Optimized Polynomial Spline Basis Function Design for Quasi-Bandlimited Classical Waveform Synthesis

Jussi Pekonen, *Member, IEEE*, Juhan Nam, *Student Member, IEEE*, Julius O. Smith, *Member, IEEE*, and Vesa Välimäki, *Senior Member, IEEE*

***Abstract*—Classical geometric waveforms used in virtual analog synthesis suffer from aliasing distortion when simple sampling is used. An efficient antialiasing technique is based on expressing the waveforms as a filtered sum of time-shifted approximately bandlimited polynomial-spline basis functions. It is shown that by optimizing the coefficients of the basis function so that the aliasing distortion is perceptually minimized, the alias-free bandwidth of classical waveforms can be expanded. With the best of the case examples given here, the generated impulse-train and sawtooth waveform are alias-free up to fundamental frequencies over 10 kHz when the sampling rate is 44.1 kHz.**

***Index Terms*—Acoustic signal processing, antialiasing, audio oscillators, interpolation, music, signal synthesis.**

I. Introduction

DIGITAL modeling of the sound production technique of the analog synthesizers from the 1960s and 1970s is an important topic in the music technology business nowadays. Unfortunately, the classical geometric waveforms, such as the sawtooth and rectangular waves, typically used in analog synthesis suffer from aliasing distortion when they are digitally generated by algorithms which compute exact samples of the continuous analog waveforms. In the past few years, more and more research effort has been devoted to finding oscillator algorithms that remove or suppress aliasing [1]–[6].

A popular and efficient antialiasing oscillator methodology is based on expressing the continuous-time waveforms as a sum of time-shifted non-bandlimited basis functions. By lowpass filtering this function train an approximately bandlimited waveform is obtained [1], [3], [7]–[9]. The digital bandlimited waveforms are constructed by sampling the lowpass-filtered continuous-time signal. This approach results in signals that have aliasing mainly at high frequencies where human hearing is least sensitive, and algorithms that are based on this idea are called quasi-bandlimited oscillator algorithms.

Manuscript received October 17, 2011; revised December 11, 2011; accepted December 22, 2011. Date of publication January 06, 2012; date of current version January 26, 2012. This work was supported in part by the Academy of Finland, under Project 122815. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Rudolf Rabenstein.

J. Pekonen was with the Department of Signal Processing and Acoustics, Aalto University School of Electrical Engineering, FI-00076 Aalto, Espoo, Finland. He is now an independent researcher (e-mail: Jussi.Pekonen@gmail.com).

J. Nam and J. O. Smith are with the Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, Stanford, CA 94305 USA (e-mail: juhan@ccrma.stanford.edu; jos@ccrma.stanford.edu).

V. Välimäki is with the Department of Signal Processing and Acoustics, Aalto University School of Electrical Engineering, FI-00076 Aalto, Espoo, Finland (e-mail: Vesa.Valimaki@tkk.fi).

Digital Object Identifier 10.1109/LSP.2012.2183123

Two well known algorithms using the above technique are the "BandLimited Impulse Train" (BLIT) [1] and "BandLimited unit stEP" (BLEP) [3] methods. The BLIT method is based on the observation that time derivatives of the classical waveforms yield a sequence of impulses [1]. The BLEP method expresses a classical waveform as a sum of a time-shifted bandlimited step functions [3]. The bandlimited step function can be decomposed to a sum of non-bandlimited step function and a correction (residual) function, which effectively cancels the aliased components. For both techniques, the ideal bandlimited basis function can be expressed in closed form [1], [9].

However, ideal BLIT and BLEP synthesis is impossible because the ideal basis functions are infinitely long. This issue is related to the problem addressed in interpolation because the basis function of the BLIT method is the sinc function [1] that is the ideal bandlimited interpolation function [10]. Therefore, the techniques that have been applied in the approximation of ideal bandlimited interpolation can, in principle, also be applied for the BLIT basis function design. However, so far only a few traditional approaches have been used [9].

The ideal basis functions of the BLIT and BLEP techniques, or the residual function in the case of BLEP, are typically windowed and tabulated [1], [3], [7]–[9], which means that they can be regarded as having been designed by the window method of FIR filter design. Alternatively, they can be designed using convex optimization methods [11].

When synthesizing bandlimited classical waveforms digitally, the basis-function approximations of the BLIT and BLEP techniques are usually generated via interpolated table-lookup. The look-up tables must be oversampled [1], [3], [7]–[9]. With sufficiently high table sampling density, linear interpolation suffices [10]. However, higher-order polynomial interpolation has also been used to reduce table size further.

For best aliasing reduction performance the tabulated basis functions need to be quite long [7]–[9], [12]. The computational complexity of the BLIT and BLEP algorithms is directly proportional to the number of samples that the algorithm has to modify around the discontinuities [7]–[9], [12]. To reduce computational complexity while maintaining interpolation quality, alternative basis functions for the BLIT and BLEP algorithms have been developed [8], [13], [14]. These approaches modify only a few samples around the discontinuities, and they provide better perceptual error measures than the table-based approaches of the same computational complexity.

In [12] it was shown that the aliasing-reduction performance of table-based BLIT and BLEP methods can be improved for short look-up tables by optimizing the table values under a perceptually informed error criterion. In [8] and [9] the alias-reduction performance was shown to be improved by designing table-free polynomial-based algorithms. In this paper, these two approaches are combined, and an optimized polynomial-based algorithm having minimal perceived aliasing and low computational cost is proposed.

Polynomial approximation of the ideal basis functions of the BLIT and BLEP algorithms is extended in Section II by introducing an optimized polynomial-spline basis function that minimizes perceived aliasing distortion for a given polynomial order under certain constraints. Section III presents a few low-order examples and evaluates the obtained polynomials using a computational measure. Section IV concludes the letter.

II. Optimal Polynomial-Spline Basis Function

Our choice of BLIT or BLEP *residual* is a *polynomial spline* $g(t)$ constructed from L consecutive polynomial segments:

$$g(t) = \sum_{k=0}^{K_l} g_{l,k} t^k, \quad t \in [t_{l-1}, t_l], \quad l = 1, 2, \ldots, L \tag{1}$$

where t is the time variable, K_l is the order and $g_{l,k}$ are the coefficients of the polynomial in the lth subrange. Outside the definition range $t \in [t_0, t_L]$ $g(t)$ is defined to be zero. In the polynomial spline domain [15], the interval endpoints t_l are called "knot points."

The polynomial spline $g(t)$ can be sampled for digital synthesis to create an FIR filter whose order is $N = \lceil t_L - t_0 - 1 \rceil$, where $\lceil \cdot \rceil$ denotes the ceiling function. The filter coefficients b_n for $n = 0, 1, \ldots, N$ are given by $b_n(d) = g(t_0 + n + d)$, where $d \in [0, 1)$ is the fractional delay which defines the sample point within the sampling interval.

A. Optimization Problem

The coefficients $g_{l,k}$ for the optimal polynomial-spline basis functions are obtained from an optimization procedure that minimizes the computational load of the algorithm while keeping aliasing distortion inaudible. A similar problem has been previously addressed in image processing where the objective has been to find short kernel functions that have aliasing properties suitable for image interpolation, see e.g. [16], [17]. However, in this paper the objective is to find polynomial functions with which one can generate classical audio waveforms that are perceptually free from aliasing.

The optimization problem discussed here can be formulated the same way as for the table-based approaches [12]. However, similarly to the look-up table optimization [12], the constraints depend on the fundamental frequency f_0 and the polynomial spline $g(t)$, resulting in an optimal basis function solution only for the given fundamental frequency. Again, the general solution can be found as the maximum of the minimized costs in the range of interesting fundamental frequencies. However, finding a feasible solution to the general optimization problem is a complex task due to the nonlinear and explicit formulation, and the existence of a solution is not guaranteed [12]. Therefore, this general optimization approach needs to be reformulated to one that has a solution.

The aliasing problem can be approached from a different point of view. Since the spectral envelope of the synthesized waveform using the BLIT or BLEP method is defined by the amplitude response of the underlying basis function [8], [9], [12], the level of aliasing is also determined by the amplitude response of the basis function. Therefore, the optimal polynomial coefficients can be obtained by minimizing the weighted amplitude response of the polynomial spline $g(t)$ above the Nyquist limit under a constraint on the computational cost. In practice, the computation cost constraint can be defined as an upper limit by setting the definition range and the polynomial orders K_l in advance.

Now, the optimization objective can be expressed as a weighted minimization of a norm $\|G\|_{W,p}^p$ of the amplitude response $G(\omega)$ above the Nyquist limit:

$$\min_{g_{l,k}} \int_{\pi}^{\infty} W(\omega) \, |G(\omega)|^p \, \mathrm{d}\omega \tag{2}$$

where $W(\omega)$ is a non-negative, real-valued, and perceptually informed weighting function and $G(\omega)$ is the Fourier transform of (1) rectangularly windowed to the interval $[t_0, t_L]$.

Note that for each BLIT function there is a related BLEP function that is obtained by integrating the BLIT function with respect to time [9]. If a BLIT function is the solution to the BLIT-related optimization problem, then its integral is also the solution to the BLEP-related optimization problem when the weighting function $W(\omega)$ is kept the same. This can be proved by analyzing the conditions of the optimal solution and by applying a property of the Fourier transform. Therefore it suffices to find only the optimal BLIT functions as the related optimal BLEP functions are trivially obtained by integrating the result.

B. Optimization Constraints

From (2) it can be seen that the optimization problem requires at least one constraint. Unconstrained optimization yields zero as the global minimum of the objective function, which is trivially obtained with $g(t) = 0$ for all t.

The optimization constraints are expressed both in time- and frequency-domains. The time-domain constraints can be given for each subrange or for the general shape of the basis function. The frequency-domain constraints are typically given for the frequencies that do not alias, and they can also be given for a set of subbands or for the baseband as a whole.

1) Time-Domain Constraints: When the resulting polynomial-spline basis function is desired to be linear-phase, a case where the phases of the nonaliased components will be those of the ideally bandlimited waveform, it is required that the spline is symmetric with respect to its midpoint with respect to its midpoint $t_{\mathrm{m}} = (t_L + t_0)/2$: $g(t_{\mathrm{m}} - \tau) = g(t_{\mathrm{m}} + \tau)$, where $\tau \in [0, t_{\mathrm{m}}]$. Moreover, the overall amplitude change obtained by the impulse function needs to be one [12]:

$$\int_{t_0}^{t_L} g_{\mathrm{BLIT}}(t)\mathrm{d}t = G(0) = 1. \tag{3}$$

2) Frequency-Domain Constraints: Ideally, the amplitude response of the basis function at the baseband is one in the BLIT synthesis. However, this ideal response can be obtained only with infinitely long basis functions, which is not practical for real-time synthesis. Therefore, the baseband response needs to be allowed to deviate from the ideal response:

$$|\Theta_i(\omega) - G_{\mathrm{dB}}(\omega)| \leq \Delta_i(\omega), \quad \omega \in \Omega_i \subseteq [0, \pi] \tag{4}$$

where $G_{\mathrm{dB}}(\omega)$ is the amplitude response of $g(t)$ in dB, and $\Theta_i(\omega)$ and $\Delta_i(\omega)$ are the target response and the allowed deviation, respectively, in dB in the ith subband Ω_i.

There are two approaches to defining the amplitude-response constraint. If the target response is the ideal response, the allowed deviation directly describes how much the amplitude response can differ from the ideal response. With this approach, once $g(t)$ is obtained, the deviation can be compensated with a post-equalizing filter that reduces the difference between the amplitude response and the target response as much as desired [9], [12].

Alternatively, the compensation process can be reversed. The post-equalizing filter is defined first and the inverse of its amplitude response is considered as the target response. Now, the allowed deviation defines how much the resulting waveform can differ from the ideal response in the output of the synthesizer. Using this approach, the deviation at the output can be controlled more easily than in the first approach in which the difference depends on how well the filter response compensates the deviation.

These approaches can be developed further by embedding the post-equalization filter design process into the optimization problem. In this case, one can add the deviation from the target response to the objective function as a penalty function [18] and give constraints that set the maximum allowed deviation and ensure filter stability when the filter to be designed is an IIR filter. However, one should note that this approach includes yet another nonlinear constraint, thus increasing the complexity of the optimization problem.

III. Case Examples of Optimized Splines

To illustrate the power of the optimization approach, the minimization task was applied to a few low-order polynomial splines. Only the $\|G\|^2_{W,2}$ norm and symmetric and zero-phase, i.e. $t_0 = -t_L$, splines were considered. As the weighting function $W(\omega)$, a rectangular pulse function given by

$$W(\omega) = \begin{cases} 1, & \omega \in [2k\pi - \omega_{\mathrm{W}}, 2k\pi + \omega_{\mathrm{W}}], \quad k \in \mathbb{N}, \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

was used. This weighting function takes into account only the frequencies that fold back between dc and the design bandwidth $\omega_{\mathrm{W}} \in [0, \pi]$. In addition, the second approach in the compensation filter design was applied, and a one-pole filter whose pole is at -0.9 was used. The allowed baseband deviation was set to 1.5 dB for frequencies below $300\pi/441$, which corresponds to 15 kHz for sampling rate of 44.1 kHz.

First, optimal two-segment ($L = 2$) first-, second-, and third-order splines for $t_L = 1$ were sought. A linear interpolation polynomial can be considered as the reference polynomial spline for this case example [7]–[9]. However, the optimization yielded splines that had a worse alias reduction performance than linear interpolation especially at low frequencies. By applying a narrower design bandwidth the performance of the optimized spline could be improved, but they eventually converged towards linear interpolation when $\omega_{\mathrm{W}} \to 0$.

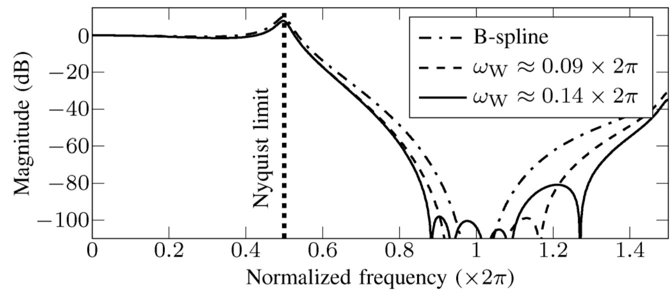

Fig. 1. Post-equalized (pole at -0.9) amplitude response of the optimized four-segment third-order BLIT polynomial-spline when the design bandwidth ω_{W} is $80\pi/441 \approx 0.09 \times 2\pi$ (dashed line) and $120\pi/441 \approx 0.14 \times 2\pi$ (solid line) for the baseband and the first two generations of aliasing. The response of the third-order B-spline polynomial (dash-dotted) is plotted for reference.

When the definition range was extended to $t_L = 2$, the optimization approach could provide four-segment ($L = 4, t_l \in \mathbb{Z}$) third-order splines that outperformed the current state-of-the-art polynomial-spline basis function, the third-order B-spline polynomial [8], [9]. Lower-order splines were also tested but they performed worse than the reference. However, the alias reduction performance depended heavily on the design bandwidth, as is illustrated in Fig. 1. When the design bandwidth is increased, the alias reduction performance gets better at middle and high frequencies (between 0.6 and 0.9 and between 1.1 and 1.4 in Fig. 1), but at low frequencies (close to 1 in Fig. 1) the performance is worse than with a narrow design bandwidth. Therefore, ω_{W} needs to be chosen carefully in order to obtain a spline that provides good alias reduction at middle frequencies and whose amplitude response at low frequencies is sufficiently small, e.g. below -100 dB.

A. Evaluation of the Alias Reduction Performance

The alias reduction performance of the optimized splines was tested by evaluating the alias reduction performance using computational models of the hearing threshold and the frequency masking phenomenon for a few values of ω_{W}. The same model has also been used in the evaluation of previously studied polynomial-spline basis function designs [8], [9].

The hearing threshold curve used by the model is given by

$$T(f) = 3.64f^{-0.8} - 6.5e^{-0.6(f-3.3)^2} + 10^{-3}f^4 \tag{6}$$

where f is frequency in kHz and the level $T(f)$ is the absolute sound pressure level (SPL) [19]. The frequency masking phenomenon is modeled as an asymmetric spreading function for a masker, expressed as

$$S(L_{\mathrm{M}}, \Delta z_{\mathrm{b}}) = L_{\mathrm{M}} + (-27 + 0.37 \max\{L_{\mathrm{M}} - 40, 0\}\theta(\Delta z_{\mathrm{b}}))\,|\Delta z_{\mathrm{b}}| \tag{7}$$

where Δz_{b} is the frequency difference between a masking component and a maskee in Bark units, L_{M} is the SPL of a masker in dB, and $\theta(\Delta z_{\mathrm{b}})$ is the step function that is one for non-negative arguments and zero otherwise [20]. The spreading functions are shifted down by 10 dB, and the maximum of the hearing threshold and the masking curves is considered the threshold

TABLE I
POLYNOMIAL COEFFICIENTS $g_{l,k}$ FOR THE OPTIMAL SYMMETRIC BLIT POLYNOMIAL-SPLINE BASIS FUNCTION (SEE (1) AND SECTION II-B-1) WHEN $t_L = 2, L = 4, t_l \in \mathbb{Z}$, AND THE DESIGN BANDWIDTH $\omega_W = 120\pi/441$

	k			
l	0	1	2	3
1	1.34261	1.94699	0.94762	0.15485
2	0.62351	−0.04817	−0.95010	−0.46625
3	0.62351	0.04817	−0.95010	0.46625
4	1.34261	−1.94699	0.94762	−0.15485

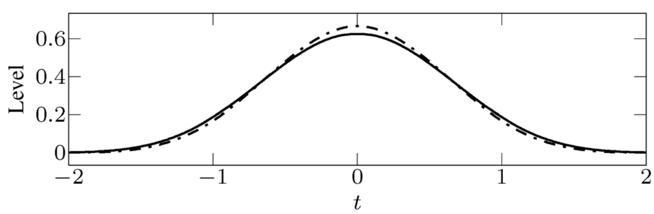

Fig. 2. Polynomial basis functions in the time-domain: B-spline (dash-dotted line) and the optimized polynomial-spline $g(t)$ obtained with $\omega_W = 120\pi/441$ (solid line).

of audibility of aliasing [8], [9]. In addition, the spectral power of the oscillator output is scaled to a reference level, and a sinusoid alternating between -1 and 1 is chosen to represent a signal having a SPL of 96 dB.

The highest alias-free fundamental frequencies of the bandlimited impulse trains generated using the optimized third-order spline for a few values of ω_W were sought. For $\omega_W = 80\pi/441$ the generated impulse train is alias-free up to 9464 Hz, for $\omega_W = 120\pi/441$ up to 10471 Hz, and for $\omega_W = 160\pi/441$ up to 9392 Hz. With these splines the alias-free bandwidth can be expanded quite much compared to 5784 Hz of the current state-of-the-art polynomial-spline basis function [8]. Furthermore, the results show that the overall best performance will be obtained with a design bandwidth close to $\omega_W = 120\pi/441$.

When the BLIT polynomial-spline for $\omega_W = 120\pi/441$, whose coefficients are given in Table I and that is plotted in Fig. 2, is integrated to a BLEP polynomial and the resulting BLEP residual is used as the correction function for a sawtooth waveform [9], the highest alias-free fundamental frequency of the sawtooth waveform is extended to 12259 Hz. Again, a great improvement is shown over the state-of-the-art polynomial-spline basis-function design, which is alias-free up to 7845 Hz [9].

IV. CONCLUSION

In this letter, polynomial-based quasi-bandlimited oscillator algorithms were improved by introducing a polynomial optimization procedure with which the resulting aliasing distortion can be minimized. The improved performance of the optimized polynomial-spline basis functions was exemplified with low-order case examples. With the best of the optimized splines and by using a one-pole equalizing filter, the highest alias-free fundamental frequency of an impulse train and a sawtooth waveform was increased to be over 10 kHz, which is well above the frequency range typically used in music (20–8400 Hz). Therefore, the proposed optimization procedure can yield polynomial-spline basis functions that provide an excellent alias reduction performance in musical sound synthesis.

REFERENCES

[1] T. Stilson and J. O. Smith, "Alias-free digital synthesis of classic analog waveforms," in *Proc. Int. Computer Music Conf.*, Hong Kong, Aug. 1996, pp. 332–335.

[2] J. Lane, D. Hoory, E. Martinez, and P. Wang, "Modeling analog synthesis with DSPs," *Comput. Music J.*, vol. 21, no. 4, pp. 23–41, Winter, 1997.

[3] E. Brandt, "Hand sync without aliasing," in *Proc. Int. Computer Music Conf.*, Havana, Cuba, Sep. 2001, pp. 365–368.

[4] V. Lazzarini and J. Timoney, "New perspectives on distortion synthesis for virtual analog oscillators," *Comput. Music J.*, vol. 34, no. 1, pp. 28–40, Spring, 2010.

[5] V. Välimäki, J. Nam, J. O. Smith, and J. S. Abel, "Alias-suppressed oscillators based on differentiated polynomial waveform," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 4, pp. 786–798, May 2010.

[6] J. Pekonen, V. Lazzarini, J. Timoney, J. Kleimola, and V. Välimäki, "Discrete-time modelling of the Moog sawtooth oscillator waveform," *EURASIP J. Adv. Signal Process.*, vol. 2011, 2011, Article ID 785103, 15 pages.

[7] V. Välimäki and A. Huovilainen, "Antialiasing oscillators in subtractive synthesis," *IEEE Signal Process. Mag.*, vol. 23, no. 2, pp. 116–125, Mar. 2007.

[8] J. Nam, V. Välimäki, J. S. Abel, and J. O. Smith, "Efficient antialiasing oscillator algorithms using low-order fractional delay filters," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 4, pp. 773–785, May 2010.

[9] V. Välimäki, J. Pekonen, and J. Nam, "Perceptually informed synthesis of bandlimited classical waveforms using integrated polynomial interpolation," *J. Acoust. Soc. Amer.*, vol. 131, no. 1, pp. 974–986, Jan. 2012.

[10] J. O. Smith and P. Gossett, "A flexible sampling-rate conversion method," in *Proc. 1984 Int. Conf. Acoust., Speech, and Signal Process. (ICASSP'84)*, San Diego, CA, Mar. 1984, vol. 9, pp. 112–115.

[11] W. Putnam and J. O. Smith, "Design of fractional delay filters using convex optimization," in *Proc. 1997 Workshop Applicat. Signal Process. Audio and Acoust.*, New Paltz, NY, Nov. 1997.

[12] J. Pekonen, J. Nam, J. O. Smith, J. S. Abel, and V. Välimäki, "On minimizing the look-up table size in quasi-bandlimited classical waveform oscillators," in *Proc. 13th Int. Conf. Digital Audio Effects (DAFx-10)*, Graz, Austria, Sept. 2010, pp. 419–422.

[13] J. Nam, V. Välimäki, J. S. Abel, and J. O. Smith, "Alias-free oscillators using feedback delay loops," in *Proc. 12th Int. Conf. Digital Audio Effects (DAFx-09)*, Como, Italy, Sep. 2009, pp. 347–352.

[14] V. Lazzarini and J. Timoney, "Theory and practice of modified frequency modulation synthesis," *J. Audio Eng. Soc.*, vol. 58, no. 6, pp. 459–471, June 2010.

[15] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Process. Mag.*, vol. 16, no. 6, pp. 22–38, Nov. 1999.

[16] T. Blu, P. Thévenaz, and M. Unser, "MOMS: Maximal-order interpolation of minimal support," *IEEE Trans. Image Process.*, vol. 10, no. 7, pp. 1069–1080, July 2001.

[17] A. Gotchev, K. Egiazarian, and T. Saramäki, "Image interpolation by optimized spline-based kernels," in *Advances in Signal Transforms Theory and Applications*, J. Astola and L. Yaroslavsky, Eds. New York: Hindawi, 2007, ch. 7, pp. 285–335.

[18] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, 3rd ed. Hoboken, NJ: Wiley, 2006, ch. 9, pp. 469–536.

[19] E. Terhardt, "Calculating virtual pitch," *Hearing Res.*, vol. 1, no. 2, pp. 155–182, Mar. 1979.

[20] M. Bosi, "Audio coding: Basic principles and recent developments," in *Proc. 6th Int. Conf. Humans and Computers*, Aiku, Japan, Aug. 2003, pp. 1–17.