

サンプル点間を考慮した非整数値遅延フィルタによる弦楽器の適応的なピッチシフト

Pitch Shifting of Stringed Instruments by Fractional Delay Filteres with Inter-sample Behaviors

京都大学 埴淵 千誉, 永原 正章, 山本 裕

Kazutaka Hanibuchi, Masaaki Nagahahara, Yutaka Yamamoto

Kyoto University

Abstract In this article, we propose a new pitch-shifting method for stringed instruments by using sampled-data H^∞ optimal fractional delay filters. Since the pitch-shifted signal length is in general different from that of the original signal, we also propose a new method for controlling the length by dividing the signal into frames on pre-specified intervals and pitch-shifting the signal on each frame.

1 はじめに

楽器のモデリングは携帯電話の着信音などの小規模なものから、多様な音を生成するシンセサイザーに至るまで広く利用されている。従来の弦楽器のモデリングとしては PCM (Pulse Code Modulation) 方式が主流である。この方式は実在音の録音再生による方式であり、波形のすべてを記録するため記憶容量が膨大になる。

そこで、ある基準となる音のみをデジタル信号として記録し、他の音階はその信号を変調することにより生成することを考える。このとき、サンプル点間の補間が必要となるが、アナログ特性を考慮したサンプル H^∞ 値制御理論を用いて設計された非整数遅延フィルタ [1] により、自然な補間と変調が可能となる。この性質を利用すれば楽音の波形を自在に伸縮すること可能になり、標本化周波数を維持したまま、高精度で自然な楽音の音高シフトが実現できる。

いっぽう、楽音データにこの手法を単純に適用すると信号の長さも変化する。この問題を解決するために、信号をフレームに分割し、その各々のフレーム上で適応的に変調を繰り返すことにより、信号の長さを調整する方法を提案する。

また数値例により、本手法の有効性を示す。

2 楽音信号の変調

楽音信号は局所的にみると同じような波形の繰り返しによる構造を持っている。この繰り返しの波形を基本波形と呼ぶことにすると、音の高低は基本波形の周期の長さによって決定される。従って信号の波形全体を時間軸方向に伸縮することによって音高が変化した信号を生成することができる。

2.1 非整数遅延フィルタによるサンプル点間の補間

本稿では非整数遅延フィルタを信号処理に適用して楽音信号の変調を行う。信号処理は計算機上で行うため、現実で演奏される楽器音の連続時間信号を離散化してからデータベースに保存しなければならず、その場合サンプル点間の情報は失われるが、非整数遅延フィルタを用いることでサンプル点間を補間することができる。図 1 は非整数遅延フィルタの設計問題である。 $F(s)$ は元の連続時間信号の特性を反映したフィルタである。連続時間信号を D だけ時間軸方向に遅延させてから標本化した信号系列に対して、元の連続時間信号を標本化してからフィルタを適用することで得られる信号系列との偏差 e_d の w からの拡大率を表す L^2 誘導ノルムを最小化することを目標として設計されるのが非整数遅延フィルタである。

$$a_0(d) = \frac{\sinh \omega_c(h-d)}{\sqrt{\omega_c} \sinh \omega_c h}, a_1(d) = e^{-\omega_c h} (e^{\omega_c d} - a_0)$$

として、連続時間信号が完全帯域制限されていると仮定すると、[1] の結果から標本化周期 h で標本化した点 $y((n-1)h)$, $y(nh)$ とサンプル点間の点 $y(nh-d)$ ($0 < d < h$) との関係

$$y(nh-d) = a_0(d)y(nh) + a_1(d)y((n-1)h)$$

が得られる。したがって $y(nh-d)$ は $y(nh)$, $y((n-1)h)$ の結合で表すことができる。この手法によって、サンプル点間上の任意の負でない時刻の点の値を補間することができる。

2.2 演奏時間の維持

楽音信号の波形全体を時間軸方向に伸縮することで音高を変えるができるが、時間長も伸縮によって変動する。信号の時間長を変えずに、楽音を変調するための手法として、信号の区間を分割し、それぞれの区間で信号の伸

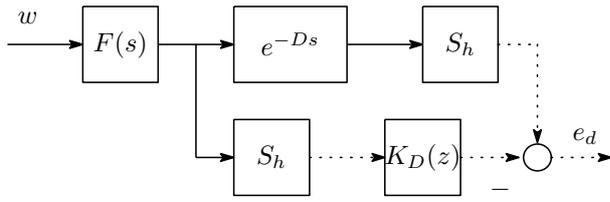


図 1: 非整数遅延フィルタの設計問題

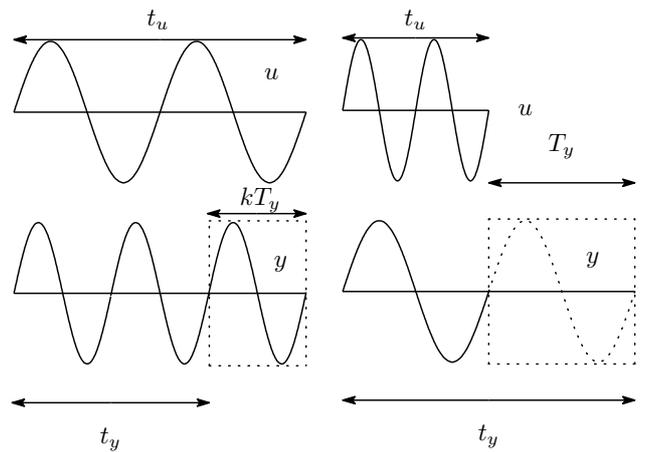
縮を行い，演奏時間の差で定義される区間上で系列を補間したり，削除する手法が挙げられる．本稿では基本波形の形が楽器の音色，基本周期の長さが音高を決定することを考慮して，時間長の差の区間を補間・削除する手法を提案する．原信号 u の周波数分析から求めたピーク周波数を f_u として，変調信号 y のピーク周波数が f_y になるように，かつ演奏時間が等しくなるように変調を行う．また $T_u = 1/f_u, T_y = 1/f_y$ とする．まず u を時間 t_u ごとに分割することを考える．また分割した信号はそれぞれについて $t_y = t_u/Fr$ になるように時間軸方向に伸縮を行う．ただし $Fr = f_y/f_u, u, y$ の分割信号の関係は図 2(a) のようになる．音高を上げる場合，再生時間の差が生じる区間が y の基本周期，またはその自然数倍になるように t_u, t_y になるように調整する．つまり $k = 1, 2, \dots$ として $t_u = t_y + kT_y$ と定める．これにより t_u, t_y が定まり，特に $t_u = kT_y Fr / (Fr - 1)$ と定めることができる．次に y の分割部の最後部時刻から時間 kT_y だけ前の時刻までの信号を複製し，演奏時間の差の区間に貼り付ける．楽器音の減衰特性を考慮すると，変調後の信号ができるだけ自然になるためには，補間部の時間が短くなるほうがよい．したがって k の値はできるだけ小さくしたい．しかし複製部には少なくとも基本周期の時間長が必要であり，つまり k は条件 $t_y \geq T_y$ を満たす必要がある．この不等式によって定まる k の条件は $k \geq Fr - 1$ である．

一方，音高を下げる場合， u の分割部を変調した y の分割部の時間 t_y は t_u より長くなる．ここで $t_y - t_u = T_y$ として， y の分割のうちこの時間区間の信号を削除する．

3 実験結果

ギターの音を原信号として録音した楽音データをもとに，数値解析ソフト MATLAB を用いてシミュレーションを行った結果を以下に記す．原信号（ピーク周波数 110.4032Hz^1 ）を $2^{5/12}$ 倍および $2^{-5/12}$ 倍に通常の変

¹元のギターの周波数が 110 Hz からずれているのは，ギターの音程調整がずれているためと考えられる．



(a) 音高を上げる場合 (b) 音高を下げる場合

図 2: 演奏時間の差の区間の補間・削除処理

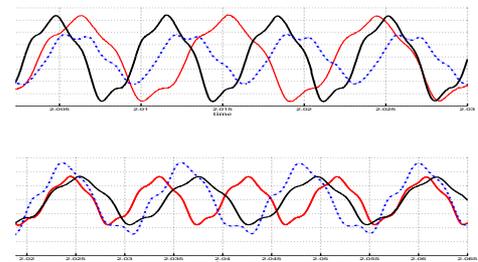


図 3: 局所的な波形の様子（点線：原信号 破線：通常変調信号 実線：提案手法）

調と提案手法によって変調させた． $2^{5/12}$ (1.3348) 倍の変調の場合，通常変調と提案手法のピーク周波数の拡大率はそれぞれ 1.3348, 1.3359． $2^{-5/12}$ (0.7492) 倍の場合はそれぞれ 0.7492, 0.7496．図 3 は原信号，変調信号（通常と提案手法）についての局所信号を表しており，変調率は上が $2^{5/12}$ 倍，下が $2^{-5/12}$ 倍になっている．いずれの場合も，提案手法の信号が通常の変調信号と同期がとれていることと，通常の変調信号の振幅の幅が独特であるのに対して，提案手法の信号は原信号と振幅の幅がほぼ同じであり，減衰特性がほぼ等しいことがわかる．このことから提案手法の信号が，時間長を維持しつつ変調している様子が読み取れる．

参考文献

- [1] M. Nagahara and Y. Yamamoto, Optimal design of fractional delay FIR filters without band-limiting assumption, *Proc. of ICASSP*, pp. 221–224, 2005.