

研究ノート

数式処理システムを用いた可聴化システムの構築

本池 巧

【要旨】 スマートフォンなどのモバイル端末が普及するにつれ、視覚情報を中心とした情報伝達が上手く機能しないことが明らかになってきている。この問題の解決策の一つとして、聴覚や触覚など他のコミュニケーションモードを活用したインターフェースの実現があげられる。本研究は、モバイル端末を介したコミュニケーションに対して有効なマルチモーダル・インターフェースの構築を目的とする。本稿では、これに向けてスマートフォンをフロントエンドに数式処理システムをバックエンドに有する可聴化システムの試作を行い、システムの構成と特徴について報告する。

【キーワード】 可聴化、数式処理システム、モバイルコンピューティング

1. はじめに

今から20年前、汎用コンピュータからパーソナルコンピュータ（PC）へと情報処理の主役が変化したことに伴い、エンドユーザコンピューティングという新しい計算機の利用形態が出現した。この流れを推進したものは、PCのハードウェア機能でも、オペレーティングシステムやアプリケーションソフトでもなく、実はマウスとビットマップディスプレイを使ったGUIと、インターネットであったのではないかと思われる。このように考えれば、新しい計算機の利用形態の出現を促すものは、システムやアプリケーションではなく、計算機と人間、計算機とネットワークのインターフェースであると考えることができる。

現在、デジタル情報へのアクセス手段の主役がPCからスマートフォンやタブレットなどのモバイル端末へと急速に移り変わろうとしている。今後、出現すると思われる新しい情報の利活用がどのような形態になるかは、モバイル端末の持つインターフェースに左右されるのではないかと思われる。この観点で、モバイル端末のインターフェースの特徴

をまとめると以下のようなになる：

- ・ 端末と人：タッチパネルを介するマルチタッチを用いた身体性の高い操作方法。
- ・ 端末とインターネット：3G、4G回線を使った、場所と時を選ばないインターネットへの持続的な接続。
- ・ 端末と実世界：カメラ、加速度センサ、デジタルコンパス、GPSなど、実世界に関する物理情報を取得する多様なセンサ類。

人と人のインターフェースがタッチパネルを使った身体性の高いものに変化したため、PCでは重要視されなかった扱う情報の実際の表示サイズという物理的特性が、新たに重要性を帯びてくる。ネットワークへの常時接続が実現されるに伴って、情報の提示は、様々な状況で行われることとなる。極端な場合、ユーザがディスプレイを見ることができない状況で、情報を表示しなければならない場合もある。端末が存在する実空間の物理情報といが新たに加わることになる。

このような扱う情報の変化に伴ない、情報サービスに於ける情報伝達方法も、ディスプレイ上の視覚情報中心のユニモーダルな伝達から、聴覚、触覚、

方覚など様々な人間のコミュニケーションモードを用いたマルチモーダルな伝達へと変化する必要があると思われる。

本研究では、モバイル端末向けに、聴覚を使ったマルチモーダル・インタフェースの実現を目的とする。幾つかの人間の知覚モードの中から、聴覚を選択した理由を以下に挙げる。

- 1) 基本となる音声信号処理についての理論が確立されており、オーディオシンセシスシステム、アルゴリズムックコンポジションシステムなど可聴化に必要なソフトウェアが揃っている。
- 2) すべてのモバイル端末には、可聴化を行うために必要な、DSP (Digital Signal Processor)、スピーカー、マイクが備わっており、仮説検証だけでなく、実証システムの構築がすぐに可能である。
- 3) 既に、常時スマートフォン+ヘッドフォンで音楽を聴くというスタイルの若者もおり、システムの検証を行うことが可能である。

モバイル端末向けの可聴化 (sonification) として、どのような手法が有効か未だ明らかとなっていない。本稿では、有効な手法を探るために、まずは、多様な可聴化が可能なシステムを、数式処理システムを中核として構築、検証した結果について述べることにする。

2. 可聴化

2.1 可聴化の定義

可視化とは、人間が直接認知することが難しい現象、事象、概念などを、視覚を通じて認知可能なものにするプロセスを意味する。これに対応して、可聴化は、直接認知できない情報を聴覚を通じて認知可能なものにするプロセスである。可聴化の一番の成功例として、ガイガーカウンター、金属探知機などがあげられる。

実用例は比較的古くからあるが、可聴化は、研究領域としては比較的若い分野であり、心理学、工学など幅広いバックグラウンドを持つ研究者が加わっ

た学際的な領域である。そのため、可聴化についての確立された定義は今の所存在しないが、この分野の研究者の間では、

“可聴化とは非言語可聴音響信号を使った情報伝達である。具体的には、可聴化とは、情報伝達や理解を目的として、データの関係性を音響信号による知覚可能な関係性に変換することである”

という定義 [1] が一般に共通の同意を得ている。作曲目的等でデータを活用してきたサウンドアーティストや音楽家が、最近、作品タイトルに可聴化という表記を用いる場合が増えてきており、可聴化の意味が曖昧なものとなってきた。そこで、T.Hermann [2] は、可聴化の定義を明確なものとするために、次の条件を満たす時に限り可聴化と呼ぶべきではないかと提案している：

- ・ (A) 音は入力データの特性 / 関係を反映している。
- ・ (B) 音への変換の方法は系統的であり、インタラクションとデータがどのように音の変化に対応するか明確に定義されている。
- ・ (C) 可聴化は再現性があること：同一のデータとインタラクション / 発音要因から生成される音は、同一の構造を持つ。
- ・ (D) システムは異なるデータに対しても意図通りに使用でき、同じデータの繰り返しについても適応可能である。

2.2 可聴化の手法

Audification

G.Kramer [3] は、時系列などのデータ列を波形データとして直接音に変換する手法を“Audification”と定義した。x-y グラフで可視化した時に波形のように見えるデータであれば、この手法で可聴化が可能である。サウンドエンジン、特別なプログラムライブラリや機器を必要としないため、可聴化の中では、最も直接的かつ単純な手法である。ただし、すべてのデータに適応可能ではなく、以下のような特性を持つデータに適している。

- 1) 要素数が大きい：CD オーディオと同等の音質

であれば、1秒あたり88200個のデータが必要である。

- 2) 波の形状をしていること：データを1次元の時間系列に変換するため、データ値の飛びやデータが存在しない領域が多数あってはならない。
- 3) 複雑な変化をすること：ゆっくりした変化、単純な変化の場合、人間はその特徴を認知することが難しい。
- 4) 低いSN比：カクテルパーティー効果のように、人間は、喧噪の中から特定の信号を取り出すことができる。Audificationは、ノイズの中から意味のある隠れた信号を取り出す作業に適している。

Auditory icons

パーソナルコンピュータ向けのデスクトップメタファを用いたGUI環境が出現した際、ボタン等のウィジェットと呼ばれるGUIパーツに、ギブソンのアフォーダンス理論を援用したビジュアル効果が適応された。例えば、ボタンをクリックした際に、ボタンの陰影表示をボタンが押されたように見せることで、ユーザはボタンを押したという行為を間接的に認知できるようにした。B. Gaver [4]は、ビジュアルアイコンに、それが持つメタファと両立する音を加えることを試みた。彼は、この音を“Auditory icon”と名付けた。Auditory iconは、我々が日常の実生活で体験する音を模倣することで、iconの意味を学ばなくとも理解することを可能としている。例えば、Windows等のデスクトップ環境では、“ゴミ箱”にファイルをドラッグ&ドロップして削除すると、同時に紙をくしゃくしゃに丸めるときの音が鳴る仕組みが組み込まれている。これは、Auditory iconの典型例である。

Earcon

Auditory iconは、実生活の中に存在する音であり、音とその意味が確立されていなければならない。ユーザインターフェース上のアクションのすべてが、実生活の中で意味を持つ行為であるわけではな

い。そのような場合であっても、何らかの音を使ってユーザインタラクションの結果を通知することは効果的である。Blattner [5]らは、“ユーザとコンピュータのインターフェースにおいて、ユーザにコンピュータの操作、インタラクションなどを伝えるための非言語の音を使ったメッセージ”をEarconと名付けた。アイコン(Icon)のIをeye-とかけて、それをearに置き換えた造語である。アイコンでも、写實的にデザインされ一目で何を表すかわかるものと、抽象的なデザインで一見しただけでは何を表すものかわからないものがあるように、ユーザインターフェースで用いられる短い音のなかで、聞いただけで何を表すか理解できる音はAuditory iconに、理解できない抽象的な音はEarconに分類される。

Parameter Mapping Sonification

Parameter Mapping Sonification(PMSon)はデータの提示のために、情報と音響パラメータ(周波数、振幅など)との間に何らかの関連を持たせる方法である。化学反応、電気反応の大半は音を伴うものが多いことから、音を通知の手段として使った例は多くある。お湯が沸くと音の鳴る笛吹きケトルはPMSonの典型例である。データを、音の様々な特性に割り当てることで、バラエティに富んだ可聴化が可能となる。この手法は、一つのデータのトレンド(温度の上昇、下降など)を表すに適切であるが、複数のデータのトレンド、データ間の関連などを表す場合、正しく認知するためにはデータと音響パラメータをどのように対応させるかが課題となる。

Model-Based Sonification

自然界では、音は何らかの活動の結果発生するものである。データを音響パラメータと結びつけるのではなく、音を発生する仮想モデルのパラメータに割り当てる。そのモデルはユーザとのインタラクションが可能となっており、ユーザのアクションに呼応して音が鳴るといった仕組みである[7]。ユーザは、起こしたアクションに対する音での反応を通じ

てデータの構造や特性を理解する。

3 可聴化システムの構築

3.1 開発システム概要

既に、可聴化の専用システムとして Sonifyer [8]、xSonify [9] などが開発されている。これらのシステムは、All-in-one でデータの可聴化を可能とする。本研究の最終目的であるモバイル端末で可聴化を実現するためには、このようなデスクトップPC向けにパッケージ化された可聴化アプリケーションはそぐわない。

GUIアプリケーション構築の際に用いられる有名なフレームワーク MVC(Model-View-Controller) に倣って、可聴化システムを、以下の三つのモジュールに分けて構成することとした(図1)。以後このフレームワークを MLC フレームワークと呼ぶこととする。

- 1) Model: 入力されたデータを可聴データに変換を行う。
- 2) Listen: Model から送られた可聴データを音に変換しユーザに提供する。
- 3) Controller: ユーザインタラクション(データ入力や操作)を、Model または Listen に伝える。

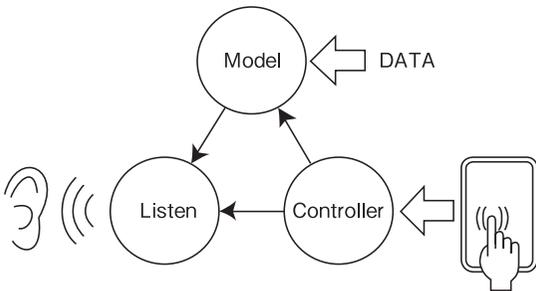


図1: MLC フレームワーク

今回は、可聴化システムの検証用のシステム構築を目的とするため、MLC フレームワークの各要素を最初から開発するのではなく、既存のアプリケーションを組み合わせて実現することとした。なお、

本研究の可聴化システムでは、端末としてスマートフォンを用いることから、各アプリケーションを連携させるために必要な通信モジュールの開発を行った。

Model

可聴化システムの Model の部分は、数式処理システム Mathematica [10] を使用することとした。その理由を以下に挙げる。

- 1) 数値データだけでなく、数式や一般的な抽象化データの取り扱いが可能である。
- 2) 可視化するための豊富な機能を備えているため、可聴化に援用可能な機能が豊富である。
- 3) プロセス間通信を用いて、外部のプログラムと連携する機能(MathLink)を持っている。
- 4) Wolfram Alpha という巨大な数値データベースとの連携機能を持っており、可聴化の検証を行うために利用可能なデータが豊富である。

なお、数式処理システムをスマートフォン上で動作させることは難しいため、数式処理システムはデスクトップPC上で動作させ、ネットワークを介してスマートフォン上のソフトウェアとデータの送受信を行う。

Listen および Controller

MLC は概念としてはわかりやすいものであるが、実際には効率性が要求されるため、Listen と Controller は、音響合成システム SuperCollider [11] でまとめて実装することとした。その理由は：

- 1) 開発ターゲットとしている iOS デバイス上で動作する。
- 2) スマートフォンなどコンピュータ資源に大きな制約のある環境でも、リアルタイムに音響合成が可能である。
- 3) ネットワークを介して他のマシン上のアプリケーションと通信可能である。

3.2 要素間の通信プロトコル

MLC の各要素（実際は、Model と Listen&Controller）は、プロセス間通信を使って連携する。今回構築するシステムでは、プロトコルとして Open Sound Control (OSC) [12] を用いることとした (図2)。

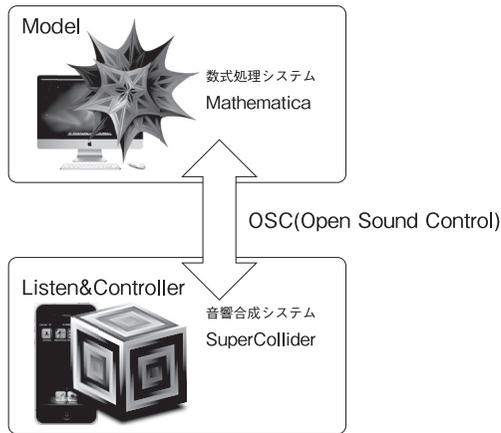


図2 システム構成図

OSC は以下のような特徴を備えており、今回の目的に適したプロトコルであると言える。

- 1) IP ネットワーク上の UDP プロトコルの上位プロトコルに位置するため、インターネットに接続可能なデバイスであれば相互に通信可能である。
- 2) 動的な URL 風 (タグ) の命名規則を採用しており、タグは自由に定義可能であり、拡張性に富む。
- 3) 数値データだけでなく記号データを送受信可能である。
- 4) 多くの、音響合成ソフト、DTM ソフトウェアが対応しているだけでなく、C 言語、Java など主要なプログラミング言語用のライブラリが揃っている。

3.3 要素間連携モジュール作成

数式処理システム Mathematica は、OSC プロトコルでの通信機能を持っていないため、今回、OSC

での入出力のためのモジュールの開発を行った。

Mathematica は MathLink という機能を介して、ユーザが作成した外部プログラムを呼び出すことが可能である。MathLink と OSC をリレーするためのブリッジモジュールを C 言語を用いて開発を行った (図3)。なお、OSC 通信には、liblo ライブラリ [12] を用いた。

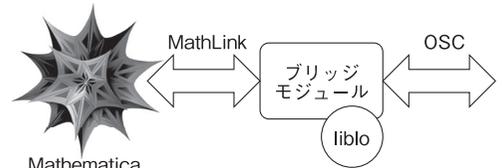


図3: MathLink-OSC ブリッジモジュール

4. 可聴化サンプル

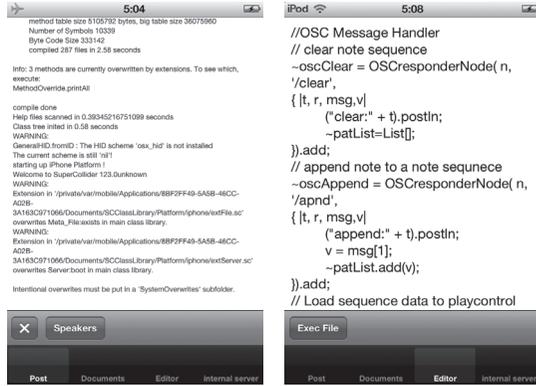
現段階では、Mathematica から送信したデータを OSC 経由で SuperCollider が受信することは実現できているが、逆に SuperCollider から受信送信したデータをリアルタイムに Mathematica で処理する際、データの取りこぼしが発生するなどの通信の不具合が解決できていない。そのため、今回は、数式処理システムから送られたデータを音響合成システム側で発音する処理について検証した結果を報告する。

無理数の可聴化

サンプルの可聴化として、有理数や無理数の 12 進数表現を作成し、各桁の数 (0 ~ 11) を 12 音階の各音にマッピングする。

円周率 π (=3.1415926535897932384626433832795028841971693993751...) を例に説明すると、この 12 進数表現は、3.184809493B918664573A6211BB151551A05729290A78...となる。12 進数表現中の記号 A、B はそれぞれ 10 進数での 10、11 を表す。0 ~ 9,A,B をそれぞれ、以下のようにハ長調での半音階ド〜シの 12 音に割り当てる。

- 0 → ド
- 1 → ド#
- 2 → レ
- 3 → レ#
- 4 → ミ
- 5 → ファ
- 6 → ファ#
- 7 → ソ
- 8 → ソ#
- 9 → ラ
- A → ラ#
- B → シ



a) サーバ起動画面 b) プログラムの編集

図5 SuperCollider の実行

この可聴化は、Model-Based Sonification の最も単純なものと思倣すことができる。

実行は、Mathematica 側より、図4に示すように、コマンド (SendNumberSeq) を実行して、数値を12進数にエンコードした結果を、iOS デバイス上の SuperCollider に送信する。なお、Mathematica は任意精度で数値を扱うことが可能であり、非常に長いデータ列の生成も可能である。

```

In[1]:= SetDirectory["~/Users/motoike/src/mathlink"];
Out[1]:= link = Install["./oscsendlist"]
Out[2]:= LinkObject[~/Users/motoike/src/mathlink/oscsendlist, 8, 8]
In[3]:= SendNumberSeq[n, len, s_] :=
Block[{temp}, OSCSendList[RealDigits[N[n, len], s][[1]]];
RealDigits[N[n, len], s][[1]]]
Out[3]:= SendNumberSeq[Sqrt[2], 40, 12]
Out[4]:= {1, 4, 11, 7, 9, 1, 7, 0, 10, 0, 7, 11, 8, 5, 7, 3,
7, 7, 0, 4, 11, 0, 8, 5, 4, 8, 6, 8, 5, 3, 5, 0, 4, 5, 6, 3, 6}
    
```

図4: Mathematica でのコマンドの実行

iOS デバイス側では、まず SuperCollider の内蔵の音響合成サーバを起動しておく (図5a)。編集画面より、OSC 経由で受信したデータの解析・発音するプログラムの作成と内蔵サーバ上で実行を行い、データ受信可能状態にする (図5b)。

このサンプル可聴化システムで生成した音のサンプルを以下の URL にて公開する。

- ・ 123/456 の可聴化
<http://www.motoike.jp/sonification/sample1.html>
- ・ $\sqrt{2}$ の可聴化
<http://www.motoike.jp/sonification/sample2.html>
- ・ π の可聴化
<http://www.motoike.jp/sonification/sample3.html>

5. まとめ

今回は、数式処理システム Mathematica を中核とする可聴化システムの構築を行った。数式処理システムの持つ柔軟なデータ処理能力、拡張可能なデータ構造という特徴と SuperCollider の持つリアルタイム音響合成能力を組み合わせさせた結果、入力データを様々な形態で可聴化することが可能なシステムの構築に成功した。

ただし、MathLink と OSC のブリッジモジュール連携に問題があるため、端末上でインタラクティブな操作が実現できていない。当面の課題として、まずは、端末とのインタラクティブな機能を実現する必要がある。この機能を実現した後、ユーザインタラクションと可聴化の有効な組み合わせ等について研究を進めていく予定である。

参考文献

- [1] G. Kramer, B. Walker, T. Bonebright, P. Cook, J. Flowers, N. Miner and J. Neuhoff, "Sonification Report: Status of the Field and Research Agenda", NSF Sonification White Paper-Master 12/13/98(1999), p.5.
- [2] T. Hermann, "Taxonomy and Definitions for Sonification and Auditory Display", in Proc. 14th Int. Conference on Auditory Display (ICAD), France, June 24-27, 2008.
- [3] G. Walker and G. Kramer, "Ecological Psychoacoustics and Auditory Display" in "Ecological psychoacoustics", edited by J. G. Neuhoff. SanDiego: Elsevier (2004), p.150.
- [4] W. W. Gaver, "Using and Creating Auditory Icons" in "Auditory Display: Sonification, Audification and Auditory interfaces" edited by G. Kramer, Addison-Wesley MA, USA, 1994, p.417.
- [5] M. M. Blattner, D. A. Sumikawa and R. M. Greenberg, "Earcons and icons: Their structure and common design principles", Human Computer Interaction, 4(1), 1989, p.11.
- [6] T. Hermann and A. Hunt, "An introduction to interactive sonification", IEEE Multimedia 12(2), 2005, p.20.
- [7] T. Bobermann, T. Herman and H. Ritter, "Tangible data scanning sonification model", Proceedings of the 12th international conference of Auditory Display (ICAD06), London, UK, p.77.
- [8] Sonifyer (Sonifyer.org):
URL: <http://www.sonifyer.org>
- [9] xSonify (NASA Goddard Space Flight Center):
URL: http://spdf.gsfc.nasa.gov/research/sonification/sonification_software.html
- [10] Mathematica (Wolfram Research Inc.):
URL: <http://www.wolfram.com>
- [11] SuperCollider
URL: <http://supercollider.sourceforge.net>
- [12] M. Wright and A. Freed, "Open Sound Control: A New Protocol for Communicating with Sound Synthesizers", International Computer Music Conference, Thessaloniki, Greece, 1997, p.101.
- [13] liblo:Lightweight OSC implementation
URL: <http://liblo.sourceforge.net>

Development of a Sonification System Employing Computer Algebra System.

by MOTOIKE, Takumi

[Abstract] As mobile devices such as smart phones become widely used, it turns out that communications based mainly on visual information do not function very well. One of the solutions for this problem is a realization of multimodal interface which utilizes other sensory modes, e.g., auditory sense, tactile sense and so on. The main purpose of this study is realizing the multimodal interface effective for the communication via mobile devices. In this paper, the development of the sonification system constituted by a mobile device on the front-end and a computer algebra system on the back-end is shown, and the configuration and characteristics of this system are reported.

[Key Words] sonification , CAS , mobile computing