

モデル生物 *C. ELEGANS* の動態可聴化におけるサウンドデザイン SOUND-DESIGN FOR THE DATA SONIFICATION OF *C. ELEGANS*' DYNAMIC MOTION

寺澤 洋子

Hiroko Terasawa

筑波大学 / JST さきがけ

University of Tsukuba / JST-PRESTO

濱野 峻行

Takayuki Hamano

JST ERATO 岡ノ谷情動情報プロジェクト

JST, ERATO OEIP

概要

この原稿では、モデル生物である線虫 *C. elegans* の動態の可聴化について述べる。野生型および遺伝子組み換え型の線虫の動態を収録したビデオデータを、動き検出とグラニューラー合成を用いて音に変換した際のサウンドデザインの詳細について報告し、実装における今後の指針について議論する。

This paper discusses the sonification of *C. elegans*' dynamic motion. The video data of wild type and transgenic type worms were sonified using motion detection and granular synthesis. We describe the technical details of the sound mapping, and discuss the directions for the further implementations.

1. はじめに

本稿では、分子生物学におけるモデル生物として頻繁に利用される線虫 *C. elegans* の動態可聴化について報告する。

近年の生命科学においては、高精細ビデオ技術の進歩と蛍光タンパク質の利用によって、生体の様子をダイナミックに観察する事が可能になっている。このような技術によって、これまでの静的データに基づいた研究手法から、動的データに基づいた研究手法への移行がなされているが、未だに目視による観察・分析が主であり、計算機を利用した情報処理あるいは視覚以外の媒体を利用したディスプレイなどには程遠いのが現状である。そこで我々の生命動態可聴化プロジェクトでは (1) 可聴化によって、生体の振る舞いを新たな角度から観察する手法を確立し、新たな知見へつなげること、そして (2) 生体の振る舞いからヒントを得て新たな音楽表現を探索すること、の二つを目的として、生命動態の可聴化を行う。そのための初期検討として、遺伝子組み換え線虫の動態のビデオデータの可聴化に取り組んだ [1]。生命科学の

研究者たちにメリットが理解しやすいよう、「目に見える動きが音として聴こえてくる」ような可聴化を目的とした。本稿では、この可聴化におけるサウンドデザインについて詳述する。

2. 背景

2.1. モデル生物 *C. elegans*

Caenorhabditis elegans (*C. elegans*) は体長 1mm 程度の透明な線虫であり、959 個の体細胞を持つ。遺伝子配列が全て解読されており、またライフサイクルが短いために実験期間が短い利点があり、遺伝子組み換え技術を利用した発達・発生、疾病、寿命に関する研究に頻繁に使われる。2002 年には、*C. elegans* をモデル生物として用いたアポトーシス (programmed cell death) に関する研究によって、ブレナー・ホロビッツ・サルストンが連名でノーベル生理学・医学賞を受賞している。

2.2. 画像・動画の可聴化

画像の可聴化には、ラスタースキャンが頻繁に使われてきた。これは静止画のピクセルデータを走査線を横一列に並べたベクトルとして表現し、それを時間波形と見なして音を出す方法である [2, 3]。また、動画の可聴化を行う際は、フレームごとに対応した時間波形を次々と再生することになる。

しかしこの方法では、画像上のオブジェクトなどの視覚的に意味ある情報が、聴覚的に理解しやすい意味ある情報として音に反映されることがない。これは、新しい音を生み出すための手法としては有意義であり、経験豊かな聴取者が気をつけて聞けば、映像と音との対応の理解も可能ではあるが、音楽経験のない聴取者が聞いて対応関係を理解するのは非常に難しい。その理由として、視覚オブジェクトが聴覚オブジェクトとして関連づけ

られていないことが挙げられるだろう。たとえば、上下に位置するピクセルは、視覚的には隣接しているが、ラスタースキャンを用いると一走査線分離れてしまい、再生時にもそれだけ時間的に離れて提示されてしまう。このような映像の二次元平面における連続性（動画の場合は、加えて時間軸における連続性も発生する）とそれによって生じるゲシュタルト認知を、音表現において保持できないことが、ラスタースキャンによる可聴化が直観的に理解されない理由であろう。

そこで、音による直観的な情報理解を促すためには、画像処理・動画処理を利用して視覚オブジェクトのサイズ・スピード・位置など、ヒトにとって直観的に意味ある情報を抽出し、それを利用して可聴化を行うことが必要になる。そのような可聴化の例として、Pelletierによって行われた画像処理による視覚オブジェクト抽出に基づく手法を挙げたい。このアプローチでは、コーナー検出法と特徴点抽出法によって、角や終点などの画像上の視覚的に重要な点を抽出し、その移動に基づいて可聴化を行っている [4]。このシステムでは様々な音合成手法を柔軟に取り入れられるが、そのなかでも、グラニューラー合成が大量のデータポイントの集合を表現するのに適した方法として紹介されている。グラニューラー合成ではウェーブレットを集めたクラスタによって、より大きな音単位を構成するので、大量のデータポイントを示す傾向がある画像データの可聴化には有効であると考えられる。

3. C. ELEGANS の動態可聴化

3.1. 映像素材と処理環境

今回の可聴化では、筑波大学生命領域学際研究センター深水研究室より提供を受けた野生型および遺伝子組み換え型の *C. elegans* の映像を素材として用いた。表 1 にビデオデータのリストを示す。また、これらのスナップショットを図 1 に示す。

表 1. 可聴化に利用したビデオデータ

<i>C. elegans</i> の種別	ビデオ
Wild (野生型)	A, C
Red Fluorescent (遺伝子組み換え・蛍光)	B
Rolling (遺伝子組み換え・Rol 表現型)	D

これらのビデオでは、*C. elegans* の個体の動態が収録されており *C. elegans* に特徴的なうねるような動きを見せている。野生型のほか、遺伝子組み換えタイプが二種準備された。蛍光タイプでは、咽頭の部分に蛍光タンパク質が発現しており、その部分だけ赤色に光る。また、Rol 表現型では、運動機能を変化させてあり、円形

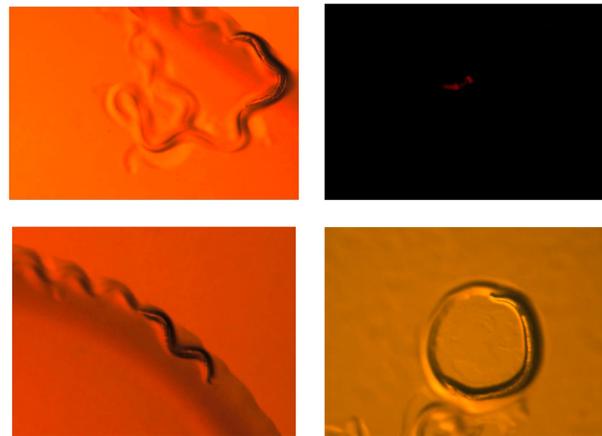


図 1. ビデオデータのスナップショット。左上から A, B, C, D に対応。

の軌跡を描いて動く様子が観察される。

これらのビデオを可聴化するにあたって、(1) 映像と音合成の処理が容易、かつ (2) 視覚的に多領域の研究者も親しみを感じやすく、議論に参加しやすいプラットフォームが必要であり、Max/MSP/Jitter を採用して実装を行った。

3.2. 動き検出を用いた動画の下処理

今回の目標である「目に見える動きが音として聴こえてくる」ような可聴化を行うにはまず視覚的に意味ある情報を抽出する必要がある。そこで、簡単な動き検出を行い、線虫の動きに対応するピクセルデータを抽出した。その流れを以下に示す。

1. フレームレートを 25fps に変換
2. フレームサイズを 320 x 240 に縮小
3. 現在のフレームと一つ前のフレームのラスタデータをグレースケール変換し、その差と絶対値を求める
4. 絶対値を 4 乗し、動きの大きいピクセルを強調
5. しきい値 (0.05 - 0.1) より低い値を示すピクセルをゼロで置き換え
6. 一次 IIR フィルタ (jit.slide オブジェクト) を用いて前後のフレーム間で円滑化を行い、突発的なノイズを抑制
7. フレームサイズを 80 x 60 に更に縮小

これらの処理のうち、(1), (2) の手順は、計算負荷を下げるために動画のサイズを小さくする工程、(3) が動き検出の工程であり、(4)-(6) はこのあとの音合成の段階で必要としないデータ (突発的ノイズあるいは値が小さすぎて音にする意味がないデータ) をそぎ落として、データ伝送の負担を減らす工程である。また、(7) では、

フレームサイズを更に小さくしその後の音合成の計算負荷が大きくなりすぎないように調整している。

これらを実装したパッチのスクリーンショットを図2に示す。右上の映像が(2)のサイズ縮小後、左上の黒い背景で解像度の細かい映像が(6)の処理後、そして左下の解像度の荒い映像が(7)の処理後に対応する。このような処理を経ることで、背景が取り除かれ線虫の動きに対応するデータが抽出されていることがわかる。

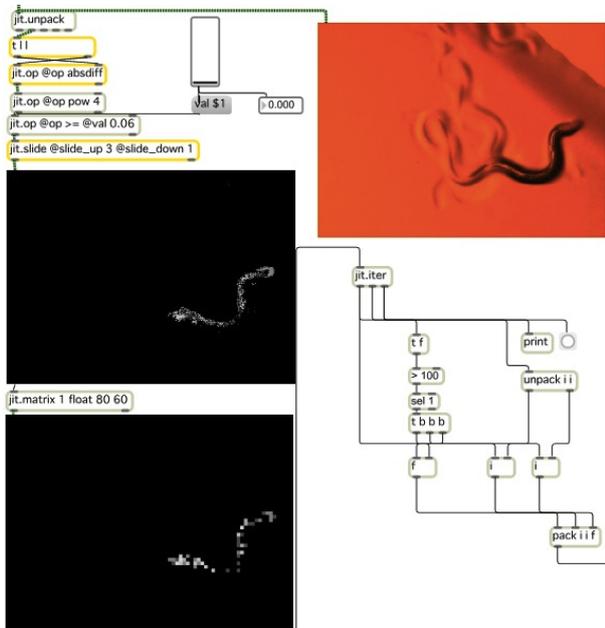


図2. 動き検出とノイズ抑圧の実装

3.3. グラニューラー合成

第2.2節でも述べたように、映像や動画の可聴化における課題は、視覚オブジェクトを聴覚オブジェクトに変換することであり、その目的にはグラニューラー合成が有益である。そこで、今回は、動き検出で得られたピクセル集合に対応するようなサウンドクラスタを生成することとした。また、マッピングに関しては、データの変化が感覚量の変化として感じられるように、音合成のパラメタを対数的に決定した。

マッピングの概略を図3に示す。フレームの上下が音高の上下に、そして左右は音のウェーブレットの立ち上がりと継続時間に対応するようにした。

サウンドクラスタの生成手順を以下に述べる。

まず、ある時刻 t でのフレームに着目すると、動き検出後のデータではピクセルごとにピクセル値 $v(x, y, t)$ が与えられる。ここで (x, y) はフレーム内での位置座標である。各ピクセルごとにウェーブレット $w(x, y, t)$ を生成したのち、フレームごとに生成された全てのウェーブレットを加算した音 $s(t)$ を合成音とする。

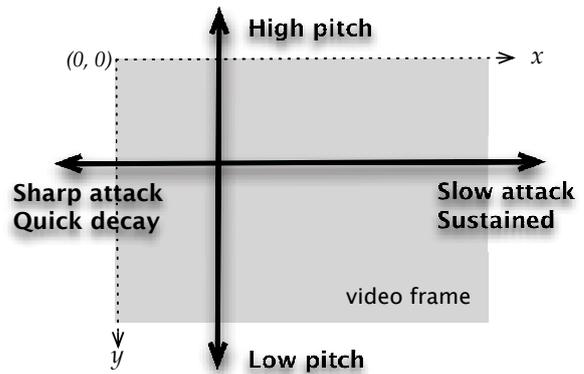


図3. グラニューラー合成のパラメタマッピング

ウェーブレットの周波数 f_w は垂直方向の座標 y によって決定される。

$$f_w(y) = 800 \cdot 2^{-y/60} \quad (1)$$

ウェーブレットの立ち上がり時間 T_a と減衰時間 T_d は水平方向の座標 x で決定される (単位: ミリ秒)。

$$T_a(x) = 10^{x/60} + 20 \quad (2)$$

$$T_d(x) = 2^{x/10+4} \quad (3)$$

周波数と立ち上がり・減衰時間では対数的なマッピングを採用したが、ウェーブレットの振幅は $v(x, y, t)$ をそのまま採用した。色の濃さに関しては、ウェーバー・フェヒナー則に基づいた対数的な知覚がされているので、色の濃さを表すピクセル値をそのまま振幅に利用しても、視覚と聴覚で同様の変化が感じ取れるためである。

ピクセル毎に生成されるウェーブレット $w(x, y, t)$ は、周波数 f_w の正弦波の振幅を変化させて作られる。正弦波の振幅は、 T_a の時間をかけて0から $v(x, y, t)$ に線形に変化し、続いて T_d の時間をかけて振幅0に戻るものとした。ウェーブレットの時間長は、 $T_a + T_d$ となる。

このように、フレーム全体から生成された全てのウェーブレット $w(x, y, t)$ を足し合わせたサウンドクラスタが時刻 t での音 $s(t)$ となる。

$$s(t) = \sum_{x=1}^{80} \sum_{y=1}^{60} w(x, y, t) \quad (4)$$

これを時間的に連続した全てのフレームに対して行くと、動きに伴って音に変化するサウンドクラスタが合成される。

実装にあたって、 $v(x, y, t)$ の値が小さい場合には、ウェーブレットの音量も小さくなるため、しきい値を設

定して $v(x, y, t)$ が十分大きい場合のみウェーブレットを生成するようにし、計算負荷を減らした。

4. 議論と今後の課題

4.1. 評価

生命科学と音響学の研究者と大学院生らに参加してもらった評価実験を行った。音をつけたビデオを提示した後、音のみを聞いて、可聴化の元になったビデオデータを A, B, C, D から予想してもらおうというタスクで、同定率はほぼ 100% であった。この結果に関する考察は文献 [1] に詳述したので、そちらを参照されたい。

4.2. Max/MSP/Jitter による実装

この研究の実施に際して、(1) 映像と音合成の処理が容易、かつ (2) 視覚的に多領域の研究者も親しみを感じやすく、議論に参加しやすいプラットフォームが必要であり、Max/MSP/Jitter を採用して実装を行った。特に後者 (2) のメリットは非常に大きく、生命系の研究者にとって安心感のある GUI であり、リアルタイムのデモも簡単にでき、途中経過の説明がわかりやすく受け入れてもらえた。円滑なコミュニケーションが図れるため、学際研究のためのプラットフォームとして非常に優れていると感じた。

しかしながら、映像の処理は計算負荷が高く、Max/MSP/Jitter では処理が間に合わないことも多かった。データ落ちが発生し、生成される音質にムラが出るなどの影響が見られた。このような問題は、第 3.2 節で述べたような、しきい値を設定して伝送するデータ量を減らすなどの工夫によってある程度は回避されたが、コンピュータの機種・状態に強く依存するため、リアルタイムのシステムを作るためには安定性と信頼性が低いと感じた。このような問題は、他の言語を利用するか、javascript によるプラグイン化を行うか、で対処する必要がある。他言語に切り替える場合は、前述した共同研究における利点が減少してしまうトレードオフが発生する。

4.3. マッピングの問題点

今回のマッピングでの問題は、左右の位置変化を立ち上がり時間と減衰時間の変化に対応させたところである。実際には線虫は右から左に少しずつ動くため、反対側に行く頃には、最初どんな音であったかわからなくなってしまい、変化があまり感じ取れなくなってしまった。画像全体における位置よりは、体のねじれなどに対応した相対的なマッピングを取り入れた方が効果が高くなる可能性を感じた。

5. まとめ

本稿では、遺伝子組み換えを行った *C. elegans* の可聴化におけるサウンドデザインについて詳細を報告した。今後の課題は多々あるものの、分子生物学への可聴化の応用は過去に例がなく、新規性が高い研究である。今後は、新しい知見を発見出来るような可聴化を達成するべく、更なる試作と検討を繰り返して行く予定である。

6. 謝辞

本研究を実施するにあたって、データ提供や試作に協力いただいた筑波大学生命領域学際センターの深水昭吉先生、廣田恵子博士、高橋裕太博士に深く感謝する。また、同センターの牧野昭二先生、山田武志先生には、様々なアドバイスと支援を頂いた。本研究は、カワイサウンド技術・音楽振興財団および科学技術振興機構さきがけの研究助成を受けて実施された。

7. 参考文献

- [1] Terasawa *et al.*, “*C. elegans* meets data sonification: Can we hear its elegant movement?” in Proceedings of Sound and Music Computing Conference, 2011.
- [2] Yeo and Berger, “Raster Scanning: A New Approach to Image Sonification” in Proceedings of the International Computer Music Conference, 2006.
- [3] Jo and Nagano, “Monalisa: See the Sound, Hear the Image” in Proceedings of International Conference on New Interfaces for Musical Expression, 2008.
- [4] Pelletier, “Perceptually Motivated Sonification of Moving Images” in Proceedings of the International Computer Music Conference, 2009.

8. 著者プロフィール

寺澤洋子 (Hiroko Terasawa)

筑波大学生命領域学際研究センター研究員。Ph.D. (Music). 電気通信大学電子工学科卒・同大学大学院修士課程電子工学専攻修了。学部在学中にフランス国立高等情報通信学校 (Telecom Paris Tech) に交換留学。NTT コミュニケーション科学基礎研究所特別研修生を経て、米国スタンフォード大学音楽学科 CCRMA にて修士課程・博士課程修了。在学中にパリ国際芸術都市アーティスト・イン・レジデンスおよび IRCAM 訪問研究者として滞在。2010 年より現職。2011 年 4 月より東京芸術大学芸術情報センター非常勤講師、2011 年 10 月より科学技術振興機構さきがけ研究員 (兼任)。