいました方々、調査にご協力下さいました方々に御礼申し上げます。誠にありがとうございました。

参考文献

- 1) 最相葉月: “絶対音感” 小学館, 1998.
- 2) 柿原彩子: “絶対音感習得プロセスに関する縦断的研究.” 教育心理学研究 47.1 (1999): 19-27.
- 3) 宮崎謙一, 石井玲子, 大串健吾: “絶対音感を持つ音楽専攻学生によるメロディの認知.” 日本音響学会誌 50.10 (1994): 780-788.
- 4) 天岩靜子, 増田桃子: “絶対音感保有者における聴覚刺激の干渉効果と脳活動: fNIRS による脳血流量の測定に基づいて.” (2013): 211-225.
- 5) 藤崎和香, 柏野牧夫: “絶対音感保持者の音高知覚特性.” 日本音響学会誌 57.12 (2001): 759-767.
- 6) 江口寿子, 江口彩子: “新・絶対音感プログラム” 全音楽譜出版社, 2001.
- 7) 柿原彩子: “なぜ絶対音感は幼少期にしか習得できないのか? 訓練開始年齢が絶対音感習得過程に及ぼす影響.” 教育心理学研究 52.4 (2004): 485-496.
- 8) 江口寿子: “相対音感プログラム・上” 全音楽譜出版社, 2000.
- 9) Hiroki Watanabe, and Tsutomu Terada, “Manipulatable Auditory Perception in Wearable Computing.” Proceedings of the Augmented Humans International Conference, 1-12, 2020.
- 10) Crypton: 初音ミクの紹介
https://ec.crypton.co.jp/pages/prod/virtualsinger/about_miku