

A SINGLE-CHIP 1,200 SINUSOID REAL-TIME GENERATOR FOR ADDITIVE SYNTHESIS OF MUSICAL SIGNALS

Fernando De Bernardinis[♦], Roberto Roncella[♦], Roberto Saletti[♦], Pierangelo Terreni[♦], Graziano Bertini[♦]

[♦]Dipartimento Ingegneria dell'Informazione: Elettronica, Informatica e Telecomunicazioni
Università di Pisa - Via Diotisalvi, 2 56126 Pisa, Italy

[♦]Istituto Elaborazione dell'Informazione IEI-CNR Via S. Maria, 46 56126 Pisa, Italy
roncella@iet.unipi.it

ABSTRACT

This paper presents a new hardware implementation of additive synthesis for high quality musical sound generation. The single-chip configuration is capable of performing 1,200 sinusoid real-time synthesis; the system is expandable to 13,200 partials by series connecting 11 chips. Each sinusoid is generated by a marginally stable second order IIR filter, and its frequency, amplitude and phase can be independently specified. The system is clocked at 60 MHz when working with a 44.1 kHz sampling rate. Two completely independent channels are available as output, and each sample relies on a 20 bit representation to achieve an SNR of at least 110 dB, thanks to the internal 24 bit word length. The IC is designed in a 0.5 μ m CMOS technology and has a core area of approximately 19 mm².

1. INTRODUCTION

Additive synthesis is one of the most powerful techniques available to perform sound synthesis, probably the most general purpose one [1]. The backing theory is quite simple and intuitive: a combination of several sinusoids, of various frequencies and amplitudes, is used to approximate the exact sound, somehow resembling a Fourier series expansion of the signal. However, the number of sinusoids is finite and they do not need to be harmonic, i.e. their frequencies may be arbitrarily set. Though extremely interesting and powerful, this technique is not very diffused yet, because of the difficulty of efficient real-time implementations. There are only few systems performing additive synthesis, but they are either not very powerful or very expensive: in fact, the maximum performance reported is 8,000 sinusoids, obtained with 100 DSPs [2]. A transputer T800 network has been proposed, but too many transputers are needed at a

sampling frequency of 44.1 kHz, since each of them can synthesize only 5 harmonics [3]. Recently an ASIC [4] to perform additive synthesis has been presented, but it can manage only 127 16 bit harmonics whose samples are interpolated from a tiny look-up table (LUT), resulting in a not negligible harmonic distortion.

This paper presents an implementation of the additive synthesis that overcomes all the previous ones, since it is capable of single-chip real-time generation of 1,200 partials with outstanding performance. Each sinusoid is generated by a second order IIR filter [5], that is realized directly in hardware and whose arithmetic is based on 24 bit internal notation. These features provide several advantages: the sinusoids are extremely pure and are not interpolated like in the case of finite-length look-up tables [6]; the absence of a large (and necessarily fast) LUT simplifies the global system design; the precision adopted provides an outstanding SNR and the possibility of setting the frequency of the harmonics very accurately. Synthesis parameters (frequency, phase and amplitude of the sinusoids) are changed every 128 samples (time frame), providing a very good response to abrupt changes in the sound characteristics, such as during the attack phase; furthermore, the amplitude of each sinusoid can be linearly changed between consecutive samples, providing an excellent smoothness to the sound produced.

The chip operation is managed by a controller (actually a DSP), which provides the new parameters in the updating phase and can control a multi-chip configuration that synthesizes up to 13,200 sinusoids. Since the chip-controller interface protocol is sensitive only to the number of partials being generated, independently of the number of chips which realize the synthesis, it is possible to obtain a simply upgradable system just providing sockets for expansion chips, that are plugged in only when necessary; with 11 chips it is possible to reach the maximum controller capacity of 13,200 sinusoids.

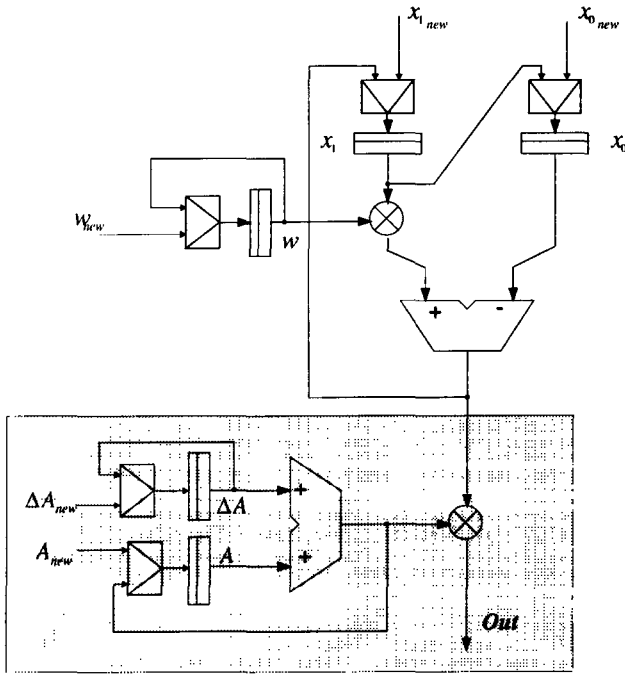


Fig. 1 - The implementation of the IIR filter generating the elemental sinusoid and of the amplitude interpolation section (shadowed). New input data are provided by an external controller at every time frame edge.

2. SINUSOID GENERATION ALGORITHM

Sinusoids are generated by means of a marginally stable second order IIR filter, the behavior of which is described in terms of state variables by the following system

$$\begin{pmatrix} x_0(k+1) \\ x_1(k+1) \end{pmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 2\cos\Omega \end{bmatrix} \begin{pmatrix} x_0(k) \\ x_1(k) \end{pmatrix} \quad (1)$$

where $\Omega = 2\pi T_s f$, f is the sinusoid frequency, T_s the sampling period and k a discrete time index.

From the above expression we have

$$X_1(z) = \frac{x_1(0) \cdot z^2 - x_0(0) \cdot z}{z^2 - 2\cos\Omega \cdot z + 1}$$

that is equal to the Z-transform

$$\cos(\Omega k + \varphi) \Leftrightarrow \frac{\cos\varphi \cdot z^2 - \cos(\varphi - \Omega) \cdot z}{z^2 - 2\cos\Omega \cdot z + 1}$$

provided that proper initial conditions x_0 and x_1 are set (φ is the initial phase).

It is thus evident that the filter can be usefully employed as a digital oscillator, where frequency is set by Ω and φ by x_0 and x_1 (amplitude is set to unity at the IIR level). Furthermore, Eq. (1) permits an advantageous

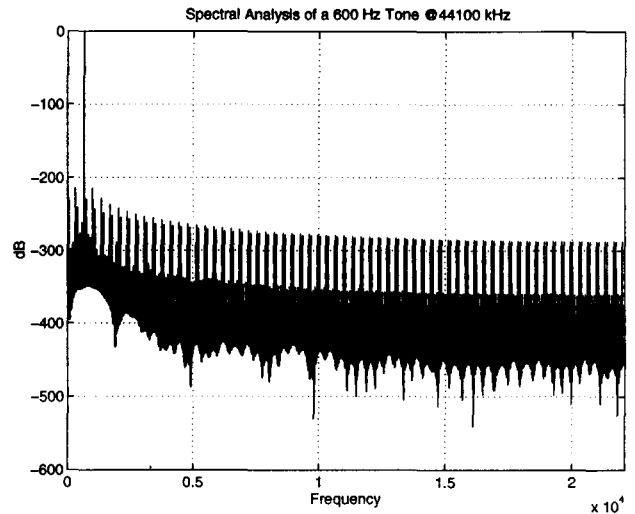


Fig. 2 - Spectral analysis of a 600 Hz sinusoid generated with the IIR filter. An outstanding spectral purity of the signal generated (more than 200 dB over spurious components) is achieved.

hardware implementation since it needs only one multiplier and one adder, as shown in the schematic reported in Fig. 1. In order to improve the filter capabilities, we added an envelope interpolation section to the basic IIR: this is obtained by specifying, along with the amplitude value A (16 bit), an interpolating factor ΔA (8 bit, so that A and ΔA fit in a 24 bit word, like all other parameters) which adds to the former at every new sample. This linear variation of the sinusoid amplitude realizes a piecewise linear approximation of the partial envelope, which is very important to smooth the output at the updating points and to reduce the granularity of the sound produced.

The signal generated can be expressed in the following way

$$y(kT_s) = \sum_{n=1}^{1200} (A_n + \Delta A_n \cdot k) \cos(\Omega_n \cdot k + \varphi_n)$$

in which parameters have a limited validity in time, since they are updated at each frame edge. We found out that a frame length of 128 (that corresponds to 2.9 ms at 44.1 kHz sampling rate) is an acceptable compromise between the smoothness of the sound and the updating frequency.

The most important parameter in the synthesis is $w = 2\cos(\Omega)$, which sets the frequency of the sinusoid. Some studies in psychoacoustic report an ear frequency discrimination threshold of 0.05 Hz @ 160 Hz [7]. Thus w must be expressed with 24 bit to achieve a good frequency resolution. x_0 and x_1 represent the state of the filter (x_1 is the output) and are subject to finite arithmetic errors, that for a marginally stable system become quite

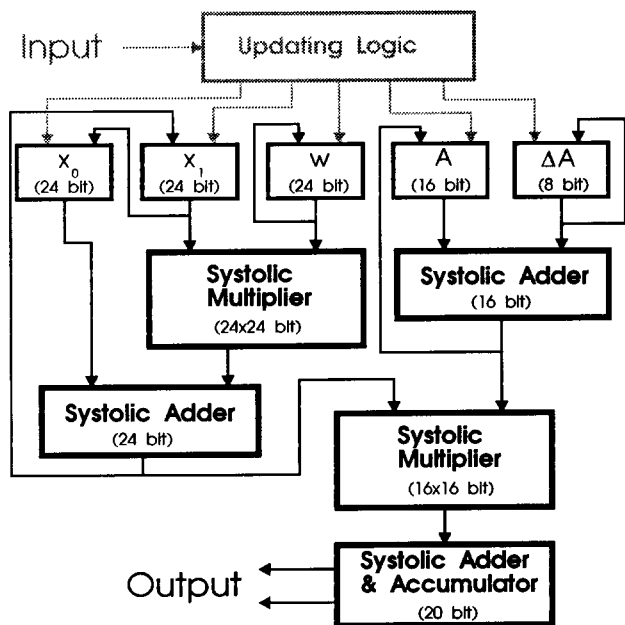


Fig. 3 - Schematic diagram of the systolic structure adopted for the chip.

relevant. This implies a 24 bit internal arithmetic that, together with the frame length, leads to an SNR at the IIR filter output of at least 120 dB and to an outstanding spectral purity (Fig. 2). The amplitude of each partial is rounded to 16 bit from the output (24 bit) of the IIR filter, while the output accumulator width is 20 bit (Fig. 3), as much as those required by the most precise DACs.

3. CHIP ARCHITECTURE

The expression (1) is very suitable to a VLSI implementation and the structure shown in Fig. 1 demonstrates an advantageous realization in terms of hardware costs. Since the generation of thousands of sinusoids can be accomplished by serially processing state register vectors on parallel, unitary throughput arithmetic blocks, we opted for a systolic approach [8]. This also satisfies the speed requirements dictated by the parallel multiplier with unitary throughput. Furthermore the whole structure can be systolic, thus avoiding large (24 bit) skewers/deskewers. Since there are only few global communications, the locality feature of this kind of architecture is well respected. The chip structure shown in Fig. 3 is analogous to that in Fig. 1 once all registers have been substituted by arrays of registers. It can take full advantage of all the positive aspects typical of this kind of solution; its principal blocks are the two systolic bit-parallel multipliers, the three systolic adders and the chain of registers forming loops around the former blocks.

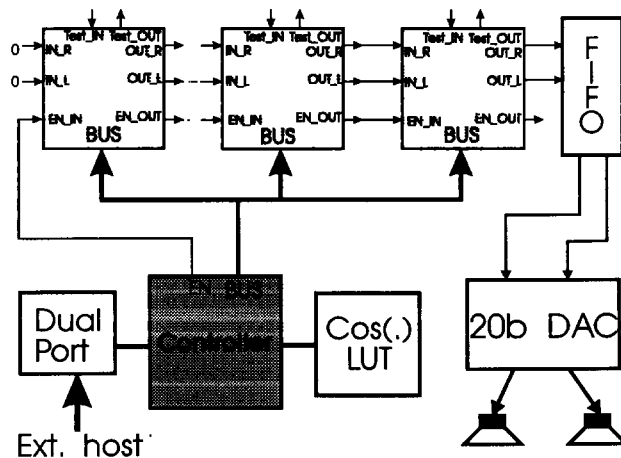


Fig. 4 - Multi-chip system configuration for 3,600 partial additive synthesis

The systolic blocks have been designed to afford extremely heavy duties and are very compact thanks to their rectangular shape. In fact most chip area is occupied by the storage elements that feed the arithmetic units.

The chip complexity is quite high, so that we inserted a testing path [9] throughout the arithmetic blocks to simplify testing. This path works in a simple but effective way increasing the observability of the system. It provides a scan-path like serial stream without interfering with the chip operation at the nominal clocking frequency, since it is realized with systolic blocks like all other sub-systems. The stream returns all the arithmetic blocks input/outputs in a systolic format. This provides a good degree of observability, while the controllability is intrinsic to the architecture used, since it is possible to set any registers to the desired values during the updating phase.

4. SYSTEM CONSIDERATIONS

Even though its performance is really remarkable, the chip has been designed to support multi-chip operation, to expand the system performance up to 13,200 partials. This is accomplished by both synchronizing chips automatically and realizing an adequate parameter updating procedure. In order to satisfy these requirements, the chip works in two different phases: synthesis and updating. In fact this leads to parameter consistency at each sample computation, and allows a simple system expandability. Since all chips share the same bus and synchronize with each other, the controller has only to manage the global number of sinusoids without knowing how many chips are present in the system (a 3,600 sinusoid system is

shown in Fig. 4). The first consequence of this choice is that samples are not generated uniformly in the frame (2.9 ms at 44.1 kHz), but there are 2.63 ms for the actual synthesis, and 0.27 ms for the updating. It follows that a FIFO buffer must be connected to the system output to get the samples exactly once every T_s . The second consequence is that the maximum number of controllable chips is limited to 11, as already stated, since this is the maximum number of updating cycles that fit in 2.9 ms. This updating period permits to have the smallest latency compatible with the system complexity, that is very important for live music performances [10].

The updating procedure is simple enough to be implemented on a general purpose DSP provided with a dual port RAM, where frequencies and amplitudes can be written by an external host. The two level structure relieves the host from providing the parameters every 128 samples, since each sinusoid frequency and amplitude are hold until new values are written in the dual port RAM. This means that the 2.9 ms is not but the shortest frame obtainable with the system.

Since the controller operations are simple and can be accomplished concurrently, a more powerful system could be obtained realizing a dedicated controller. This new system could provide initial phase setting, arbitrary length frames and automatic frequency and amplitude interpolation. In this way it would be possible to realize an abstract additive synthesis machine, interfaceable with any synthesis algorithm developed for additive synthesis.

5. CONCLUSIONS

The chip has been designed in a 0.5 μm three metal layer CMOS process. It works at 60 MHz when the sampling frequency is set to 44.1 kHz and has a core area of only 19 mm^2 . Since the computing power of the systolic blocks is not wholly utilized, more powerful versions can be obtained by simply increasing the parameter memory and the working frequency. For example, with a working frequency of 100+120 MHz, it would be possible to synthesize more than 2,000 sinusoids in a single chip with an estimated area of about 28 mm^2 .

The purpose of this work was to realize a chip performing additive synthesis powerful enough to make this technique effective and cost-affordable. The solution proposed is very simple and powerful: the system is easily expandable and performs better than any other comparable system known to the authors. The synthesis characteristics provide an excellent sound generation, and a multi-chip system in its full configuration is able to produce up to 13,200 sinusoids at a sampling frequency of 44.1 or 48 kHz (the last one, or any other one, is obtained

simply increasing, or decreasing, the system clock frequency).

The system has been designed for high quality musical syntheses, but it can easily be used in any field involving audio-range synthesis, such as speech synthesis [11], or in any system requiring spectral determination of a (audio) signal, with some hints to adaptive noise cancellation [12].

References

- [1] G. De Poli, "A Tutorial on Digital Sound Synthesis Techniques", *Computer Music Journal*, 7:4, pp. 20-23, Winter 1983
- [2] Cor Jansen, "Sine Circuitu: 10,000 high quality sine waves without detours", *ICMC Proc.* 1991, pp. 222-225
- [3] T. Itagaki, A. Purvis, P. D. Manning, "Real-time synthesis on a multiprocessor network", *ICMC Proc.* 1994, pp. 382-385
- [4] A. D. Houghton, A. J. Fisher and T. F. Malet, "An ASIC for Digital Additive Sine-wave Synthesis", *Computer Music Journal*, 19:3, pp. 26-31, Fall 1995
- [5] J. W. Gordon, J. O. Smith, "A Sine Generation Algorithm for VLSI Applications", *ICMC Proc.* 1985, pp. 165-168
- [6] Moore, "Table Look-up Noise for Sinusoidal Digital Oscillator", *Computer Music Journal*, 1:2, pp. 26-29, Spring 1977
- [7] A. Rakowski, "Pitch discrimination at the threshold of hearing", *Proceedings of the Seventh International Congress on Acoustics*, Vol. 3, 1971
- [8] R. Roncella, R. Saletti, "A VLSI Systolic Adder for Digital Filtering of Delta-Modulated Signals", *IEEE Trans. on Acoustic, Speech and Signal Processing*, pp.749-754, Vol. ASSP-37, No. 5, May 1989
- [9] N. H. E. Weste, K. Eshraghian, "Principles of CMOS VLSI design", Addison Wesley, 1993
- [10] D. A. Jaffe, "Ten Criteria for Evaluating Synthesis Techniques", *Computer Music Journal*, 19:1, pp. 76-87, Spring 1995
- [11] R. J. McAulay, T. F. Quateri, "Speech Analysis/Synthesis based on a Sinusoidal Representation", *IEEE Trans. on Acoustic, Speech and Signal Processing*, pp. 744-754, Vol. ASSP-34, N. 4, 1986
- [12] S. J. Elliott, P. A. Nelson, "Active Noise Control", *IEEE Signal Processing Magazine*, pp. 66-78, October 1993