PAN ARCHIVES OF ACOUSTICS
Vol. **39**, No. 2, pp. 231–242 (2014)

DOI: 10.2478/aoa-2014-0027

Variable Ratio Sample Rate Conversion Based on Fractional Delay Filter

Marek BLOK[(1)], Piotr DRÓZDA[(2)]

[(1)] *Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology*
G. Narutowicza 11/12, 80-233 Gdańsk, Poland; e-mail: mblok@eti.pg.gda.pl

[(2)] *ADVA Optical Networking Sp. z o.o.*
Śląska 35/37, 81-310 Gdynia, Poland

(*received February 3, 2013; accepted May 29, 2014*)

In this paper a sample rate conversion algorithm which allows for continuously changing resampling ratio has been presented. The proposed implementation is based on a variable fractional delay filter which is implemented by means of a Farrow structure. Coefficients of this structure are computed on the basis of fractional delay filters which are designed using the offset window method. The proposed approach allows us to freely change the instantaneous resampling ratio during processing. Using such an algorithm we can simulate recording of audio on magnetic tape with nonuniform velocity as well as remove such distortions. We have demonstrated capabilities of the proposed approach based on the example of speech signal processing with a resampling ratio which was computed on the basis of estimated fundamental frequency of voiced speech segments.

Keywords: sample rate conversion, offset window method, variable fractional delay filter, variable resampling ratio.

1. Introduction

A huge number of sample rate standards (AES5-2008, 2008) available today create a demand for the development of sample rate conversion (SRC) algorithms (Tarczynski *et al.*, 1994; Hermanowicz *et al.*, 2000; Evangelista, 2003). The digital resampling algorithm replaces digital-analog conversion followed by an analog signal sampling which allows for more flexible implementations. A common example used to demonstrate the usefulness of such algorithms is the conversion between compact disc (CD) with $F_s = 44.1$ kHz (multimedia standard) and digital tape (DAT) with $F_s = 48$ kHz (communications standard) (Rajamani *et al.*, 2000; Hermanowicz *et al.*, 2000). In this paper we present SRC implementation based on a variable fractional delay (VFD) filter. Using this algorithm we not only can implement an arbitrary constant resampling ratio but fluctuating changes in resampling ratio can be readily implemented as well. Based on this tool, we propose a novel application for a resampling algorithm which is a correction of signals with unintentional nonuniform sampling. For example, a correction of old recordings with distortions which resulted from nonuniform velocity of the media ("wow" distortion) (Ciarkowski *et al.*, 2005; Czyzewski, 2007; 2010), a problem which is typically addressed with interpolation techniques (Maziewski, 2006). On the other hand, with a correctly sampled signal, we might simulate such distortions using the same VFD SRC algorithm. In both cases we make use of the fact that a nonuniformly sampled signal reconstructed with uniform sampling changes its pitch inversely proportionally to the sample rate changes.

2. Fractional delay filter

The SRC algorithm investigated in this paper is based on fractional delay (FD) filters and its performance depends on the design method used to calculate coefficients of these filters. The problem is that the impulse response of the ideal FD filter with total delay τ_d

$$h_{\mathrm{id}}[n] = \mathrm{sinc}(n - \tau_d) \tag{1}$$

is infinite and non-causal. Therefore, in practice instead of using the ideal filter, a filter approximating

its frequency response defined by the following formula (Laakso *et al.*, 1996)

$$H_{\mathrm{id}}(f) = \exp(-j2\pi f\tau_d), \qquad f \in [-0.5, 0.5) \tag{2}$$

has to be used.

Because of the causality requirement high performance FD filters are characterized with nonzero integer delay $D = \mathrm{round}(\tau_d)$, which for FIR filters is usually selected close to the bulk delay $\tau_N = (N-1)/2$, where N is length of impulse response of the filter. With those two delays defined, we receive the following formula for the total delay

$$\tau_d = D + d = \tau_N + \varepsilon, \tag{3}$$

where $d \in [-0.5, 0.5)$ is the fractional delay and ε is the net delay. The value of the fractional delay d relates to the difficulty of FD filter design, the closer it is to zero the lower approximation error can be achieved. Nevertheless, in application which require variable delay, it is more convenient to use the net delay ε because in contrast to the fractional delay d it is a continuous function of the total delay τ_{d}. Since in most practical applications $\varepsilon \in [-0.5, 0.5)$, discontinuity of d is not a problem for filters of odd length N, for which the integer delay D is usually selected equal to τ_N. In such case, both fractional d and net delays ε are the same. For even N, however, fractional delay d is discontinuous at $\tau_{\mathrm{d}} = \tau_N$.

In this paper we use the FIR FD filters which approximate the ideal frequency response $H_{\mathrm{id}}(f)$ with the frequency response

$$H_N(f) = \sum_{n=0}^{N-1} h[n] \exp(-j2\pi fn), \tag{4}$$

where $h[n]$ is the impulse response. Usually the designer computes the impulse response $h[n]$ using one of several design methods offering optimal FD filters with maximally flat (MF), least squares (LS) and minimax being the most popular. The optimality criteria for all of these methods are based on complex approximation error (Laakso *et al.*, 1996)

$$E(f) = H_N(f) - H_{\mathrm{id}}(f). \tag{5}$$

For MF filters approximation error and its $N-1$ derivatives must be equal to zero at $f = 0$. In the result, the MF filter offers excellent performance, but only around zero frequency. In contrast, the LS and minimax filters allow the designer to specify the approximation band $f \in [0, f_{\mathrm{a}}]$ in which the error is minimized. The LS FD filter (Laakso *et al.*, 1996) has minimized the energy of error

$$E_{\mathrm{LS}}(f_{\mathrm{a}}) = 2\int_0^{f_{\mathrm{a}}} |E(f)|^2 \,\mathrm{d}f \tag{6}$$

while the minimax FD filter (Laakso *et al.*, 1996; Blok, 2005) has minimized the peak error (PE)

$$E_{\mathrm{PE}}(f_{\mathrm{a}}) = \max_{f\in[0,f_{\mathrm{a}}]} |E(f)| \tag{7}$$

in the approximation band.

The coefficients of the impulse response, vector $\mathbf{h} = [h[0], h[1], \ldots, h[N-1]]^{\mathrm{T}}$, of all these optimal filters, MF, LS and minimax, with fractional delay d can be found solving the following matrix equation (Laakso *et al.*, 1996)

$$\mathbf{Ph} = \mathbf{p}, \tag{8}$$

where the coefficients of matrix $\mathbf{P}$ and vector $\mathbf{p}$ depend on the optimization criterion. For MF filter matrix $\mathbf{P}$ is a Vandermonde matrix

$$\mathbf{P}_{k+1,n+1} = n^k \tag{9}$$

and vector $\mathbf{p}$ has elements

$$\mathbf{p}_{1,k+1} = {\tau_d}^k, \tag{10}$$

where $k, n = 0, 1, ..., N-1$. To find the LS filter we need only to change coefficients of matrix $\mathbf{P}$

$$\mathbf{P}_{k+1,n+1} = f_{\mathrm{a}}\mathrm{sinc}f_{\mathrm{a}}(n-k) \tag{11}$$

and column vector $\mathbf{p}$

$$\mathbf{p}_{1,k+1} = f_{\mathrm{a}}\mathrm{sinc}f_{\mathrm{a}}(k-\tau_d). \tag{12}$$

A minor modification is needed for minimax filters. First, a set of $N+1$ frequency points f_k, called extremal points, must be found using recursive complex Remez algorithm (Blok, 2002b). Then, coefficients of matrix $\mathbf{P}$ and vector $\mathbf{p}$ can be computed using the following formulas

$$\begin{aligned} \mathbf{P}_{k+1,n+1} &= \cos(2\pi f_k n) - \sin(2\pi f_k n), \\ \mathbf{P}_{k+1,N+1} &= (-1)^k \end{aligned} \tag{13}$$

and

$$\mathbf{p}_{1,k+1} = \cos(2\pi f_k \tau_d) - \sin(2\pi f_k \tau_d), \tag{14}$$

where $k = 0, 1, ..., N$ and $n = 0, 1, ..., N-1$. In this case vector $\mathbf{h}$ has one additional element with magnitude equal to peak approximation error (7).

The aforementioned design methods guarantee that each FD filter offers the best performance in respect to the selected criterion. Nonetheless, as we will demonstrate in this paper, in the case of the SRC algorithm better results can be achieved with the use of the offset window method (Yardim *et al.*, 1996; 1997; Blok, 2012b; 2013)

$$h_d[n] = h_{\mathrm{id}}[n]w(n-\varepsilon), \tag{15}$$

where a symmetric prototype window is offset by the net delay ε (3) of the designed filter. One of the advantages of this method is that the FD filter bandwidth can be readily adjusted (Blok, 2012a; 2012b). To achieve this, in formula (15) the ideal impulse response of a fullband FD filter (1) has to be replaced with its lowpass version

$$h_{\mathrm{id}}[n] = 2f_{\mathrm{c}}\,\mathrm{sinc}(2f_{\mathrm{c}}(n-\tau_d)), \tag{16}$$

where f_{c} is the assumed cutoff frequency of the filter.

Nevertheless, in order to use the offset window method efficiently two problems must be solved. First is a prototype window selection. This can be solved with window extraction method (Hermanowicz, 1998; Blok, 2012b) in which a prototype window is extracted from the impulse response of a single optimal (minimax or LS) FD filter $h_{\mathrm{opt}}[n]$

$$w_{\mathrm{ext}} = h_{\mathrm{opt}}[n]/h_{\mathrm{id}}[n]. \tag{17}$$

The second problem is window offsetting since it requires computation of new samples of the prototype window delayed by a fraction of the sampling period. Such prototype window delaying can be efficiently done using the short FD filters (Blok, 2012b), but the drawback is that the window offsetting increases computational complexity of the FD filter impulse response update. Fortunately, it is enough to compute several impulse responses beforehand and on such basis implement a VFD filter using the Farrow structure presented in Sec. 5.

3. Sample rate conversion using VFD filter

The classic three rate sample rate conversion algorithm is presented in Fig. 1. Input signal with sample rate F_{s1} is upsampled by means of insertion of $L-1$ zeros between each pair of consecutive input samples. Next, at the intermediate sample rate, the lowpass interpolation filter with upper frequency $f_u = \min(0.5/L, 0.5/M)$ prevents the aliasing and removes spectral images resulting from upsampling. At last stage, the sample rate is reduced to the desired one by means of decimation, only every M-th sample remains in the output signal.

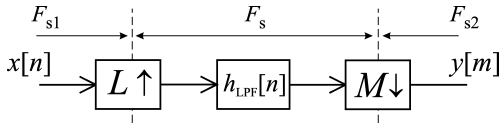

Fig. 1. Classic three-rate sample rate conversion algorithm by rational factor L/M.

The classic approach, because of its simplicity, is well suited for simple cases requiring a constant sample ratio with small factors L and M. In other cases the intermediate sampling ratio is very high, which results in an extremely narrow passband of the interpolation filter. In the result, computational costs increase drastically and, what is more important, the interpolation filter becomes very difficult to design. Therefore for arbitrary resampling ratios the VFD filter is used (Fig. 3) (Tarczynski *et al.*, 1994; Hermanowicz *et al.*, 2000; Evangelista, 2003).

As we can see in Fig. 2, each output sample $y[m]$ is no farther from the nearest input sample than the fraction of the input sampling period. This distance is the fractional delay $d[m]$ and can be computed using the following recursive formula (Blok, 2002a)

$$d[m] = d[m-1] - r[m] + \Delta n[m], \tag{18}$$

where $r[m]$ is the inverse of the instantaneous resampling ratio defined as the ratio of assumed input and output sampling rates sampled at output instants

$$r[m] = F_{s1}[m]/F_{s2}[m] = T_{s2}[m]/T_{s1}[m] \tag{19}$$

and

$$\Delta n[m] = \mathrm{round}(r[m] - d[m-1]) \tag{20}$$

is a number of new samples expected in the input buffer before the next output sample can be computed. For the rational resampling ratio (Fig. 1)

$$r[m] = M/L \tag{21}$$

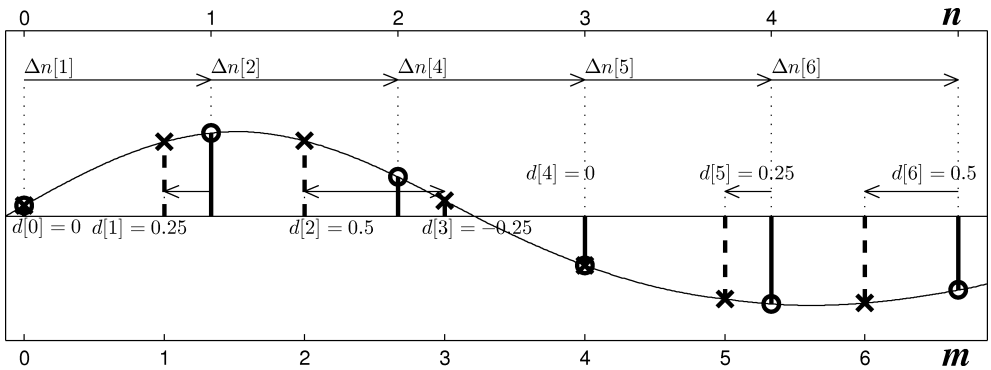

Fig. 2. Illustration of sample rate conversion by $L/M = 5/4$. Input samples $x[n]$ – ∘, output samples $y[m]$ – ×.

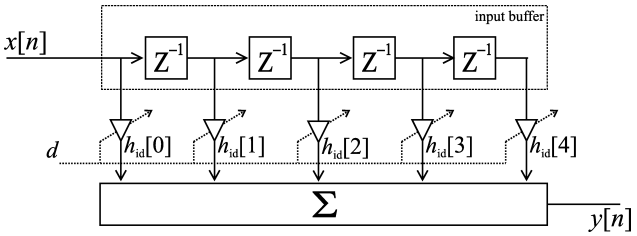

Fig. 3. General VFD filter structure applied to SRC.

a limited number of fractional delays has to be implemented, since sequences $d[m]$ and $\Delta n[m]$ are periodic with period L. In the other cases we need to be able to implement any fractional delay, which can be achieved with the VFD filter. In such case the resampling ratio can be an arbitrary positive number and, moreover, can change in time. Nevertheless, if we want to avoid nonlinear distortions of the resampled signal, the ratio $r[m]$ must be limited by the instantaneous oversampling ratio of the input signal. Alternatively, the band of the input signal has to be limited to assure that instantaneous oversampling ratio is large enough, for example, by adjusting the cutoff frequency of the overall interpolation filter.

With two parameters defined in (18) and (20) the resampling algorithm is following (Fig. 4):

1. Fill the buffer with zeros and start with $d[0] = 0$ and $\Delta n[0] = 0$,
2. wait for $\Delta n[m]$ new samples in the input buffer,
3. compute a value of the output sample $y[m]$ using FD filter with the fractional delay $d[m]$,
4. calculate $\Delta n[m]$ and $d[m]$ for the next m and go back to step 2.

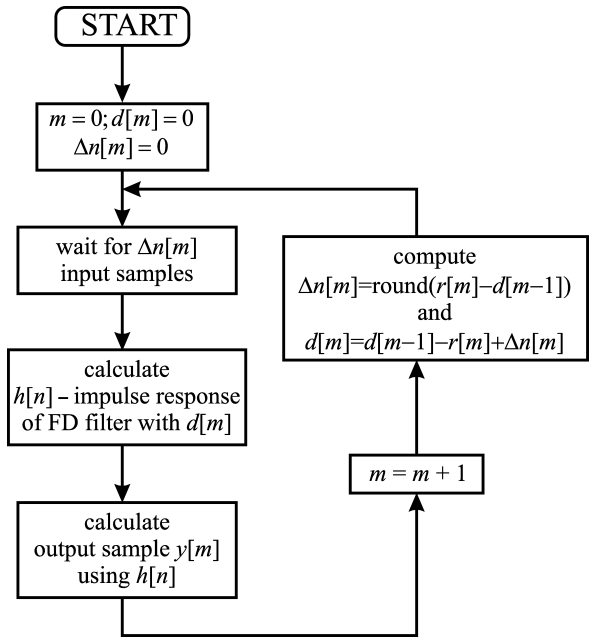

Fig. 4. Diagram of SRC algorithm based on VFD filter.

For every output sample the resampling algorithm requires different fractional delay (18). This means that for each output sample we need to compute a new impulse response of the FD filter. Since for rational resampling ratios we only need L impulse responses, we can store them in a look-up-table (LUT) (Hermanowicz *et al.*, 2000). Conversely, when the ratio is arbitrary, and additionally changing in time, the filters needed in the resampling cannot be specified beforehand. They have to be computed during runtime which can be done readily using the Farrow structure (Farrow, 1988; Harris, 1997; Hermanowicz, 2004; Blok, 2005).

4. FD filter design for SRC

The selection of FD filters for the SRC algorithm implementation might seem simple, since we can use optimal FD filters such as minimax or LS FD filters (Laakso *et al.*, 1996; Blok, 2002a; 2002b). The problem is that it is not sufficient to minimize the approximation errors of FD filters within approximation band specified by the designer to guarantee the best overall performance of such SRC algorithm.

Let us notice that for the rational resampling ratio M/L the SRC algorithm based on FD filters (Fig. 3 and 4) is equivalent to the classic approach (Fig. 1) (Blok, 2002a) if the interpolation filter is replaced with the overall filter $h_o[n]$ composed of all FD filters $h_{d[m]}[n]$ used in the resampling (Blok, 2002a). Since for the rational resampling rate equal to M/L only L different FD filters are used regardless of M we have

$$h_o[m + nL] = h_{d[m]}[n], \quad m = 0, 1, \ldots, L-1, \quad (22)$$

where fractional delays $d[m]$ are arranged in the decreasing order

$$d[m-1] = d[m] + 1/L, \quad m = 1, \ldots, L-1. \quad (23)$$

Using the overall filter (22) we can readily analyze distortions introduced by the SRC algorithm based on FD filters since this filter must fulfill the same requirements as the interpolation filter in the classic approach (Fig. 1). Such analysis based on the overall filter concept can be applied directly only for rational resampling ratios, but the conclusions can be readily adapted to any arbitrarily selected resampling ratio.

In Fig. 5 we can see magnitude responses of the overall filters obtained for FD filters of length $N = 17$ designed using two different approaches: the optimal filter design minimizing peak error (7) (minimax FD filters) with $f_a = 0.4$ and the offset window method with window extracted from minimax filter with $N = 17$, $d = -0.5$ and $f_a = 0.4$. The window in the offset window method is offset using MF filter of length $N_{off} = 5$. Additionally in case of the offset window method, the lowpass impulse response (16) with

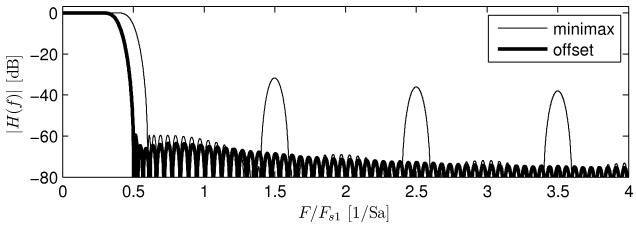

Fig. 5. Comparison of magnitude responses of overall filters composed of minimax FD filters ($L = 9$) and FD filters designed using offset window method.

$f_c = 0.4$ is used instead of the impulse response of the fullband ideal FD filter (1).

As we can notice, the problem with the SRC implementation based on minimax filters, is that the overall filter demonstrates large lobes in stopband which can also be observed if LS and MF filters are used (Blok, 2012a). These large lobes, which may result in aliasing when input signal has components above f_a, are related to properties of the overall window. This overall window is obtained from the interpolation filter using the formula for window extraction (17) based on the assumption that the interpolation filters is a lowpass filter designed using the window method. When minimax, LS or MF FD filters are used to compose the overall filter, the overall window exhibits periodic discontinuities with period equal to input sampling period (Blok, 2012b). These discontinuities result in large side lobes in frequency response of the overall window, which are directly related to the large lobes in the stopband of the overall filter. Moreover, the location of the transition band of the overall filter cannot be adjusted (Blok, 2012a).

To overcome disadvantages of minimax and LS FD filters the offset window method (Yardim *et al.*, 1996; 1997; Blok, 2012b) can be used. Since in this approach each window used in the FD filter design is a delayed version of a single prototype window, the resulting overall window is smooth. In the result, when filters designed using the offset window method are used in the SRC algorithm, the large lobes in the stopband of the overall filter are eliminated (Fig. 5). Additionally with the window method we can readily adjust the location of the transition band of the overall filter using the lowpass impulse response (16) as the prototype FD filter.

For example, selection of $f_c = 0.4$ for the offset method in the example presented in Fig. 6 results in the overall filter which, in comparison to the overall filter obtained from minimax FD filters, has narrower passband.

Although overall filters presented in Fig. 5 are composed of just $L = 9$ filters, properties of their magnitude responses, discussed in this section, are maintained for any resampling ratio. In Fig. 6 we can see the results of testing the SRC algorithm with ratio 160/147

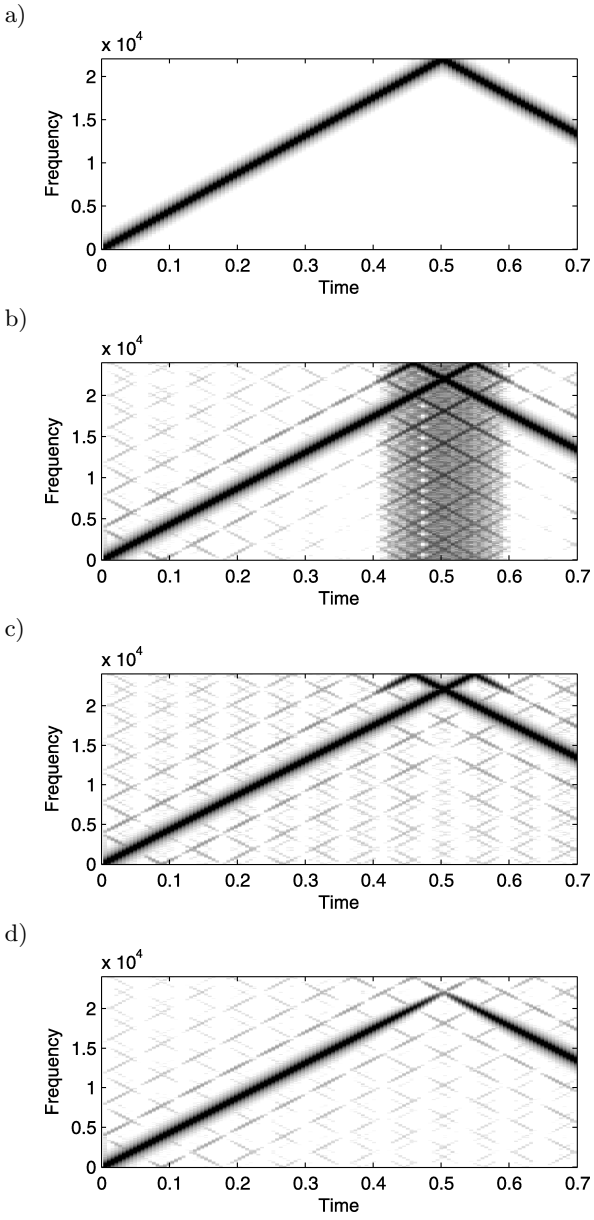

Fig. 6. Spectrograms illustrating resampling with constant resampling ratio 160/147 using VFD filters from Fig. 5 implemented with the Farrow structure of order $q = 5$: a) input chirp signal, b) signal resampled using minimax FD filters ($f_a = 0.4$), c) signal resampled using fullband ($f_c = 0.5$) FD filters designed using offset window, d) signal resampled using FD filters designed using offset window for $f_c = f_a$.

using the constant amplitude LFM chirp, which is an excellent signal for performing such tests. As we can see, the signal converted using minimax FD filters demonstrates high distortions when input signal frequency exceeds f_a. On the other hand, the signal converted using FD filters designed with offset window has no components caused by nonlinear distortions larger than −60 dB. This is because the offset window method allows us to adjust the cutoff frequency f_c.

In all cases presented in Fig. 6, the attenuation of resampling distortions, which is equal to stopband attenuation of the overall filter, is directly related to the peak error (7) of the worst FD filter used in the SRC algorithm (Blok, 2002a) and can be adjusted with a proper selection of FD filter specification, its length and width of approximation band.

5. VFD filter implementation

Since the VFD filter needs to be able to change its delay for each output sample, high computational costs related to solving matrix Eq. (8), or costs of offsetting a window in the offset window method (15) become a significant problem. The most popular solution of this problem is the Farrow structure (Farrow, 1988; Harris, 1997; Hermanowicz, 2004; Blok, 2005). This structure is based on a concept that each sample of the impulse response can be approximated with a separate polynomial of order q dependent on fractional delay d

$$h[n] = \sum_{m=0}^{q} c_m[n] d^m. \tag{24}$$

Now, the output samples of the FD filter can be expressed with the following formula

$$y[n] = \sum_{k=0}^{N-1} h[n]x[n-k] = \sum_{m=0}^{q} y_m[n] d^m, \tag{25}$$

where

$$y_m[n] = \sum_{k=0}^{N-1} c_m[k]x[n-k]. \tag{26}$$

Formulas (25) and (26) define the Farrow structure presented in Fig. 7 where each row of coefficients implements separate filter with impulse response $c_m[n]$.

To find coefficients $c_m[n]$ we need to compute only a few impulse responses of FD filters with fractional delays d uniformly spread in range $[-0.5, 0.5]$. In practice, to compute approximation polynomials it is enough to use only $q+1$ impulse responses. The polynomials of order q equal to 6 or 7 offer performance adequate for high quality FD filters with approximation error about -100 dB (Blok, 2005). The resulting structure is equally efficient for both the optimal (minimax or LS) FD filters and FD filters designed using the offset window method (Blok, 2012a).

The drawback of the implementation based on the Farrow structure is that, although its coefficients can be computed beforehand for selected f_c in case of the offset window method, each change of f_c requires a new set of structure coefficients. Thus for this structure only a preselected f_c value can be used which has to be a compromise between a retained band of the processed signal and a level of aliasing distortions.

6. Resampling with variable ratio

In this section we will demonstrate SRC with changing ratio. The first simple, yet spectacular example is the conversion of sinusoidal signal with constant frequency F_{in} into a chirp signal with linear instantaneous frequency

$$F_{\text{out}}[m] = F_0 + m\Delta F, \tag{27}$$

where F_0 is the initial output chirp frequency and ΔF is the chirp rate. In this example the input sample rate is constant and the output sample rate must change in such way that the oversampling ratio changes linearly according to the output time index m.

$$F_{s2}[m] = F_{\text{in}}\overline{F_{s2}}/F_{\text{out}}[m], \tag{28}$$

where $\overline{F_{s2}}$ is the constant output sample ratio for which the resampled signal can be observed as a chirp with the assumed instantaneous frequency (27). Since in this example the input sampling frequency F_{s1} is con-

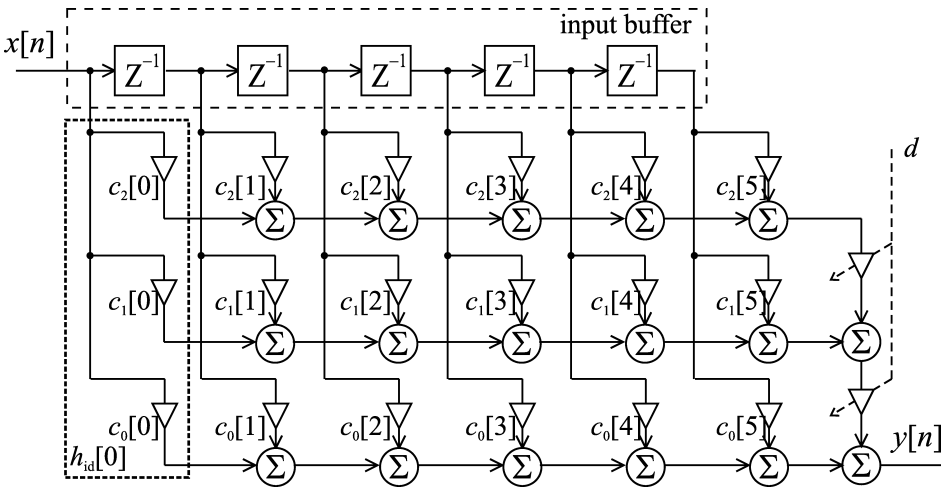

Fig. 7. Farrow structure of order $q = 2$ implementing VFD filter of length $N = 6$.

stant, the formula for the inverse of the instantaneous resampling ratio can be readily obtained

$$r_{\text{chirp}}[m] = F_{s1}/F_{s2}[m] = \frac{F_{\text{in}}(F_0 + m\Delta F)}{F_{s1}\overline{F_{s2}}}. \qquad (29)$$

The effects of resampling are presented in Fig. 8. Let us notice that when the output signal is reconstructed using the variable output sample rate (28) the input (Fig. 8a) and output (Fig. 8c) signals represent the same analog signal. However, if we assume the constant output rate $\overline{F_{s2}}$, the output samples of SRC algorithm represent the chirp signal (Fig. 8b) which we wanted to obtain.

The process described above can be reversed using the same algorithm. The chirp signal obtained in the previous step can be converted back into sinusoid, but the selection of the ratio $r[m]$ is now more difficult. In the first scenario the sample rate and the frequency of the input signal are constant, which simplifies the derivations. This time we need to assume that either the frequency of the input signal or its sample rate is changing. Since we want to demonstrate how to reverse the resampling process we will assume that the signal frequency is constant with variable distance between input samples.

From (28) we know the sample ratio at sampling instants corresponding to samples of the input sig-

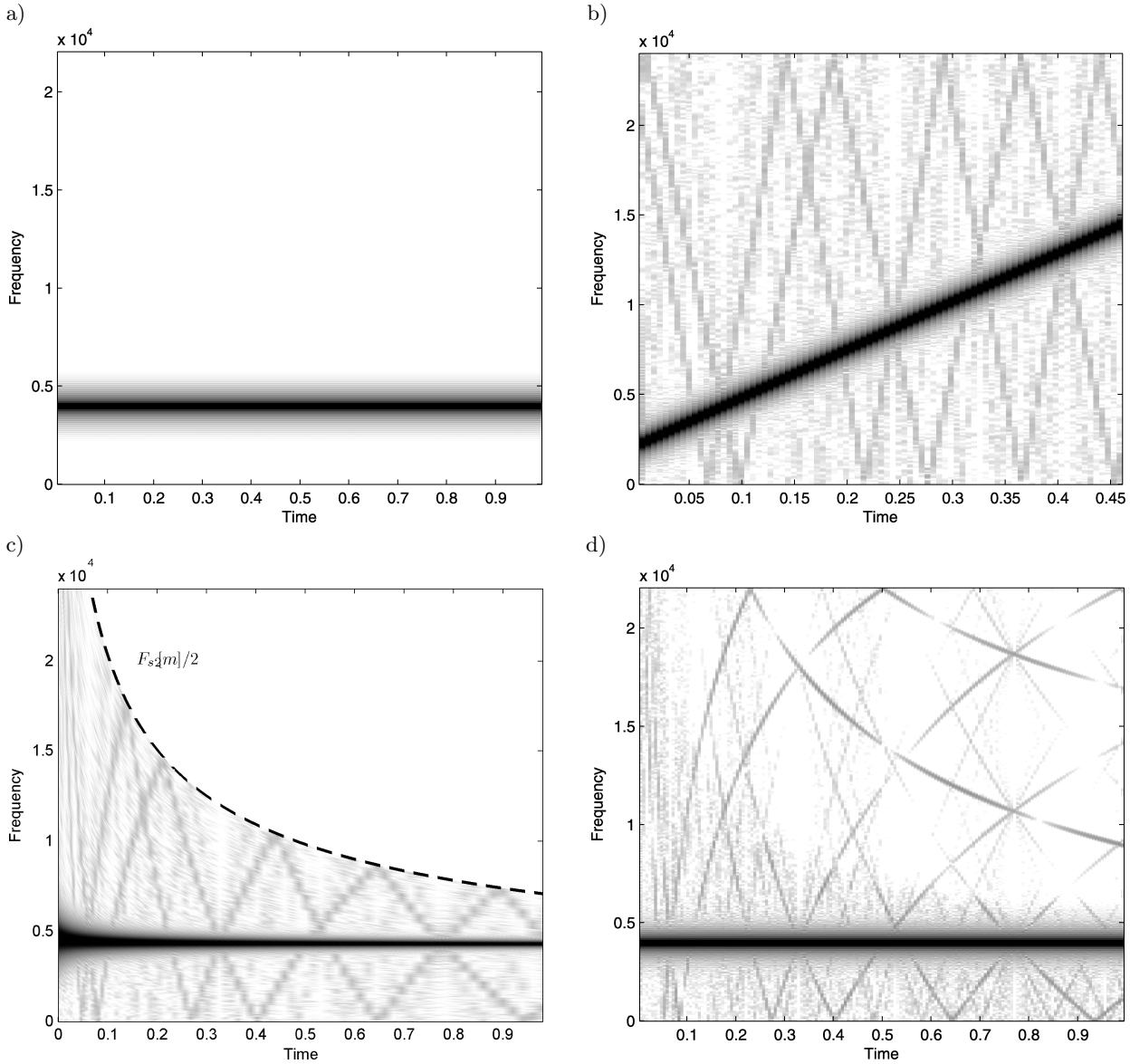

Fig. 8. Conversion from sinusoidal signal into chirp and back using SRC based on VFD filters. VFD filter with $f_a = 0.4$ of length $N = 17$ and the Farrow structure order $q = 5$ designed using window offset using MF FD filter of length $N_{\text{off}} = 5$: a) spectrogram of input sinusoidal signal, b) spectrogram of chirp signal obtained from sinusoidal signal. Assumed constant sample rate $\overline{F_{s2}} = F_{s1}$, c) spectrogram from Fig. 8b reshaped with accordance to variable sample rate. Dashed line indicates the folding frequency, d) spectrogram of restored sinusoidal signal.

nal but need to find the ratio $r[m]$ (19) specified in equidistant output instants m. Let us assume that we know the instantaneous input sample rate $F_{s1}[n] = 1/T_{s1}[n]$ sampled in the same instants as the input signal $x[n]$ and the instantaneous output sample rate $F_{s2}[m] = 1/T_{s2}[m]$ sampled in the same instants as the output signal $y[m]$. Now we are looking for a ratio $r[m] = T_{s2}[m]/\widehat{T}_{s1}[m]$ where $\widehat{T}_{s1}[m]$ is the input sampling period corresponding to the output time instant m.

Assuming that we know the positions of input sampling instants, which are our output instants from the previous problem

$$t_{\text{in}}[n] = \sum_{i=1}^{n} T_{s1}[n] = t_{\text{in}}[n-1] + T_{s1}[n] \tag{30}$$

and output sampling instants, which are our input instants from the previous problem

$$t_{\text{out}}[m] = \sum_{i=1}^{m} T_{s2}[m] = t_{\text{out}}[m-1] + T_{s2}[m]. \tag{31}$$

We propose the following algorithm for computation of the inverse of the instantaneous resampling ratio $r[m]$.

1. Start with input and output discrete time indexes $n := 0$ and $m := 0$ with corresponding continuous time instants $t_{\text{in}} := 0$ and $t_{\text{out}} := 0$.
2. Compute distance from the current output time instant to the current and the next input time instant:

   $\Delta t := t_{\text{out}} - t_{\text{in}}$,

   $\overline{\Delta t} := T_{s1}[n] - \Delta t$.
3. If $\overline{\Delta t} >= 0$ then
   (a) if the previous output sample is located in the same input sampling interval ($\Delta r = 0$ and $\Delta t > 0$) then

   $r[m] := T_{s2}[m]/T_{s1}[n]$,

   otherwise

   $r[m] := \Delta t/T_{s1}[n] + \Delta r$,

   (b) If $\overline{\Delta t} < T_{s2}[m]$ then

   $\Delta r := \overline{\Delta t}/T_{s1}[n]$

   and move to the next input instant

   $t_{\text{in}} := t_{\text{in}} + T_{s1}[n]$,

   $n := n + 1$,

   otherwise

   $\Delta r := 0$,

   (c) move to the next output instant

   $t_{\text{out}} := t_{\text{out}} + T_{s2}[m]$,

   $m := m + 1$,
4. otherwise ($\overline{\Delta t} < 0$)
   (a) $\Delta r := \Delta r + 1$,
   (b) move to the next input instant

   $t_{\text{in}} := t_{\text{in}} + T_{s1}[n]$,

   $n := n + 1$,
5. Go to point 2.

Figure 9 presents a modified version of the above algorithm with modifications which eliminate the problem of continuous increase of input and output times which eventually would lead to roundoff errors.

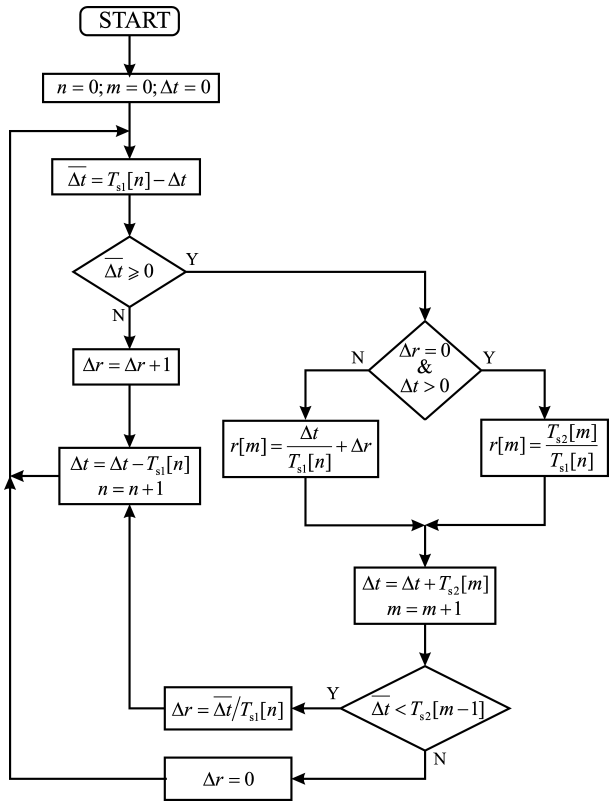

Fig. 9. Diagram for computation of $r[m]$ (19) based on instantaneous input and output sample rates.

With the ratio $r[m]$ calculated using the proposed algorithm the chirp signal presented in Fig. 8b can be resampled back into sinusoid. A spectrogram of obtained sinusoidal signal is presented in Fig. 8d. We can observe the nonlinear distortions but their level can be controlled with the selection of the overall filter attenuation in the stopband (Fig. 5).

7. Examples of audio signals resampling

The algorithm for calculation of the $r[m]$ ratio presented in previous section is universal and can be used for computation of an instantaneous ratio $r[m]$ for any variable input and output sample rate. In Fig. 10 an example demonstrating introduction and removal of sinusoidal sample rate changes into piano music are

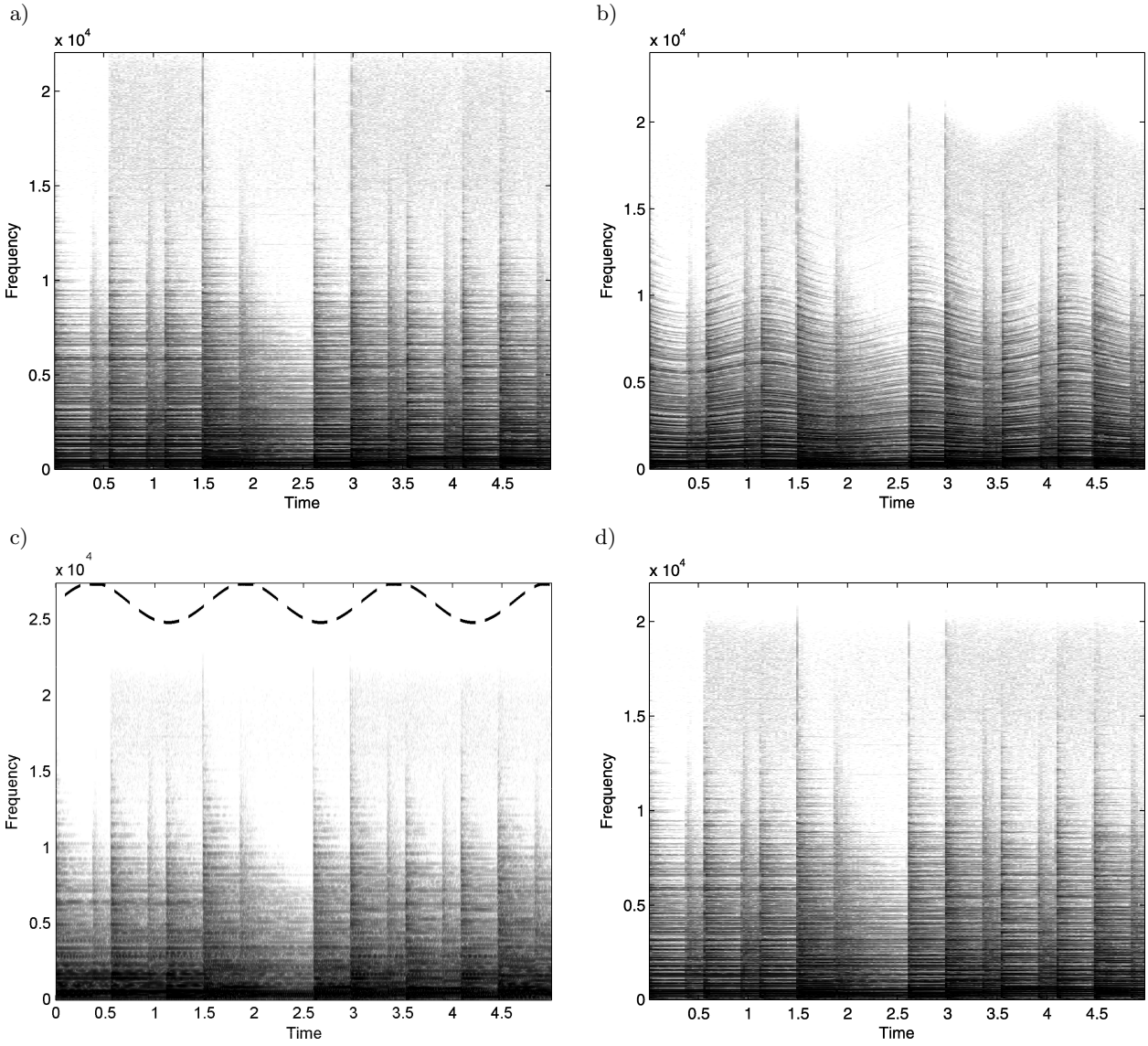

Fig. 10. Manipulations on piano music. Introduction and removal of sinusoidal fluctuations into instantaneous sample ratio: a) spectrogram of the input piano signal, b) spectrogram of piano signal with introduced sinusoidal fluctuations of sample rate, c) spectrogram from Fig. 10b reshaped with accordance to variable sample rate, d) spectrogram of the restored signal.

presented. Resampling have been performed using the VFD filter of length $N = 47$ with $f_a = 0.45$ implemented using the Farrow structure of order $q = 6$. FD filters were designed using windows offset with MF FD filters of length $N_{off} = 6$. Processing presented in Fig. 10 simulates changes of velocity of magnetic tape resulting from mechanical problems. On the other hand, the proposed solution can be used efficiently to remove such distortion (Fig. 10d), but only if we are able to find out how the velocity/sampling ratio changes (Ciarkowski *et al.*, 2005; Czyzewski *et al.*, 2007; 2010).

Figure 11 presents an example of speech signal (Fig. 11a) resampling with the resampling ratio computed on the basis of fundamental frequency (F_0). Based on the estimate of F_0 frequency (Fig. 11c) computed using algorithm presented in paper (Blok *et al.*, 2004), in each voiced segment a frequency modulation of excitation have been suppressed. This effect is clearly visible in Fig. 11e where frequency components of voiced segment have constant frequency. In this case resampling have been performed using the VFD filter of length $N = 33$ with $f_a = 0.4$ implemented using the Farrow structure of order $q = 5$. FD filters were designed using a window offset with MF FD filter of length $N_{off} = 5$ for cutoff frequency $f_c = 0.94$.

To achieve this result we have to assume that the instantaneous sample rate of the input signal is following

$$F_{s1}[n] = \overline{F_{s1}F_0}/F_0[n] \tag{32}$$

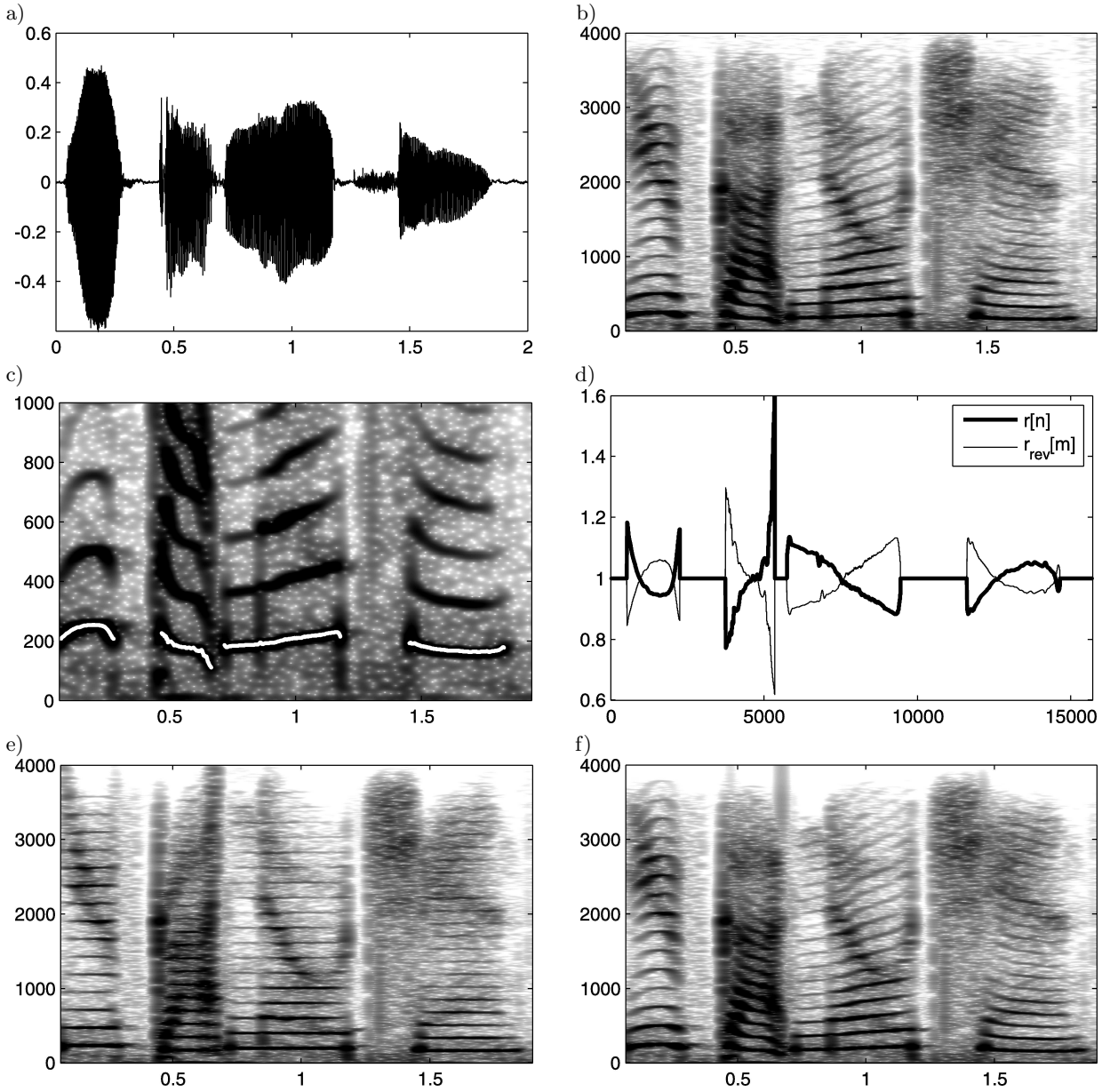

Fig. 11. Variable sample rate conversion of speech signal (phrase "You've got no chance"). Example demonstrating removal and restoration of modulation of fundamental frequency: a) input speech signal, b) spectrogram of input signal, c) zoomed spectrogram of input signal with marked F_0 frequency (white lines), d) resampling ratios, e) spectrogram of resampled signal with constant fundamental frequency, f) spectrogram of restored signal.

and the output sample rate is constant $F_{s2}[m] = \overline{F_{s1}}$. On this basis the reversed instantaneous resampling ratio $r[m]$ (Fig. 11d) is computed using the algorithm presented earlier (Fig. 8). Additionally, in order to keep the length of the speech signal unchanged, for each voiced segment we have selected such constant value of the output fundamental frequency $\overline{F_0}$ which results in average value of $r[m]$ equal to one. The resampled speech signal sounds synthetic but is more melodious.

This process can be reversed and the excitation frequency modulation can be restored (Fig. 11f). This time computation of $r[n]$ is simple since the input sample rate is now constant and we already know the variable output sample rate sampled at output sampling instants.

$$r_{\text{rev}}[n] = F_0[n]/\overline{F_0}. \tag{33}$$

The signal is reconstructed with almost no noticeable distortions (Fig. 11f). The only problems are a slightly limited signal bandwidth and minor aliasing distortions that can be noticed at $t = 0.7$. This aliasing is the result of a large $r[m] > 1.5$ which introduced distortions during the first stage of processing (Fig. 11e). These distortions can be reduced by decreasing f_c, but this would result in significant bandwidth limitation in other segments of the processed signal. Another solu-

tion is to use in the formula (16) a variable f_c changing in accordance to the ratio $r[m]$ but this would significantly increase computational cost since we would not be able to use the Farrow structure.

8. Conclusion

We have demonstrated that a VFD filter implemented using the Farrow structure can be applied to audio signal resampling with continuously changing sample rate ratio. The proposed approach can be used to simulate signal distortions, for example to change or remove speech intonation, as well as to correct old recordings distorted due to non-constant media velocity, e.g. magnetic tape. Full utilization of the proposed tool requires that the properties of recorded distortions, as in the case of reconstruction of old recordings, are automatically measured (Ciarkowski *et al.*, 2005; Czyzewski *et al.*, 2007; 2010). In this paper we have demonstrated the capabilities of the proposed solution based on the example of speech signal processing. Resampling ratio have been selected based on estimated fundamental frequency of voiced speech segments. On this basis, modulation of voiced excitation has been suppressed and then reintroduced into the processed signal. The primary limitation of the proposed approach is the need to select a constant cutoff frequency of the overall interpolation filter. It would be better if we could change the cutoff frequency in accordance with the instantaneous resampling ratio. Therefore further research should focus on the application of variable cutoff frequency and a search for an efficient structure to implement FD filters designed using the offset window method allowing for such changes during processing.

The demonstration code allowing for the reproduction of experiments presented in this paper and audio sample files are available on the web page http://pg.edu.pl/26fd45817c_marek.blok/VR-SRC-AofA.

References

1. AES5-2008 (2008), *AES recommended practice for professional digital audio – preferred sampling frequencies for applications employing pulse-code modulation*, Technical report, Audio Engineering Society, revision of AES5-1997.

2. Blok M. (2002a), *Collective filter evaluation of an FSD filter-based resampling algorithm*, OSEE 2002, TechOnLine, Bedford, Massachusetts, USA, www.eetimes.com/design/signal-processing-dsp/4017905.

3. Blok M. (2002b), Optimal fractional sample delay filter with variable delay, *OSEE 2002*, TechOnLine, Bedford, Massachusetts, USA, www.eetimes.com/design/analog-design/4018005.

4. Blok M. (2005), *Farrow structure implementation of fractional delay filter optimal in Chebyshev sense*, Proceedings of SPIE, volume 6159, 61594K, Wilga, Poland.

5. Blok M. (2012a), *Fractional delay filter design for sample rate conversion*, Proceedings of FedCSIS'2012, pp. 701–706, Wroclaw, Poland.

6. Blok M. (2012b), *Fractional delay filter design with extracted window offsetting*, Proceedings of MixDes'2012, pp. 489–494, Warsaw, Poland.

7. Blok M. (2013), *Comments on "Closed Form Variable Fractional Time Delay Using FFT"*, IEEE Signal Processing Letters, **20**, 8, 747–750.

8. Blok M., Rojewski M., Sobociński A. (2004), *A new algorithm for pitch period estimation* [in Polish], Zeszyty Naukowe, Faculty of ETI, Gdańsk University of Technology, Technologie Informacyjne, **3**, 125–134.

9. Ciarkowski A., Czyzewski A., Dziubinski M., Kaczmarek A., Kulesza M., Maziewski P., Kostek B. (2005), *Methods for detection and removal of parasitic frequency modulation in audio recordings*, AES Conference: 26th International Conference: Audio Forensics in the Digital Age, paper no. 3–3.

10. Czyzewski A., Ciarkowski A., Kaczmarek A., Kotus J., Kulesza M., Maziewski P. (2007), *DSP techniques for determining „wow" distortion*, Journal of the Audio Engineering Society, **55**, 4, 266–284.

11. Czyzewski A., Maziewski P., Kupryjanow A. (2010), *Reduction of parasitic pitch variations in archival musical recordings*, Signal Processing, **90**, 4, 981–990.

12. Evangelista G. (2003), *Design of digital systems for arbitrary sampling rate conversion*, Signal Processing, **83**, 2, 377–387.

13. Farrow C.W. (1988), *A continuously variable digital delay element*, Proceedings of ISCAS'88, pp. 2641–2645, Espoo, Finland.

14. Harris F.J. (1997), *Performance and design of Farrow filter used for arbitrary resampling*, Proceedings of DSP'97, **2**, pp. 595–599, Santorini, Greece.

15. Hermanowicz E. (1998), *A nearly optimal variable fractional delay filter with extracted Chebyshev window*, Proceedings of ICECS'98, **2**, pp. 401–404, Lisboa, Portugal.

16. Hermanowicz E. (2004), *On designing a wideband fractional delay filter using the Farrow approach*, Pr6oceedings of EUSIPCO'2004, pp. 961–964, Austria.

17. Hermanowicz E., Rojewski M., Blok M. (2000), *A sample rate converter based on a fractional delay filter bank*, Proceedings of ICSPAT 2000, Dallas, Tx, USA.

18. Laakso T.I., Välimäki V., Karjalainen M., Laine U.K. (1996), *Splitting the unit delay – tools for fractional delay filter design*, IEEE Signal Processing Magazine, **13**, 1, 30–60.

19. Maziewski P. (2006), *Wow defect reduction based on interpolation techniques*, Bulletin of the Polish Academy of Science: Technical Sciences, **54**, 469–477.

20. Rajamani K., Yhean-Sen L., Farrow C.W. (2000), *An efficient algorithm for sample rate conversion from CD to DAT*, IEEE Signal Processing Letters, **7**, 10, 288–290.

21. Tarczynski A., Kozinski W., Cain G.D. (1994), *Sampling rate conversion using fractional-sample delay*, Proceedings of ICASSP'94, pp. 285–288, Adelaide, Australia.

22. Yardim A., Cain G.D., Henry P. (1996), *Optimal two-term offset windowing for fractional delay*, Electronics Letters, **32**, 6, 526–527.

23. Yardim A., Cain G.D., Lavergne A. (1997), *Performance of fractional-delay filters using optimal offset windows*, Proceedings of ICASSP'97, **3**, pp. 2233–2236, Munich, Germany.