

生命動態可聴化の現状と課題

寺澤 洋子^{1,2,a)} 松原 正樹¹ 牧野 昭二¹

概要：聴覚は視覚よりも時間変化の理解に優れているため、データ可聴化（本来、音でないデータを音に変換する）は時系列情報の表現に本質的に適する。生命動態の研究では、可視化や統計を用いた手法が主流であり、可聴化などそれ以外のメディアを用いたデータ観察は殆どされていない現状がある。こういった中、可聴化がデータ観察のための手法として普及していくには、可聴化のデザイン理論の確立と、説得ある具体例の提示が必要不可欠である。そこで、本稿では、可聴化をデザインするとはどういうことか、そして説得力ある手法であると認められる為にはどのような評価が必要であるか、のお互いに関連する二点について具体例を提示しながら検討する。

キーワード：可聴化、評価手法、聴覚ゲシュタルト、デザイン理論、時系列データ、生命科学

Current Issues in Biological Dynamic Data Sonification

TERASAWA HIROKO^{1,2,a)} MATSUBARA MASAKI¹ MAKINO SHOJI¹

Abstract: Since auditory perception is superior to visual perception in noticing dynamic changes, data sonification is suitable to represent time-series data. While current research methods in life sciences strongly rely on visualization and statistical data analysis, methodologies using other modalities, such as data sonification, are less investigated. In such a situation, sonification design theories and more convincing examples are desirable in order to encourage the further use of data sonification. In this paper, we discuss the principles of sonification design and evaluation methods using examples from the sonification of dynamic biological information.

Keywords: sonification, evaluation methods, auditory gestalt formation, design theory, dynamic data, life science

1. はじめに

聴覚は視覚よりも時間変化の理解に優れているため、データ可聴化（本来、音でないデータを音に変換する）は時系列情報の表現に本質的に適する。そこで、我々のプロジェクトでは「時間変化する事象」の代表例として生命動態をとりあげ、その可聴化に取り組んでいる。生命動態の研究では、可視化や統計を用いた手法が主流であり、可聴化などそれ以外のメディアを用いたデータ観察は殆どされてい

ない現状がある。こういった中、可聴化がデータ観察のための手法として普及していくには、可聴化のデザイン理論の確立と、多数の説得ある具体例の提示が必要不可欠である。そこで、本稿では、可聴化をデザインするとはどういうことか、そして説得力ある手法であると認められる為にはどのような評価が必要であるか、のお互いに関連する二点について具体例を提示しながら検討する。

2. 可聴化のデザインと評価

可聴化とはどういうことであろうか、まずは定義から紹介していきたい。以下は、少々古いが、データ可聴化において古典とされる“The Sonification Report”[1]からの引

¹ 筑波大学生命領域学際研究センター
Life Science Center of Tsukuba Advanced Research Alliance,
University of Tsukuba

² JST さきがけ
JST PRESTO

^{a)} terasawa@tara.tsukuba.ac.jp

用である(原文*1)。

可聴化とは、コミュニケーションや解釈を容易にするために、データにおける関係性を、音響信号において知覚される関係性に変換することである。

つまり、可聴化では、本来は音ではないデータを音に変換する時に、「データにおける関係性」が音の関係性として「知覚される」ように変換する必要がある。しかし、たとえ関係性が知覚されたとしても、音への変換が随意に(例えば「テキトーな」「毎回違う」などのマッピングで)なされていたら、知覚された関係性への信頼性が揺らいでいく。信頼性を確保するためには、科学的なデータ観察に必要とされる「客観性」「再現性」「正確さ」などの要件を満たしている必要がある。これらの観点を含めた Thomas Hermann による定義は以下のようなになる(原文*2) [2]。

データを入力とし、音響信号を生成する技術(制御を伴う場合もある)は、以下の条件を満たすと可聴化と呼ばれる。(C1) 音は入力データの客観的な特性や関係性を反映する。(C2) 音への変換はシステマティックである。つまり、データ(およびインタラクションによる制御)と音の変化の因果関係が明確に定義されている。(C3) 可聴化は再現可能である。全く同じデータと全く同じ制御によって生成された音は、構造的に等しくなければならない。(C4) 可聴化のシステムは、他のデータを入力としても利用でき、そして同じデータを繰り返しても利用できる。

これらの定義から読み取れるのは、可聴化とデータを用いる音楽生成は、同じようにデータから音を生成する行為であっても、目的が本質的に異なるということである。厳密には可聴化でなくとも、データから生み出された音楽は“data-inspired music”あるいは“data-controlled music”として区別されるべきだと Hermann は述べる [2]。また、データによって生み出される音響効果なども同様に可聴化とは区別されるべきであろう。

音楽を生み出す目的は多様である。しかし作曲は基本的には芸術的な創作であり、客観性や再現性は制作者の美学として採用されるかもしれないが、多くの聴衆をはじめと

する音楽を評論する立場からすれば、必ずしも必要とされる条件ではない。しかし、可聴化においては「音を聴く人」は、客観性や再現性などを前提として「データを観察する人」であり、コンサートで音楽を聴く聴衆とは音を通じたデータとの関わり方が明らかに異なる。

データを観察する場合、データの特徴を理解することが目的となる。例えば、グラフによってデータを理解したい場合、抽出したい特徴に合わせて異なったグラフ化の手法(x-y プロット、棒グラフ、パイチャート、離散図など)を用いて、データの特徴を浮かび上がらせるというプロセスをふむ [6]。つまり、グラフ化の手法自体に、ある特定の条件を満たしたデータ構造を提示するフィルタのような機能が備わっており、観察者はその機能を(意識的あるいは無意識的に)理解した上で、グラフを読み取ってデータの解釈と理解を行う。

しかし可聴化研究では、多くの場合、プロットのように既に一般的に普及している手法ではなく、新規に音合成の手法やマッピングを考案して行われ、データ解釈・理解に関しても新規に検証される必要がある。その場合、

- (1) システムの動作原理を論理的に検証する
- (2) 生成された音が期待する機能を満たしているかどうかを検証する

の大別して二つのアプローチが可能であろう。前者は演繹的であり、体系化するとデザインセオリーの確立につながる。また、後者は帰納的であり、多数の事例が検討されることで、可聴化が社会に受け入れられやすくなる。

演繹と帰納は論証において補いあう働きをするため [7], [8], 可聴化が説得力をもって受け入れられるためには、この二つのアプローチを併用することが有益であると考えられる。しかし、数々の文献 [1], [3], [4], [5] で指摘されているように、演繹的な検証が可聴化の研究において欠落する傾向があり、特に auditory gestalt と auditory scene の生成と制御に関する演繹的研究は殆ど見受けられない。反面、帰納的な評価実験はいくつかの事例が検討されており、実験デザインの類型も見受けられる。これらの類型について次節で紹介する。

3. 評価方法の類型

本節では可聴化における効果の検証のあり方について述べる。国内においても [9], [10], [11], [12] の事例が報告されており、世界的にも様々な事例が展開されている EMG*3 信号データを用いた音響合成を対象とし、評価実験の方法の類型について検討する。Tanaka によれば EMG 信号の持つ自然で連続的な包絡信号の特徴が音合成時の音色のコントロールに適しているため音楽パフォーマンスにも応用が可能であると述べている [13]。これらの研究は data-inspired

*1 Sonification is the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation.

*2 Definition: A technique that uses data as input, and generates sound signals (eventually in response to optional additional excitation or triggering) may be called sonification, if and only if (C1) The sound reflects objective properties or relations in the input data. (C2) The transformation is systematic. This means that there is a precise definition provided of how the data (and optional interactions) cause the sound to change. (C3) The sonification is reproducible: given the same data and identical interactions (or triggers) the resulting sound has to be structurally identical. (C4) The system can intentionally be used with different data, and also be used in repetition with the same data.

*3 Electromyogram(筋電位)

music であるものの最近ではこうした音楽制作に対しても心理学的評価を行っていることから [14], 今後も心理計測による評価が期待できる。

一方で EMG 信号の可聴化によって身体の特徴的な構造を理解させることを目的とした研究がある。その目的は、バイオフィードバック療法時のリハビリテーションへの応用 [15], [16], 身体技能の熟達支援 [17], [18], [19], 個人の身体に適した癒しの音楽生成 [20] などと多様である。

これらの研究における有効性の検証には多くの研究が帰納的アプローチを採用しており、被験者による評価実験を行っている。評価のためのデータは多岐に渡るが、データ取得方法は主に以下の 3 種類に分類することができる。

外部観測 被験者の身体の状態など客観的に計測可能な数値情報をもとに、可聴化による支援システムの導入によってパフォーマンスが向上したかどうか評価を行う ([17], [18], [19])。外部観測を行うタイミングは、姿勢の時間的変化のようにリアルタイムに計測する (例えば [19]) 場合と、可聴化実験を通じたスポーツ能力の変化など一定の時間を置いて計測する場合 (例えば [17]) がある。客観的に数値を評価することでパフォーマンスの促進度合いを検証できるものの、生成された音の理解と実際に現れたパフォーマンスに因果関係があるかどうか注意する必要がある。

アンケート 街頭アンケートなど大人数の被験者に対し、理解の満足度などあらかじめ決められた項目についてアンケート調査を行い、数値的評価を行う ([16], [20])。決められた項目しか調べることができないものの手軽に大人数の理解度を評価することができる。アンケートを行うことによって影響がないよう統計的に検証を行う必要がある。

自由回答 データ可聴化を通じて理解したかどうかを測るために、被験者本人の内観を口述・回答してもらい (例えばインタビューや Think a loud 法? を用いて) 評価を行う ([17])。理解の結果としての行動を測ることが難しいデータなどに対し、主観的ではあるが被験者が理解したかどうかを口述により計測できるというメリットがあるが、評価のために労力がかかるという欠点を持つ。アンケートよりも自由回答が得られるので、可聴化による予想外の効果を発見することができる。

これらの類型からもわかるように、可聴化研究における有効性検証では被験者実験による帰納的なアプローチを用いることが多いが、演繹的アプローチを採用した論文は皆無と言って良いほど見当たらなかった。次節以降では、演繹的アプローチと帰納的アプローチの両方の事例を紹介し、それぞれの特長について概観する。

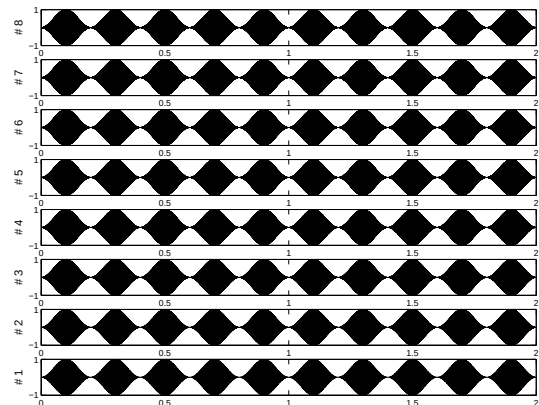


図 1 サウンド 1: 各行に倍音の振幅変調パターンの波形表示を示す。下から上に向かって、基音から第 8 倍音となる。この例では 8 個の倍音すべてで同じ振幅変調パターンを有する。

Fig. 1 Sound 1. Each row in the figure shows the amplitude pattern over time of each harmonic, from the first to the 8th harmonics from the bottom to the top row, respectively. This example has the same sinusoidal amplitude pattern for all the eight harmonics.

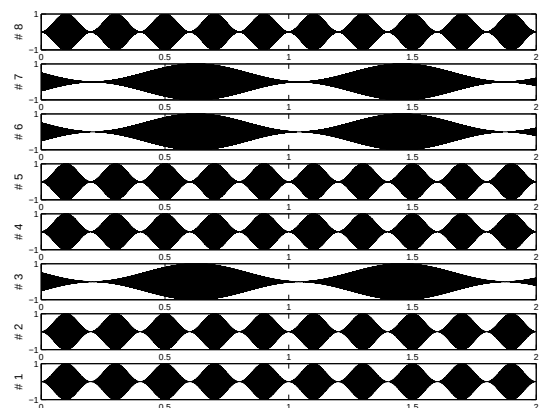


図 2 サウンド 2: 第 3, 6, 7 倍音は低い周波数で変調。

Fig. 2 Sound 2. The 3rd, 6th, 7th harmonics are modulated with slower frequency.

4. 研究事例紹介：Auditory Gestalt の生成による動態可聴化

第2節では、可聴化の有効性検証における演繹的アプローチと帰納的アプローチの概念を紹介したが、本節では我々が実施した二つの研究事例を紹介する。一つ目は、演繹的アプローチに重点を置いた例で、二つ目は、帰納的アプローチについて重点を置いた例である。どちらも例においても、音の組織化を考慮して可聴化がされている。

4.1 演繹的アプローチ：ECoG 脳波信号の可聴化

演繹的アプローチの例として、ECoG^{*4}脳波信号の可聴化研究 [22] を挙げる^{*5}。脳波信号の可聴化は既に多数報告されているが、この研究では相関のあるマルチチャンネル時系列信号から auditory gestalt (組織化された音像) [24] を生成する方法として、脳波の変動を倍音の振幅にマッピングする overtone mapping という手法を提案したものであり、auditory gestalt の生成・制御まで踏み込んだ部分に新規性がある。

この方法は、動作原理は単純だが、信号に相関がある場合は複数の倍音が融合して一つの auditory gestalt として認知され、相関が無い場合は複数の倍音が別々の auditory gestalt として認知される特徴を持つ。

合成の基本的な仕組みは、加算合成 [25] とほぼ同じであるが、マッピングにおいて、ゲシュタルト原理の「共有 (common fate)」の要因を利用し音の組織化が行われる。まず基本周波数を決定し、それから信号のチャンネル数と同じ数のハーモニクスを持つ、正弦波による倍音列を作る。信号のチャンネルと倍音列の倍音は対応関係にあり、信号の n 番目のチャンネルを用いて、 n 番目の倍音の振幅変調を行う。生成される音は、これらの倍音を足し合わせたものとなる。

これを擬似的なデータで実践したのが、図1～図3である。これらの図において、それぞれの行は振幅変調のパターンを表す。これらすべて、振幅変調には正弦波のパターンを用いているが、生成された音像は異なって認知される。図1の場合では、全ての倍音が全く同じ正弦波を用いた振幅変調であるが、図2の場合は、一部の倍音がより低い周波数の正弦波で振幅変調される。そして、図3の場合は、一部の倍音は同じ周波数であるが、異なった位相で変調されている。

聴取した場合は、図1のパターンによる音は、全ての倍音が融合して一つの組織化された音像として認知されるが、図2のパターンによる音では明らかに分離して認知さ

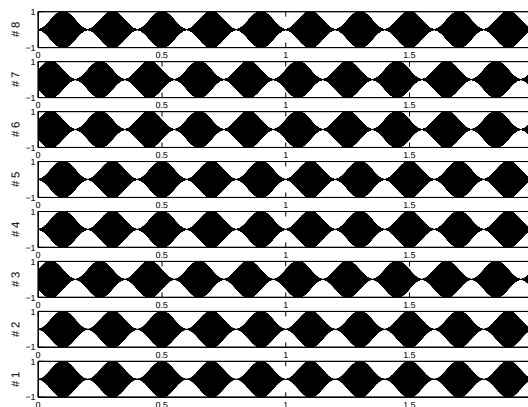


図3 サウンド3: 第3, 6, 7倍音は異なる位相で提示。

Fig. 3 Sound3: The 3rd, 6th, 7th harmonics differ only by their phase from the rest of the harmonics.

れる。また、図3のパターンによる音は、分離の度合いが低くなるが、完全に融合してはいない、いわば陰のついたような、にじんだような音像が生まれる。

この手法を用いることで、図4に示すような多チャンネルECoGデータにおいても、相関した信号は融合して認知される。このデータはてんかんの発作の際の信号を56チャンネルで計測したものであり、各チャンネルではシナプスを介した脳神経ネットワークによる相関した信号が得られている。この信号間で相関が強いチャンネルを3つのグループにわけて、それらに対応する部分音をプロットしたものが図5になる。これらのグループ毎の音を聴くと、複数の倍音が一つの音像を形成するものの、音像のなかの音色ゆらぎとして、各チャンネルの様相を聴き取る事ができる。これは岩宮らのフレームワークによれば動的音色と呼ばれるように、音の時間的な変化が音色を特徴付ける例である [26]。

この研究では、評価実験は行わず、演繹的アプローチに徹した有効性検証を行った。評価実験を行わないことには違和感もあったが、その分、auditory gestalt の生成と制御の議論に焦点をあてることができた。Auditory gestalt のコンセプトについては Bregman の文献 [24] に詳しく、またそれに着想を得た音源分離手法も多く提案されているが [27], [28], 合成音の制御に関しては、当たり前と思えるような概念でもいまだに殆ど論文化されていないのが現状である。これらの検証とドキュメンテーションを丁寧に行うことで、auditory gestalt の生成・制御に基づいた可聴化のデザイン理論を確立することが可能になると考える。

4.2 帰納的アプローチ：遺伝子組み換え線虫の動態可聴化

もう一つの例、遺伝子組み換え線虫の動態可聴化 [23] では、帰納的なアプローチに重点を置いた。

この可聴化では、遺伝子組み換えした線虫と、野生型

^{*4} Electrocorticography (皮質脳波検査)

^{*5} この研究に使用したデータはスタンフォード大学倫理委員会の監督のもと、スタンフォード大学病院において測定され、被験者の了解を受けた上で、筆者 (寺澤) の在学中に研究用に提供を受けたものである。

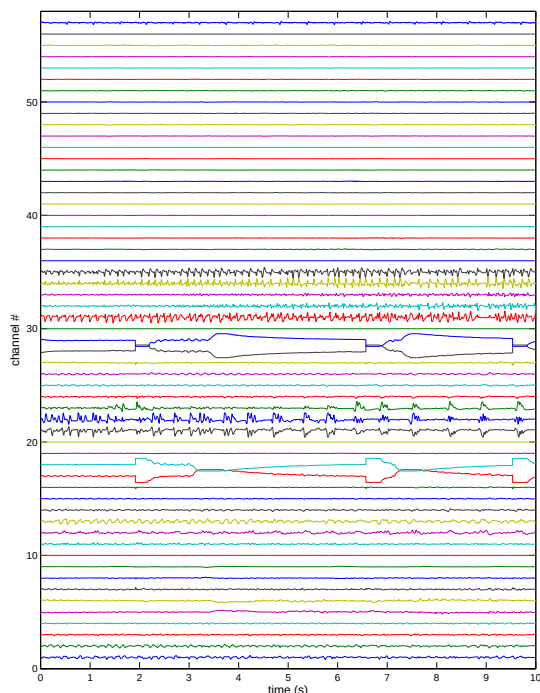


図 4 ECoG 波形データのプロット。各ラインはそれぞれのチャンネルに対応し、下から上にチャンネル 1～56 となる。

Fig. 4 Plot of 56 channel ECoG data for 10 seconds. Each line shows the signal for each channel, from the bottom to the top showing channel 1 to 56, respectively.

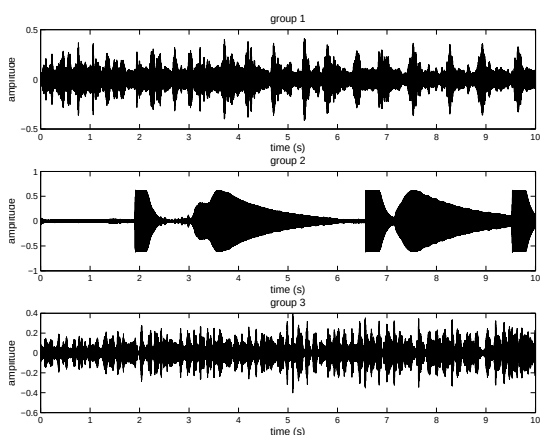


図 5 3つの相関の強い信号グループに対応する部分音の波形プロット

Fig. 5 Waveform plot of subset-tones of sonification: Group 1 (top), group 2 (middle), and group 3 (bottom).

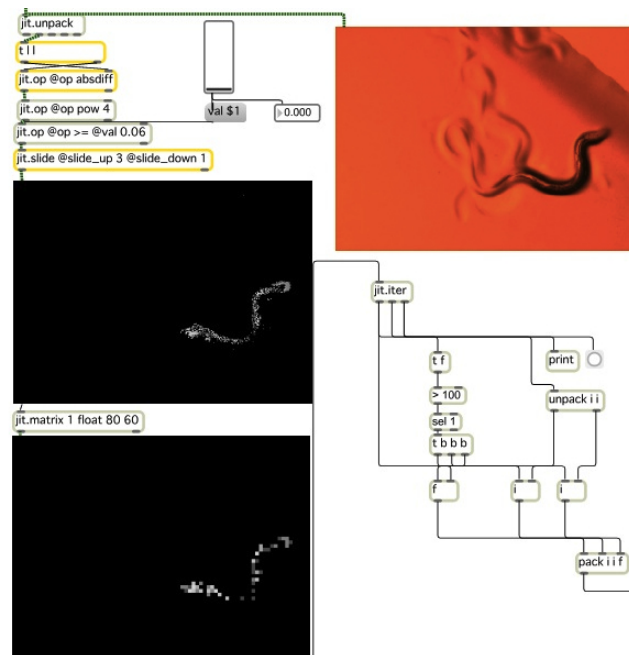


図 6 線虫の動態可聴化における動き検出。

Fig. 6 Motion detection in *C. Elegans* movement sonification.

の線虫のビデオデータを用いて、動き検出を行った後に、granular synthesis[29]によって可聴化を行った(図6)。動きが見られたピクセルクラスタに対応した音クラスタが生成され、ゲシュタルト原理の「近接 (proximity)」の要因によって音クラスタは融合した auditory gestalt として認知される。この研究では、演繹的な説明よりも評価実験に重点をおいて有効性を検証した。

評価実験では「線虫の動きが聴き取れるかどうか」に焦点を絞り、4つのビデオデータを用いて、練習の後に音からビデオデータを同定出来るかどうかを判定した。結果としては、ほぼ100%に近い正答率であった。同時に、インタビューによる質問も行い、どのような音の変化が同定のヒントになったかを検証した。

この研究の場合は、評価実験に重きをおいたために、理論的な検討が手薄になってしまった印象はあったが、評価実験によって有用可能性を数値化して訴求できたのは大きなメリットであった。可聴化研究では「本当にわかるのか」といった疑問が付きまとうものであるが、数値評価を行う事で、そういった疑念を払拭して建設的な議論を行う事が可能になった。

5. まとめと今後の課題

本稿においては、可聴化におけるデータ理解のプロセスについて考察し、その評価のあり方について演繹と帰納の両面から考察した。既存研究における評価手法について概観し、本プロジェクトでの可聴化における auditory gestalt の生成と制御に関する研究から、演繹的アプローチと帰納的アプローチを用いた例を紹介した。

今後の課題としては、演繹的アプローチと帰納的アプローチの両方を十分に組み合わせた評価をしていくことを挙げる。帰納的アプローチが可聴化の実用性を示すには有効であるが、デザインセオリーの確立にはやはり演繹的アプローチは欠かせず、両者が補いあってこそ、可聴化が説得力あるデータ提示方法として受け入れられる未来が生まれるだろう。しかし、本プロジェクトの事例をみてもわかるように、どちらかの評価手法に偏りがちであり、しかも、既存研究を見ても、演繹的アプローチはあまり追求されていない現状がある。これを打開することが、可聴化の更なる普及につながると考える。

謝辞 本研究は JST さきがけの支援を受けて行われている。また共同研究者のスタンフォード大学 Josef Parvizi 先生, Chris Chafe 先生, および筑波大学生命領域学際研究センター深水昭吉先生, 廣田恵子先生, 高橋悠太さんのご協力に深く感謝致します。脳波信号をご提供くださった匿名の被験者様に感謝申し上げます。

参考文献

- [1] Kramer, G., Walker, B., Bonebright, T., Cook, P., Flowers, J., Miner, N. and Neuhoff, J.: *The sonification report: Status of the field and research agenda*, Technical Report, Prepared for the National Science Foundation by members of the International Community for Auditory Display Editorial Committee and Co-Authors, (1999).
- [2] Hermann, T.: *Taxonomy And Definition For Sonification And Auditory Display*, in Proceedings of the 14th International Conference on Auditory Display, pp. 1–8, (2008).
- [3] Barrass, S. and Kramer, G.: *Using sonification*, Multimedia Systems, vol. 7, pp. 23–31, (1999).
- [4] Flowers, J. H.: *Thirteen years of reflection on auditory graphing: Promises, pitfalls, and potential new directions*, in Proceedings of the 2005 International Conference on Auditory Display, Ireland, (2005).
- [5] Hermann, T., Hunt, A. and Neuhoff, J. G. (Eds.): *The Sonification Handbook*, Logos Verlag, Berlin, Germany, (2011).
- [6] Tufte, E. R.: *The Visual Display of Quantitative Information*, Graphics Press, (2001).
- [7] Hawthorne, J.: *Inductive Logic*, in: Zalta, E. N. (Eds.), *The Stanford Encyclopedia of Philosophy*, (2012).
- [8] Deductive and Inductive Arguments, <http://www.iep.utm.edu/ded-ind/> (2012.5.7).
- [9] Nagashima, Y.: *BioSensorFusion: New Interfaces for Interactive Multimedia Art*, in Proceedings of 1998 International Computer Music Conference, (1998).
- [10] Nagashima, Y.: *Interactive Multimedia Art with Biological Interfaces*, in Proceedings of 17th Congress of the International Association of Empirical Aesthetics, (2002).
- [11] Nagashima, Y.: *Combined Force Display System of EMG Sensor for Interactive Performance*, in Proceedings of 2003 International Computer Music Conference, (2003).
- [12] Nagashima, Y.: *Bio-Sensing Systems and Bio-Feedback Systems for Interactive Media Arts*, in Proceedings of 3rd International Conference on New Interfaces for Musical Expression, (2003).
- [13] Tanaka, A.: *Musical Performance Practice on Sensor based Instruments*, in: Wanderley, M., Battier, M. (Eds.) *Trends in Gestural Control of Music*, pp. 389–405, IRCAM, Paris. (2000).
- [14] 長嶋 洋一: 身体に加わる加速度とサウンドの音像移動に関する心理学実験報告 (2/4), 情報処理学会音楽情報科学研究会, MUS-95-13, (2012).
- [15] Tsubouchi, Y. and Suzuki, K.: *Biotones: A Wearable Device for EMG Auditory Biofeedback*, in Proceedings of Annual International Conference of the IEEE EMBS, pp. 6543–6546, (2010).
- [16] Pauletto, S. and Hunt, A.: *The sonification of EMG data*, in Proceedings of the International Conference on Auditory Display, (2006).
- [17] Schaffert, N., Mattes, K. and Effenberg, A. O.: *A Sound Design for Acoustic Feedback in Elite Sports*, in: Ystad, S., Aramaki, M., Kronland-Martin, R. and Jensen, K. (Eds.): *Auditory Display. CMMR/ICAD 2009, Lecture Notes in Computer Sciences (LNCS) Vol. 5954*. Springer-Verlag Berlin, pp. 143–165, (2010).
- [18] Dozza, M.: *Biofeedback Systems for Human Postural Control*, Ph.D Thesis, University of Bologna, (2006).
- [19] Chiari, L., Dozza, M., Cappello, A., Horak, F. B., Macellari, V. and Giansanti, D.: *Audio Biofeedback for Balance Improvements: An Accelerometry-Based System*, IEEE Eng Med Biol Mag, 52, No.12, 2108–2111, (2005).
- [20] 伊藤 英則: 個人生体信号から癒しの映像・音楽を生成するシステム, 人工知能学会誌, Vol.23, No.3, pp.334–341, (2008).
- [21] Ericsson, K. A. and Simon, H. A.: *Protocol analysis: Verbal reports as data*, MIT Press, (1993).
- [22] Terasawa, H., Parvizi, J. and Chafe, C.: *Sonifying ECoG Seizure Data with Overtone Mapping: a Strategy for Creating Auditory Gestalt from Correlated Multichannel Data*, in Proceedings of ICAD 2012 (to appear in July), (2012).
- [23] Terasawa, H., Takahashi, Y., Hirota, K., Hamano, T., Yamada, T., Fukamizu, A. and Makino, S.: *C. elegans meets data sonification: Can we hear its elegant movement?*, in Proceedings of the 8th Sound and Music Computing Conference, (2011).
- [24] Bregman, A.: *Auditory Scene Analysis: The Perceptual Organization of Sound*, MIT Press, (1994).
- [25] Roads, C. 著, 青柳 龍也ら訳: コンピュータ音楽, 歴史, テクノロジー, アート, 東京電器大学出版局, (2001).
- [26] 岩宮 眞一郎: 音色の感性学 音色・音質の評価と創造, 日本音響学会 (Eds.), コロナ社, (2010).
- [27] Kameoka, H., Nishimoto, T. and Sagayama, S.: *A Multipitch Analyzer Based on Harmonic Temporal Structured Clustering*, IEEE Trans. on Audio, Speech and Language Processing, Vol. 15, No. 3, pp.982–994, (2007).
- [28] Kitahara, T., Goto, M., Komatani, K., Ogata, T. and Okuno, H. G.: *Instrument Identification in Polyphonic Music: Feature Weighting to Minimize Influence of Sound Overlaps*, EURASIP Journal on Applied Signal Processing, Special issue on Music Information Retrieval Based on Signal Processing, Vol.2007, (2007).
- [29] Roads, C.: *Microsound*, The MIT Press, (2002).