

# 倍音分析によるいい声作りの支援アプリ開発に向けて

菅原衣織<sup>1</sup> 伊藤貴之<sup>1</sup>

**概要:** 声色の印象が与える効果はコミュニケーションにおいて約4割を占めるという考え方がある。これは声色が言葉以上にものを語る可能性を示唆している。本研究では倍音という観点から音声进行分析し、声の傾向を視覚化し提示することで、ユーザの目指す理想の声に近づけるような支援システムの開発を目標にしている。その前段階として、専門知識の無いユーザでも自分の声の傾向を把握できるプロトタイプのアプリケーションを提案する。

## Toward Application Development for Assisting Impressive Voice Tones Applying Overtone Analysis

IORI SUGAHARA<sup>1</sup> TAKAYUKI ITOH<sup>1</sup>

**Abstract:** There is a report that impression of tone of voice covers about 40% of the impression in communication. This result suggests that the tone of voice may impress rather than words. The goal of our study presented in this paper is the development of a support system for training to archive ideal voices, by analyzing overtones of the voices, and visualizing the analysis results. As a preliminary step, this paper discusses the application prototype that can identify spectral property of our own voice for novice users.

### 1. はじめに

私たちが生活する上で切っても切り離せないコミュニケーション。そのコミュニケーションを円滑にし、私たちの人柄をことば以上に伝えてくれるのが、声である。明るく明瞭な声で話している人にはいい印象を受けることができるように、逆にぼそぼそとはっきりしない声で話す人には好印象を受け難い。また、メラビアンの法則 [1] として従来から知られているように、コミュニケーションにおいて人の行動が他人にどのように影響を及ぼすかという実験では、言語情報が7%、聴覚情報が38%、視覚情報が55%を占めるという考え方がある。このように、声の印象が自己の評価に与える影響は少なくない。

もし自己の声の印象を自在に変えることが可能になれば、日常生活を豊かにする重要なツールのひとつになるだろう。企画や研究発表の際、ライブやコンサートの際、聴衆や観客にいい印象を与えるために髪型や化粧、洋服、小道具など身嗜みに気をつける。それに加えて声もいい印象

を与えるための重要な要素であると考えられる。声は自分を飾るアクセサリのような存在でもありと考えられ、これをうまく魅せられるかどうかは重要である。例えば学会や社内でのプレゼンテーションの際、内容や話術もさることながら、声で聴衆を惹きつけることができれば、内容の評価だけではなく話者自身への興味にも繋がり、思いがけないコネクションができるかもしれない。また例えば好意を寄せている相手を口説く際、変に緊張していたり、照れ隠しからかはっきりしない物言いをする人よりも、無駄に色気たっぷりのいい声で迫る方が成功するかもしれない。また別の例として、動物の求愛行動においても、鳥や蟬のオスは特徴的な声色で鳴くことによって縄張りを主張しメス呼び寄せる。このように声の使い方が恋の成就に少なからず関与しているのは周知の事実である。

近年ではボイストレーニングなどの技術が発達している。人間の声帯を根本から変えることは不可能に近いが、どんな人の声でもトレーニングによって「いい声」になれる可能性を持っている。しかしボイストレーニングなどの専門家の診断に頼らずに、日常的に自分の声の状態を知り、日

<sup>1</sup> お茶の水女子大学大学院人間文化創成科学研究科

常に自分の声の改善方法を知ることには簡単ではない。

本研究では、声の印象、つまり声色の観点から「いい声」を追求し、これを支援するシステムの作成を目標にした。一般的に「いい声」と判断するためには、滑舌や話の速度、声の大きさなどいくつかの要素が複雑に絡んでくると考えられる。しかし、滑舌は早口言葉の練習を積むことで良くなり、声の大きさは拡声器を使うことで補正が可能であるように、いくつかの要素は既存の手法を用いて改善される。そこで本研究では声色の印象に絞って議論を進める。

声色の印象を左右する要素として、我々は倍音に着目した。音楽理論などにおいて倍音とは、楽音の音高とされる周波数に対して、2以上の整数倍の周波数を持つ音の成分のことを指す。逆に楽音の音高とされる成分を基音と呼ぶ。基音と倍音はフーリエ変換等の周波数変換手法を用いることで検出が可能である。以上のように定義される狭義の倍音を「整数次倍音」と呼ぶのに対して、基音に対して非整数倍の周波数を持つ音の成分は「非整数次倍音」と呼ばれている。図1は整数次倍音が大きく含まれる声、図2は非整数次倍音が大きく含まれる声に対して、横軸を周波数、縦軸を音量として周波数分布を示したものである。尺八奏者の中村は著書 [2] の中で、整数次倍音が強いとカリスマ性や明朗性、豊かさが印象付けられ、非整数次倍音が強いと情緒や親密感が印象付けられると述べている。本研究ではこれらの特徴を分析することで声色の印象を評価する。また本報告ではこれ以降、非整数次倍音を含む広義の倍音を「倍音」と称する。

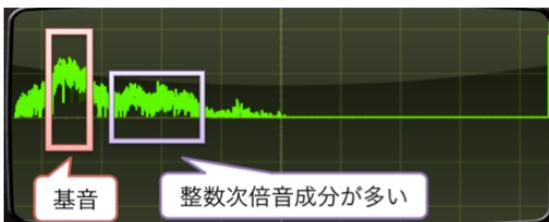


図 1 整数次倍音が大きく含まれる声

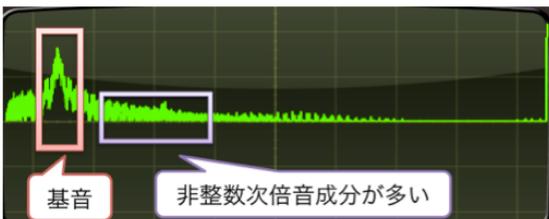


図 2 非整数次倍音が大きく含まれる声

## 2. 関連研究

### 2.1 いい声に関する研究

人間の発声に着目し、ユーザがいい声を出せるように身体的スキルの向上を促すメディアの提案を目的とする発声のメタ認知促進システム“いい声マイク”の提案 [3] では、いい声を「音としてしっかりと発せられていて、響いている声」と定義している。いい声であるかを判定するアルゴリズムの構築のために、バットの素振り音と俳優と新人の朗読を用いた予備調査を経て、第1、第2フォルマント領域（人間の声領域と言われる周波数領域）のパワーの分散が低いこと、第3フォルマント領域以上の直線回帰残差が低いことを挙げている。

また、モテ声診断 VQ チェッカー [4] というウェブサイトおよびスマートフォン用の人気アプリケーションがある。このサービスでは「声の高さ」「声の大きさ」「一音の長さ」「耳への入りやすさ」「滑舌の良さ」の5つの項目から声の良さを定義している。

文献 [3] ではいい声の判定基準として第1、第2フォルマント領域のパワーの分散が低いことなどをあげていたり、文献 [4] では「声の大きさ」や「滑舌の良さ」などを挙げている。これらと本研究には一定の関連性を有するが、本研究では全体的な意味での声量には制約を加えず、またフォルマント領域といった特定の周波数領域に限定せず、倍音の観点から声色を分析し、いい声とは何なのかを追求するものである。

### 2.2 声色の変化に関する研究

The Throat III [5] は役割に応じて自身の声の表現を変えることが必要なオペラにおいて、歌唱技術を具現化して拡張し、オペラ歌手を支援するツールとして開発された。

また VocaListener2 [6] は、ユーザの歌唱音声からその声色変化を真似て歌声合成するシステムである。ユーザ歌唱の音の高さと音量を真似た多様な歌声を合成して声色空間を構成し、その結果を用いてユーザ歌唱の声色変化を反映し合成することで声色変化を実現している。

声色の変化に関する研究はいくつか見られるが、歌唱を対象としたものが多く、特に文献 [5] のように歌唱改善のためにユーザにフィードバックを返すことや、文献 [6] のように歌唱用の音声合成に用いている例が代表的である。それに対して本研究は、手軽に扱えるスマートフォンのアプリとして実装し、日常生活でのいい声作りを支援することを目標としている。

## 3. 事前調査

倍音が声の印象にどのような影響を及ぼすのかを知るために、以下の調査を実施した。

### 3.1 倍音の効果の検証

#### 3.1.1 調査内容

前述したように中村は著書 [2] で、整数次倍音はカリスマ性や明朗性が印象付けられ、非整数次倍音は情緒や親密感が印象付けられると述べている。そして、整数次倍音の強い芸能人としてタモリ氏や黒柳徹子氏、非整数次倍音の強い芸能人として桑田佳祐氏やビートたけし氏を挙げている。彼らの特徴的な声が視覚的にどのように周波数分布に現れるのかを観察するために、動画サイトの YouTube に掲載されていた演説中のタモリ氏の声と、ラジオでトーク中の桑田佳祐氏の声と比較してみた。

#### 3.1.2 調査結果

図 3(左) がタモリ氏の声、図 3(右) が桑田佳祐氏の声の周波数分布である。前者では基音に対してほぼ整数倍の周波数をもつ倍音をはっきり確認することができる。一方で後者では、基音の整数倍の周波数に目立った点は見られず、むしろ非整数次倍音のほうの特徴的であるといえる。この周波数分布はある短時間の声を抜粋した結果に過ぎないが、話している間ほぼ終始に渡って同様の傾向が見られた。以上の調査により、声の印象の差異が倍音に表れているといえる。

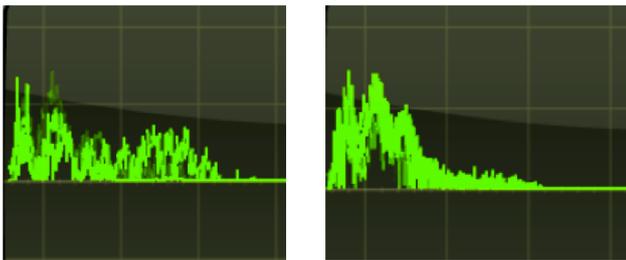


図 3 芸能人の声における倍音の特徴

### 3.2 感情が与える影響

#### 3.2.1 調査内容

次に、感情の変化が周波数にどのような影響を与えるのかを調査するために、被験者に「あいうえお」のことばを「通常」「喜び」「怒り」「哀しみ」の 4 つの状態で作声してもらい、これを分析した。「あいうえお」のことばを採用した理由は、母音であるため特徴が掴みやすく比較が容易であることと、意味のあることばを採用してしまうとアクセントやイントネーションに差が生じると考えたためである。

#### 3.2.2 調査結果

被験者は 20 代の男性 6 名、女性 10 名である。

図 4 は「あいうえお」のうち「あ」の音に対して、それぞれの感情における基音の周波数を記録したものである。緑色が通常の状態、黄色が喜び、赤色が怒り、青色が哀しみの状態を表している。また、縦軸は周波数 (Hz) を、横軸において M と表記のあるものが男性 (Man)、W と表記のあるものが女性 (Woman) の被験者を表す。図 4 のグラ

フから、通常の状態と比較すると、喜びや哀しみの感情を発するときには基音周波数が高くなる傾向にあり、逆に怒りの感情を発するときには基音周波数が低くなる傾向にあることがわかる。また、哀しみの感情を表すグラフにおいて個人差が激しい理由として、哀しみを「落ち込み」で表現するのか「嘆き」で表現するのかといった解釈の相違が考えられる。

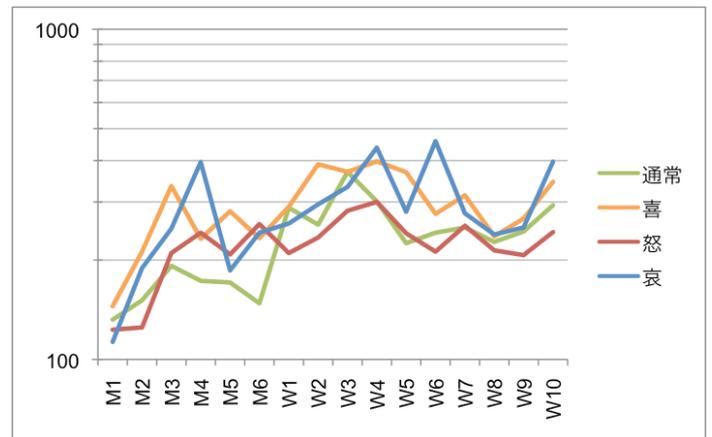


図 4 感情の変化に伴う基音周波数の音量の変化

また、図 5、図 6 は男性の被験者の「あ」の音の周波数分布を示し、図 7、図 8 は女性の被験者の「あ」の音の周波数分布を示したものである。横軸は周波数 (Hz) を表し、右へいくほど高くなる。縦軸は音量 (dB) を表し、上に行くほど音量が大きくなるということになる。

ここで基音ではなく倍音に着目する。喜びや怒りの感情は声を張る傾向にあるためか、広い周波数領域にわたって全体的に音量が大きくなり、通常の声と比較して周波数の高い部分で非整数次倍音が大きくなる傾向が見られる。逆に哀しみの感情は通常よりも整数次倍音のピークをはっきり確認できる傾向にあるという予想外の知見を得ることになった。我々はこの傾向に関して、文献 [2] で提唱されてきた「整数次倍音が豊富であればカリスマ性があり明瞭な声」という傾向に合致するとは言い切れないのではないかと直感的に考える。この点について、さらに検討を進めたい。

また一般的に、基音の 10 倍、20 倍にわたるまで整数次倍音が大きく含まれる声は遠くまでよく通る声であり、逆に高い周波数まで整数次倍音が大きく含まれない声はか細く聞こえる、という特徴が知られている。今回の調査でもその傾向が観察された。我々は主観的に、本調査で分析した 4 種類の声のうち、「喜び」の声が最も明瞭に聞こえ、「悲しみ」の声がか細く聞こえるであろうと仮定した。そして実際に、喜びの声の周波数分布は、多くの場合において他の 3 種類の周波数分布よりも高い周波数にわたって倍音が大きく含まれていた。そして哀しみの声の周波数分布は他の 3 種類の周波数分布よりも断続的になっており、か細い声の特徴に近いといえる。

以上の調査により、感情の変化がどのように周波数や倍音に影響を及ぼすのかを観察することができた。この結果を踏襲し、より詳細な分析を経て、いい声作りへの支援アプリを開発したい。

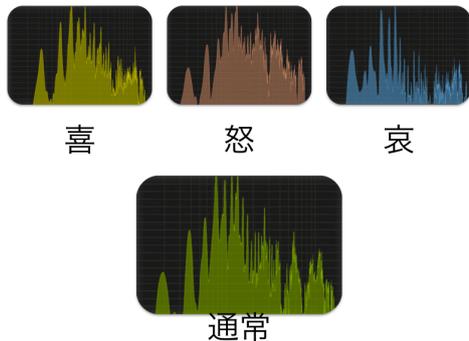


図 5 被験者男性の「あ」の音

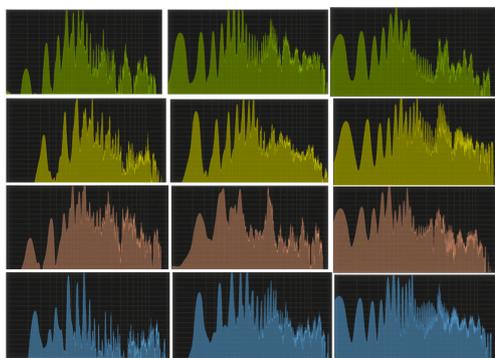


図 6 被験者男性の「あ」の音

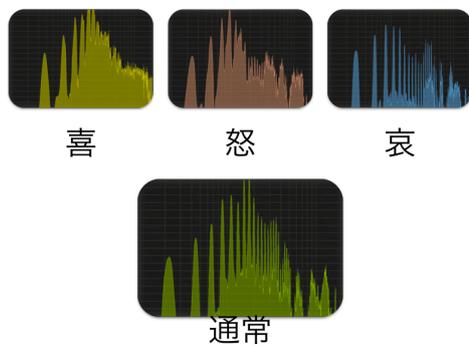


図 7 被験者女性の「あ」の音

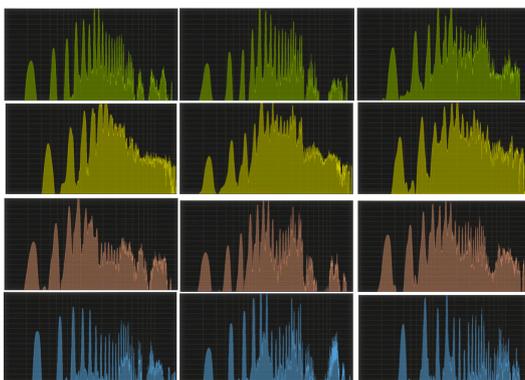


図 8 被験者女性の「あ」の音

#### 4. いい声作り支援のためのアプリ開発に向けて

本研究の目標は「いい声作り支援のためのアプリケーション開発」である。そのための最初のプロトタイプとして現在、ユーザの声の特徴を把握できるアプリケーションを開発中である。入力した声に対して、整数次倍音と非整数次倍音のどちらの成分が優位であるかを、周波数のグラフではなく、「あなたの声はカリスマ性がある」「あなたの声は親密性がある」といった言葉や、声の特徴が似ている芸能人の画像を用いて表現する。このアプリケーションによって、エンタテインメント性が高く親しみやすいインタフェースによっていい声作りを支援できると考えられる。

また事前調査の結果より、感情の変化によって異なる周波数分布を確認することができた。このことから、通常時つまりリラックスしている状態に記録した周波数分布と、新たに記録した声を比較することで、現在の状態が緊張状態にあるのか興奮状態にあるのかを判別することが可能になるのではないかと考える。さらに、ただ倍音が強調されているだけではなく、高い周波数にわたっても倍音を含んでいることがいい声と判定するためのひとつの判断基準になり得るのではないかと考えた。これらを考慮に入れ今後の開発に取り組んでいきたい。

将来的には、情報可視化の手法などを採用して、単に声の特徴を知らせるだけでなく、もっと自在にユーザが自分の声色を観察して制御できるためのアプリ開発を目指したい。

#### 5. まとめ

本報告では、いい声作り支援のためのアプリ開発に向けて、事前調査結果を報告し、今後のアプリ開発への指針を述べた。本研究では、声色の観点からのいい声を判定する指標として倍音を採用した。事前調査の結果より、特定の芸能人に対して、彼らの声の特徴が倍音に表れていることが明らかになった。また感情がどのように周波数に影響するのかを調べた結果、基音の周波数の差異や倍音の継続に見られる特徴など、多くの知見を得ることができた。

今後はアプリの開発に力を入れ、意見を仰いでいきたい。

#### 参考文献

- [1] A. Mehrabian: *Silent messages*, Wadsworth, Belmont, California (1971).
- [2] 中村明一: 倍音 音・ことば・身体の文化誌, 春秋社 (2010).
- [3] 矢島佳澄, 笈康明, 諏訪正樹: 発声のメタ認知促進システム“いい声マイク”の提案, 情報処理学会インタラクシオン 2011 (2011).
- [4] 声総研: モテ声診断 VQ チェッカー, <http://www.koesouken.com/vqchecker/>(2001).
- [5] C. Unander-Scharin, K. Hook, L. Elblaus: *The Throat III - Disforming Operatic Voices Through a Novel In-*

*teractive Instrument*, CHI 2013 Extended Abstracts on Human Factors in Computing Systems, pp. 3007-3010 (2013).

- [6] 中野倫靖, 後藤真孝: VocaListener2: ユーザ歌唱の音高と音量だけでなく声色変化も真似る歌声合成システムの提案, 情報処理学会研究報告, 音楽情報科学, Vol.2010-MUS-86, NO. 3, pp. 1-10 (2010).