

多重ベータ分布を用いた音色形状の数理モデリングによる楽器音生成*

中鹿 亘, 滝口 哲也, 有木 康雄 (神戸大・工)

1 はじめに

デジタルオーディオの分野において, 楽器音を電子的手法により生成するミュージックシンセシスが盛んである. この分野では物理モデリングやPCM音源による楽器音の生成が代表的であるが, 当研究では多重ベータ分布と呼ばれる分布関数を用いて楽器音の時間調波構造を数式的に表現し, 逆にモデル化された分布から倍音加算方式 [2] で楽器音を生成する手法を提案する. これは分布モデリングといった視点から楽器音のシンセサイズを行うといった意味で新しい手法である.

2 分布モデリングによる楽器音生成

本研究で提案する楽器音生成手法の概要について述べる. 手法の大まかな流れを Fig. 1 に示す.

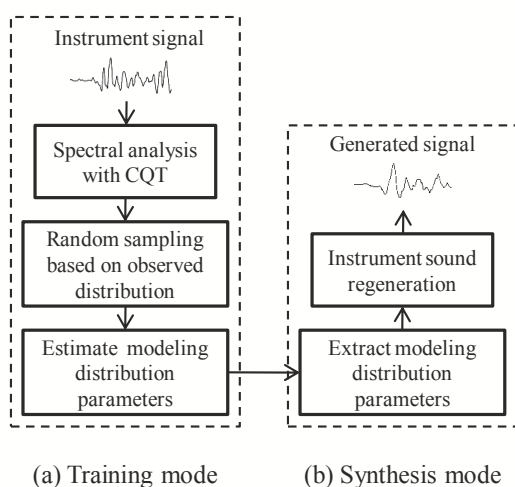


Fig. 1 提案手法の流れ

Fig. 1 のように, 提案手法は大きく分けて学習モードと合成モードの 2 つのセクションから成る. まず学習モードでは, 単音で鳴らした楽器音の音響信号をスペクトル解析し, その後分布形状に基づく乱数サンプリング, モデル分布パラメータ推定という手順を経て楽器ごとの分布モデリングを行う. 合成モードでは逆にこのモデル分布のパラメータを用いて楽器音を倍音加算方式で再合成する.

ここで分布モデリングとは, 各々の倍音に相当する強度時間変化を, ある分布関数を用いて近似するこ

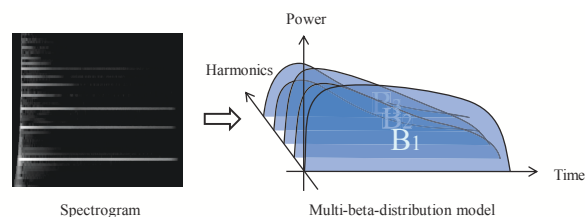


Fig. 2 音色形状の多重ベータ分布によるモデル化

とを言う. 例えば我々が従来研究で考案した多重ベータ分布によって楽器音の調波構造をモデル化する概念図を Fig. 2 に示す. この分布モデリングを行う際には次の 3 つの前提条件が必要である. まず 1 つ目に, 楽器音はいずれの音高においても, 調波時間構造は変わらないと仮定する. また, 楽器が発音されている間ピッチの変動がないものとする (つまりビブラート等はないものとする). 最後に, 楽器音は基音とその倍音のみから構成され, ノイズ等その他の周波音は出現しないものとする. これらの制約は, 少ないパラメータ数で楽器音を実現する上での妥当な簡単化と言える.

次章以降で CQT によるスペクトル解析, 多重ベータ分布の定義, 倍音加算方式による楽器音の合成法について説明する.

3 CQT による音響信号の解析

CQT (Constant Q Transform) とは, 周波数解析手法の一つである [1]. この CQT の最大の特徴は, 注目したい対数周波数だけを解析することができる点である. つまり A4 や C5 といった特定の音階に対応する周波数を離散的に解析できる. このため, 楽器音にはピッチ変動がなく倍音のみで構成されると仮定している本研究において, CQT による周波数の離散的解析は非常に有効な手法だと言える.

4 多重ベータ分布によるモデル化

多重ベータ分布とは, 複数のベータ分布が別の次元の軸上に重なり合っているような分布のことを言う. それぞれのベータ分布は, $p(t; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} t^{\alpha-1} (1-t)^{\beta-1}$ のように表される確率密度関数である. この 2

*Synthesis of Musical Instrument Sound by Modeling Timbre Structure with Multi-beta-distribution, by Toru Nakashika, Tetsuya Takiguchi, Yasuo Ariki (Kobe University, Engineering).

つのパラメータ α, β の値を様々に変化させることで、非常に多様な分布形状を表現することが可能である。これを別の次元軸に複数並べたモデルを用いれば、2次元的な形状を容易に表現できることが期待される。多重ベータ分布を以下のように定義する。

$$q(x, t; \alpha, \beta, c) = \sum_n \delta_{x, (\log n) - 1} c_n p_n(t; \alpha_n, \beta_n) \quad (1)$$

$$p_n(t; \alpha_n, \beta_n) = \frac{1}{B(\alpha_n, \beta_n)} t^{\alpha_n - 1} (1 - t)^{\beta_n - 1} \quad (2)$$

$$B(\alpha_n, \beta_n) = \int_0^1 t^{\alpha_n - 1} (1 - t)^{\beta_n - 1} dt \quad (3)$$

ただし、

$$\sum_n c_n = 1, \quad \forall n, \alpha_n, \beta_n, c_n > 0, \quad (4)$$

を満たし、 $\delta_{i,j}$ はクロネッカーのデルタを表す。また x, t, n は、それぞれ対数周波数、時間、倍音インデックスを指す。

多重ベータ分布の各パラメータの推定法についてはここでは省略するが、最尤法によってパラメータを推定することができる。

5 分布パラメータからの楽器音合成

多重ベータ分布のパラメータから倍音加算方式によって楽器音を合成する手法について述べる。倍音加算方式では、合成される楽器音信号 $s(t)$ は (5) 式で表される [2]。

$$s(t) = \sum_n a_n(t) \sin\left(\frac{2\pi f_n t}{T}\right) \quad (5)$$

n はハーモニクスを表わすインデックス、 f_n は n 倍音の周波数、 T は発音長である。ここで $a_n(t)$ を (6) 式のようにおけば、学習済みの分布パラメータからの楽器音合成が可能である。

$$a_n(t) = c_n \cdot p\left(\frac{t}{T}; \alpha_n, \beta_n\right) \quad (6)$$

ただし、 p は c_n, α_n, β_n をパラメータとする多重ベータ分布の部分分布である。

6 評価実験

ピアノの単音 (A4 : 440 Hz) をサンプリングレート 22.05 kHz で録音した音響データに対し、CQT を用いてスペクトル解析を行った。多重ベータ分布のパラメータを推定規則により楽器音構造をモデル化し、最後に楽器音の生成を行った。

多重ベータ分布による楽器音モデリングの実験結果は Fig. 3 のようになった。左図が入力された元の音響信号の強度分布であり、右図が実験結果で得られ

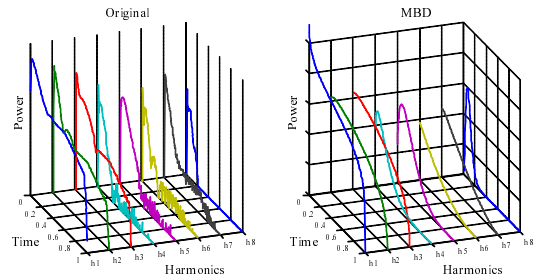


Fig. 3 多重ベータ分布によるピアノ音のモデル化

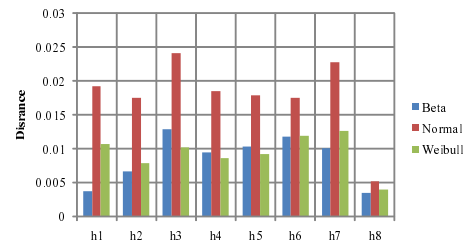


Fig. 4 各分布による DP マッチングの結果比較

た多重ベータ分布によるモデル形状を示す。また、多重ベータ分布によるモデル化の有効性を確かめるため、多重ベータ分布の他に正規分布、Weibull 分布によって同様に楽器構造をモデル化し、オリジナル分布との DP マッチングをとったときの比較を Fig. 4 に、各倍音の DP 距離の総和をとったときの分布性能比較を Table. 1 に示す。

これらの図表より、多重ベータ分布が最もピアノ音の楽器構造をよく近似できていることが分かる。これはベータ分布の形状多様性がピアノ音の急峻な立ち上がり (アタック音) を、他の分布よりもよく表現できたからだと考えられる。

Table 1 オリジナル分布との総和距離

Beta	Normal	Weibull
0.068063	0.142527	0.074789

7 おわりに

本研究では楽器の調波時間構造をモデル分布によって表現し、再合成を行う手法を提案した。実験結果により、多重ベータ分布を用いた場合が一番よく音色形状を近似できていることが分かった。

参考文献

- [1] Derry Fitzgerald et al., "Towards an inverse constant Q transform," 120th AES Convention, pp.12-16, 2006.
- [2] Xavier Rodet, "Musical Sound Signal Analysis/Synthesis: Sinusoidal+Residual and Elementary Waveform Models," TFTS'97, 1998.