

コトダマの可視化に関する一手法

岩淵 勇樹^{1,2,3} 小松原 宏識³ 西田 涼麻^{2,3} 中丸 潤^{2,3}

概要: 本稿は、音声信号を元として、言葉に宿る魂を意味する「コトダマ」の可視化の基本技術を提案する。解析信号を利用し、音高を図形の色に割り当てることにより、コトダマの色彩豊かな表現が可能となった。提案法を用いて、声と楽曲の可視化を試み、その結果について考察した。この技術を利用して、ライブ演出やインタラクティブ作品への応用が可能となる。

A Method on Kotodama Visualization

YU(U)KI IWABUCHI^{1,2,3} HIROSHI KOMATSUBARA³ RYOMA NISHIDA^{2,3} JUN NAKAMARU^{2,3}

Abstract: We propose a method of “kotodama” (word soul) visualization. It is possible to express colorful kotodama, using analytic signal and mapping color of shape from pitch of sound. We tested visualizing from voices and music, also considered its results. Using this technology, It is possible to apply live performance or interactive art works.

1. はじめに

コトダマとは、映画『きみの声をとどけたい』^{*1}の作品内で登場する重要なキーワードで、言葉に宿る魂を意味する。本研究による音声の可視化はこの作品から着想を得ている。

本研究では、この作品に登場するコトダマを信じる人々が主体となり、ファン同士での制作と交流を目的とした“コトダマ部”の活動の一環として、音声を用いた視覚表現によるライブ演出やインタラクティブ作品の制作を目的とした基礎技術を提案する。

2. 音声と視覚の関係性

人間がもつ五感のうち、視覚と聴覚は特に重要な役割を占めており、歴史的に見ても、学術、芸術、商業等あらゆる面で、視覚情報と聴覚情報に関する技術が中心的役割を担っている。また、映画における劇伴やミュージックビデ



図 1 コトダマの映像表現の一例

Fig. 1 Example picture expression about “Kotodama” [1]

オに代表されるように、現代のエンターテインメントに視覚表現においては音楽、聴覚表現においては映像が必要不可欠と言っても過言ではない。コミュニケーションの根幹である「言葉」についても、手話や点字等はあくまでも聴覚や視覚に障がいがある場合の代替手段に過ぎず、文字と話し言葉が最も一般的である。

視覚と聴覚は本来、視覚表現や音楽表現が与える感情変化や、文字と話し言葉が脳内で相互に影響を与えてはじめて関係するものであるが、ごくまれに、視覚情報が聴覚として、聴覚情報が視覚として直接想起される共感覚という

¹ 面白法人カヤック

KAYAC inc., Yokohama, Kanagawa 220-0011, Japan

² カマコン

KAMACON

³ コトダマ部

Kotodamabu

^{*1} <http://kimikoe.com/movie/>

特殊な感覚が生じうる*2。音（音高、音階など）に対して色を想起する例 [4] は共感覚の中でも多く報告されており、音楽を聴くと絵画的なイメージが想起され、それをほぼイメージのまま絵画として表現できるという画家も実在する。コトダマを見る行為は、言葉の持つ魂と、それを見る能力、あるいは見られるほどの強さが前提となる。言葉に魂が宿っているかどうかはまた別の議論が必要であるが、声の実体のように感じられる現象はおとぎ話の世界の中だけのことと言い切ることはできないであろう。

岩淵は、心理学的アプローチとは異なり、工学的アプローチで図形と音声を結びつける研究に取り組み、解析信号という技術が有用であることを示した [2]。

図形を基にした音色生成の研究のひとつとして、Levinらのメディアアート作品が挙げられる [3]。この研究では、シルエット図形の面積や周囲長などの特徴量を音声に割り当てている。

3. 原理と実装方法

解析信号を用いて、音声の原信号を実部、ヒルベルト変換を虚部としてもつ図形を描く [2]。

解析信号の性質に従って、音量が図形の大きさ、音色が図形の形状に反映される。

解析信号の実部と虚部を複素平面上に投影したグラフは、潰れている時間軸上の音高情報が欠落しているため、音の三要素である音量（音の大きさ）、音高（音の高さ）、音色 [5] のうち、残っている音高を今回は色に割り当てた。また、具体的な色へのマッピング方法として、ドを赤として1オクターブで色相環を1周する設計を行った。この音高のオクターブと色相環の関連付けは、周波数が2倍になると1周するという知覚の特徴において妥当であると考えられる。ただし、音と色のつながり [4] は人によって捉え方が大きく分かれるため、この割り当ては必ずしも自然とはいえないことに注意されたい。なお、色相を扱うにはHSV表色系が簡便である。ただし、今回は彩度と明度は100%とし、可変のものとしては扱っていない。

4. システム実装

本研究のデモはHTML5技術*3*4を用いて実装した。音声情報処理技術としてWeb Audio API *5を採用した。可視化技術としてはCanvas APIとSVGを検討し、本デモでは視覚的表現の試行錯誤が容易であったためSVGを採用した。

デモおよびソースコードはWeb上に公開している*6。

*2 本稿では視覚聴覚感の共感覚について述べるが、触覚と聴覚など別の組み合わせの感覚が結びつくという共感覚者も存在する。

*3 WHATWG <https://whatwg.org/>

*4 W3C <https://www.w3.org/>

*5 <https://www.w3.org/TR/webaudio/>

*6 <https://kotodamabu.github.io/interaction-ipsj-2018/>

5. 検証

以下に声を可視化した例と音楽を可視化した例を示す。

表 1 声の可視化結果

Table 1 Result of voice visualization.

「あ」			
「お」			
「い」			
	ド (C2)	ミ (E2)	ソ (G2)

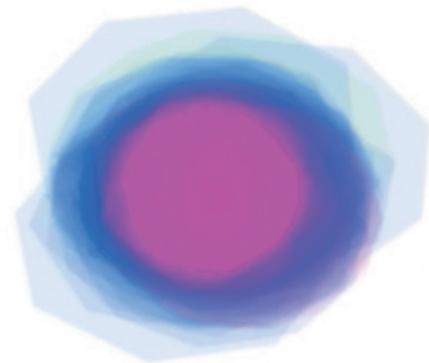


図 2 楽曲の可視化結果

Fig. 2 Result of music visualization.

表 1 は、同一の男声について 130 Hz (ド: C2), 164 Hz (ソ: E2), 195 Hz (ミ: G2) の 3 つの音高で、数秒の持続音として発生した「あ」「お」「い」の 3 つの母音を可視化し、代表的な瞬間を切り取ったものである。

図 2 は、伴奏入りの女性ボーカル曲を PC の内蔵マイク経由で入力した音声を可視化し、得られた動画像から代表的な静止画を切り出したものである。

6. 結果と考察

声については母音・音高ともに明確に識別できるような結果を期待していたが、見た瞬間に識別できる程度の結果は得られなかった。同じ母音同士について形状が異なることは、声質が不変であっても音高によってスペクトル成分が変化することによる形状位変化が確認されているが、形状に大きな違いが現れた原因はピッチと声量のぶれの他、PC内蔵マイクの補正等も影響しているものと考えられる。同じ音高同士について色が異なることは、今回扱った色の範囲が純色に限られたため、表現される色数が乏しく、音高が大きく変化していてもほぼ同一の色として表現されていたり、逆に小さなピッチのぶれが色に大きな変化を与えているのではないかと考えられる。

一方、楽曲については同時に様々な色が出現するような可視化がなされた。これまでの研究では、解析信号は主に単音を扱った場合に形状に特徴が現れ、複数の音が重なった場合には複雑な形状を示すため、複数の音を扱うことは視野に入れられなかったが、今回の色へのマッピングを加えることにより、色彩豊かな結果が見て取れた。当初の見通しでは、楽曲の可視化は色や形状の変化が激しいことを予想しており、見栄えについての懸念があったが、全体を通して見ると円形に近い塊が次々と現れるような効果が生じたため、著者の印象としては汚さや不快さは感じなかった。

7. 考えられる応用

短い単音の発声については、ピッチに対応する色は安定しなかったものの、はっきりとした単色の塊となりやすいため、発声ごと、あるいは秒数ごとに1つの音のブロックを生成し、それを複数個集めて鍵盤のようにタイミングよくタッチして音楽を奏でるなど、ゲーム要素を加えたインタラクティブ作品への応用を検討したい。

歌唱については、今回の色のマッピングにおいては色相しか用いなかったが、音声の三要素以外に抑揚やビブラートなどの特徴量を、明度、彩度、あるいは映像フィルタなどに反映させ、より直感的な表現の実現に取り組みたい。その他、これまで深く考察することができなかった子音発声時の過渡的な変化も取り入れていきたい。

また、複数人の歌唱によるライブ演出として、各マイクの入力音声の可視化結果と全体音の可視化結果を個別に提示することにより、ステージ全体と歌唱者それぞれに注目して音の変化を見て楽しむような効果も期待される。

なお、本研究のデモは、HTML5技術によって開発されたため、PC、スマートフォン、タブレット、VRヘッドセット等で動作するWebアプリケーションへの応用が容易である。

8. まとめ

本研究では、コトダマの可視化を試み、解析信号を用いて一つの可視化方法を提示した。解析信号の原理および、音声の三要素と共感覚の知見に従い、音声における音量、音色、音高をそれぞれ図形の大きさ、形状、色（色相）に対応付けた。結果として、解析信号が視覚表現を目的とした可視化技術としても有用であることが示された。

今後は、コトダマ部としてライブイベントやデモ展示を行い、ファンが作るアート、エンターテインメントの可能性を広げる所存である。

謝辞 こしごえ聖地化プロジェクトの町田隆彦氏、小笠原弘高氏、武富直之氏、メンバーのみんな、いつもありがとう。コトダマ部の皆さん、日ノ坂町民の皆さん、NOW ON AIRの皆さん、そして映画キミコエを生み出してくださったすべての人に、感謝のコトダマが届きますように。

参考文献

- [1] 映画『きみの声をとどけたい』本予告
<https://youtu.be/MuZL3QPRtrQ>
- [2] 岩淵勇樹, 秋田純一, 北川章夫: 閉曲線を利用した音色の操作法, 芸術科学会論文誌 (2011).
- [3] Levin, G, Lieberman, Z: *Sounds from shapes: Audiovisual performance with hand silhouette contours in the manual input sessions*, Proceedings of 5th International Conference on New Interfaces for Musical Expression (2005).
- [4] 長田典子: 音を聴くと色が見える: 共感覚のクロスモダリティ, 日本色彩学会誌 (2010)
- [5] JIS Z 8106:2000 音響用語, 日本規格協会, https://webdesk.jisa.or.jp/books/W11M0090/index/?bunsho_id=JIS%20Z%208106:2000