

閉曲線図形に基づいた音色生成方法の検討

岩淵 勇樹^{*1}

秋田 純一^{*1}

北川 章夫^{*1}

A Study on the Sound Synthesis from Closed Curves

Yuuki Iwabuchi^{*1}, Junichi Akita^{*1} and Akio Kitagawa^{*1}

Abstract – In this paper, we propose the method of sound synthesis which output audio signal based on given closed curve. First, we derive the condition that similar figures constantly represent the same timbre. And we try to convert curves into analytic signals, that is the simplest representation that satisfy such condition. Through this research, it can be created an intuitive musical interface.

Keywords : sound synthesis, timble, Hilbert transform, analytic signal

1. はじめに

電子楽器を用いることにより、今までにない新たな音色を作ることが可能となっており、音色を生成する方法も数多く考案されている。しかしながら、それらのほとんどは正弦波や矩形波などの単純な波形あるいはPCM録音によってサンプリングされた波形をそのまま用いたり、それらに加算・乗算などの演算やフィルタリングを施したものであるが、多種多様な音色を作るには多くのパラメータ制御を必要とするため、直感的な操作が難しく、特性を掴むのに多くの時間を要する。

最近の新しい入力インターフェイスを持つ楽器の例として、KORG社による「Kaossilator」[1]があり、これはタッチパッドを使った平面的な入力により従来のボタンやツマミを用いる操作方法よりも直感的かつ動的に扱うことができる。しかしながらこれは水平軸と垂直軸で独立した2つのパラメータ入力同時に行っているに過ぎず、その点では旧来の音色入力方法とさほど変化がない。

その他GUIの特性を活かした例として、一部の効果音作成ソフトウェア等では波形をマウスで描画して音色を作る機能を備えているものもあり、パラメータ入力よりは直感的に操作できるが、生の音声波形は視覚的に音色を想起させるに優れているとも一概には言い難い。

一方、メディアアートにおいては、図形を基に

した音色生成の先行研究のひとつとして Levin らの作品 [2] が挙げられる。この作品では図形と音色の調和を意識して構成されているため直感的ではあるが、面積や周囲長などの画像の特徴量をパラメータ化しているに過ぎないため、形状情報が音色に寄与する割合は低く、表現力にも限界がある。

このように、直感的な音色入力は発展途上であり、GUIに適した音色入力方法を探ることは今後において重要な課題といえる。

本研究では、マウスやタッチパッドなどの入力デバイスによって直感的な音色入力インターフェイスを構築するために、閉曲線から周期信号を生成するオシレーターの基本原理を考える。具体的には、平面上の閉曲線図形について、相似図形同士が同じ音色と対応するような変換方法を検討する。

2. 音声信号と曲線

音声信号を $x(t)$ とし、そのフーリエ変換を $X(\omega)$ とする。

このとき、 $X(\omega)$ にあるフィルタ $F(\omega)$ を掛けて

$$Y(\omega) = F(\omega) \cdot X(\omega) \quad (1)$$

となる $Y(\omega)$ を与えると、その逆フーリエ変換 $y(t)$ と原信号 $x(t)$ を用いて

$$z(t) = x(t) + iy(t) \quad (2)$$

という複素関数が定まる (i は虚数単位)。 $F(\omega)$ が定数の場合などを除けば、 $z(t)$ は複素平面 $x + iy$ 上の曲線となり、特に $x(t)$ が周期関数かつ $x(t)$ および $y(t)$ が区分的に滑らかで常に有限値の場合は閉曲線となる。

^{*1}: 金沢大学大学院 自然科学研究科

butchi@merl.jp, {akita, kitagawa}@is.t.kanazawa-u.ac.jp

^{*1}: Graduate School of Natural Science & Technology, Kanazawa University

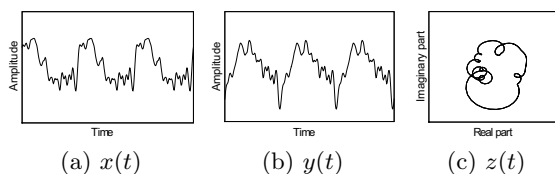


図1 楽器音と解析信号

Fig. 1 Instrumental sound and analytic signal

2.1 ヒルベルト変換

$F(\omega)$ を以下のように定める．

$$F(\omega) = \begin{cases} i & (\omega < 0) \\ 0 & (\omega = 0) \\ -i & (\omega > 0) \end{cases} \quad (3)$$

このとき、 $y(t)$ は $x(t)$ のヒルベルト変換であるという．また、 $z(t)$ は負の周波数成分を持たないという性質があり、このような性質を持つ信号のことを解析信号と呼ぶ．なお、 $x(t)$ は $y(t)$ の逆ヒルベルト変換であるといい、 $F(\omega)$ の代わりに $-F(\omega)$ を用いた時の $x(t)$ と $y(t)$ の関係に等しい．

$z(t)$ が解析信号であるとき、以下の等式が成り立つ (\mathcal{F} はフーリエ変換を表す)．

$$|\mathcal{F}[Re[e^{i\theta}z(t)]]| = |X(\omega)| \quad (4)$$

ところで、人間が周期的な音声信号を知覚する際、聞こえに影響するのはほとんど振幅スペクトルの形状の違いであり、位相が与える影響はごく僅かである．そのため、式 (4) より、周期関数 $x(t)$ および $Re[e^{i\theta}z(t)]$ は同等の音色をもつとみなす．すなわち、複素平面上の閉曲線が与えられたとき、その実部を音声信号に対応させるとすると、閉曲線が解析信号であった場合はその回転図形も同等な音色をもつことになる．

また、音声信号の定常成分は音色に全く影響を与えないため、定常成分のみ異なる音声信号もそれぞれ同等な音色をもつと見なせる．

図1 にヒルベルト変換および解析信号の例を示す．原信号 $x(t)$ はバイオリンの音色である．

2.2 曲線図形の多様性

任意の単調増加写像 h によってパラメータ変換された曲線 $z(h(t))$ の軌跡は、 $z(t)$ の軌跡と等しい．よって、閉曲線が与えられたとき、パラメータ変換によって解析信号を生成することが可能ならば、任意の閉曲線図形の入力に対して回転しても生成される音色が不変な音声信号生成が可能といえる．さらに、図形の大きさは音声信号の振幅、つまり

音量に相当するため、相似図形については同一の音色が対応することになり、形状特徴に固有な変換といえる．

しかしながら、筆者の予想では、解析関数が閉曲線図形として表現できる範囲には限界があり、そのようなパラメータ変換がほとんどの場合において不可能だと考えられる．

そのため、式 (4) が恒等的に成立するための条件を求め、閉曲線の表現範囲を拡張する．

2.3 回転によって音色が不変である条件 任意の複素数 w について

$$Re[w] = \frac{w + \bar{w}}{2} \quad (5)$$

が成立することを用いると、 $Re[e^{i\theta}z(t)]$ は以下のように変形できる．

$$\begin{aligned} Re[e^{i\theta}z(t)] &= \frac{e^{i\theta}z(t) + \overline{e^{i\theta}z(t)}}{2} \\ &= \frac{e^{i\theta}\{x(t) + iy(t)\} + e^{-i\theta}\{x(t) - iy(t)\}}{2} \\ &= \frac{(e^{i\theta} + e^{-i\theta})x(t) + i(e^{i\theta} - e^{-i\theta})y(t)}{2} \\ &= \cos\theta \cdot x(t) - \sin\theta \cdot y(t) \end{aligned} \quad (6)$$

そして、これのフーリエ変換の絶対値を2乗したものは以下のように変形できる．

$$\begin{aligned} &|\cos\theta \cdot X(\omega) - \sin\theta \cdot Y(\omega)|^2 \\ &= \{\cos\theta \cdot X(\omega) - \sin\theta \cdot Y(\omega)\} \\ &\quad \{\cos\theta \cdot \overline{X(\omega)} - \sin\theta \cdot \overline{Y(\omega)}\} \\ &= \cos^2\theta |X(\omega)|^2 + \sin^2\theta |Y(\omega)|^2 \\ &\quad - \cos\theta \sin\theta \{X(\omega)\overline{Y(\omega)} + \overline{X(\omega)}Y(\omega)\} \\ &= \cos^2\theta |X(\omega)|^2 + \sin^2\theta |F(\omega)X(\omega)|^2 \\ &\quad - \cos\theta \sin\theta \\ &\quad \{X(\omega)\overline{F(\omega)\overline{X(\omega)}} + \overline{X(\omega)}F(\omega)X(\omega)\} \\ &= |X(\omega)|^2 \{\cos^2\theta + \sin^2\theta |F(\omega)|^2\} \\ &\quad + \cos\theta \sin\theta \{Re[F(\omega)]\} \end{aligned}$$

そして、上式は恒等的に $|X(\omega)|^2$ になる必要があるため、以下の条件を満たさなければならない．

$$\begin{cases} |F(\omega)| = 1 \\ Re[F(\omega)] = 0 \end{cases} \quad (7)$$

これはつまり絶対値が1である純虚数、 $\{i, -i\}$ のみを値域としてもつ関数である．また、 $x(t)$ と

$y(t)$ はともに実関数であるため、 $F(\omega)$ は奇関数である必要がある。

この関係より、任意の $\omega' > 0$ について $X(\omega') + iY(\omega')$ か $X(-\omega') + iY(-\omega')$ のどちらかは打ち消しあって 0 になるため、最終的に $z(t)$ のスペクトルは「周波数毎に正か負どちらかみの周波数成分をもつ」ものとなる。

また、周期関数であるため、複素係数列 a_n と符号列 $s_n (\in \{-1, 1\})$ を用いて以下のような級数として表現することができる。

$$\sum_{n=1}^{\infty} a_n e^{(s_n i n \omega_0 t)} \quad (8)$$

3. 関数の近似

以上により曲線がもつ条件が拡張されたが、なおも任意の閉曲線図形を表現できるかどうかは自明でない。いずれにせよ、実装の際には式 (8) のような無限和を扱えないため、近似解を求める必要がある。

今回は、前節の拡張表現については考えず、閉曲線を解析信号に漸近させる方法について検討する。

シミュレーションの初期曲線には、約 15 ピクセル角のシルエット画像を 8-連結の境界追跡によって求めた輪郭線を用いる。初期曲線は $m + 1$ 個の点列 $z_0(t) = (z_0(0), z_0(1), \dots, z_0(m))$ によって構成され、 $z_0(m) = z_0(0)$ とする。フーリエ変換時にはこの重複する末尾の点を無視して FFT を行う。

初期曲線には図 2 に示す曲線 1 と曲線 2 を用いる。それぞれのスペクトルは (c)(d) のようになる。

3.1 方法 A: 解析信号対の漸近による方法

解析信号 $z(t)$ では、実部解析信号 $\xi(t) = x(t) + i\mathcal{H}[x(t)]$ と虚部解析信号 $\eta(t) = \mathcal{H}^{-1}[y(t)] + iy(t)$ は当然等しくなる。なお、 $\mathcal{H}, \mathcal{H}^{-1}$ はそれぞれヒルベルト変換、逆ヒルベルト変換を表す。

以下の手順で $\xi_{tmp}(t) - \eta_{tmp}(t)$ の値を最小化させ、入力曲線を解析信号に漸近させるを試みる。

1. $n = 0$ とおく
2. $z_n(k)$ と $z_n(k + 1)$ の間の位置に $(z_n(k) + z_n(k + 1))/2$ を挿入し、 $z_{tmp}(t)$ とおく
3. $z_{tmp}(t)$ について、 $\xi_{tmp}(t) - \eta_{tmp}(t)$ の 2 乗和が最小となる k を求め、そのときの $z_{tmp}(t)$ を $z_{n+1}(t)$ とする
4. n を $n + 1$ に置き換え、2. 以降を繰り返す

この方法での $n = 100$ におけるシミュレーション結果を図 3, 4 に示す。

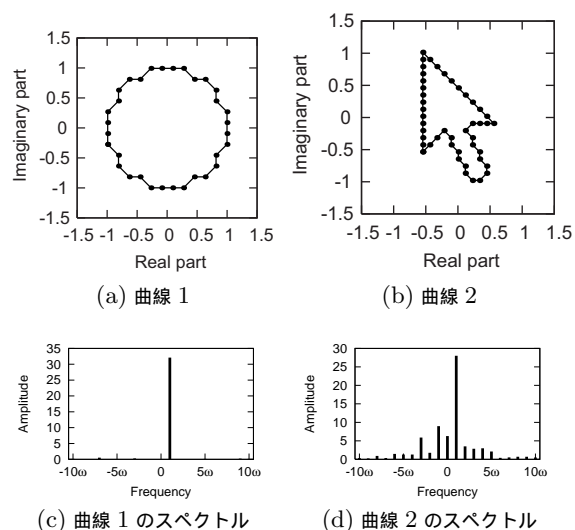


図 2 入力図形とそのスペクトル
Fig. 2 Input curves and its spectrum

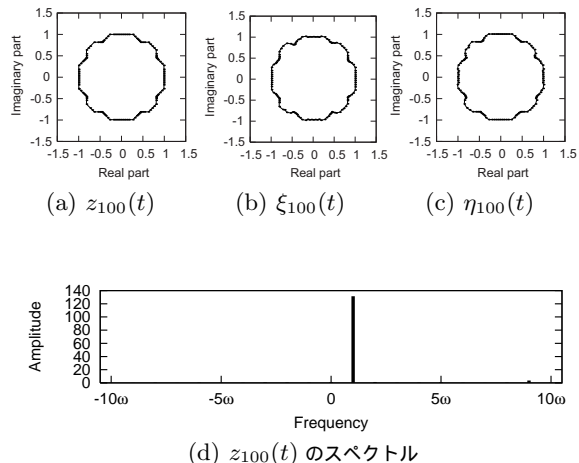


図 3 方法 A のシミュレーション結果 (曲線 1)
Fig. 3 The results from method A (curve 1)

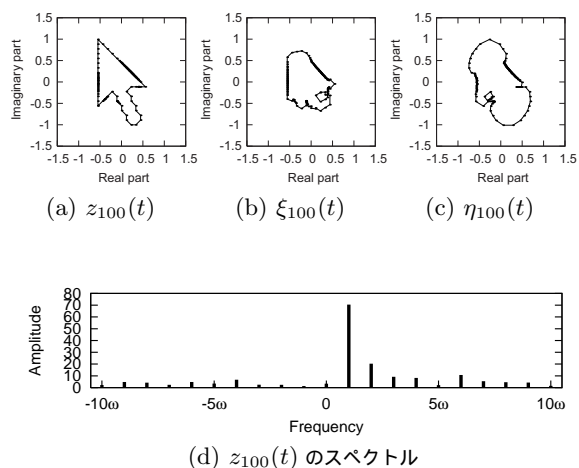


図 4 方法 A のシミュレーション結果 (曲線 2)
Fig. 4 The results from method A (curve 2)

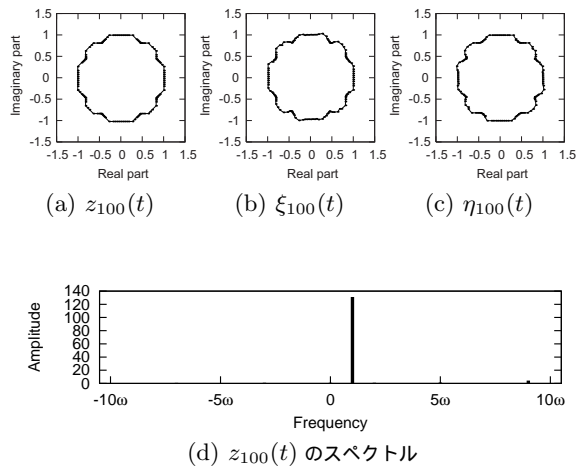


図5 方法 B のシミュレーション結果 (曲線 1)
Fig. 5 The results from method B (curve 1)

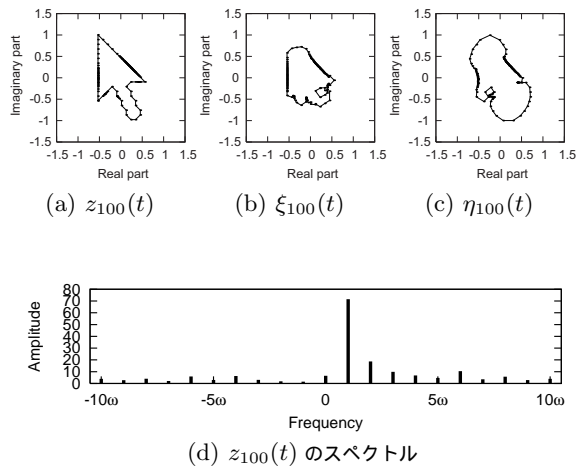


図6 方法 B のシミュレーション結果 (曲線 2)
Fig. 6 The results from method B (curve 2)

3.2 方法 B: 負周波数成分除去による方法

既に述べた通り，解析信号はその周波数成分のうち負周波数成分は全て 0 である．また，定数成分も含まないと考える．そのため，上記の方法のうち 3. における評価値を $z_{tmp}(t)$ の 0 以下の周波数成分の 2 乗和に置き換える．

この方法での $n = 100$ におけるシミュレーション結果を図 5, 6 に示す．

3.3 考察

以上，2 種類の図形の解析信号近似を試みた結果，2 つの方法でほとんど同じ結果が得られた．

曲線 1 については解析信号にかなり近い形となったが，これは元々入力曲線が解析信号に近かったことが大きな要因であると考えられる．曲線 2 については，0 以下のスペクトル成分をある程度減ら

すことは出来たが， $\xi_{100}(t)$ および $\eta_{100}(t)$ を入力曲線と見比べると，形状の見た目としては少しは似ているものの，特徴が大きく失われている．

今回のように曲線の点列に 1 点ずつ標本点を追加する場合，挿入点は総当たりによって決定されるため，点が増えるとともに計算量も増加することになり，効率も良くない．

また，挿入点に隣り合う値の平均値を挿入値としているため，これは入力点列に対する線形補間と捉えることもできるが，線形補間は理想的な補間方法とはいえず，滑らかにならず結果として高周波成分を多く含むことになる．自然な音色は高周波成分を多く含まないため，これは望ましくない．今回用いたシルエット画像のように解像度の低い画像を用いる場合は，補間する点にある程度の自由度があるとのおよびと考えられる．

今回の方式を改良するならば，点の追加と削除を交互に行うなどして，標本点の個数を固定した上で最適化を行うことができればより理想的な結果が得られるのではないかと考えられる．

4. まとめと展望

本研究では，合同図形について音色が同一であるような音声信号を持たせられるような曲線もつ条件を求めた．また，その中で最も単純な構造を持つ解析信号について，閉曲線を解析信号に変換する方法についてのシミュレーションを行った．

今後は式 (8) を満たす，解析信号を拡張した曲線を求める近似解の導出方法について検討し，ソフトウェア実装する予定である．

参考文献

- [1] KAOSILATOR: <http://www.korg.co.jp/Product/Dance/kaossilator/>
- [2] G. Levin and Z. Lieberman, "Sounds from Shapes: Audiovisual Performance with Hand Silhouette Contours in The Manual Input Sessions," Proceedings of NIME '05, Vancouver, BC, Canada. May pp.26-28, 2005.