

Otonona: マルチモーダルな拡張現実を用いた絶対音感体験の提案

岡あゆみ^{†1} 栗原一貴^{†1}

絶対音感の一部の人が6歳頃までに身につける能力である。程度に差はあるものの、この能力を持つ人は一般に音の名前を言い当てる事が出来ると言われている。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つけることは難しい。そこでマルチモーダルな拡張現実を用いて絶対音感を体験するためのシステム Otonona を提案する。

1. はじめに

一部の人が6歳頃までに身につける絶対音感という能力がある。この能力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は音名を判定することが出来る。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が、大人になってからの訓練で身につけることは難しい。榊原によると”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。”[1]という。そこでシステムによるフィードバックにより音感を補助することを考え、視覚と聴覚によりユーザに音名を提示する Otonona を提案する。本稿では Otonona の概要、設計方針、システム構成、音を変換する仕組み、機能を紹介し、最後に今後の展望を挙げる。

2. 関連研究

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。本来の音に編集を加えた上で音を聴覚で提示するシステムは Watanabe らのシステム[2]で実装されている。

3. Otonona

3.1 Otonona 概要

本研究で実装した Otonona は音を音名で把握出来るようにユーザの音感を補助するシステムである。ユーザへのフィードバックは聴覚と視覚によってもたらされる。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、ボコーロイドの初音ミク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ている。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。

骨伝導ヘッドホンと一緒に使えば、耳を塞がず、周りの音と音名がどちらも聞こえるのでよりリアルに絶対音感を体験することができると考えられる。図1に Otonona の概要を示す。

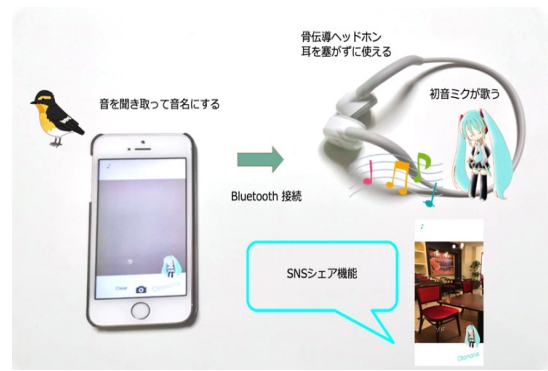


図1 Otonona の概要

3.2 設計方針

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る。また、外界の音に音名を重畳して提示するために骨伝導ヘッドホンを用いる。ユーザが Otonona での発見をシェア出来るようソーシャルメディアシェア機能を実装した。

3.3 システム構成

Otonona のシステム構成を図2に示す。

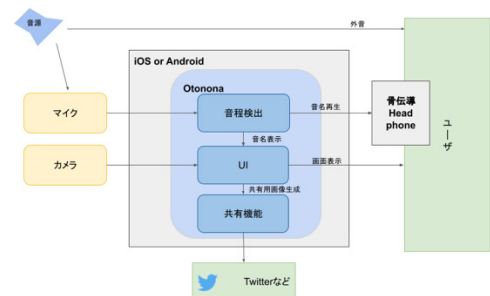


図2 Otonona のシステム構成

^{†1} 津田塾大学
Tsuda University.

図2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。この機能はユーザがソーシャルメディアにその画像をシェアすることを促進する。

3.4 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。図3はスペクトラムの内、大きな値を取っている周波数を取得する簡易的なイメージである。また一定の値以下を無音とすることで、ノイズの拾いかたを調整する。図3では点線以下を無音と定義する。

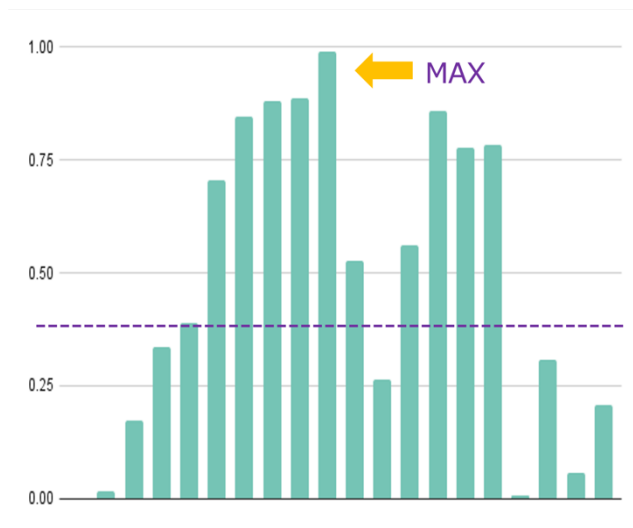


図3 簡易的に表現した音名変換のイメージ

3.5 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スクリーンショットを撮影することができる。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4は画像の一例である。



図4 スクリーンショットの一例

3.6 骨伝導ヘッドホン

Otonona の使用方法として、骨伝導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。それに加え、本システムと骨伝導ヘッドホンを併用すれば外界の音と音名を同時に聞くことが出来るため、絶対音感の体験としてより自然なものになると考えられる。図5は骨伝導ヘッドホンの使用イメージである。



図5 骨伝導ヘッドホン

4. 課題と展望

4.1 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した。自然界のあらゆる音を対象とするときこの提案手法には限界がある。例えば、鳥の声を

Otonona で音名に変換すると人間の耳に納得感のある結果が得られていない。この理由として自然界の音は一つの極大値を持つスペクトラムとは限らないことが挙げられる。具体例を挙げると、キジバトの声を音の分析ソフトウェア Sonic Visualiser[3]で確認すると図6のようになる。キジバトの声は複数の極大値をもつ複雑なスペクトラムとなっている。本システムは最大値をとる極大値の周波数を優先して提示するため人間の感覚と異なる可能性がある。よってこの差を埋めるためには、スペクトラムの内の最大値をとっている周波数だけではなく、2番目、3番目の極大値と候補を取得しユーザへ音名をフィードバックする際に重ねて表示、再生するということが有効であるかもしれない。また、提示方法として本稿では聴覚と視覚を選んだが他の提示方法も検討する余地がある。また今回はスマートフォンのアプリケーションとして実装したが、使用するシチュエーションなどに合わせディスプレイ方法に関しても議論が必要である。

<https://www.sonicvisualiser.org/>

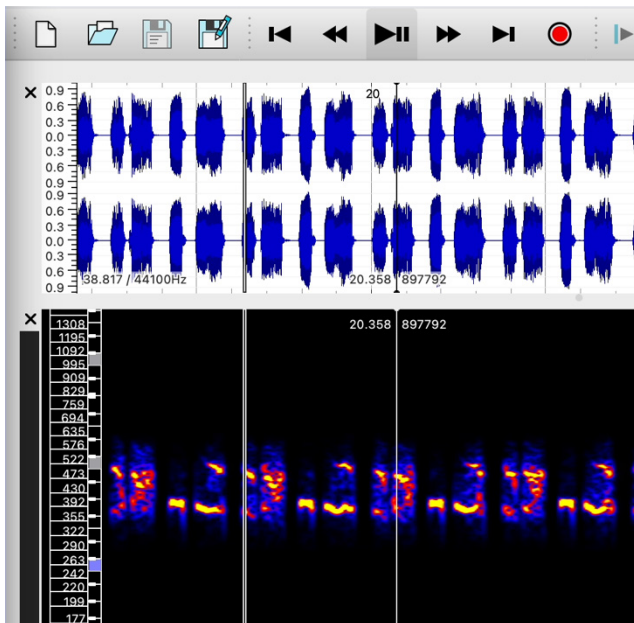


図 6 キジバトの声の分析

謝辞 本研究を進めるに当たり、終始助言を賜り丁寧な指導をして下さいました栗原一貴先生を始め、栗原研究室の皆様、技術面でのサポートを下さいました方、意見を交換して下さいました皆様に感謝を申し上げます。

参考文献

- 1) 榊原彩子: "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.
- 2) Hiroki Watanabe, and Tsutomu Terada: Manipulatable Auditory Perception in Wearable Computing. Proceedings of the Augmented Humans International Conference. 2020.
- 3) Sonic Visualiser: