

コンピュータ音楽で用いる珍しい音合成方式 Uncommon Sound synthesis used in computer music

小坂 直敏
Naotoshi Osaka

東京電機大学
Tokyo Denki University

〒101-8457 東京都千代田区神田錦町 2-2
osaka@im.dendai.ac.jp

あらまし コンピュータ音楽あるいはマルチメディアアートを目的とした音合成方式の中で信号を直接合成する信号モデルでは、与えられた音信号を表現する能力のある分析 / 合成方式を基にして発展した加工方式と、こうした能力のない非分析 / 合成方式に基づくモデルとがある。非分析 / 合成方式の信号モデル、あるいは物理的実態と対応しない物理モデルは、音をデザインとして扱うこの分野独特の合成方式である。一方、実証的に音響現象を追求する音楽音響を含む一般の音響研究者や音声研究者には、これらの方式は実証研究分野への応用の可能性が少ないために、あまり知られていない。本発表では、これらの代表としてウェーブシェーピング、細粒合成、地表面軌道合成を取り上げ基本的な仕組みを紹介する。

1. はじめに

コンピュータ音楽、あるいはマルチメディアコンテンツのための音合成には数多くの手法が報告されている。文献[1]はこれらを幅広く紹介したものである。これらの方式の多くは情報通信分野での音声表現方法を基にさまざまな加工方法を加えたものが多い。チャンネルポコーダ、フェーズポコーダ、LPCなどがその例である。

一方、FM音源方式[3]、物理モデル[4]による方式など音声表現方法からの応用でなく、元から音楽のための音合成方式として独自に発展した方式も数多い。これらは逆に音声表現方法としては興味を持たれなかったものである。その理由は、分析 / 合成方式としては確立しておらず、与えられた音声を表現する能力がないことが最も大きな理由である。しかし、この枠組みは、合成できる音色の広範なことから、楽音合成方式として広く愛好されてきた。

本稿では、これらのいずれにも属さない方式として、筆者から見てまだ発展の余地のある方式の

基礎をいくつか紹介したい。それらは、1)FM音源と同等な表現能力がありながら、その使用方法がよく知られておらず、あまり普及していないウェーブシェーピング方式、2)分析 / 合成方式としては理論的には知られているが、音声情報処理としては興味を持たれず、そのコンセプトゆえに非分析 / 合成方式の楽音合成方式としてしか利用されない細粒合成方式、また、3)さらに80年代に検討されながらも、その後応用を広げるための検討がなされてこなかった地表面軌道合成の各方式である。細粒合成以外はコンピュータ音楽の研究者あるいは制作者にとってもあまり馴染みのない方式であると考えられる。表題の「珍しい」との主観的用語は、この様子を表現したものである。以下では各方式について簡単に紹介する。

2. ウェーブシェーピング

ウェーブシェーピング、あるいは波形テーブル制御は、FM合成と類似したスペクトルモデルである。これは、分析 / 合成を目的としておらず、真

に創造的な合成をねらいとしたものである。

以下は、[5]の提案した方法で、非線形フィルタによる高調波への変換関数である。今、 $T_n(x)$ をChebychevの第一種の多項式とし、入力を $x = \cos \theta$ 、すなわち、一つの調波とすると、

$$T_n(x) = \cos(n\theta) \quad (1)$$

と書くことができる。これにより、 n 次の高調波が生成できる。 $T_n(x)$ は以下の漸化式により算出できる。

$$T_0(x) = 1 \quad (2)$$

$$T_1(x) = x \quad (3)$$

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x) \quad (4)$$

なお、ここに、 $-1 \leq x \leq 1, -1 \leq T_n(x) \leq 1$ である。振幅および、変換特性が定常的な音について考えると、加算合成と同様、

$$f(x) = \frac{h_0}{2} + \sum_{l=1}^N h_l T_l(x) \quad (5)$$

の形式で算出できる。さらに動的なスペクトル表現をするため、指数 $a; 0 \leq a \leq 1$ を導入し、 $f(ax)$ の振る舞いも検討されている。非線形式であるため、 $T_n(ax) \neq aT_n(x)$ となり、 $f(ax)$ を

$$f(ax) = \frac{h_0}{2} + \sum_{l=1}^N h_l(a) T_l(x) \quad (6)$$

としたときの $h_l(a)$ を解析的に求めた解は単純ではない。図1に入力信号にウェーブシェーピングをかけて出力信号を作までのブロックダイヤを示した。

ウェーブシェーピングは豊富な調波信号(あるいは非調和信号)を表現できる。例えば一つの調波から、このエネルギーを周波数上に散乱させるべく、多くのサイドバンドを作成することもで

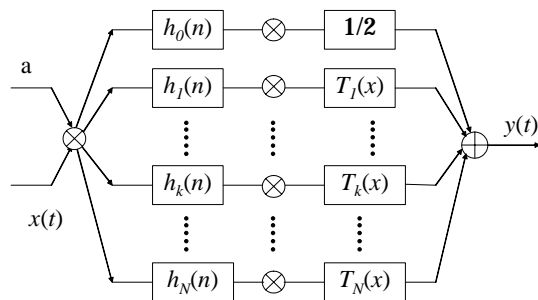


図1 ウェーブシェーピングのブロックダイヤ

きる。特に指数 a により動的なスペクトルを扱うことができ、FM合成と同様、豊富な音色を合成できる。Le Brunは[5]の中でFM合成もウェーブシェーピングの枠組みで記述できることを示している。

3. 細粒合成

細粒合成はグラニューラ合成とも呼ばれる。理論的にはウェーブレット変換と類似した分析/合成である。基本的な考え方は、いかなる音も聴覚上の単位“粒子(grain)”より成り立っている、とするものである。通常、粒子の継続時間は5~10msecである。Gaborはこの基本的な概念を[6]で提案した。図2は粒子を示したものである。ガウス窓のかかった正弦関数である。Bastiansはこの考えを数学的に拡張した[7]。こうして、本来分析/合成としての枠組みとして考案されたが、実際のコンピュータ音楽へ応用では、ほとんど分析/合成として厳密に使われることはなく、原音の復元はできない使い方が一般的である。

Xenakisはグラニューラ合成を音楽作品に採り入れた最初の作曲家である。また、Roadsはこの考えを採り入れ、粒子として窓関数付き短時間正弦波を用いた[8][9]。合成時は、個々の10-60msecの時間フレーム毎にスペクトルパターンに応じた粒子が貼り付けられる。分析/合成手法は文献[10]に述べられている。モデルのこれら二つの使い方は完全に異なっており、Roadsの手法はどちらかというとなり分析/合成的であるのに対し、Xanakisの手法は自然音を擬似、あるいはこれらを変形しようとする意図はまるでなく、完全に音楽応用目的の合成である。

このように、細粒合成は他のアルゴリズムと比して「粒子状の波形の重ね合わせで音合成を行

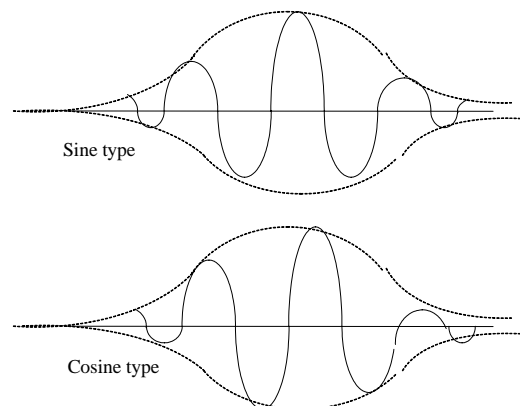


図2 粒子波形(基本信号) [6]より

う、とより抽象的なレベルで定義されているため、その実現にあたってユーザの自由度が大きい。分析/合成に基づく最も単純な手法として、オーバーラップ加算方法がある。

この手法は窓をかけた波形データをシフトしながら加算していくとき、三角窓では窓長の1/2、またハミング窓のように余弦関数(三角関数)の1周期で定義された窓では、窓長の1/4ずつシフトしながら窓掛けされた波形を加算すると原波形に戻る、というものである。これは、上記のようにシフトするとき、窓関数に含まれている余弦波を1/4周期ずつずらした窓の総和が定数になることに起因している。

オーバーラップ加算による加工では、個々の窓かけされた波形をSTFT(短時間フーリエ変換)し、周波数領域で加工した後にIFFT(逆フーリエ変換)を行い、オーバーラップ加算を行う。

一方、細粒合成では、時間軸の操作を基本とし、再合成時に時間軸上に再配置する時に、原波形の時間構造とは異なる加工を施す。具体的には時間の伸縮、ランダム配置、逆配置などである。

また、オリジナルな分析/合成に近い方法は、フィルタバンクで音声帯域分割しておき、さらに個々のチャンネルに関してオーバーラップ加算を行うものである。帯域分割はフェーズボコーダなどの利用が可能である。

ここでは、前者のオーバーラップ加算による時間領域での加工による方法をもって細粒合成を実現する。図3に示すように、入力波形からの窓による切り出し方に二つのパラメータを用いる。窓のシフト量(サンプル数)と次の窓とのオーバーラップ量(サンプル数)である。窓長は両者の加算値である。一方、再合成時に音信号の伸縮率を指定して再配置を行う。これにより時間領域特有のエフェクトが実現できる。

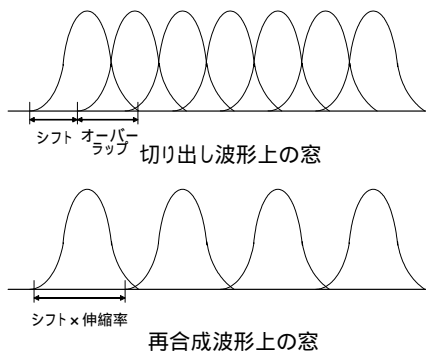


図3 オーバーラップ加算方式を用いた細粒合成

4. 地表面軌道合成

耳慣れないこの方式はWave Terrain Synthesisの邦訳である[1]。R. Goldらが最初に同方式を提案し[11]、Mitsuhashiも基礎的な検討を行った[12]。この方式は2変数をパラメータとし、これによる軌道を定義する。一方、2変数の関数を別途定義しておき、この軌道の上で描く関数値を波形値とする方式である。すなわち、例えていえば、山の形を定義しておき、そのまわりを時変の軌道で回るモーターバイクの高さを音波形値とする、ということである。

これらの検討は80年代に行われたが、FM音源と類似したモデル音とまりで、実際の音響的な音合成、楽音合成の域に達した、という報告はなく、まだ深い検討がなされていない状況である。

軌道を変化させる代わりに、軌道を固定しておき、山の形状を変化させることも考えられる。それがScanned Synthesisで、Max mathewsらが近年提案している[13],[14]。彼らは必ずしも地表面軌道合成との関連で同方式を捉えているわけではなく、簡単な力学系の物理モデルを基に、その動きを音に変換するために、オーディオレートで走査することにより音波形を得ている。この報告は物理モデルによる実時間の音合成制御、という視点が強く、多様な音色を追求する、という視点はやや弱い。

ここでは最も初歩的な関数として、

$$z = x^2 + y^2 \quad (7)$$

を選びこの軌道として、以下の簡単なリサージュ図形を考えよう。地表面 z と軌道関数を図4に示す。

$$x = a \cos(2\pi f_0 t + \theta) \quad (8)$$

$$y = b \sin(2\pi f_0 t) \quad (9)$$

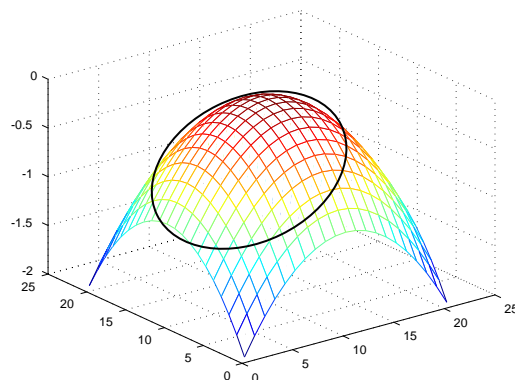


図4 楕円軌道と地表面関数

この軌道は、 $a>0, b>0, a \neq b$ のとき楕円であり、軌道上の地表面 z は x, y の値を代入すると、 $2f_0$ の周波数を有す正弦波が合成される。軌道を少しずつ複雑にすることにより、各種変調方式が得られる。

図5は軌道関数を下式のリサージュ図形としたものである。

$$x = a \cos(2\pi f_0 t + \theta) \quad (10)$$

$$y = b \sin(2\pi f_1 t) \quad (11)$$

ここに、 $a=1, b=0.5, \theta=1/3, f_0=440\text{Hz}, f_1=450\text{Hz}$ とした。これは変調周波数 $f_1 - f_0$ Hz の AM (振幅変調) 音が含まれている。図6はさらに複雑に三角関数からなる地表面をいくつにも構成したもので、以下の式からなる。

$$z = A \sin(2\pi f_3 x) \cdot \sin(2\pi f_3 y) \quad (12)$$

これに10,11式と同様のリサージュ図形の軌道を定義すると、FM (周波数変調) 音が合成される。

このような基本関数のみならず、より複雑な関数を得ることでより豊かな音色が得られる期待が生ずる。また、地表面および軌道は関数のみならず、自然界からの形状データを用いることでさらに現実的な音色が得られる可能性がある。

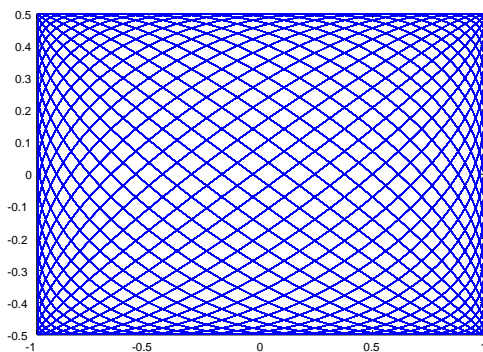


図5 時変軌道の例(リサージュ図)

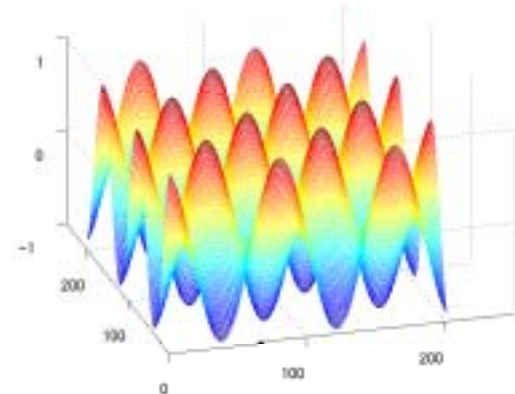


図6 2次元の正弦関数による地表面関数

5. あとがき

ウェーブシェーピング、細粒合成、知表面軌道合成の3方式についてその基礎を概説した。さらなる応用については現在検討中で、本稿では詳細は省略した。これらの手法はもう少し多くの研究者に興味を持たれ発展してほしい、と筆者は考えている。

参考文献

- [1] 青柳龍也, 小坂直敏, 平田圭二, 堀内靖雄, “コンピュータ音楽 歴史・テクノロジー・アート,” 東京電機大学出版局, 2001 1.
- [2] 小坂直敏, “コンピュータ音楽の歴史と現状,” 音響学会誌 Vol. 52., No. 2, pp.155-160, 2001.
- [3] J. M. Chowning; The synthesis of complex audio spectra by means of frequency modulation, *Journal of the Audio Engineering Society*, vol. 21, No. 7, Sep. 1973
- [4] T. Hikichi, N. Osaka and F. Itakura “Time - domain simulation of sound production of the sho,” *J. Acoust. Soc. America* Vol.113, No.2, 1092-1101 (2003).
- [5] M. Le Brun: Digital Waveshaping synthesis, *Journal of Audio Engineering Society*, vol. 27, No. 4, pp. 250-266, Apr. 1979
- [6] D. Gabor: Acoustical Quanta and the Theory of Hearing, *Nature*, 159, No. 4044, pp. 591-594, may, 1947.
- [7] M. J. Bastiaans: Gabor's Expansion of a Signal into Gaussian Elementary Signals, *Proc. Of the IEEE*, vol. 68, No. 4, pp. 538-539. Apr. 1980.
- [8] C. Roads: Automated Granulay Synthesis of Sounds, *Computer Music Journal* 2(2): pp. 61-62, 1978.
- [9] C. Roads and J. Strawn: Granulay Synthesis of Sound, *Foundations of Computer Music*, Cambridge, Massachusetts, 1987. (updated version of [8]).
- [10] J. S. Lienard: Speech analysis and reconstruction using short-time elementary waveforms, *IEEE ICASSP '87*, pp. 948-951, 1987.
- [11] Bischoff, J. , R. Gold and J. Horton, “A micro computer-based network for live performance,” *Computer Music Journal* 2(3), pp. 24-29, 1978.
- [12] Mitsuhashi, Y., “Audio Signal Synthesis by Functions of Two Variables,” *Journal of the Audio Engineering Society* 30(10), pp. 701-706. 1982.
- [13] Boulanger, R. “Scanned Synthesis & Csound @Csounds.com”, <http://www.csounds.com/scanned>.
- [14] Boulanger, R. and Smaragdis, P. « Scanned Synthesis : An Introduction and Demonstration of a New Synthesis and Signal processing Technique,” *Proc. of ICMC*, Berlin, 2000.