

微分方程式モデルに基づく音響センシングと 信号処理に関する研究

早川 大智^{1,†1,a)} 安藤 繁^{1,b)}

概要：楽音・音声信号を表現する新しいモデルとして、微分方程式モデルを提案する。微分方程式は、楽器や音声の物理挙動に基づいて構築され、シンプルな式ながら非定常信号を表現可能である。また、微分方程式モデルのパラメータ推定の枠組みも提案する。パラメータ推定は、微分方程式に荷重積分法を適用することで代数方程式を導出し、代数方程式を解くことによって実施される。ビブラートを含む楽音信号に対して提案手法を適用した。結果、各楽器特有のビブラートの特徴を抽出できることを示す。また、提案手法で得られたパラメータを基に波形合成を行った結果も報告する。

Acoustic Sensing and Signal Processing Based on Differential-Equation Modeling

Abstract: We propose a differential equation model for musical sound signal and its parameter estimation method. The differential equation is derived from physical models of musical instruments. The differential equation model is simple and can represent various non-stationary signals with its parameters. The parameters in the differential equation model are estimated by weighted integral method. In the weighted integral method, the differential equation is integrated in the observed time region with a set of weight functions to obtain algebraic equations that include the weighted integrals as the coefficients and the differential equation parameters as the unknowns. The parameters are estimated by solving these algebraic equation. The proposed model and parameter estimation method are applied to musical sound signal with vibrato. We show that the differential equation model can extract a characteristic of vibrato inherent in each musical instrument. We also show that the waveform synthesized from estimated parameters is hard to distinguish perceptually from the original signal.

1. はじめに

楽音分析や音声分析は、音声や楽音の合成および変調といった信号処理問題に対して様々な知見を与える。

楽音や音声分析で広く用いられているモデルの一つに、各調波成分を振幅と周波数が非定常の正弦波で近似する正弦波モデルが挙げられる。正弦波モデルを用いた分析では、各調波成分に対応するパラメータ、すなわち振幅、周波数、位相を正確に推定することが必要不可欠である。文献 [1] では、分析区間において信号が定常であるという仮

定のもと、パワースペクトルのピークにおける周波数、振幅、位相に基づいてパラメータを推定する手法が提案されている。また、近年では、分析区間における信号の非定常性を表現すべく、瞬時周波数を利用する手法 [2]、正弦波成分の周波数と振幅の非定常性を簡単な時変関数で近似する手法 [3] が提案されている。しかし、これらの手法の殆どでは、短時間フーリエ変換 (Short Time Fourier Transform, STFT) で得られた複素スペクトルまたはパワースペクトルに基づいて、パラメータが推定されている。そのため、分析時間幅が小さい場合の周波数分解能の低下あるいは分析時間幅が大きい場合の時間分解能の低下による推定性能劣化の可能性があり、パラメータ推定の性能は分析時間幅の調整に左右されるという問題を孕んでいる。したがって我々は、STFT による解析を前提とした手法ではなく、物理的な音の発生モデルに立ち戻った楽音や音声の捉え方、およびセンシングの枠組みが必要であると考え、

¹ 東京大学情報理工学系研究科
University of Tokyo, Graduate School of Information Science and Technology

^{†1} 現在、株式会社東芝 研究開発センター
Presently with R&D center, Toshiba Corporation, Kawasaki, Kanagawa, Japan

^{a)} daichi1.hayakawa@toshiba.co.jp

^{b)} Shigeru.Ando@ipc.i.u-tokyo.ac.jp

本報告では、各調波成分が時変係数の微分方程式に従うとして近似する時変微分方程式モデル（以下、微分方程式モデル）を提案する。本報告では、微分方程式モデルの初期検討の題材として、ビブラートを含む擦弦楽器及び歌声を対象としたモデル化に焦点を当てる。微分方程式は、楽音や音声の発生機構を単純化した物理モデルに基づいて導出される。導出された微分方程式は、シンプルな式でありながら、パラメータに応じて様々な種類の非定常信号を表現可能である。

同時に本報告では、微分方程式モデルを用いた分析におけるパラメータ推定の枠組みも提案する。微分方程式のパラメータ推定には、荷重積分法 [4] を用いる。荷重積分法は数学的に厳密であり、周波数推定でクラメル・ラオの下限に近い推定有効性を達成する [4] など、微分方程式のパラメータ推定に非常に有効であることが示されている [5], [6]。荷重積分法の枠組みでは、各調波成分が従う微分方程式からパラメータを未知数とする代数方程式を導出する。その後複数の代数方程式を連立して解くことにより、パラメータを推定する。

本報告では、ビブラートを含む擦弦楽器および歌声の物理モデルに基づく微分方程式モデルの構築、荷重積分法によるパラメータ推定手法を提案し、実楽音に対する微分方程式モデルの適用実験結果を報告する。

2. ビブラートを含む楽器の物理モデル

2.1 擦弦楽器

弓で擦られた弦は、弦が弓と固着し静止摩擦力を受けて弓の速度で移動する状態（固着）と、弓とは独立に運動する状態（滑り）を交互に繰り返す。ビブラートは、弦を抑える位置および張力を周期的に変動させることにより生じる。本報告では、文献 [7] の固着・滑りモデルに、張力の周期的変動を加えたモデルに基づき、調波成分に対応する固有モードが従う微分方程式を導出する。

長さ L 、線密度 ρ の弦が $0 \leq x \leq L$ に張力 T の大きさで張られており、弦上の微小領域 $x_0 \leq x \leq x_0 + \Delta x$ を弓で擦っていると仮定する。また、弦が弓を滑るときに働く動摩擦力を 0 と仮定する。張力が $T(t) = T_0 + \Delta T \cos \alpha t$ のように変動すると仮定すると、弦の横波の運動方程式は、

- 滑り状態

$$\rho \frac{\partial^2 u}{\partial t^2} = T(t) \frac{\partial^2 u}{\partial x^2}, \quad (1)$$

- 固着状態

$$\rho \Delta x \frac{\partial^2 u}{\partial t^2} = T(t) \Delta x \frac{\partial^2 u}{\partial x^2} + \theta(x - x_0) \theta(x_0 + \Delta x - x) F \quad (2)$$

となる。ただし $u(x, t)$ は弦の縦方向の変位、 F は弓からの力、 $\theta(x)$ はヘヴィサイド関数である。

2.1.1 滑りの状態での固有モードが従う微分方程式

$\Delta T = 0$ のときの弦の正規固有関数を用いて、固有モー

ド別の微分方程式に変形する。 $\Delta T = 0$ のときの弦の正規固有関数は、 $X_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}$ であり、固有振動数は、 $\omega_n = \frac{n\pi}{L} \sqrt{\frac{T_0}{\rho}}$ である。(1) 式の解を、正規固有関数 $X_n(x)$ で級数展開して、

$$u(x, t) = \sum_{n=1}^{\infty} X_n(x) \psi_n(t) \quad (3)$$

と表す。(3) 式を、(1) 式に代入すれば、

$$\rho \sum_{n=1}^{\infty} X_n(x) \ddot{\psi}_n(t) - (T_0 + \Delta T \cos \alpha t) \sum_{n=1}^{\infty} \frac{d^2 X_n(x)}{dx^2} \psi_n(t) = 0 \quad (4)$$

となる。ただし、 $\cdot$ は時間微分を表す。(4) 式の両辺に第 m 次の正規固有関数 $X_m(x)$ をかけて、 $[0, L]$ の区間で x について積分すると、正規直交関数の性質

$$\int_0^L X_m(x) X_n(x) dx = \delta_{mn},$$

$$\int_0^L X_m(x) \frac{d^2 X_n(x)}{dx^2} dx = - \left(\omega_m \sqrt{\frac{\rho}{T_0}} \right)^2 \delta_{mn}$$

より、以下が成り立つ。

$$\ddot{\psi}_m(t) + \omega_m^2 \left(1 + \frac{\Delta T}{T_0} \cos \alpha t \right) \psi_m(t) = 0. \quad (5)$$

2.1.2 固着の状態での固有モードが従う微分方程式

2.1.1 節と同様、(3) 式を (2) 式に代入し、両辺に正規固有関数 $X_m(x)$ をかけて、 $[0, L]$ の区間で x について積分すると、正規直交関数の性質、および

$$\int_0^L \theta(x - x_0) \theta(x_0 + \Delta x - x) X_m(x) dx \simeq X_m(x_0) \Delta x$$

より、

$$\ddot{\psi}_m(t) + \omega_m^2 \left(1 + \frac{\Delta T}{T_0} \cos \alpha t \right) \psi_m(t) = \frac{F}{\rho} X_m(x_0). \quad (6)$$

$\psi_m(t) = \frac{1}{\omega_m^2} \frac{F}{\rho} X_m(x_0) + \psi'_m(t)$ とおけば、

$$\ddot{\psi}'_m(t) + \omega_m^2 \left(1 + \frac{\Delta T}{T_0} \cos \alpha t \right) \psi'_m(t) = 0. \quad (7)$$

つまり $\psi_m(t)$ から直流成分を除いた $\psi'_m(t)$ は、弦が弓を滑っているときと同様の時変係数微分方程式に従う。

2.1.3 擦弦楽器の反響特性

実際に観測される音は、弦の振動そのものではなく、振動する弦が楽器の胴体に伝わり、胴体から放射されたものである。ビブラートによって周波数が正弦的に変動する場合、音圧の強度変化も対数スケールでほぼ正弦的に変動する [8] と近似すると、 m 版目の調波成分は $f_m(t) = A e^{-\frac{k_m}{2} \sin \alpha t} \psi_m(t)$ と表せる。ただし、 A, k_m は定数である。これを (5) 式に代入し、 $\omega_m \gg \alpha$ と仮定すると、 m 版目の調波成分は

$$\ddot{f}_m(t) + k_m \alpha \cos(\alpha t) \dot{f}_m(t) + \omega_m^2 \left(1 + \frac{\Delta T}{T_0} \cos \alpha t \right) f_m(t) = 0 \quad (8)$$

の微分方程式に従う。

2.2 歌声

歌声は声道形状に固有の音響特性により特徴付けられるが、声道形状は複雑であり、その形状から厳密な音響特性を計算するのは困難である。そこで、声道を断面積が変化する一本の管とみなして、平面波伝播を仮定できる低い周波数帯域で音響特性を計算する [9]。ビブラートは、声道を振動させることにより生じる [10] という考え方にに基づき、声道の断面積が周期的に変わるものとしてモデル化する。

声道の管に沿って座標 x をとる。 $p(x, t)$ を圧力、 $A(x, t)$ は位置 x における断面積、 ρ は空気密度、 K を体積弾性率と置く。空気の粘性を無視して考えると、 $p(x, t)$ と $A(x, t)$ について以下の式が成り立つ [11]。

$$\frac{\partial}{\partial x} \left(\frac{A(x, t)}{\rho} \frac{\partial p(x, t)}{\partial x} \right) = \frac{1}{K} \frac{\partial A(x, t)}{\partial t} \frac{\partial p(x, t)}{\partial t} + \frac{A(x, t)}{K} \frac{\partial^2 p(x, t)}{\partial t^2}. \quad (9)$$

$A(x, t) = A_0 + \Delta A \sin \alpha t$ とおき、(9) 式を整理すると、

$$\frac{K}{\rho} \frac{\partial^2 p}{\partial x^2} = \frac{\Delta A \alpha \cos \alpha t}{A_0 + \Delta A \sin \alpha t} \frac{\partial p(x, t)}{\partial t} + \frac{\partial^2 p(x, t)}{\partial t^2}. \quad (10)$$

ここで、 $A_0 \gg \Delta A$ と仮定し、

$$\frac{\Delta A \alpha \cos \alpha t}{A_0 + \Delta A \sin \alpha t} = \frac{1}{A_0} \frac{\Delta A \alpha \cos \alpha t}{1 + \frac{\Delta A}{A_0} \sin \alpha t} \simeq \frac{\Delta A \alpha \cos \alpha t}{A_0}$$

と近似すると、(10) 式は、以下のように変形できる。

$$\frac{K}{\rho} \frac{\partial^2 p}{\partial x^2} = \frac{\Delta A \alpha \cos \alpha t}{A_0} \frac{\partial p(x, t)}{\partial t} + \frac{\partial^2 p(x, t)}{\partial t^2}. \quad (11)$$

2.1 節と同様に、 $p(x, t)$ を正規固有関数 $X_n(x)$ で展開し

$$p(x, t) = \sum_{n=1}^{\infty} X_n(x) \psi_n(t)$$

と置いて、(11) 式に代入すると、 $\psi_m(t)$ は、

$$\ddot{\psi}_m(t) + \frac{\Delta A \alpha \cos \alpha t}{A_0} \dot{\psi}_m(t) + \omega_m^2 \psi_m(t) = 0 \quad (12)$$

の微分方程式に従う。ただし、 ω_m は固有振動数である。

m 版目の調波成分が $f_m(t) = C \psi_m(t)$ (C は定数) として表されるならば、(13) 式より、 m 版目の調波成分は

$$\ddot{f}_m(t) + \frac{\Delta A \alpha \cos \alpha t}{A_0} \dot{f}_m(t) + \omega_m^2 f_m(t) = 0 \quad (13)$$

の微分方程式に従う。

3. 微分方程式モデル

(8) 式、(13) 式に基づき、楽音信号 $s(t)$ に含まれる各調波成分 $f_m(t)$ が、以下に示す時変係数の微分方程式に従うとして近似する微分方程式モデルを提案する。

$$\begin{aligned} s(t) &= \sum_m f_m(t) \\ \forall m, \quad \ddot{f}_m(t) &+ k_m \cos(\alpha t + \psi_m) \dot{f}_m(t) \\ &+ a_m^2 (1 + \epsilon_m \cos(\alpha t + \phi_m)) f_m(t) = 0 \end{aligned} \quad (14)$$

調波成分ごとの微分方程式のパラメータは $a, \alpha, k, \epsilon, \psi, \phi$ の 6 種類となる。以下、それぞれのパラメータが持つ意味を説明する。

3.1 $a, \alpha, \epsilon, \phi$

(14) 式で $k = 0$ として考えた微分方程式

$$\ddot{f}(t) + a^2 (1 + \epsilon \cos(\alpha t + \phi)) f(t) = 0 \quad (15)$$

は、マシュー方程式 [12] と呼ばれている。マシュー方程式は係数励振系の運動方程式の一種であり、駆動項が 0 で、係数の周期的な変動により振動が発生する。後述の通り、 $a, \alpha, \epsilon, \phi$ の組み合わせにより、様々な波形が表現できる。

【 $\alpha = 2a$ の場合】 $\phi = 0$ のとき、指数増大波形になり、 $f(t) = e^{\lambda t} \cos(\alpha t)$ で近似できる [13]。 λ は ϵ に依存する定数であるため、 ϵ は増大の度合いを調節する項として機能する。また、 ϕ を変化させることにより、指数増大波形だけでなく指数減衰波形を作り出すこともできる。

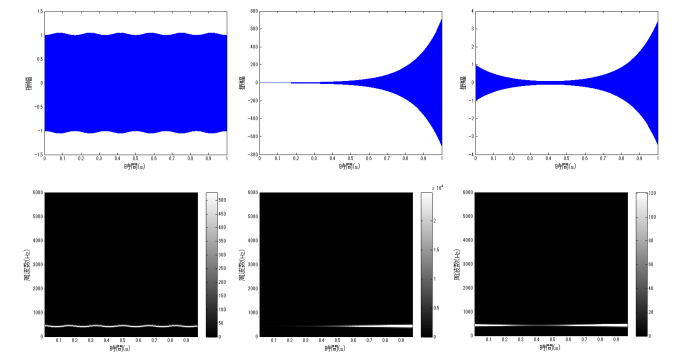
【 $a \gg \alpha$ の場合】 初期位相 ϕ 、角振動数 α の速さで、固有振動数 a が周期的に変調された波形となる [13]。 ϵ は変調の大きさを調整する項として機能する。すなわち、ビブラートに伴う周波数変調をモデル化することに相当する。

図 1(a)~(c) に、Matlab による微分方程式数値解法 (初期値は $f(0) = 1, f'(0) = 0$, ルンゲクッタ法) で生成した波形とスペクトログラムの例を示す。パラメータの組み合わせを変えることにより、周波数が周期的に変動する波形、指数増大・減衰波形が表現できていることがわかる。

3.2 k, ψ

$f(t) = C e^{-\frac{1}{2} k \sin(\alpha t + \psi)}$ (C は定数) を (14) 式に代入すると、

$$\begin{aligned} \ddot{g}(t) &+ \{a^2 (1 + \epsilon \cos(\alpha t + \phi)) - \frac{1}{4} \alpha^2 k^2 \cos^2(\alpha t + \psi) \\ &+ \frac{1}{2} \alpha^2 k \sin(\alpha t + \psi)\} g(t) = 0 \end{aligned} \quad (16)$$



(a) $a = 880\pi, \epsilon = 0.01$ (b) $a = 880\pi, \epsilon = 0.01$ (c) $a = 880\pi, \epsilon = 0.01$, $\alpha = 12\pi, \phi = 0$ $\alpha = 1760\pi, \phi = 0$ $\alpha = 1760\pi, \phi = \pi/2$

図 1 各パラメータの組み合わせにおける波形とスペクトログラム
Fig. 1 Waveforms and spectrograms for each combination of the parameters.

が成立する． $a \gg \alpha$ の場合， $(\frac{\alpha}{a})^2 \simeq 0$ と近似すると，(14) 式は (15) 式と一致する．これは，(14) 式における $f(t)$ は，(15) 式の $f(t)$ を振幅の周期的変調が加わった $\frac{1}{C}e^{\frac{1}{2}k\sin(\alpha t+\psi)}f(t)$ に置き換えたものと近似できることを意味する． ψ は振幅の変調の初期位相， k は振幅の変調の大きさを調整する項として機能する．すなわち，ビブラートに伴う振幅変調をモデル化することに相当する．

以下， a は固有振動数， α は係数振動周波数， k は減衰定数， ϵ は変調率， ψ は減衰初期位相， ϕ は係数振動初期位相と呼称する．

4. 荷重積分法によるパラメータ推定

荷重積分法では，各調波成分が従う微分方程式からパラメータを未知数とする連立方程式を導出し，その後連立して解くことで，パラメータを推定する．本報告では，係数振動周波数 α が既知という前提の下，分析区間の長さを係数振動周期の p 倍にとり $a, k, \epsilon, \psi, \phi$ を推定する．

4.1 代数方程式の導出

時間区間 $[0, T]$ において各調波成分が (14) 式に従う場合，(14) 式の両辺に任意の荷重関数 $w(t) = e^{-jn\Delta\omega t}$ (n は整数) を乗じて両辺を時間区間 $[0, T]$ で積分した

$$\int_0^T \{ \dot{f}(t) + k \cos(\alpha t + \psi) \dot{f}(t) + a^2(1 + \epsilon \cos(\alpha t + \phi)) f(t) \} e^{-jn\Delta\omega t} dt = 0 \quad (17)$$

も n ごとに成立する．ただし， j は虚数単位， $\Delta\omega = \frac{2\pi}{T}$ である．

分析区間の長さが係数振動周期の p 倍の場合， $\alpha = p\Delta\omega$ となるため，

$$g_n = \int_0^T f(t) e^{-jn\Delta\omega t} dt, \\ c_{n,p} = g_{n-p} + g_{n+p}, \quad d_{n,p} = g_{n-p} - g_{n+p}, \\ F_0 = f(T) - f(0), \quad F_1 = \dot{f}(T) - \dot{f}(0)$$

とおくと，(17) 式は以下の代数方程式に変換できる．

$$\forall n \geq 0, AB = n^2 \Delta\omega^2 g_n. \quad (18)$$

ただし，転置を $'$ として，

$$A = \begin{bmatrix} 1 \\ jn\Delta\omega \\ j\Delta\omega(nc_{n,p} - pd_{n,p}) \\ \Delta\omega(-nc_{n,p} + pd_{n,p}) \\ g_n \\ c_{n,p} \\ jd_{n,p} \end{bmatrix}', \quad B = \begin{bmatrix} F_1 + kF_0 \cos \psi \\ F_0 \\ \frac{k}{2} \cos \psi \\ \frac{k}{2} \sin \psi \\ a^2 \\ \frac{a^2\epsilon}{2} \cos \phi \\ \frac{a^2\epsilon}{2} \sin \phi \end{bmatrix}.$$

である．ここで， g_n は分析区間で有限時間フーリエ変換し

たときの周波数成分， n は周波数のインデックスに相当する．すなわち， A の要素は観測により得られる量である．また， B はパラメータのみから構成される行列である．したがって，パラメータを推定する問題は， A が得られた下で未知数である B の要素を推定する問題に帰着される．

4.2 パラメータ推定

荷重積分方程式 1 つに対して，未知数は B の要素の 7 種類存在する．よって，相異なる 4 つ以上の n における (18) 式を選択すれば，実部と虚部計 8 個以上の式を連立し，最小二乗法によって解くことで，行列 B の値が得られる．

パラメータ推定に用いる (18) 式は以下のように選択する．単一の調波成分の寄与は全周波数成分の絶対値のピーク付近に集まるため，ピーク付近の連続した少数個の周波数成分で推定を行えば，その中に含まれる調波成分のパラメータを推定することができる．そこで本稿では，ピークをとる n 周りの 5 個の (18) 式を用いて推定を行う．

最小二乗法によって得られた B の要素を上から順にそれぞれ b_i ($1 \leq i \leq 7$) とおくと，

$$F_0 = b_2, F_1 = b_1 - 2b_2b_3, \\ k = 2\sqrt{b_3^2 + b_4^2}, \psi = \arg(b_3 + ib_4), a = \sqrt{b_5}, \\ \epsilon = \frac{2\sqrt{b_6^2 + b_7^2}}{b_5}, \phi = \arg(b_6 + ib_7)$$

より，未知パラメータ $k, \psi, a, \epsilon, \phi$ が求まる．

4.3 係数振動周波数 α の設定

$a \gg \alpha$ に設定すると，微分方程式が α の速さで周波数と振幅が変調する波形となる．すなわち，ビブラートが表現できる可能性がある．一方， $\alpha = 2a$ に設定すると，微分方程式モデルは，固有振動数 a の指数増大波形あるいは指数減衰波形を表現できる．すなわち，楽器の立ち上がり，鳴り終わりを表現できる可能性がある．そこで本報告では，ビブラート存在区間では α をビブラートの速さに設定し，不在区間では $\alpha = 2a$ と設定してパラメータ推定を行う．

5. 実楽器へのモデルの適用実験

5.1 実験条件

5.1.1 分析対象の信号

文献 [14] に収録されている，ビブラートが含まれたヴァイオリンの $F\#_5$ 音 (約 727 Hz)，歌声の B_4 音 (約 461 Hz) を選んだ．いずれの楽音も無響室で録音されたものであり，サンプリングレートは 44100Hz である．

5.1.2 ビブラート区間の推定

ビブラート存在区間と不在区間で異なる係数振動周波数 α を設定するため，前準備としてビブラート存在区間を推定する．

図 2 に，ビブラートが存在する区間の推定の手順を示

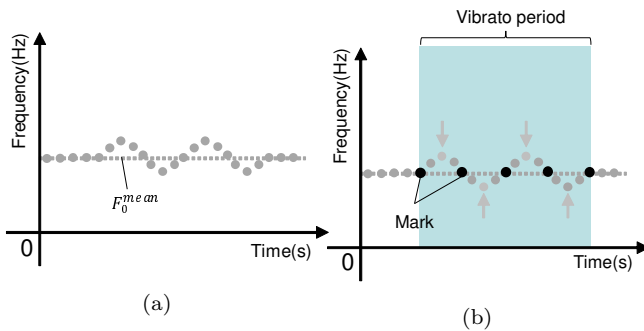


図 2 ビブラート区間の推定手順.

Fig. 2 Estimation procedure of vibrato period.

す。まず、分析区間 256 点、シフト 64 点で、文献 [4] の方法により F_0 成分の周波数を分析区間ごとに推定した (図 2(a))。次に、推定された F_0 系列のピーク及びボトムが F_0 系列の平均 F_0^{mean} から 0.5% 以上変動している箇所を検出し、検出されたボトムまたはピーク前後で、 F_0 系列の平均 F_0^{mean} に最も近づく分析区間をマークした図 2(b))。最後に、マークされた系列の始端から終端までをビブラート存在区間とし、それ以外の区間をビブラート不在区間とした。

5.1.3 パラメータ推定

ビブラート不在区間においては、 m 番目の調波成分における係数振動周波数 α_m を $\alpha_m = mF_0^{mean}$ 、 $p = 5$ に設定して、調波成分ごとにパラメータ推定した。一方、ビブラート存在区間においては、5.1.2 節でマークされた時間に基づき、ビブラート一周周期ごとに分析区間をとり、 $p = 1$ に設定してパラメータ推定を行った。

5.2 演奏におけるビブラートの特徴抽出

ビブラートが含まれた楽音は、時間が経過するにつれて周波数の変動幅が大きくなる傾向があること、周波数の変動幅と振幅の変動幅には相関があることが指摘されている [8], [15]。本報告では、振幅の変調の大きさに対応する減衰定数 k 、周波数の変調の大きさに対応する変調率 ϵ を、ビブラート一周周期ごとに推定し、得られたパラメータからこれらの傾向を計測できるか検証した。

ヴァイオリンおよび歌声の基本音で推定したビブラート一周周期ごとの k および ϵ の時間変化を、それぞれ図 3、図 4 に示す。横軸の番号は検出されたビブラートのインデックス、縦軸は k または ϵ の値である。図 3、図 4 から、ヴァイオリンと歌声のどちらも、時間経過とともに k および ϵ が大きくなり、鳴り終わり（歌い終わり）直前で小さくなっていることがわかる。すなわち、時間が経過するにつれ周波数の変動幅が大きくなっていること、周波数の変動幅と振幅の変動幅には相関があること、鳴り終わり（歌い終わり）においてはビブラートのかけ方を小さくする傾向があることが、計測によって確認できる。

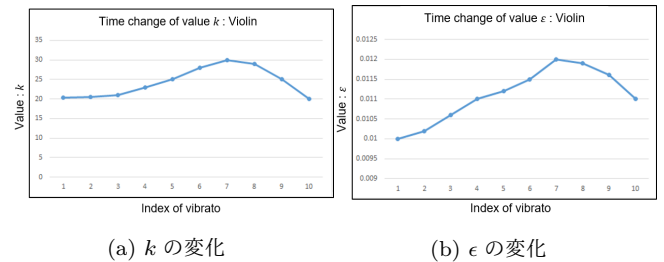


図 3 ヴァイオリンの基本音で推定したビブラート一周周期ごとの k および ϵ 。

Fig. 3 Each vibrato's k and ϵ estimated from fundamental component of violin.

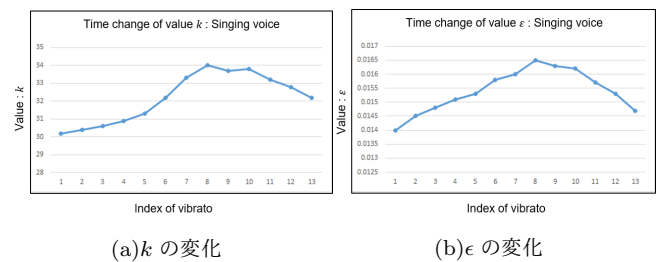


図 4 歌声の基本音で推定したビブラート一周周期ごとの k および ϵ 。

Fig. 4 Each vibrato's k and ϵ estimated from fundamental component of singing voice.

5.3 波形合成実験

本報告では、各分析区間で推定された調波成分ごとのパラメータから、微分方程式の数値解法によって波形を生成し、合成する実験を行った。数値解法は Matlab に実装されているルンゲクッタ法を用いた。

5.3.1 合成手順

微分方程式の解を求めるためには各パラメータを推定するだけでなく、初期値も推定する必要がある。しかし、荷重積分法の枠組みで推定できるパラメータは、 $F_0 = f(T) - f(0)$ 、 $F_1 = \dot{f}(T) - \dot{f}(0)$ であり、初期値 $[f(0), \dot{f}(0)]$ のみを推定することは不可能である。

本報告では妥協策として、最初の分析区間では、正弦波モデルを適用して高調波ごとにパラメータの推定を行い、パラメータを基に生成した各調波成分の波形から調波成分ごとの初期値 $[f(0), \dot{f}(0)]$ を得た。その後、求められた $[f(0), \dot{f}(0)]$ を用いて、Matlab の数値解法により各調波成分の波形を生成した。同時に、パラメータ推定によって得られた F_0, F_1 の情報から $[f(T), \dot{f}(T)]$ を計算した。次の分析区間では、最初の分析区間によって求められた $[f(T), \dot{f}(T)]$ を次の分析区間の初期値 $[f(0), \dot{f}(0)]$ に流用して、調波成分ごとに波形を生成した。当該分析区間における初期値 $[f(0), \dot{f}(0)]$ が直前の分析区間における $[f(T), \dot{f}(T)]$ と一致させることで、分析区間の境界で連続した波形を生成することができるためである。以後、この手順を繰り返して、調波成分ごとに波形の生成を行い、合成した。

5.3.2 結果

図 5 に、ヴァイオリンの元波形と推定された微分方程式

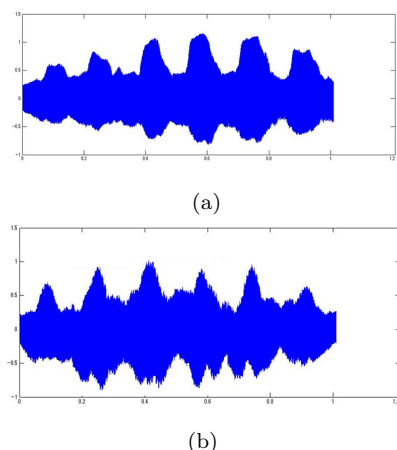


図 5 ヴァイオリンの (a) 元波形と (b) 合成波形.

Fig. 5 (a) Original waveform and (b) synthesized waveform of violin.

のパラメータに基づいて合成した波形, 図 6 に, 歌声の元波形と推定された微分方程式のパラメータに基づいて合成した波形を示す.

いずれの実験でも, 初期値と微分方程式のパラメータを与えるだけで, 元波形に近い波形が合成できている. とくに, ヴァイオリンでは, 聴覚上の主観評価であるが, 観測音と非常に近い波形が得られた. ピブラートが含まれている楽音の性質を微分方程式モデルがよく表現できていること, また, 高い精度でパラメータが推定できていることがわかる. これらの結果は, 楽音の立ち上がりやピブラートの性質が, 駆動項が 0 の微分方程式モデルにおける係数励振の作用の枠組みで表現できる可能性があることを示唆している.

本報告における合成手法では, 当該分析区間における初期値 $[f(0), \dot{f}(0)]$ は直前の分析区間における $[f(T), \dot{f}(T)]$ と等しく設定しているため, 直前の分析区間における推定誤差がそのまま当該分析区間に伝搬する構造となっている. 実際, 歌声では, 2 秒以降のピブラートが含まれる区間において振幅に乱れが生じ, その乱れが時間経過とともに増幅されている. そのため, より高質な楽音合成を達成するためには, 適切な時刻ごとに初期値を変化させる, 微分方程式のパラメータを修正するなどの対処が必要と考えられる.

6. 結論

楽音・音声信号を表現する新しいモデルとして, 楽器や音声の物理挙動に基づき, 微分方程式モデルを提案した. また, 荷重積分法を用いた微分方程式モデルのパラメータ推定の枠組みも提案した. 係数振動周波数が既知という条件の下, ピブラートを含む楽音信号に対して提案手法を適用した結果, 各楽器特有のピブラートの特徴を抽出できた. また, 提案手法で得られたパラメータを基に波形合成を行った結果, 初期値と微分方程式のパラメータを与える

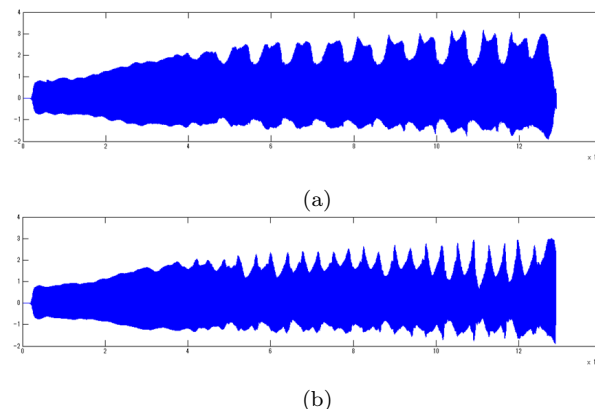


図 6 歌声の (a) 元波形と (b) 合成波形.

Fig. 6 (a) Original waveform and (b) synthesized waveform of singing voice.

だけで, 元波形に近い波形が合成できることを示した.

謝辞 実験において貴重なコメントをいただいた東京大学情報理工学系工学科の奈良 高明教授, 実験のための音源を提供いただいた首都大学東京システムデザイン学部の小野 順貴教授に, 感謝申し上げます.

参考文献

- [1] R. J. McAulay and T. F. Quatieri, "Speech analysis/synthesis based on a sinusoidal representation," IEEE Trans. Acoust. Speech Signal Processing 34, pp. 744-754 (1986).
- [2] T. Abe and M. Honda, IEEE Trans. Audio Speech Language Processing 14, pp. 1292-1300 (2006).
- [3] M. Ito and M. Yano, "Sinusoidal modeling for nonstationary speech based on a local vector transform," J. Acoust. Soc. Am. 121, pp. 1717-1727 (2007).
- [4] S. Ando and T. Nara, "An exact direct method of sinusoidal parameter estimation derived from finite Fourier integral of differential equation," IEEE Trans. Signal Processing, 57(9), pp. 3317-3329 (2009).
- [5] 安藤繁, 信学技報, vol. 115, no. 182, SIP2015-55, pp. 21-28 (2015).
- [6] 安藤繁, 信学技報, vol. 119, no. 115, EA2019-17, pp. 87-92 (2019).
- [7] 吉岡大二郎, "振動と波動," 東京大学出版会 (2005).
- [8] N. H. Fletcher and T. D. Rossing, "The Physics of Musical Instruments" Springer; 2nd ed. (1998).
- [9] J. L. Flanagan, "Speech analysis synthesis and perception," Second, Expanded Edition, Springer (1970).
- [10] J. Sundberg, "The Science of the Singing Voice," Northern Illinois University Press (1987).
- [11] A. G. Webster, Proc. Natl. Acad. Sci. USA, 5, pp. 275-282 (1919).
- [12] 松尾恒, 奈良高明, 安藤繁, "Mathieu 方程式系のパラメータ推定と自励振動型計測への応用," 計測自動制御学会論文誌, no.11, pp.785-793 (2015).
- [13] G. Arfken, H. Weber, and F. E. Harris, "Mathematical Methods for Physicists, 7th Edition" Academic Press (2012).
- [14] 日本建築学会編, "建築と環境のサウンドライブラリ," 技報堂出版, (2004).
- [15] J. Bretos and J. Sundberg, Jour. of Voice, 17(3), pp. 343-352 (2003).