

PREVIEW

Victor Oh

15 de mayo de 2020

Capítulo 1

Síntesis de Sonidos mediante Modelos Físicos

1.1. Introducción

Se utilizó el modelo de Karplus-Strong para sintetizar el sonido de instrumentos de cuerda percutida u otros tipos de percusión. Este algoritmo, creado por Kevin Karplus y Alexander Strong en 1983 para sintetizar sonidos con pocos recursos y a tiempo real.

En este trabajo se analizaron el modelo básico para la síntesis de cuerdas percutidas y el modelo modificado para la síntesis de instrumentos de percusión.

1.2. Modelo Conceptual

En principio, el sistema se trata de un sistema lineal excitado con una secuencia aleatoria de longitud finita L . El sistema consiste de una línea de retardo de L muestras realimentadas a través de un filtro, tal como se ve en la figura 1.1.

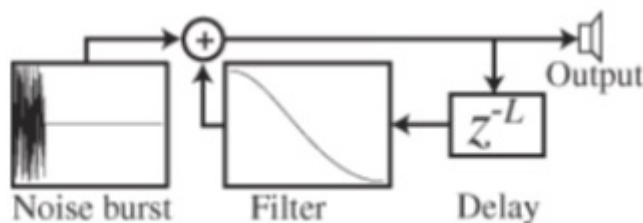


Figura 1.1: Diagrama Conceptual del Modelo de Karplus-Strong

La línea de retardo simula la cuerda, la cual los distintos armónicos de una nota recorren, se atenúan y decaen en el tiempo. En el modelo principal que se plantea en trabajo original de Kevin Karplus y Alex Strong, a la ausencia de algún filtro a la salida de la línea de retardo, la salida del sistema sería una repetición periódica de la señal entrante, lo que generaría un tono puro con una frecuencia de f_s/L , donde f_s es la frecuencia de sampleo del sistema.

La modificación que inventa Alex Strong es promediar dos muestras consecutivas del sistema, generando un decaimiento lento de la amplitud en el tiempo. Matemáticamente, esta operación puede expresarse como en la expresión (1.1).

$$Y_t = 0,5 \times (Y_{t-L} + Y_{t-L-1}) \quad (1.1)$$

Como consecuencia de esta modificación, la frecuencia del tono generado estará dada por la expresión (1.2). Este tono tiene un sonido similar al de una cuerda percutida. Independientemente de las condiciones iniciales del sistema, el espectro decaerá hacia un tono puro, y eventualmente, a un valor constante (el silencio)

$$f = \frac{f_s}{L + 1/2} \quad (1.2)$$

Dado que para producir un sonido realista, se desea que el impulso inicial del sistema contenga una gran variedad de armónicos. Para ello existen distintas opciones. La primera es la propuesta por el mismo trabajo de Karplus-Strong: una señal aleatoria bi-nivel cuya expresión matemática está dada por (1.3), donde A es la amplitud de la nota. Otras opciones para la excitación inicial con ruido es utilizar señales de ruido con distribución uniforme o gaussiana. Como estas muestras iniciales son repetidas periódicamente, no generarán ruido en el fondo.

$$Y_t = \begin{cases} +A & \text{probabilidad 0.5} \\ -A & \text{probabilidad 0.5} \end{cases} \quad \text{para } -L \leq t \leq 0 \quad (1.3)$$

Finalmente, el modelo utilizado para simular la cuerda percutida será el de la figura 1.2. $x(n)$ será la señal inicial de entrada, por donde entrará la señal de ruido de longitud de muestras L .

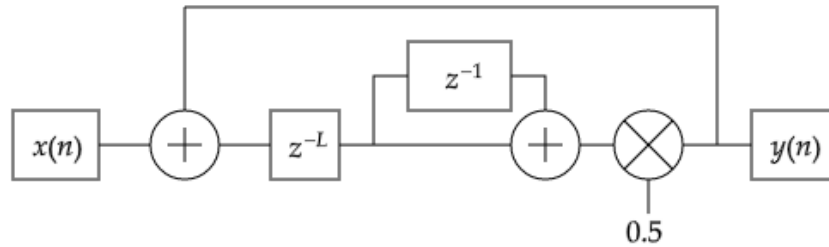


Figura 1.2: Modelo Original de Karplus Strong

Modificando el modelo también es posible sintetizar instrumentos de percusión. Si se cambia el signo del promedio de las muestras aleatoriamente, como se ve en la figura 1.3, matemáticamente tendrá un comportamiento expresado en (1.4). b es llamado el factor de mezcla. Cuando $b = 1$, el sistema se comporta normalmente y sigue simulando una cuerda. Cuando $b = 0$, la señal es negada cada $L + 0,5$ muestras, generando un sonido parecido a un arpa.

Finalmente, cuando $b = 0,5$ se pierde el "tono" del sonido y este se asemejará más a un sonido percutido. L ya no controla el tono, sino el tiempo de decaimiento. Dado que la naturaleza aleatoria ya se encuentra dentro del sistema, es suficiente con inicializar el sistema con una señal constante en lugar de utilizar ruido.

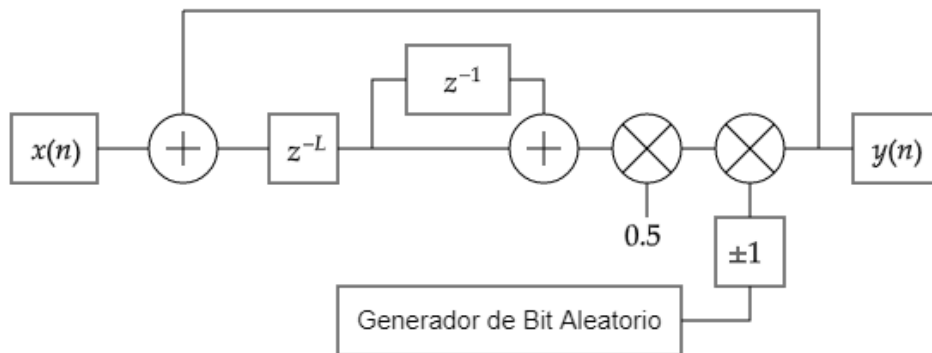


Figura 1.3: Modelo Modificado de Karplus Strong para instrumentos de percusión

$$Y_t = \begin{cases} 0,5 \times (Y_{t-L} + Y_{t-L-1}) & \text{probabilidad } b \\ -0,5 \times (Y_{t-L} + Y_{t-L-1}) & \text{probabilidad } 1 - b \end{cases} \quad (1.4)$$

1.3. Transferencia del Modelo

Se calculó la función transferencia del modelo original de Karplus-Strong. La transferencia del promedio está dada por

$$H_a(z) = \frac{1 + z^{-1}}{2} \quad (1.5)$$

y la transferencia de la línea de retardo está dada por

$$H_b(z) = z^{-L} \quad (1.6)$$

Con esto, se obtiene que la transferencia de todo el sistema está dada por

$$H(z) = \frac{1}{1 - H_a(z)H_b(z)} = \frac{1/2 \cdot (1 + z^{-1})}{1 - 1/2 \cdot z^{-L}(1 + z^{-1})} = \frac{z + 1}{2z^{L+1} - z - 1} \quad (1.7)$$

A partir de la expresión (1.7) se puede calcular que existe un único cero en $z = -1$, que corresponde a la frecuencia de Nyquist $f_s/2$. Los polos pueden ser encontrados calculando las raíces de la expresión $2z^{L+1} - z - 1 = 0$. Estas se pueden aproximar si la expresamos de la forma $2z^{L+1/2} = z^{1/2} + z^{-1/2}$. Reemplazando $z = ae^{j\omega}$ se obtiene:

$$2a^{L+1/2}e^{j\omega(L+1/2)} = a^{1/2}e^{j\omega/2} + a^{-1/2}e^{-j\omega/2} = \sqrt{a + a^{-1} + 2\cos\omega} \cdot e^{j\theta} \quad (1.8)$$

Analizando las partes imaginarias, se obtiene que

$$\sqrt{a + a^{-1} + 2\cos\omega} \sin\theta = (a^{1/2} - a^{-1/2}) \sin \frac{\omega}{2}$$

Aproximando $e^{j\omega(L+1/2)} \approx 1$ se despeja que $\omega = 2\pi n/(L + 1/2)$. Aproximando $2a^{p+1/2} \approx 2\cos(\omega/2)$ se puede despejar la constante del tiempo de caída del n -ésimo armónico, siendo

$$a = \left[\cos\left(\frac{2\pi n}{2L+1}\right) \right]^{1/(L+1/2)} = \left[\cos\left(\pi n \frac{f}{f_s}\right) \right]^{f/f_s} \quad (1.9)$$

Por lo tanto, los polos estarán ubicados en $z = ae^{j\omega}$ donde

$$a = \left[\cos\left(\frac{2\pi n}{2L+1}\right) \right]^{1/(L+1/2)} \quad \omega = \frac{2\pi n}{L + 1/2}$$

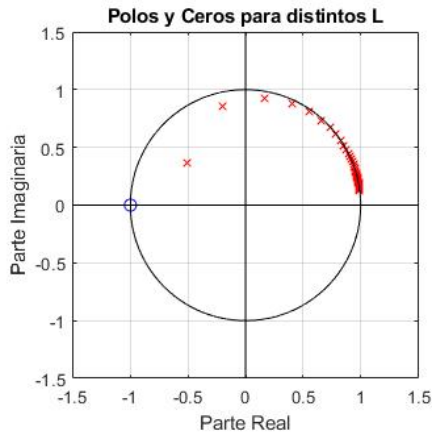


Figura 1.4: Polos y Ceros para distintos L

En la figura 1.4 se puede observar la distribución de los polos y ceros para distintos valores de $L \in [2, \infty)$, recorridos en sentido horario. Esto muestra que a valores pequeños de L , los polos se alejan de la aproximación asumida para su cálculo. Mientras tanto, el cero siempre está ubicado sobre $z = -1$.

La condición para la estabilidad es que, dado que este es un sistema causal y por lo tanto la región de convergencia no puede incluir a $z = 0$, todos los polos deben estar ubicados dentro de la circunferencia de radio $|z| = 1$. Tomando la expresión (1.9) se puede determinar que entonces

$$\cos(x) < 1 \Rightarrow 0 < \pi n/(L + 1/2) < \pi/2$$

Dada la condición de Nyquist, se confirma entonces que este filtro será estable para cualquier longitud de retardo L .

1.4. Afinamiento: selección de L

A partir de la expresión (1.7), la frecuencia de la nota generada estará dada por:

$$f_1 = \frac{f_s}{P_a(z) + P_b(z)} = \frac{f_s}{L + 1/2} \quad (1.10)$$

donde $P_a(z) = 1/2$ es la fase generada por el bloque promediador y $P_b(z) = L$ es la fase generada por el bloque de retardo. Dado que L es un número natural, las frecuencias fundamentales que se pueden generar con este modelo se verán cuantizadas. Sin embargo, como serán inversas a L habrán diferencias en el error de la frecuencia generada dependiendo del tamaño de L . Para L grande, esto no presentará un serio problema, ya que para dos L consecutivos, las frecuencias estarán próximas unas a otras. Sin embargo, para L chicos, las frecuencias estarán demasaido separadas.

Para solucionar este problema, es necesario, para altas frecuencias, poder cambiar el valor de la fase sin afectar la ganancia del filtro. Para esto se puede agregar un filtro pasa-todo de la forma (1.11). Para que sea estable, $|C| < 1$.

$$H_c(z) = \frac{C + z^{-1}}{1 + Cz^{-1}} \quad P_c(z) \approx \frac{1 - C}{1 + C} \quad (1.11)$$

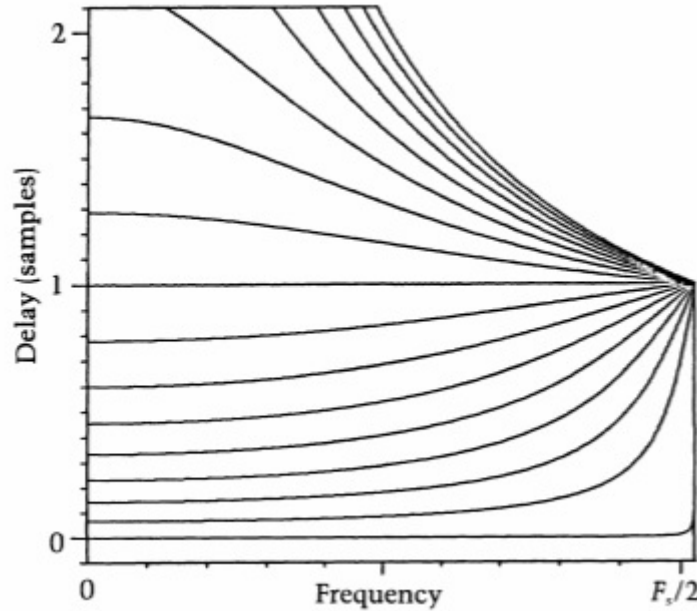


Figura 1.5: Retardo de Fase del filtro Pasa-Todo

Agregando este filtro, la frecuencia de la nota generada estará dada por

$$f_1 = \frac{f_s}{P_a(z) + P_b(z) + P_c(z)}$$

En la figura 1.5 se puede observar cómo para distintos valores de C la fase varía. Dado que la situación de retardo cero es difícil de obtener por errores de redondeo, es preferible que el retardo del filtro esté dado por $\varepsilon \leq P_c \leq 1 + \varepsilon$ para ε chico. No es deseable superar el retardo de 1 dado que este deja de ser constante a altas frecuencias y retardos mayores.

Se puede calcular entonces la fase

$$L = \left\lfloor \frac{f_s}{f_1} - \frac{1}{2} - \varepsilon \right\rfloor$$

$$P_c = f_s/f_1 - N - 0,5$$

Para bajas frecuencias, se puede aproximar

$$C \approx \frac{1 - P_c}{1 + P_c}$$

De esta manera se puede afinar más precisamente la nota que se desea sintetizar.

1.5. Posibles mejoras al modelo

En el trabajo "Extensions of the Karplus Strong Algorithm" se detallan varias maneras en que se puede generar un sonido más convincente.

1.5.1. Control del tiempo de decaimiento

La primera mejora que presenta el trabajo es en el decaimiento de la señal. Esto se debe a que el tiempo de decaimiento de frecuencias bajas es demasiado distinta al de frecuencias altas.

Para acortar este tiempo, se puede agregar un factor de pérdida ρ a la salida del promediador, de manera que el sistema quedaría caracterizado por:

$$y_n = x_n + \rho \frac{y_{n-N} + y_{n-(N+1)}}{2}$$

De esta manera, la amplitud de la envolvente de la señal será proporcional a:

$$\alpha_f(t, \rho) = |\rho \cos(\pi f T_s)|^{f_1 t} = |\rho|^{f_1 t} \alpha_f(t)$$

Y el tiempo de decaimiento quedará descrito por

$$\tau_1(\rho) = -\frac{1}{f_1 \ln |\rho \cos(\pi f_1 T_s)|}$$

Este método no sirve para alargar el decaimiento, dado que para mantener estabilidad $\rho \leq 1$.

Para ello se deberá implementar una modificación para el estiramiento de la nota. Esto se logra modificando la función de promedio por un promedio ponderado (H_a) de manera que:

$$H_a(z, S) = (1 - S) + Sz^{-1}$$

donde $S \in (0, 1)$ es el factor de estiramiento.

La ganancia de este filtro estará dada por

$$G_a(f, S) = \sqrt{(1 - S)^2 + S^2 + 2S(1 - S) \cos \omega T_s}$$

Cuando $S = 0,5$ se tiene el caso de la guitarra normal.

Sin embargo, agregar esta modificación altera la fase, de manera que ahora $P_a(f, S) \approx S$, el cálculo para generar la nota deberá ser $f = f_s/(N + S)$.

1.6. Conclusión

El modelo de Karplus Strong es uno fácil de implementar y altamente efectivo en la síntesis del sonido de una guitarra. Además, existen varias modificaciones menores al modelo que permiten agregar distintos efectos dinámicos y realismo a las notas, y generar timbres de otros instrumentos, como los de percusión.

Capítulo 2

Efectos de Audio

2.1. Flanger

Flangeo es un efecto basado en retardos, cuyo origen se traza a las cintas magnéticas. Dos copias de una misma cinta eran alternadamente retardadas y deceleradas para reducir su frecuencia y crear un pequeño retardo entre ambas. No demasiado, porque de lo contrario la señal producida por la suma resultaría en dos copias distinguibles. En este trabajo se implementó el modelo básico.

Matemáticamente, este se puede describir de la forma

$$y(n) = x(n) + gx(n - M(n)) \quad (2.1)$$

donde $M(n)$ es un retardo variable en el tiempo y $g \in [0, 1]$ es el la profundidad del efecto de flangeo. Su función transferencia estará dada de la forma

$$H(z) = 1 + gz^{-M(n)} \quad H(\omega) = 1 + g \cdot e^{-j\omega M(n)} \quad (2.2)$$

y su respuesta en frecuencia será dada por

$$|H(\omega)| = \sqrt{1 + 2g \cos(\omega M(n)) + g^2} \quad (2.3)$$

Para controlar $M(n)$ se utiliza un modulador de baja frecuencia, de manera que el largo de retardo se encuentre entre $1ms$ y $10ms$.

El retardo estará dado por

$$M(n) = M_0 + \frac{M_w}{2} \left[1 + \sin \left(2\pi \frac{f_{LFO}}{f_s} n \right) \right]$$

En este caso, se eligió una senoidal como señal moduladora por simplicidad en la implementación.

Un último parámetro de este efecto es la inversión de fase, la cual define si se invertirá el signo de g . En algunos casos esto es deseable para no anular la señal continua y otras frecuencias más bajas.

2.2. Chorus

El “Chorus”, es otro efecto que requiere una implementación similar, pero el rango de retardos estará entre $20 - 30ms$. Esto es porque se desea que la señal retardada y la original sean distinguibles. También se utilizan frecuencias más bajas para el modulador.

Se resumieron los parámetros de ambos efectos en la Tabla 2.1

Efecto	Flange	Chorus
Delay	$1 - 10ms$	$20 - 30ms$
f_{LFO}	$0,1 - 10Hz$	$0,1 - 3Hz$

Tabla 2.1: Características de los efectos