

多重ベータ分布による音色形状モデルを用いた 多重楽音の解析*

中鹿 亘, 滝口 哲也, 有木 康雄 (神戸大・工)

1 はじめに

近年情報処理技術の応用例として、音楽を対象にした研究が盛んになっている。中でも音楽音響信号からの楽音特徴量（音高，強度，発音時刻，音価など）の抽出は最も関心の高い課題の一つであり，この分野の研究は過去に数多く見られる。本研究では楽器の特徴を表すための多重ベータ分布に関して定義し，それを用いて多重楽音の楽音特徴量を抽出する新しい手法を提案する。

2 多重楽音の解析手法

2.1 提案手法による楽音解析の概要

本研究で提案する楽音特徴量の抽出手法の概要について述べる。解析手法の大まかな流れは Fig. 1 のようになる。

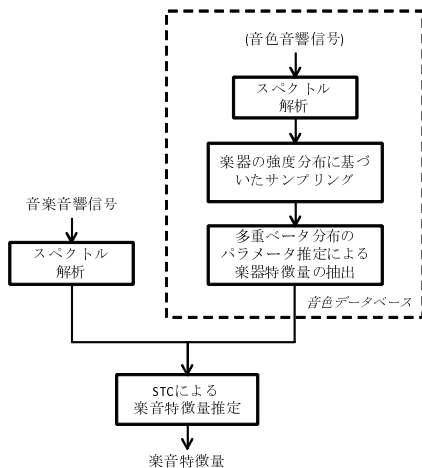


Fig. 1 解析手法のフローチャート

入力された音楽音響信号は予め音色ごとに用意されたデータベースを用いて，STC(Spectro-Temporal structured Clustering)[1] で定式化された数式を解くことで，各音響オブジェクトの楽音特徴量を同時に推定することができる [2]。本研究ではデータベースとして蓄積される楽器の特徴量の表現手法に焦点を当てているので，STC による問題の解法についてはここでは省略する。

学習される楽器ごとの音響信号には，ある楽器がある音高で一定時間演奏された，正規化された音響

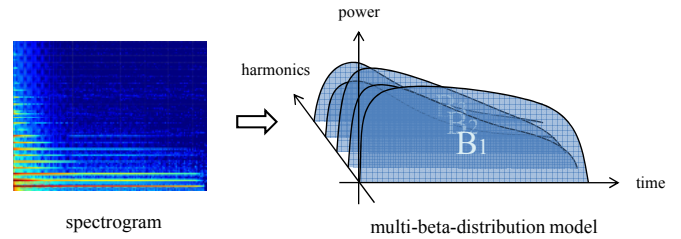


Fig. 2 音色形状の多重ベータ分布によるモデル化

データを用いる。この音色音響信号をスペクトル解析すると，基音とその倍音の強度分布を読み取ることができる。提案手法では，Fig. 2 のように，この強度分布を倍音ごとに離散的に取り出し，それぞれの強度分布を，複数のベータ分布の集合体として数式的に表現する。

3 多重ベータ分布の定式化

3.1 多重ベータ分布の定義

本研究で提案する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。

* Analysis of mixed music sounds using multi-beta-distribution model for harmonic-timbre structure, by Toru Nakashika, Tetsuya Takiguchi, Yasuo Ariki (Kobe University, Engineering).

3.2 パラメータの推定法

3.2.1 パラメータ α, β の推定

ベータ分布のパラメータ α, β は、最尤法を用いて推定できる [3] . 今、 $X_1 \leq \dots \leq X_k$ が或るベータ分布 $p(t; \alpha, \beta)$ からサンプリングされたとすると、対数尤度関数は (5) 式で表される .

$$\begin{aligned} l(\alpha, \beta) &= \log \prod_{i=1}^k p(X_i) \\ &= -k \log [B(\alpha, \beta)] + (\alpha - 1) \left[\sum_{i=1}^k \log X_i \right] \\ &\quad + (\beta - 1) \left[\sum_{i=1}^k \log(1 - X_i) \right] \end{aligned} \quad (5)$$

ここで $\frac{\partial l}{\partial \alpha} = 0, \frac{\partial l}{\partial \beta} = 0$ を解くと、(6)(7) 式を得る .

$$\begin{cases} \Psi(\alpha) - \Psi(\alpha + \beta) = \frac{1}{k} \sum_{i=1}^k \log X_i \\ \Psi(\beta) - \Psi(\alpha + \beta) = \frac{1}{k} \sum_{i=1}^k \log(1 - X_i) \end{cases} \quad (6)$$

$$\Psi(\beta) - \Psi(\alpha + \beta) = \frac{1}{k} \sum_{i=1}^k \log(1 - X_i) \quad (7)$$

ただし、 $\Psi(x) = \frac{d}{dx} \log \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}$ は digamma 関数を表す . (6)(7) 式の右辺は既知であるので、両式を満たす α と β を近似的に求めることが可能である .

以上の議論は、多重ベータ分布のパラメータ $\alpha \in R^n, \beta \in R^n$ の各要素についても同様のことが言える .

3.2.2 パラメータ c の推定

c は、複数のベータ分布間の重みパラメータである . 楽器構造の各倍音の平均強度は、基音の平均強度との比率として表現されるのが自然な発想である . よって、学習に用いる音響信号から、パラメータ $c \ni c_n$ を (8) 式のように決定できる .

$$c_n = \frac{K_n}{\sum_n K_n} \quad (8)$$

$$K_n = \frac{m_n}{m_1} \quad (9)$$

m_n は、第 n 番目の倍音の強度分布の平均を表す . ただし便宜上、基音は $n = 1$ の倍音を表すものとする .

4 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、窓幅 2,048 の短時間フーリエ変換を用いてスペクトル解析を行った . それぞれの倍音の強度分布に従って学習用のデータをサンプリングし、多重ベータ分布のパラメータを推定した .

実験により得られた結果は Fig. 3 のようになった . 左図が元の音響信号の強度分布であり、右図が実験

結果で得られた多重ベータ分布によるモデルを示す . また可視性を上げるため、同様に 2 倍音の分布を 2 次元上にプロットしたものを Fig. 4 に示す .

これらの図より、分布のピークや立下り位置など、楽器の倍音構造の大まかな形状が、多重ベータ分布を用いて表現できていることが分かる .

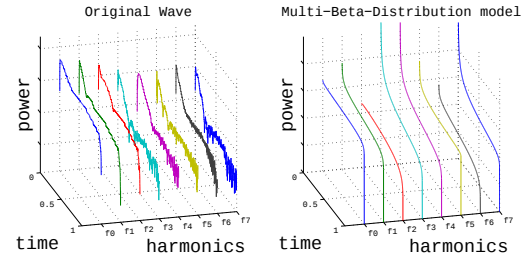


Fig. 3 実験結果

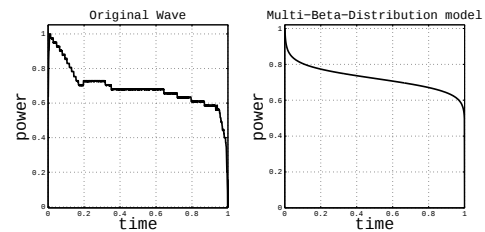


Fig. 4 2 倍音の比較結果

5 おわりに

本研究では楽器の各倍音の強度分布を多重ベータ分布によって表現する手法を提案し、その有効性を確かめるための評価実験を行った . 実験結果により、提案手法の多重ベータ分布で音色形状を数式で表現できることが分かった .

今後はさらに楽器の種類を増やして評価実験を行っていきたい .

参考文献

- [1] 亀岡 他, “調波時間構造化クラスタリング (HTC) による音楽音響特徴量の同時推定,” 情報処理学会研究会 MUS, 71-78, 2005.
- [2] K. Miyamoto *et al.*, “Harmonic-Temporal-Timbral Clustering (HTTC) for the analysis of multi-instrument polyphonic music signals,” ICASSP 2008, 113-116, 2008.
- [3] R. GNANADESIKAN *et al.*, “Maximum likelihood Estimation of the Parameters of the Beta Distribution from Smallest Order Statistics,” Technometrics 9, 607-620, 1967.