H^∞-Optimal Fractional Delay Filters with Application to Pitch Shifting

Masaaki Nagahara * **Yutaka Yamamoto** **

* *Graduate School of Informatics, Kyoto University, Kyoto, 606-8501, JAPAN (e-mail: nagahara@ieee.org).*
** *Graduate School of Informatics, Kyoto University, Kyoto, 606-8501, JAPAN (e-mail: yy@i.kyoto-u.ac.jp).*

Abstract: Fractional delay filters are digital filters to delay discrete-time signals by a fraction of the sampling period. Since the delay is fractional, the intersample behavior of the original analog signal becomes crucial. In contrast to the conventional designs based on the Shannon sampling theorem with the band-limiting hypothesis, the present paper proposes a new approach based on the modern sampled-data H^∞ optimization which aims at restoring the intersample behavior beyond the Nyquist frequency. By using the lifting transform or continuous-time polyphase decomposition, the design problem is equivalently reduced to a discrete-time H^∞ optimization, which can be effectively solved by numerical computation softwares. Moreover, a closed-form solution is obtained under an assumption on the original analog signals. Using this closed-form solution, we introduce a sampling rate conversion with arbitrary conversion rate, and propose a new pitch shifting method for digital sound synthesis. Design examples are given to illustrate the advantage of the proposed method.

Keywords: Fractional delay filter, sampled-data control, H^∞ optimization, digital signal processing, sampling rate conversion, pitch shifting.

1. INTRODUCTION

Fractional delay filters are digital filters that are designed to delay discrete-time signals by a fractional amount of the sampling period. Such filters have wide applications in signal processing, including sampling rate conversion [Ramstad (1984); Smith and Gossett (1984)], nonuniform sampling [Johansson and Löwenborg (2002); Prendergast et al. (2004)], wavelet transform [Yu (2007)], digital modeling of musical instruments [Lehtonen and Laakso (2007); Välimäki et al. (2006)], to name a few. For more applications, see survey papers [Laakso et al. (1996); Välimäki and Laakso (2000)].

Conventionally, fractional delay filters are designed based on the Shannon sampling theorem [Shannon (1949); Unser (2000)] for strictly-bandlimited analog signals. By this theory, the optimal filter coefficients are obtained by sampling a delayed sinc function. This ideal filter is however not realizable because of its non-causality and instability, and hence many studies have focused their attention on approximating the ideal filter by, for example, windowed sinc function [Cain et al. (1995); Selva (2008)], maximally-flat FIR approximation [Hermanowicz (1992); Pei and Wang (2001); Samadi et al. (2004); Hachabiboglu et al. (2007); Shyu and Pei (2008)], all-pass approximation [Jing (1987)], weighted least-squares [Tarczynski et al. (1997); Shyu and Pei (2008)], and minmax (Chebyshev) optimization [Putnam and Smith (1997)].

Although these studies are based on the Shannon paradigm, no real analog signals are fully band-limited, and hence the assumption is not realistic. It is, therefore, necessary to design a filter that takes account of high-frequency components beyond the Nyquist frequency and the intersample behavior.

In our recent study [Yamamoto et al. (2012)], we have proved that *sampled-data H^∞ control theory* provides an optimal platform to overcome the frequency limitation enforced by the Shannon sampling theorem. Based on this study, we formulate the design of fractional delay filters as a sampled-data H^∞ optimization problem. That is, we design a filter that minimizes the H^∞ norm of the error system between the ideal fractional delay and an approximated one. In particular, the closed-form formula for the H^∞ optimal fractional delay filter is given under the assumption that the underlying frequency characteristic of the continuous-time input signal is governed by a low-pass filter of first order.

By using the closed-form formula, we propose a new pitch shifting method for digital sound synthesis [Roads (1996)]. Pitch shifting is a technique for raising or lowering the original pitch of audio signals. This is often used in synthesizing musical tone from a recorded signal of a musical instrument with a fixed fundamental frequency [Roads (1996)]. We show by simulation that the proposed method outperforms the conventional phase-vocoder method [Portoff (1976); Ellis (2002)].

⋆ This research is supported in part by the JSPS Grant-in-Aid for Scientific Research (B) No. 2136020318360203, Grant-in-Aid for Exploratory Research No. 22656095, and the MEXT Grant-in-Aid for Young Scientists (B) No. 22760317.

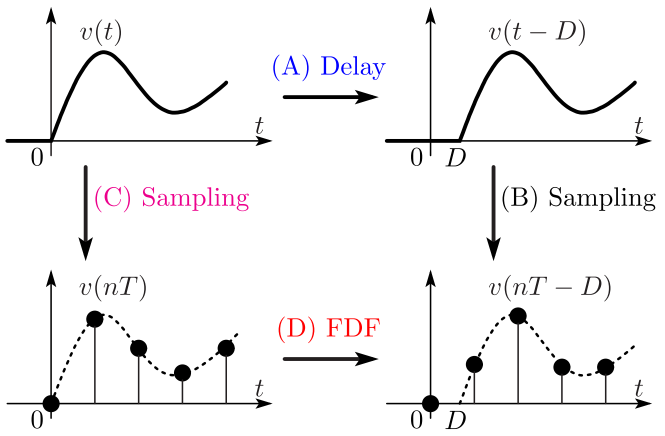


Fig. 1. Fractional delay process: (A) a continuous-time signal $v(t)$ (top left) is delayed by $D > 0$. (B) the delayed signal $v(t-D)$ is sampled at $t = nT$, $n = 0, 1, \ldots$. (C) the signal $v(t)$ is sampled at $t = nT$, $n = 0, 1, \ldots$. (D) digital filtering (fractional delay filter, FDF) to produce (or estimate) the sequence $\{v(nT - D)\}$ from the sampled-data $\{v(nT)\}$.

2. DESIGN PROBLEM OF FRACTIONAL DELAY FILTERS

In this section, we review fractional delay filters with conventional design methods based on the Shannon sampling theorem. Then, we formulate the design problem as a sampled-data H^∞ optimization problem.

2.1 Fractional delay filters

Consider a continuous-time signal $v(t)$ shown in Fig. 1 (top-left figure). Assume $v(t) = 0$ for $t < 0$ (i.e., it is a causal signal). Delaying this signal by $D > 0$ gives the delayed continuous-time signal $v(t - D)$ shown in Fig. 1 (top-right in Fig. 1). Then by sampling $v(t - D)$ with sampling period T, we obtain the discrete-time signal $\{v(nT-D)\}_{n\in\mathbb{Z}}$ as shown in Fig. 1 (bottom-right in Fig. 1).

Next, let us consider the directly sampled signal $\{v(nT)\}_{n\in\mathbb{Z}}$ of the original analog signal v as shown in Fig. 1 (bottom-left in Fig. 1). The objective of fractional delay filters is to reconstruct or estimate the delayed sampled signal $\{v(nT - D)\}_{n\in\mathbb{Z}}$ directly from the sampled data $\{v(nT)\}_{n\in\mathbb{Z}}$ when D is not an integer multiple of T. We now define the *ideal* fractional delay filter.

Definition 1. The *ideal fractional delay filter* K^{id} with delay $D > 0$ is the mapping which produces $\{v(nT - D)\}_{n\in\mathbb{Z}}$ from $\{v(nT)\}_{n\in\mathbb{Z}}$, that is,

$$K^{\text{id}} : \{v(nT)\}_{n\in\mathbb{Z}} \mapsto \{v(nT - D)\}_{n\in\mathbb{Z}}.$$

Assume for the moment that the original analog signal v is fully band-limited up to the Nyquist frequency, that is [1],

$$V(\mathrm{j}\omega) = 0, \quad |\omega| \geq \Omega_{\text{Nyquist}} := \frac{\pi}{T}, \tag{1}$$

where V is the Fourier transform of v. Then the impulse response of the ideal fractional delay filter is obtained by [Laakso et al. (1996)]:

[1] The symbol $:=$ means that the left-hand side is defined by the right-hand side.

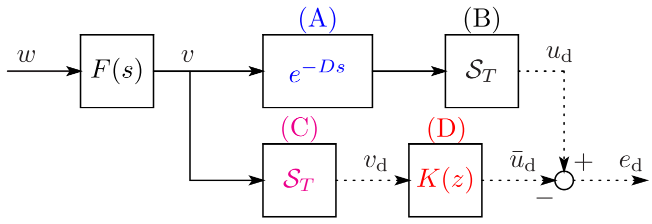


Fig. 2. Error system $\mathcal{E}$ for designing fractional delay filter $K(z)$. (A)–(D) correspond to those in Fig. 1.

$$\begin{gathered} k^{\text{id}}[n] = \frac{\sin \pi(n - D/T)}{\pi(n - D/T)} = \operatorname{sinc}(n - D/T), \\ n = 0, \pm 1, \pm 2, \ldots, \quad \operatorname{sinc}(t) := \frac{\sin(\pi t)}{\pi t}. \end{gathered} \tag{2}$$

The frequency response of this ideal filter is given in the frequency domain as

$$K^{\text{id}}(e^{\mathrm{j}\omega T}) = e^{-\mathrm{j}\omega D}, \quad \omega \leq \Omega_{\text{Nyquist}}. \tag{3}$$

Since the impulse response (2) does not vanish at $n = -1, -2, \ldots$ and is not absolutely summable, the ideal filter is noncausal and unstable, and hence the ideal filter is not physically realizable. Conventional designs thus aim at approximating the impulse response (2) or the frequency response (3) by a causal and stable filter via a window method, maximally-flat FIR approximations, weighted least-squares approximation, and so forth, as mentioned in Section 1.

These methods rely upon the band-limiting assumption (1). In practice, however, real analog signals always contain frequency components beyond the Nyquist frequency, and hence (1) never holds. To overcome this, we formulate the design problem of fractional delay filters without such an assumption by introducing the notion of sampled-data H^∞ optimization [Yamamoto et al. (2012)].

2.2 Design problem of fractional delay filters

Let us consider the error system shown in Fig. 2.

$F(s)$ is a stable and strictly proper transfer function which defines the frequency-domain characteristic of the original analog signal v. More precisely, we assume that the analog original signal v is in the following subspace of L^2:

$$FL^2 := \left\{v \in L^2 : v = Fw,\ w \in L^2\right\}.$$

The upper path of the diagram in Fig. 2 is the ideal process of the fractional delay filter (the process (A) $\to$ (B) in Fig. 1), that is, the continuous-time signal v is delayed by the continuous-time delay e^{-Ds}, and then sampled by the ideal sampler $\mathcal{S}_T$ with period $T > 0$ to become an ℓ^2 signal $u_{\rm d} := \mathcal{S}_T e^{-Ds} v$, or

$$u_{\rm d}[n] := \left(\mathcal{S}_T e^{-Ds} v\right)[n] = v(nT - D), \quad n \in \mathbb{Z}_+.$$

On the other hand, the lower path represents the real process ((C) $\to$ (D) in Fig. 1), that is, the continuous-time signal v is directly sampled with the same period T to produce a discrete-time signal $v_{\rm d} \in \ell^2$ defined by

$$v_{\rm d}[n] := (\mathcal{S}_T v)[n] = v(nT), \quad n \in \mathbb{Z}_+.$$

This signal is then filtered by $K(z)$ to be designed, and we obtain an estimation signal $\bar{u}_{\rm d} = K(z)\mathcal{S}_T v \in \ell^2$.

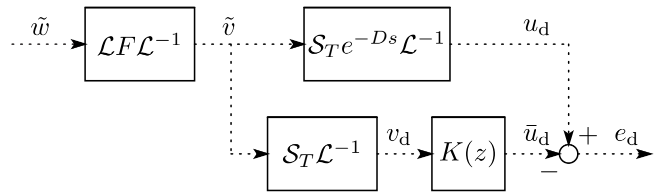


Fig. 3. Lifted error system $\tilde{\mathcal{E}}$

Put $e_{\mathrm{d}} := u_{\mathrm{d}} - \bar{u}_{\mathrm{d}}$ (the difference between the ideal output u_{d} and the estimation $\bar{u}_{\mathrm{d}}$), and let $\mathcal{E}$ denote the error system from $w \in L^2$ to $e_{\mathrm{d}} \in \ell^2$ (see Fig. 2). Symbolically, $\mathcal{E}$ is represented by

$$\mathcal{E} = \left(\mathcal{S}_T e^{-Ds} - K(z)\mathcal{S}_T\right) F(s).$$

Then our problem is to find a filter $K(z)$ that minimizes the H^∞ norm of the error system $\mathcal{E}$:

Problem 1. Given a stable, strictly proper $F(s)$, a delay time $D > 0$, and a sampling period $T > 0$, find the digital filter $K(z)$ that minimizes

$$\|\mathcal{E}\|_\infty = \left\|\left(\mathcal{S}_T e^{-Ds} - K(z)\mathcal{S}_T\right) F(s)\right\|_\infty = \sup_{\substack{w\in L^2\\ w\neq 0}} \frac{\|\mathcal{E}w\|_{\ell^2}}{\|w\|_{L^2}}. \tag{4}$$

3. DESIGN OF FRACTIONAL DELAY FILTERS

The error system $\mathcal{E}$ in Fig. 2 contains both continuous-time and discrete-time signals, and hence the system is not time-invariant; in fact, it is T-periodic [Chen and Francis (1995)]. In this section, we introduce the continuous-time lifting technique [Yamamoto (1994); Chen and Francis (1995)] to derive a norm-preserving system transformation from $\mathcal{E}$ to a time-invariant finite-dimensional discrete-time system. In this section, we fix the delay $D > 0$. For variable delay case, see Section 3.3.

3.1 Lifted model of sampled-data error system

Let $\{A, B, C\}$ be a minimal realization [Rugh (1996)] of $F(s)$:

$$\frac{dx_F(t)}{dt} = Ax_F(t) + Bw(t),\ v(t) = Cx_F(t),\ t \in \mathbb{R}_+. \tag{5}$$

We assume $A \in \mathbb{R}^{\nu\times\nu}$, $B \in \mathbb{R}^{\nu\times 1}$, $C \in \mathbb{R}^{1\times\nu}$, and $x_F(0) = 0 \in \mathbb{R}^\nu$ (ν is a positive integer). Let $D = mT + d$ where $m \in \mathbb{Z}_+$ and d is a real number such that $0 \le d < T$. First, we introduce the lifting operator $\mathcal{L}$ [Yamamoto (1994); Chen and Francis (1995)]:

$$\mathcal{L} : L^2[0,\infty) \ni f \mapsto \{\tilde{f}[k](\theta)\}_{k=0}^\infty \in \boldsymbol{\ell}^2 := \ell^2(\mathbb{Z}_+, L^2[0,T)),$$
$$\theta \in [0,T),\ \tilde{f}[k](\cdot) := f(kh + \cdot) \in L^2[0,T).$$

This operator transforms a continuous-time signal in $L^2[0,\infty)$ to an ℓ^2 sequence of functions in $L^2[0,T)$. We apply lifting to the continuous-time signals w and v and put $\tilde{w} := \mathcal{L}w$, $\tilde{v} := \mathcal{L}v$. By this, the error system in Fig. 2, is transformed into a time-invariant discrete-time system $\tilde{\mathcal{E}}$ shown in Fig. 3. Since the operator $\mathcal{L}$ is an isometry, we have

$$\|\mathcal{E}\|_\infty = \|\tilde{\mathcal{E}}\|_\infty := \sup_{\substack{\tilde{w}\in\boldsymbol{\ell}^2\\ \tilde{w}\neq 0}} \frac{\|\tilde{\mathcal{E}}\tilde{w}\|_{\ell^2}}{\|\tilde{w}\|_{\boldsymbol{\ell}^2}}. \tag{6}$$

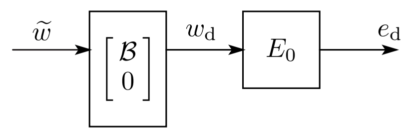


Fig. 4. Lifted sampled-data system $\tilde{\mathcal{E}}$

Then a state-space realization of the lifted error system $\tilde{\mathcal{E}}$ is given by

$$\begin{aligned} x[n+1] &= A_{\mathrm{d}}x[n] + \begin{bmatrix}\mathcal{B}\\0\end{bmatrix}\tilde{w}[n],\\ e_{\mathrm{d}}[n] &= C_e x[n] - \bar{u}_{\mathrm{d}}[n], \quad v_{\mathrm{d}}[n] = C_y x[n],\\ \bar{u}_{\mathrm{d}} &= Kv_{\mathrm{d}}, \end{aligned} \tag{7}$$

where the pertinent operators A_{d}, $\mathcal{B}$, C_e and C_y are given as follows: First, $\mathcal{B}$ is a linear (infinite-dimensional) operator defined by

$$\mathcal{B} : L^2[0,T] \to \mathbb{R}^{\nu+1},$$
$$\tilde{w} \mapsto \mathcal{B}\tilde{w} = \begin{bmatrix} \int_0^T e^{A(T-\tau)}B\tilde{w}(\tau)d\tau \\ \int_0^{T-d} Ce^{A(T-d-\tau)}B\tilde{w}(\tau)d\tau \end{bmatrix}$$

The other operators in (7) are all matrices defined by

$$A_{\mathrm{d}} := \begin{bmatrix} e^{AT} & 0 & 0 \\ Ce^{A(T-d)} & 0 & 0 \\ 0 & B_m & A_m \end{bmatrix} \in \mathbb{R}^{(\nu+m+1)\times(\nu+m+1)},$$
$$C_e := [0, 0, C_m] \in \mathbb{R}^{1\times(\nu+m+1)},$$
$$C_y := [C, 0, 0] \in \mathbb{R}^{1\times(\nu+m+1)},$$

where $\{A_m, B_m, C_m\}$ is a realization of the discrete-time delay z^{-m}.

3.2 Norm-equivalent finite dimensional system

The discrete-time state-space realization (7) involves an infinite dimensional operator $\mathcal{B} : L^2[0,T) \to \mathbb{R}^{\nu+1}$ where ν is the dimension of the state space of $F(s)$ defined by (5). Introducing the dual operator [Yamamoto (1993)] $\mathcal{B}^* : \mathbb{R}^{\nu+1} \to L^2[0,T)$ of $\mathcal{B}$, and composing this with $\mathcal{B}$, we can obtain a norm-equivalent finite dimensional system of the infinite dimensional system (7).

For the state space equations (7) of the lifted system $\tilde{\mathcal{E}}$, put

$$w_{\mathrm{d}} := \begin{bmatrix}\mathcal{B}\\0\end{bmatrix}\tilde{w}$$

and let E_0 denote the system from w_{d} to e_{d} (see Fig. 4). Note that the system E_0 is a finite-dimensional discrete-time system. Since $\mathcal{B}\mathcal{B}^*$ is a positive semi-definite matrix [Nagahara and Yamamoto (2003)], there exists a matrix B_{d} such that $\mathcal{B}\mathcal{B}^* = B_{\mathrm{d}}B_{\mathrm{d}}^\top$. Define a finite-dimensional discrete-time system by

$$E_{\mathrm{d}} := E_0 \begin{bmatrix} B_{\mathrm{d}} \\ 0 \end{bmatrix}.$$

See Fig. 5 for the block diagram of E_{d}. Then the discrete-time system E_{d} is equivalent to the sampled-data system $\mathcal{E}$ with respect to their H^∞ norm as described in the following theorem [Nagahara and Yamamoto (2003)]:

Theorem 1. Assume that the sampled-data system $\mathcal{E}$ is in $\mathbf{B}(L^2, \ell^2)$, the set of all bounded linear operators of L^2 into ℓ^2. Then, we have $E_{\mathrm{d}} \in \mathbf{B}(\ell^2, \ell^2)$ and $\|\mathcal{E}\|_\infty = \|E_{\mathrm{d}}\|_\infty$.

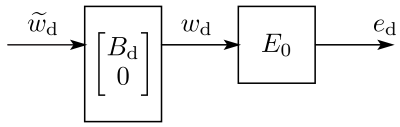


Fig. 5. Discrete-time system E_d

A state space realization of E_d is obtained by replacing $\mathcal{B}$ in (7) with B_d.

Thus the sampled-data H^∞ optimization (Problem 1) is equivalently reduced to discrete-time H^∞ optimization to find the optimal K that minimizes

$$\|E_d\|_\infty = \|G_1 - KG_2\|_\infty,$$

where

$$G_1(z) := C_e(zI - A_d)^{-1} \begin{bmatrix} B_d \\ 0 \end{bmatrix},$$
$$G_2(z) := C_y(zI - A_d)^{-1} \begin{bmatrix} B_d \\ 0 \end{bmatrix}.$$

We can easily find the optimal filter $K(z)$ by standard softwares such as MATLAB Robust Control Toolbox [Balas et al. (2005)]. Moreover, the design problem is reducible to an LMI (linear matrix inequality) by assuming that the filter $K(z)$ is a fixed-order FIR filter [Nagahara and Yamamoto (2005)].

3.3 Closed-form solution under first-order assumption

The H^∞ design method in the previous section requires the delay $D = mT + d$ to be fixed. In some applications, a filter with a variable delay (i.e., the delay can be changed without redesigning the filter). We here give design methods for the filter $K(z)$ having delay D (or m and d) as an adaptive parameter.

To design variable fractional delay filters, we first assume that the filter $F(s)$ is a first-order low-pass filter with cutoff frequency $\omega = \omega_c > 0$:

$$F(s) = \frac{\omega_c}{s + \omega_c}. \tag{8}$$

Under this assumption, we have the following theorem.

Theorem 2. Assume that $F(s)$ is given by (8). Then the optimal filter $K(z)$ is given by

$$K(z) = a_0(d)z^{-m} + a_1(d)z^{-m-1}, \tag{9}$$

where

$$a_0(d) := \frac{\sinh(\omega_c(T-d))}{\sinh(\omega_c T)}, \quad a_1(d) := e^{-\omega_c T}(e^{\omega_c d} - a_0). \tag{10}$$

The proof is found in [Nagahara and Yamamoto (2003)]. For higher order $F(s)$, one can obtain the H^∞ optimal filters for fixed delay parameters, say, $d_1 < d_2 < \cdots < d_M$, via numerical optimization with a linear matrix inequality (LMI) [Nagahara and Yamamoto (2005)]. Then one can obtain an approximate filter for arbitrarily $d \in [d_k, d_{k+1}]$ by linear combination of the kth and $k+1$th filters. See [Nagahara and Yamamoto (2012)] for more detail.

4. APPLICATION TO PITCH SHIFTING

In this section, we introduce sampling rate conversion by using the H^∞-optimal fractional filter given in the previous section, and propose a new pitch shifting method by the sampling rate converter.

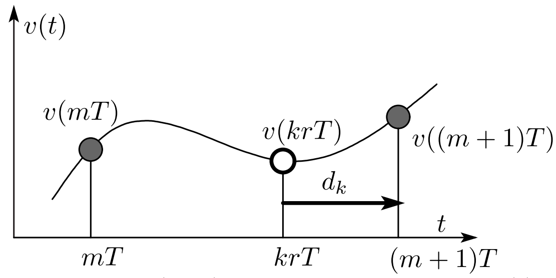


Fig. 6. The value $v(krT)$ is given by shifting $v(t)$ by $d_k = (m+1)T - krT$ and sampling at $t = (m+1)T$.

4.1 Sampling rate conversion by fractional delay filter

Let us consider a continuous-time signal $\{v(t)\}_{t\in\mathbb{R}_+}$. Assume that we are given sampled data $v[m] := v(mT)$, $m \in \mathbb{Z}_+$ where $T > 0$ is a sampling period. Then we execute sampling rate conversion on this discrete-time signal. By r, we denote the conversion rate. We assume r is a positive real number. Then sampling rate conversion aims at estimating the values of $\{v(krT)\}_{k\in\mathbb{Z}_+}$.

For this purpose, we adopt a sampling rate conversion by using fractional delay filters [Ramstad (1984)]. In this conversion, we use the sampled-data H^∞ optimal fractional delay filter given in Theorem 2. Let us consider estimation of the value $v(krT)$ where k is a positive integer. Assume that the time krT satisfies $mT < krT \leq (m+1)T$ where m is a non-negative integer. Let $d_k := (m+1)T - krT$. Then we have

$$\begin{aligned} v(krT) &= v\left((m+1)T - (m+1)T + krT\right) \\ &= v\left((m+1)T - d_k\right) \\ &= v(t - d_k)|_{t=(m+1)T}. \end{aligned}$$

That is, the value $v(krT)$ is obtained by delaying $v(t)$ by d_k and sampling at time $t = (m+1)T$ (see Fig. 6). Therefore, the estimation $v_r[k]$ for $v(krT)$ can be obtained by the fractional delay filter given in (9) as

$$v_r[k] = a_0(d_k)v\left((m+1)T\right) + a_1(d_k)v(mT),$$

where $a_0(\cdot)$ and $a_1(\cdot)$ are given in (10). Note that this filter is a two-tap FIR filter and the estimation needs much fewer computation than the conventional upsampler/filter/downsampler scheme [Vaidyanathan (1993)]. Also we emphasize that the computation load is the same for arbitrary real rate r while that of the conventional scheme depends on r. This is an advantage over conventional methods in the case of real-time processing. In addition, while conventional design of the digital filter in sampling rate conversion depends on the band-limiting assumption mentioned above, our design can take account of analog characteristic of input signals.

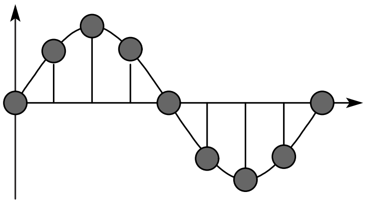
The algorithm of the proposed sampling rate conversion is shown in Algorithm 1. In this algorithm, we define $v[m] := v(mT)$, $m = 0, 1, 2, \ldots$.

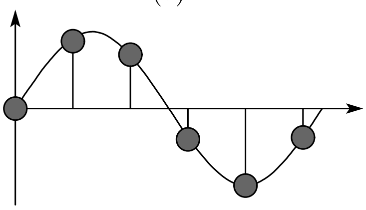
4.2 Pitch shifting

Pitch shifting is a technique for raising or lowering the original pitch of audio signals. To accomplish this, we can use a sampling rate converter; convert the sampling

Algorithm 1 Sampling rate conversion

$v_r[0] := v[0]$
$k := 1$
for $m = 0, 1, 2, \ldots$ **do**
  while $kr \leq m + 1$ **do**
    $d := (m + 1)T - krT$
    $v_r[k] := a_0(d)v[m + 1] + a_1(d)v[m]$
    $k := k + 1$
  end while
end for


(a)


(b)

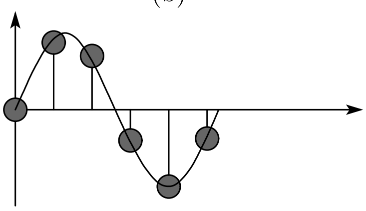

(c)

Fig. 7. Pitch shifting by sampling rate converter: (a) original digital signal, (b) sampling rate conversion, and (c) pitch sifting by processing the signal in (b) with the same sampling rate as (a).

period T to rT with $r > 0$, and play the converted signal with the original sampling period T. By this process, we have a raised (if $r > 1$) or lowered (if $r < 1$) pitch. For example [2], the pitch 110 Hz, that is A_2 note is converted to D_3 note ($110 \times 2^{5/12} \approx 146.83$ Hz) with $r = 2^{-5/12}$. This is analogous to listening to music by fast- or slow-forwarding a cassette tape. This process is effectively done by the closed-form solution in Theorem 2 if the sound is given by digital data. We illustrate this process in Fig. 7. For details on the algorithm for pitch shifting, see [Nagahara et al. (2011)].

We then show an example of pitch shifting. By using the proposed fractional delay filter, we shift the pitch of a guitar sound with A_2 (110 Hz) to D_3 note ($110 \times 2^{5/12} \approx 146.83$ Hz). For this, we take $r = 2^{-5/12}$. Fig. 8 shows the frequency responses of the original signal with fundamental frequency 110 Hz (dots) and the pitch-shifted signal. The frequency response of the shifted signal shows that the processed signal has a valid fundamental frequency $110 \times 2^{5/12} \approx 146.83$ Hz and the harmonics with $110 \times 2^{5/12} \times n$, $n = 2, 3, 4$.

[2] In this article, we consider *equal temperament* [French (2009)].

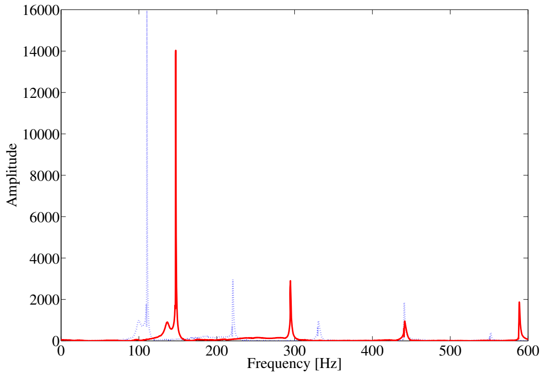
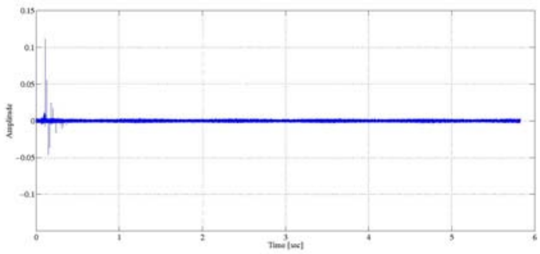
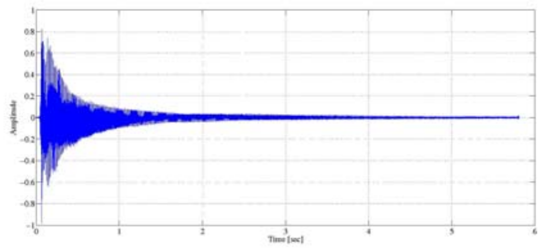


Fig. 8. Frequency response: original signal with fundamental frequency 110 Hz (dots) and pitch-shifted signal with fundamental frequency $110 \times 2^{5/12} \approx 146.83$ Hz (solid).



(a) Proposed method



(b) Phase vocoder

Fig. 9. Reconstruction error

We show another example of pitch shifting. We shift the note A_2 (110Hz) used above to A_3 (220 Hz). Then the shifted signal is again shifted down to the original note A_2. We measure the error between the original sound and the processed one. Figure 9(a) shows the reconstruction error by the proposed method. For comparison, we processed by the phase vocoder [Portoff (1976); Ellis (2002)], which is widely used in pitch shifting. Figure 9(b) shows the error of the phase vocoder. We can see that the phase vocoder produces much larger errors than the proposed method. This shows effectiveness of our method.

One can listen to the processed sounds at the following web page:

http://www-ics.acs.i.kyoto-u.ac.jp/~nagahara/ps/

5. CONCLUSION

We have presented a new method of designing fractional delay filters via sampled-data H^∞ optimization. An advantage here is that an *optimal analog performance* can be attained. The optimal design problem can be equivalently transformed to discrete-time H^∞ optimization, which is easily executed by standard numerical optimization toolboxes. We have also given the H^∞ optimal filter having delay time variable D as a variable parameter. In particular, a closed-form solution is given when the frequency distribution of the input analog signal is modeled as a first-order low-pass filter. Using this closed-form solution, we have introduced a sampling rate conversion with arbitrary conversion rate, and proposed a new pitch shifting method for digital sound synthesis. Examples have been given to illustrate the advantage of the proposed method.

REFERENCES

Balas, G., Chiang, R., Packard, A., and Safonov, M. (2005). *Robust Control Toolbox, Version 3*. The Math Works.

Cain, G.D., Yardim, A., and Henry, P. (1995). Offset windowing for FIR fractional-sample delay. In *IEEE ICASSP'95*, 1276–1279.

Chen, T. and Francis, B.A. (1995). *Optimal Sampled-data Control Systems*. Springer.

Ellis, D.P.W. (2002). A phase vocoder in Matlab. URL `http://www.ee.columbia.edu/ dpwe/resources/matlab/pvoc/`.

French, R.M. (2009). *Engineering the Guitar*. Springer.

Hachabiboglu, H., Gunel, B., and Kondoz, A. (2007). Analysis of root displacement interpolation method for tunable allpass fractional-delay filters. *IEEE Trans. Signal Processing*, 55(10), 4896 –4906.

Hermanowicz, E. (1992). Explicit formulas for weighting coefficients of maximally flat tunable FIR delayers. *Electronics Letters*, 28, 1936–1937.

Jing, Z. (1987). A new method for digital all-pass filter design. *IEEE Trans. Acoust., Speech, Signal Processing*, 35, 1557–1564.

Johansson, H. and Löwenborg, P. (2002). Reconstruction of nonuniformly sampled bandlimited signals by means of digital fractional delay filters. *IEEE Trans. Signal Processing*, 50(11), 2757 – 2767.

Laakso, T.I., Välimäki, V., Karjalainen, M., and Laine, U.K. (1996). Splitting the unit delay. *IEEE Signal Processing Mag.*, 13, 30–60.

Lehtonen, V.V.H.M. and Laakso, T.I. (2007). Musical signal analysis using fractional-delay inverse comb filters. In *Proc. of the 10th Int. Conf. on Digital Audio Effects*, 261–268.

Nagahara, M., Hanibuchi, K., and Yamamoto, Y. (2011). Pitch shifting by H^∞-optimal variable fractional delay filters. *IFAC 18th World Congress*, 4386–4391.

Nagahara, M. and Yamamoto, Y. (2003). Optimal design of fractional delay filters. *Proc. of 35th Conf. on Decision and Control*, 6539–6544.

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

Nagahara, M. and Yamamoto, Y. (2012). H^∞-optimal fractional delay filters. submitted to *IEEE Trans. Signal Processing*.

Pei, S.C. and Wang, P.H. (2001). Closed-form design of maximally flat FIR Hilbert transformers, differentiators, and fractional delayers by power series expansion. *IEEE Trans. Circuits Syst. I*, 48, 389–398.

Portoff, M.R. (1976). Implementation of the digital phase vocoder using the fast Fourier transform. *IEEE Trans. Acoust., Speech, Signal Processing*, 24(3), 243–248.

Prendergast, R., Levy, B., and Hurst, P. (2004). Reconstruction of band-limited periodic nonuniformly sampled signals through multirate filter banks. *IEEE Trans. Circuits Syst. I*, 51(8), 1612–1622.

Putnam, W. and Smith, J. (1997). Design of fractional delay filters using convex optimization. In *Applications of Signal Processing to Audio and Acoustics, 1997 IEEE ASSP Workshop on*.

Ramstad, T.A. (1984). Digital methods for conversion between arbitrary sampling frequencies. *IEEE Trans. Acoust., Speech, Signal Processing*, 32(3), 577–591.

Roads, C. (1996). *The Computer Music Tutorial*. The MIT Press.

Rugh, W.J. (1996). *Linear Systems Theory*. Prentice Hall.

Samadi, S., Ahmad, O., and Swamy, M.N.S. (2004). Results on maximally flat fractional-delay systems. *IEEE Trans. Circuits Syst. I*, 51(11), 2271–2286.

Selva, J. (2008). An efficient structure for the design of variable fractional delay filters based on the windowing method. *IEEE Trans. Signal Processing*, 56(8), 3770 –3775.

Shannon, C.E. (1949). Communication in the presence of noise. *Proc. IRE*, 37(1), 10–21.

Shyu, J.J. and Pei, S.C. (2008). A generalized approach to the design of variable fractional-delay FIR digital filters. *Signal Processing*, 88(6), 1428–1435.

Smith, J.O. and Gossett, P. (1984). A flexible sampling-rate conversion method. In *IEEE ICASSP'84*, 19.4.1–19.4.4.

Tarczynski, A., Cain, G.D., Hermanowicz, E., and Rojewski, M. (1997). WLS design of variable frequency response FIR filters. In *Proc. IEEE Int. Symp. Circuits Syst.*, 2244–2247.

Unser, M. (2000). Sampling — 50 years after Shannon. *Proc. IEEE*, 88(4).

Vaidyanathan, P.P. (1993). *Multirate Systems and Filter Banks*. Prentice Hall.

Välimäki, V. and Laakso, T.I. (2000). Principles of fractional delay filters. In *IEEE ICASSP'00*, 3870–3873.

Välimäki, V., Pakarinen, J., Erkut, C., and Karjalainen, M. (2006). Discrete-time modelling of musical instruments. *Rep. Prog. Phys*, 69(1), 1–78.

Yamamoto, Y. (1993). On the state space and frequency domain characterization of H^∞-norm of sampled-data systems. *Syst. Control Lett.*, 21, 163–172.

Yamamoto, Y. (1994). A function space approach to sampled-data control systems and tracking problems. *IEEE Trans. Automat. Contr.*, 39, 703–712.

Yamamoto, Y., Nagahara, M., and Khargonekar, P.P. (2012). Signal reconstruction via H^∞ sampled-data control theory — Beyond the Shannon paradigm. *IEEE Trans. Signal Processing*, 60(2), 613–625.

Yu, R. (2007). Characterization and sampled-data design of dual-tree filter banks for Hilbert transform pairs of wavelet bases. *IEEE Trans. Signal Processing*, 55(6), 2458–2471.