PHYSICAL AND BEHAVIORAL CIRCUIT MODELING OF THE SP-12 SAMPLER

David T. Yeh, John Nolting, Julius O. Smith
Center for Computer Research in Music and Acoustics (CCRMA)
Stanford University, Stanford, CA

ABSTRACT

Aliasing, usually considered an artifact of discrete time systems to be avoided, is found to be an aesthetic feature of the E-MU SP-12 sampler/drum machine. This paper presents the steps in modeling the SP-12 as a signal processing system. Measurements of the characteristics of the SP-12 are presented. The signal path is analyzed to produce a physically based model of the circuit. Circuit analysis in SPICE provides transfer functions, which are converted into digital filters by system identification. Aliasing is implemented using interpolation and downsampling. The results of the algorithm are compared to samples from the original system.

1. INTRODUCTION

Physical modeling is an approach to deriving efficient algorithms to simulate various signal processing circuits. Sometimes a physical model is too involved to implement directly, but its insights are used to derive a behavioral model that approximates the correct response. This is sometimes termed virtual analog circuit modeling. A combination of reverse engineering and circuit analysis allows the systematic formulation of an algorithm that faithfully reproduces the character of the original system. In this paper, these techniques are applied to the E-MU SP-12 sampler/drum machine to create an algorithm that reproduces the sound of this device, which is no longer being manufactured.

1.1. Features of the SP-12

The SP-12 (Sampling Percussion), introduced in the mid-80's, is a sampling drum machine with a sequencer to lay down drum tracks. It features 8 velocity-sensitive pads, 8 control slides, and 8-voice polyphony through 8 independent outputs. It samples at a low 12 bits and 27.5 kHz rate. The Turbo version features 192 kB of sample memory which is about 5 seconds, but each sample can only be a maximum of 2.5 seconds. It features MIDI interfacing and SMPTE synchronization. There are 24 internal drum samples stored in ROM, and 8 samples that the user can record. Each output channel features a different equalization. The interface allows users to edit samples by looping, truncation, and adjusting the decay.

Figure 1. Control panel of the E-MU SP-12 sampler.

The SP-12 is used for hip hop beats, to give drum sounds a hard edge and grit. A rudimentary pitch shifter can detune sounds, but also adds a gritty character that comes from aliasing. It also features a warm low-pass equalization that musicians desire.

1.2. Background material

Literature in virtual analog often discusses alias reduction. Efficient methods of generating bandlimited waveforms for subtractive synthesis are described in [7, 9].

The canonical virtual analog example is the Moog filter [8, 1]. This work extends upon that earlier work, integrating the analysis of the components into the modeling of an entire system.

2. OVERVIEW OF THE SP-12 SIGNAL PATH

The SP-12's signal path consists of an anti-aliasing filter based on operational amplifiers (opamp), a sample and hold at 27.5 kHz, a 12-bit successive approximation quantizer, time-domain digital signal processing, a zero-order hold (see 3.4), and a choice between six optional equalization filters to attenuate spectral aliases from the zero-order hold. Two of these filters employ the SSM-2044 Voltage Controlled Filter (VCF) chip as a 4-pole lowpass with time-varying cutoff frequency.

The time-domain processing of the SP-12 features variable-decay time-enveloping of the signal, and a rudimentary pitch shifting algorithm common to digital synthesis systems of that era [7].

3. DIGITAL IMPLEMENTATION OF THE SP-12

3.1. Assumptions

The following assumptions, while not strictly true, are made to make the problem more tractable:

- Frequency range of input and output is 0 – 20 kHz. The spectrum outside this range gives insight into the operation of the device, but is not reproduced by the algorithm.
- Sample rate is 27.5 kHz with no jitter.
- Operational amplifiers are ideal (infinite bandwidth and gain, zero output impedance, no distortion)
- Filters are linear.
- Ideal 12-bit A/D conversion, all quantization steps have uniform size, yielding a 72-dB noise floor [4].

3.2. Anti-alias filter model

Because the SP-12 involves aliasing effects, one must ensure that the digital implementation aliases properly. Some of the aliasing comes from the non-idealities in the anti-aliasing filters.

The circuits for the analog filters are based upon opamp circuits, for which transfer functions can be found in analytic form. This requires a familiarity with circuit analysis and the formulae are quite involved. An approach more suited to the toolbox of computer musicians is to find the complex frequency response of the circuit by AC analysis in SPICE [1], a simulator that analyzes circuits given a description of the schematic. Then frequency domain system identification [3] is used to find the digital filter coefficients. Practically, this is done with `invfreqz.m` in MATLAB (Mathworks, Cambridge, MA).

Both the input anti-aliasing and the output equalization filters were designed this way. The fitted filters are plotted with the transfer function from SPICE in Fig. 2. An oversampled rate, f_s=96 kHz greatly improved the fitting for the input filters. The output filters fit well using f_s = 48 kHz.

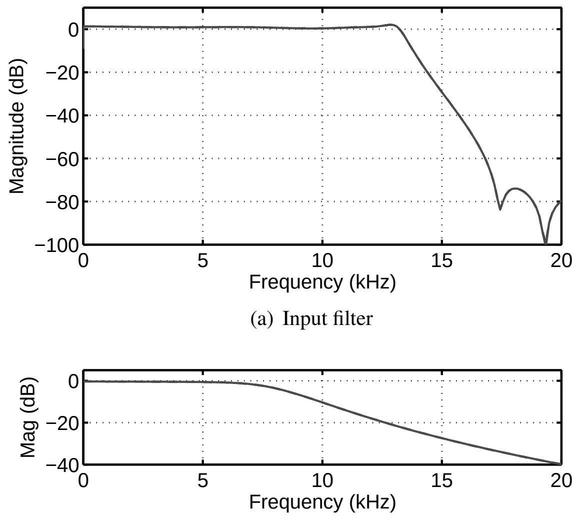
3.3. Resampling at 27.5 kHz with correct aliasing

To simulate aliasing accurately using a digital implementation, the discrete-time signal is ideally interpolated to the time-grid corresponding to the sampling rate of the SP-12.

Methods to approximate the ideal interpolation include using a variable delay filter [2], or resampling[5] to a multiple of the SP-12's sampling rate and then downsampling to the SP-12 rate.

Linear interpolation is found to generate significant spurious artifacts when viewed on a sine sweep. In particular, it has excessive sidelobe levels.

[1] http://bwrc.eecs.berkeley.edu/Classes/IcBook/SPICE/



(a) Input filter

(b) Output filter

Figure 2. Magnitude response of input and output anti-aliasing filters. Shown on linear scale to demonstrate good high frequency match. Input filter is order 11, f_s=96 kHz. Output filter is order 5, f_s=48 kHz. Fitted filter and SPICE filter are perfectly overlaid

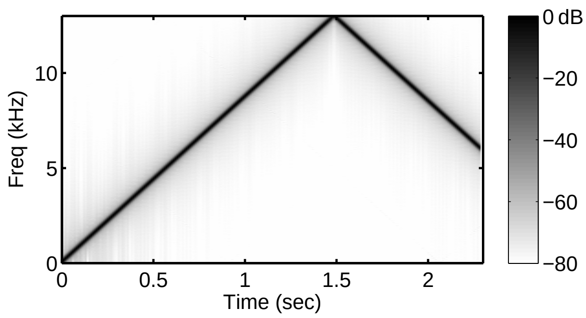


Figure 3. Linear sine sweep from 0 to 20 kHz of the resampling process shows proper aliasing behavior.

Piecewise spline produces much cleaner results, but requires four times oversampling to push unwanted aliasing at high input frequencies into the noise floor. This painstaking process is necessary because the SP-12 is used as a drum machine and sounds with high frequency content such as cymbal crashes are often sampled.

These methods can be viewed as resampling using a particular type of interpolation filter[5] and then downsampling. In general, resampling can be done with better or longer filters to improve interpolation accuracy for a particular sampling rate. This is found to be more efficient in practice. In this implementation, the upsampled signal at 96 kHz is resampled to twice the SP-12 frequency and downsampled by two to generate the correct aliasing. The response of this process to a linear sine sweep is shown in Fig. 3.

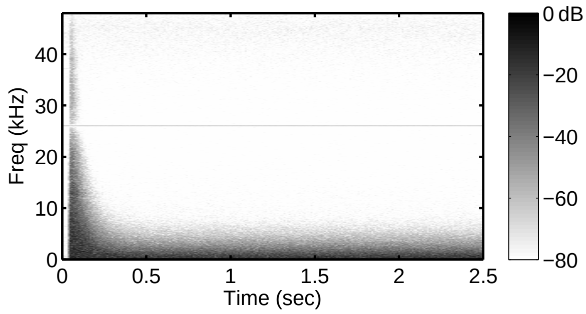


Figure 4. Response of the VCF channel to white noise.

3.4. Zero-order hold

Digital to analog conversion typically involves a zero-order hold (ZOH), which is equivalent to convolving the discrete signal with a rectangular pulse that is one period wide and delayed by a half period[4]. A digital ZOH is accomplished by repeating each sample N times. Its frequency response is an aliased sinc[6]. However, the inherent oversampling by N causes the part of the frequency response that deviates from that of an analog ZOH to be outside the audio band. It is found that a ZOH with $N = 4$ has negligible error in the frequency response between 0 and 20 kHz. Resampling to an audio rate, f_s=48 kHz, eliminates the errors outside the audio range.

3.5. Drum tuning

Sine sweeps and single sine tone inputs to the SP-12 reveal that the tuning settings are in increments of half steps on the chromatic scale. Tuning the sample changes its length. Certain tuning intervals were found to introduce heavy aliasing. It was surmised that the tuning was done by reading the table at rates proportional to the tuning interval with no interpolation.

To shift the pitch up, the table is read with a fractional increment to the index greater than one. In this implementation, the fractional part is truncated when indexing the wave table. To shift the pitch down, the table is read with a fractional increment less than one. This implies that a sample may be repeated several times.

If the tuning ratio is irrational, the irregular skips in reading the table cause significant and complicated aliasing.

3.6. Voltage Controlled Filter

The output channels featuring the VCF 4-pole lowpass can be analyzed by exciting the SP-12 with white noise (Fig. 4). This excitation reveals that the function of the VCF is to change the cutoff frequency of the lowpass according to a programmed schedule triggered by the start of the sample. The schematic indicates that the bandwidth (Q) control of the filter is not varied, while the cutoff frequency follows an approximately exponential trajectory.

This VCF can be implemented digitally in the same manner as the Moog VCF, also a 4-pole filter.

4. RESULTS: COMPARISON WITH SP-12

The results from the actual SP-12 are compared with the results from the model implemented in Matlab. An exponential sine sweep from 20 Hz to 20 kHz lasting 1.5 sec is used as an excitation for all plots shown. The output of the SP-12 is recorded at 96 kHz into a computer audio interface.

Notice the tone in Fig. 5(a) at 26 kHz, indicating that f_s is 26 kHz, not 27.5 kHz as in the specifications. The algorithm has been modified accordingly. The anti-aliasing filter attenuates the response above 15 kHz; therefore, aliasing due to sampling is a minor contribution to the sound. However, the zero-order-hold for the output channel with no anti-imaging filter contributes aliases (spectral images) seen above the 13-kHz Nyquist frequency of the SP-12. Quantization shows up as harmonic distortion to the sinusoidal input, which is very faint in the figure because, at 72 dB, it is just above the floor of the 80-dB dynamic range shown. The model aliases and cuts off at the correct frequencies within the audio band, 0 – 20 kHz.

Figs. 6 and 7 compare the results of setting the tuning control to the highest and lowest pitches. It is demonstrated that the tuning affects the length of the sample. When the ratio of frequency to the original frequency is rational, simple aliasing patterns are created. These plots demonstrate the validity of the tuning algorithm.

5. CONCLUSIONS

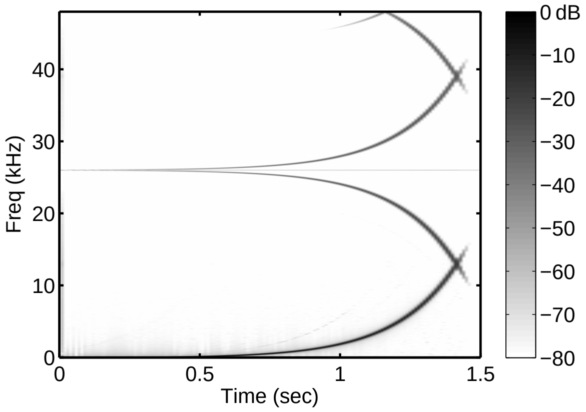
It is shown that the primary sonic characteristic of the SP-12 is aliasing due to its poor output filters and its rudimentary tuning algorithm. The SP-12 sampler was modeled using SPICE as an analytical tool and system identification as a filter synthesis tool. Appropriate tests reveal the behavior of the tuning algorithm and the VCF, which are not fully described by the schematic. Implementing a system such as a sampler, which produces aliasing as a feature, requires careful consideration in the design of the resampling algorithms. Modeling the circuit using this procedure creates algorithms that closely reproduce the behavior of the original system.

6. ACKNOWLEDGMENTS

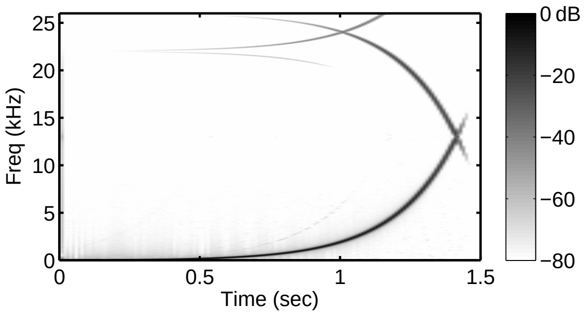
David Yeh is supported by the Stanford, NDSEG and NSF graduate fellowships. Thanks to Jonathan Abel for sleuthing the method of tuning.

7. REFERENCES

[1] A. Huovilainen. Nonlinear Digital Implementation of the Moog Ladder Filter. In *Proc. Int. Conf. Digital Audio Effects (DAFx-04)*, Naples, Italy, Oct. 5–8, 2004.

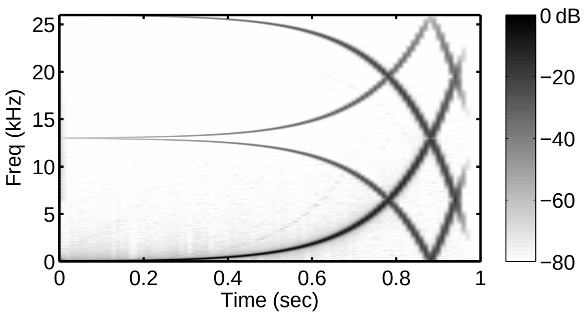


(a) SP-12

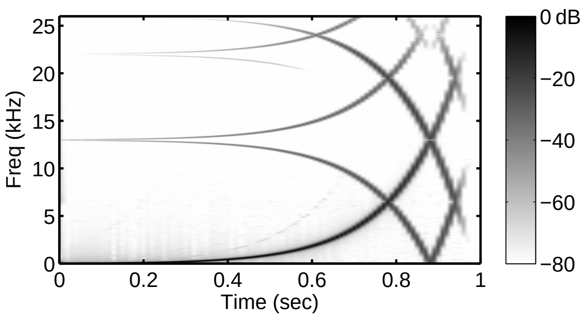


(b) Model

Figure 5. Exponential sweep (20 Hz – 20 kHz) of (a) SP-12 and (b) Model. Model is accurate to 20 kHz.

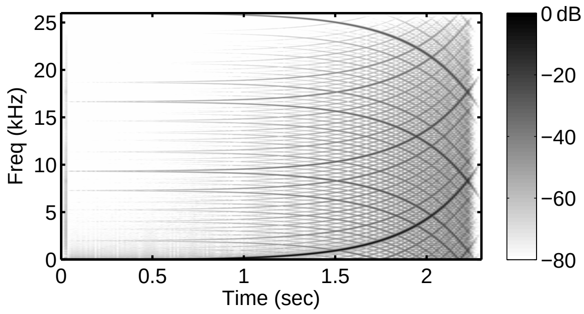


(a) SP-12

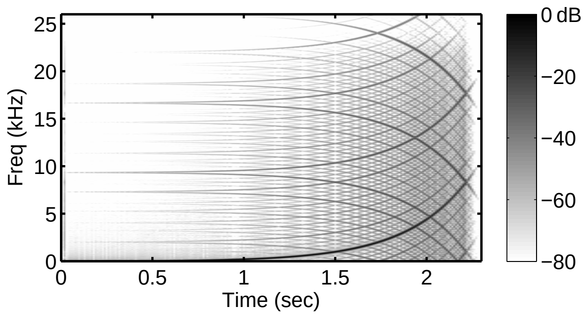


(b) Model

Figure 6. Exponential sweep (20 Hz – 20 kHz). Tuned to highest pitch, $2/3$ times original length of 1.5 sec.



(a) SP-12



(b) Model

Figure 7. Exponential sweep (20 Hz – 20 kHz). Tuned to lowest pitch, 1.56 times original length of 1.5 sec.

[2] T. I. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine. Splitting the unit delay. *IEEE Sig. Proc. Mag.*, 13(1):30–60, 1996.

[3] L. Ljung. Some results on identifying linear systems using frequency domain data. In *Proc. 32nd IEEE Conf. Dec. Contr.*, pages 3534–3538, San Antonio, USA, 1993.

[4] A. Oppenheim and R. Schafer. *Discrete-time Signal Processing*. Prentice Hall, 1999.

[5] J. O. Smith. *Digital Audio Resampling Home Page*. http://www-ccrma.stanford.edu/˜jos/resample/, Nov. 11, 2006.

[6] J. O. Smith. *Spectral Audio Signal Processing, March 2007 Version*. Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, Mar. 2007.

[7] T. Stilson and J. Smith. Alias-free digital synthesis of classic analog waveforms. In *Proc. Int. Comput. Music Conf. (ICMC-96)*, San Francisco, USA, pages 332–335, 1996.

[8] T. Stilson and J. Smith. Analyzing the Moog VCF with considerations for digital implementation. In *Proc. Int. Comput. Music Conf. (ICMC-96)*, San Francisco, USA, pages 398–401, 1996.

[9] V. Välimäki and A. Huovilainen. Oscillator and filter algorithms for virtual analog synthesis. *Computer Music Journal*, 30(2):19–31, 2006.