Musical instrument synthesis by nonregenerative nonlinear processing

James W. Beauchamp

School of Music and Department of Electrical and Computer Engineering,
University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

Abstract: With the method of nonlinear distortion described in this paper a sine wave is progressively distorted by a nonlinear function to produce a series of harmonics whose relative amplitudes can be controlled in two ways: First, when the amplitude of the sine is at a "target value", a target output spectrum is produced. Second, as the amplitude of the sine wave increases from zero to the target value and beyond, the bandwidth (or centroid) of the output spectrum increases. When the target spectrum is chosen to be representative of a musical instrument, this behavior can be used to mimic the time-variant spectral behavior of that instrument. A high pass filter approximating the radiation characteristic of the instrument is used to improve the spectral evolution of the partials during the attack and decay of a sound.

NONLINEAR PROCESSOR DESIGN

When a sine wave of amplitude $a$ is applied to a distorting device, harmonics are produced. For the special case where the distortion can be described by a polynomial function, i.e.,

$$F(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

and $x(t) = a \cos(\omega t)$, the output can be expressed in terms of harmonics as

$$y(t) = F(a \cos(\omega t)) = A_0(a) + A_1(a) \cos(\omega_1 t) + A_2(a) \cos(2\omega_1 t) + \ldots + A_n(a) \cos(n\omega_1 t)$$

where each $A_k(a)$ is a function which depends on the sine wave amplitude $a$.

Using Tchebycheff polynomials $T_k(x)$ having the property that $T_k(\cos(\omega_1 t)) = \cos(k\omega_1 t)$, we can construct a polynomial

$$F(x) = A_0(1) T_0(x) + A_1(1) T_1(x) + A_2(1) T_2(x) + \ldots + A_n(1) T_n(x)$$

such that Equation 3 reduces to Equation 2 when $a = 1$. $\{A_k(1)\}$ is the specified or "target" spectrum.

Collecting like power terms in Equation 3 allows the calculation of the polynomial coefficients from the target amplitudes:

$$a_k = 2^{k-1} \sum_{i=0}^{k} \binom{k}{i} (-1)^i \frac{(k+2i)!}{i!k!} A_{k+2i}(1)$$

So, if a unit amplitude sinusoid is applied to this $F(x)$, the output spectrum will be precisely $\{A_k(1)\}$.

CONTROL OF THE TIME-VARIANT SPECTRUM

Whereas the nonlinear processor function $F(x)$ will produce a specific target spectrum for $a = 1$, we would like to know how the harmonic amplitudes vary with $a$ as implied by Equation 2. Using well known trigonometric identities for $\cos^k(\omega_1 t)$ we can predict the amplitudes in terms of $a$ and the polynomial coefficients:

$$A_k(a) = 2^{k-1} \sum_{i=0}^{k} \binom{k}{i} \frac{(k+2i)!}{i!(k+i)!} (-\alpha/2)^{k+2i} a_{k+2i}$$
Substituting Equation 4 into Equation 5 we have

\[ A_k(\alpha) = \left( \sum_{i=0}^{K-k} \frac{\alpha^{k+2i}}{i!(k+i)!} \left( \sum_{j=0}^{i} (-1)^j \frac{(k+2i+j)!}{j!} \right) \right) A_{k+2i+2j}(1) \]  

Equation 6 shows the dependence on \( \alpha \): If \( k \) is odd (even), \( A_k(\alpha) \) depends on only odd (even) powers of \( \alpha \), starting with \( k \). Rearranging, gives us

\[ A_k(\alpha) = \left( \sum_{i=0}^{K-k} \frac{\alpha^{k+2i}}{i!(k+i)!} \left( \sum_{j=0}^{i} (-1)^j \frac{(k+2i+j)!}{j!} \right) \right) A_{k+2i+2j}(1) \]

Equation 7 shows that if \( k \) is odd (even), \( A_k(\alpha) \) depends on only odd (even) target harmonic amplitudes with numbers equal to \( k \) and above up to \( K \) or \( K-1 \). Thus, the odd and even terms are completely independent.

It is highly desirable that the relations between \( A_k(\alpha) \) and \( \alpha \) be monotonic. Unfortunately, this is not guaranteed, because of the minus signs in Equations 6 and 7. However, we find that this is much more probable if the amplitudes roll off as \( k \) increases, although we have not yet discovered the minimum roll-off required.

**MATCHING MUSICAL INSTRUMENT TIME-VARYING SPECTRA**

Matching an original sound’s time-varying spectrum (TVS) [1,2] is accomplished by first selecting a target spectrum from the sound’s TVS. Then, this spectrum is divided by the radiation response of the instrument (this is the response of a simple digital filter which may be approximated from measurements on the instrument or just estimated) to yield the \( \{A_k(1)\} \). Since the radiation response is a high pass function, this generally yields a target spectrum that rolls off quite well. To find a good variation of \( \alpha \) we attempt to match the normalized spectral centroid (NSC) of the instrument spectrum with that of the nonlinear processor. We do this by first constructing the NSC vs. \( \alpha \) function for the nonlinear process and find the range (0 to \( \alpha_{max} \)) over which this relation is monotonic. We then invert this function so that NSC can be used to predict a corresponding value of \( \alpha \). Note that \( \alpha_{max} \) corresponds to NSC_{max} and that the former should be larger than 1; also, NSC_{max} should be equal to or larger than the max NSC of the instrument spectrum. Then, for synthesis, measured values of NSC from the instrument spectrum are used to predict \( \alpha \) values on an instantaneous basis, using the NSC vs. \( \alpha \) function. The amplitude output of the nonlinear processor can also be predicted, and this is corrected by multiplying the output by the factor \( \beta \). Thus, in addition to the function \( F(x) \), the radiation filter, and the \( \alpha \) vs. NSC function, functions for NSC vs. time and \( \beta \) vs. time are needed. The actual synthesis proceeds at each time instant by 1) using NSC to predict \( \alpha \), 2) using \( \alpha \) to control the amplitude of a sine wave (at the desired frequency), 3) processing that by the nonlinear function \( F(x) \), 4) passing the result through the radiation filter, and 5) multiplying by the \( \beta \) value.

The advantages of this nonlinear/filter approach are that reasonable matches to instrument behaviors can be made over a fairly wide range of dynamics, pitches, and attack/decay while keeping \( F(x) \), the radiation filter, and the \( \alpha \) vs. NSC function fixed. NSC vs. time and \( \beta \) vs. time functions can be tailored to the specific articulation style and durations desired. This method has proved successful for such instruments as cornet, saxophone, clarinet, and piano.

It is possible to improve the accuracy of matching by parallel combination of more than one nonlinear processor each with separate \( F(x) \) nonlinearities and separate \( \alpha \) and \( \beta \) controls [3]. These functions are found by an iterative process of subtracting the TVS of the first nonlinear output from the actual instrument TVS and then approximating the result with a nonlinear module. This process can be repeated until suitable matching has been obtained.

**REFERENCES**