

1 Algoritmo de Karplus Strong

1.1 Introducción

El algoritmo de Karplus-Strong fue introducido en 1983 por Kevin Karplus y Alex Strong, y consiste en la aplicación de ciertas técnicas digitales para poder simular sonidos de instrumentos de manera sencilla y con pocos recursos tanto físicos como digitales. En particular en este trabajo práctico, se analizaran los modelos básico (también conocido como punteo de guitarra), y una modificación para lograr reproducir sonidos de instrumentos de percusión.

1.1.1 Modelo conceptual

El modelo conceptual del algoritmo de Karplus-Strong consiste en excitar un sistema lineal con una secuencia aleatoria de longitud finita. El sistema consiste en una línea de retardo de L muestras, realimentada mediante un filtro como podemos observar en la figura 1.

La idea principal es simular como los armónicos que recorre una cuerda de longitud L, se van atenuando a medida que circulan por la misma. Combinados con el filtro pasabajos, logramos que las frecuencias más altas se atenúen aún más rápido y logramos mejorar la calidad del audio que se percibe.

La primera aproximación al diagrama de la figura 1 es consiste en promediar dos muestras consecutivas, lo que produce una atenuación progresiva de la señal entrante al sistema. Matemáticamente se expresa de la siguiente manera:

$$\begin{cases} Y_n = \frac{1}{2} \cdot (Y_{n-p} + Y_{n-p-1}) & \text{con } p \in \mathbb{N} \\ Y_n = \begin{cases} +A & \text{con probabilidad } 0.5 \\ -A & \text{con probabilidad } 0.5 \end{cases} & \text{para } -p \leq t \leq 0 \end{cases}$$

Este algoritmo logra emular perfectamente como el sonido de una cuerda decae en el tiempo. Posee un período de $p + \frac{1}{2}$ muestras ($\frac{f_s}{p+1/2}$). Sin importar la forma que toma la señal ni su contenido armónico, eventualmente siempre tiende a un valor constante (manifestado en “silencio”). No obstante, el sonido resultante va a variar según las condiciones iniciales dadas. Para lograr un mejor efecto, es conveniente que la señal tenga una gran cantidad de armónicos de alta frecuencia. Para ello, Y_n se genera con valores aleatorios de valores $\pm A$, con igual probabilidad, lo que hace que la señal tenga un gran contenido armónico pero además esta libre de ruido.

1.1.2 Variación para instrumentos de percusión

Teniendo el modelo básico desarrollado anteriormente, es posible aplicarle una variante para lograr sintetizar instrumentos de percusión. Esta variante consiste en mezclar los valores de las muestras medios con su recíproco para así obtener el sonido. Matemáticamente, puede expresarse de la siguiente manera:

$$Y_n = \begin{cases} +\frac{1}{2} \cdot (Y_{n-p} + Y_{n-p-1}) & \text{probabilidad } b \\ -\frac{1}{2} \cdot (Y_{n-p} + Y_{n-p-1}) & \text{probabilidad } 1 - b \end{cases}$$

El valor b, conocido como factor de mezcla, es el que determina la proporción de componentes realimentadas positiva y negativamente. Para valores de $b \in [0.1; 0.6]$ se logran sonidos similares a instrumentos de percusión. Notese que si $b=1$, obtenemos el algoritmo clásico del sonido de guitarra.

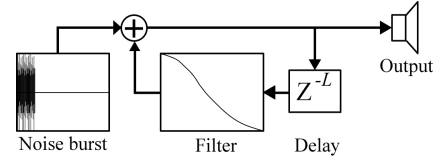


Figura 1: Diagrama conceptual del modelo básico de Karplus-Strong.

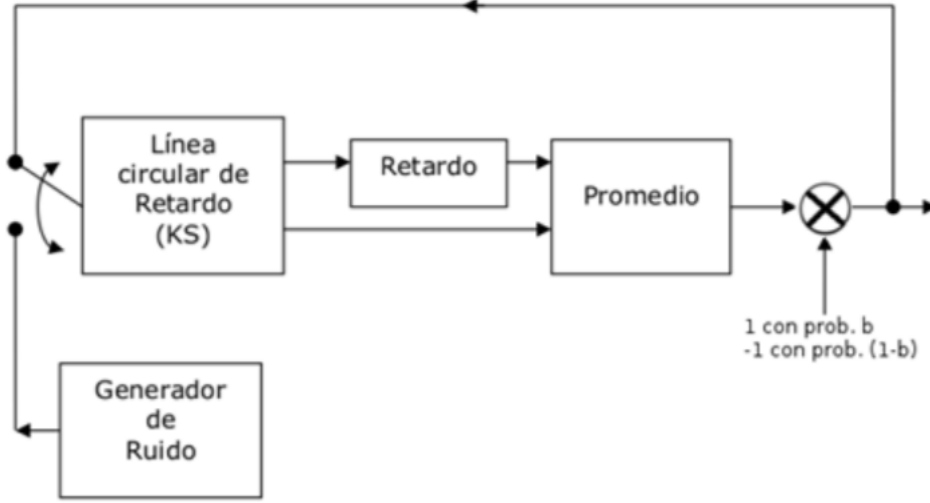


Figura 2: Diagrama conceptual del modelo extendido.

1.2 Análisis de ambos modelos

Como a cualquier filtro digital, para el algoritmo Karplus-Strong es posible encontrar su función transferencia (tanto para el filtro convencional como a su modificación).

Para el modelo clásico, se puede obtener la función transferencia Z mediante relaciones de recursión, obteniéndose:

$$H(Z) = \frac{\frac{1}{2} \cdot (z^{-1} - 1) \cdot z^{-p}}{1 - \frac{1}{2} \cdot (z^{-1} - 1) \cdot z^{-p}} = \frac{1+z}{2z^{p+1} - z - 1}$$

Los polos se pueden encontrar igualando la expresión $2z^{p+1} - z - 1 = 0$, mientras que existe un único cero en $z = -1$, correspondiente a la frecuencia de Nyquist $\frac{1}{2f_s}$.

Aproximando la fórmula de la expresión de los polos a $2 \cdot z^{p+\frac{1}{2}} = z^{\frac{1}{2}} + z^{-\frac{1}{2}}$, y asumiendo que los polos son de la forma $a \cdot e^{j\omega}$, la expresión resultante es:

$$2 \cdot a^{p+\frac{1}{2}} e^{j\omega(p+\frac{1}{2})} = a^{\frac{1}{2}} e^{j\omega/2} + a^{-\frac{1}{2}} e^{-j\omega/2}$$

Aproximando $e^{j\omega(p+1/2)} \approx 1$, podemos despejar $\omega = \frac{2\pi n}{(p+\frac{1}{2})}$. Aproximando $2 \cdot a^{p+1/2} \approx 2\cos(\frac{\omega}{2})$ podemos despejar el tiempo de caída del n-esimo armónico siendo $a = (\cos(\frac{2\pi n}{2p+1}))^{\frac{1}{p+\frac{1}{2}}}$.

Para el modelo de percusión, necesitamos otro tipo de análisis, puesto que la señal que produce es aperiódica. Descomponiendo el sistema en etapas, podemos expresar la función transferencia de la siguiente manera:

$$H(z) = \frac{1}{1 - H_a(z) * H_b(z)}, \text{ con } H_a(z) = \frac{1+z^{-1}}{2}, H_b(z) = z^{-n}$$

Sabiendo que $H(j\omega) = H(z = e^{j\omega})$, podemos expresar las funciones de ganancia y fase de la siguiente manera:

$$\begin{cases} G_a(f) = |\cos(\frac{\omega}{4\pi f_s})| \\ G_b(f) = 1 \end{cases} \quad \begin{cases} P_a(f) = \frac{1}{2} \\ P_b(f) = N \end{cases}$$

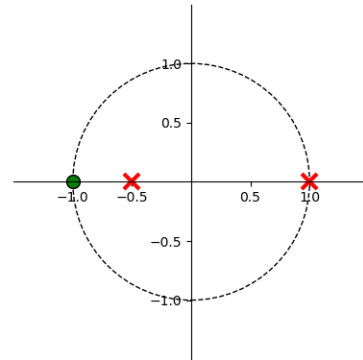


Figura 3: Constelación de polos del modelo simple.

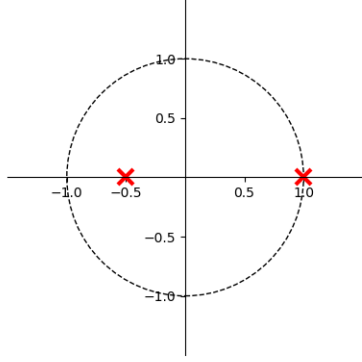


Figura 4: Constelación de polos del modelo extendido.

Con estos parámetros, podemos calcular la ganancia de lazo $G = G_a \cdot G_b = \cos(\pi f T_s)$ y la longitud de lazo que es igual a $P_a + P_b = N + \frac{1}{2}$.

1.3 Modelo conceptual básico

1.3.1 Estabilidad

Conociendo la función transferencia del sistema, la condición para que el mismo sea estable es que todos sus polos estén dentro del círculo unitario, puesto que la función es causal.

Sabiendo que los polos son de la forma $a \cdot e^{j\omega}$, entonces si $|z| < 1 \rightarrow |a \cdot e^{j\omega}| < 1 \rightarrow |a| < 1$

Como $a = (\cos(\frac{2\pi n}{2p+1}))^{\frac{1}{p+\frac{1}{2}}}$ entonces $a \neq 1$ lo que implica que $\cos(c) \neq 1$ lo que significa que $c \neq k\pi$, por lo que la expresión final queda como:

$$\frac{2\pi n}{2p+1} \neq k\pi \Rightarrow n \neq k \cdot (p + \frac{1}{2})$$

siendo n la longitud de la cuerda.

1.3.2 Sintonización

Uno de los problemas que el modelo básico acarrea en sí, es el hecho de que la longitud de la cuerda L debe ser un número entero. Esto ocasiona que las frecuencias fundamentales de la cuerda que se quiera generar estén cuantizadas. Siendo $f = \frac{f_s}{N+0.5}$, a medida que N es más chico, la diferencia entre N y N+0.5 es mucho más notoria, lo que termina manifestandose en el dibujo y empeorando la calidad del sonido.

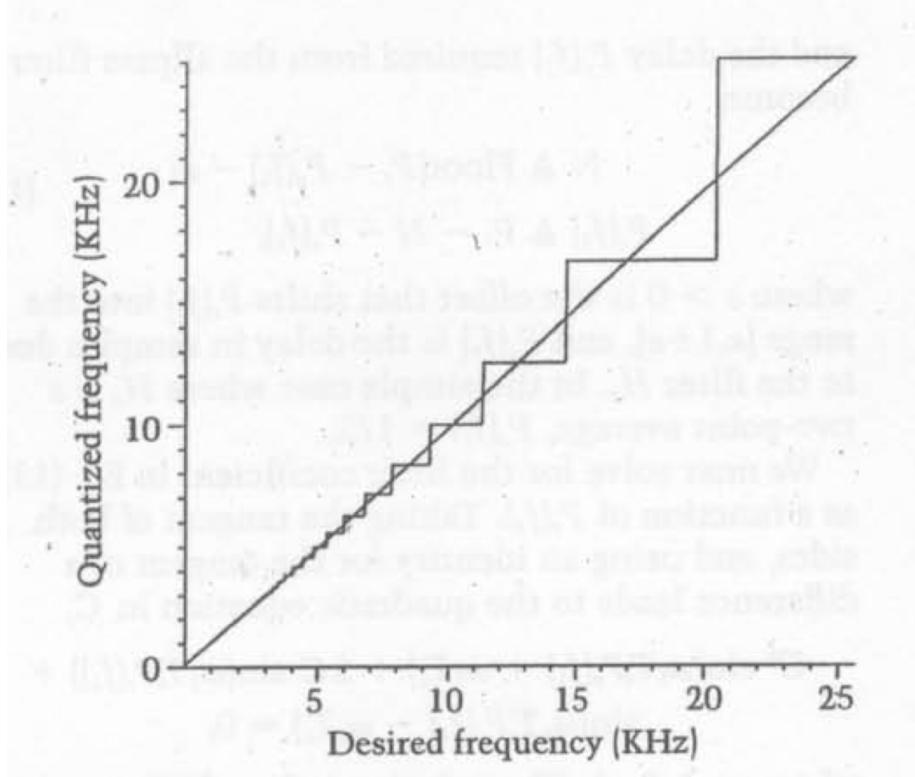


Figura 5: Gráfico de la cuantización de la frecuencia.

Una de las posibles soluciones a este inconveniente, es la de colocar un filtro pasatodo en la realimentación del sistema para aplicar un pequeño retardo de fase sin alterar la ganancia del mismo, el cual esta dado por la siguiente ecuación:

$$H_c(z) = \frac{C+z^{-1}}{1+C \cdot z^{-1}}$$

donde C es una constante que define el nuevo retardo del lazo, y se calcula mediante la siguiente expresión:

$$C = \frac{\sin(2\pi \cdot f_1 T_s) - \sin(2\pi \cdot f_1 T_s \cdot P_c(z))}{\sin(2\pi \cdot f_1 T_s + 2\pi \cdot f_1 T_s \cdot P_c(z))} = \frac{\sin(2\pi \cdot f_1 T_s - 2\pi \cdot f_1 T_s \cdot P_c(z))}{\sin(2\pi \cdot f_1 T_s + 2\pi \cdot f_1 T_s \cdot P_c(z))}$$

donde $P_c(z)$ es la fase del filtro pasatodos, que se puede expresar como $P_c(z) = N - \text{floor}(N - \varepsilon)$

1.4 Modelo extendido

1.4.1 Fase en función de la probabilidad b

Teniendo en cuenta que por el análisis hecho previamente conocemos la función transferencia del sistema, podemos hallar la fase en función de la probabilidad b. Sabiendo que $H(j\omega) = H(z = e^{j\omega})$, y utilizando la igualdad de Euler para números complejos podemos llegar a que la fase $\Phi(f)$ se puede expresar como:

$$\Phi(f) = \tan\left(\frac{b \cdot \sin\left(\frac{2\pi}{f_s} \cdot n + \pi \cdot \frac{1}{f_s}\right) \cdot \cos\left(\pi \cdot \frac{1}{f_s}\right)}{1 - b \cdot \cos\left(\frac{2\pi}{f_s} \cdot n + \pi \cdot \frac{1}{f_s}\right) \cdot \cos\left(\pi \cdot \frac{1}{f_s}\right)}\right)$$

1.4.2 Aplicaciones del paper “Extensions of the Karplus Strong Algorithm”

El paper realiza una serie de mejoras las cuales modifican el diseño del algoritmo original para perfeccionarlo y lograr un sonido aún más realista. Una de las series de mejoras que implementa es el filtro dinámico, que controla la intensidad con la que se toca una cuerda. Cuanto mayor es la intensidad en la misma, más energía poseen las frecuencias más altas, efecto que no se ve reflejado en el algoritmo original. Para lograrlo, se sugiere el siguiente filtro:

$$y(n) = (1 - R) \cdot x(n) + R \cdot y(n - 1)$$

La función transferencia del filtro queda conformada:

$$H(z) = \frac{1-R}{1-R \cdot z^{-1}}$$

Donde R es función de la frecuencia deseada y la longitud del arreglo de muestras, y puede ser aproximado a $R = e^{-\pi L}$.

Otra de las mejoras que ofrece el paper, es la posibilidad de alargar o acortar los tiempos de decaimiento del sonido. Para acortar el tiempo de decaimiento, basta con multiplicar al algoritmo por una constante $\rho < 1$, tal que $\rho = 0.99 \cdot \frac{1 - (1000 \cdot \text{duración})^{-1}}{1 + (1000 \cdot \text{duración})^{-1}}$. Ahora, el sistema pasa a ser:

$$Y_n = \rho \cdot \frac{1}{2} \cdot (Y_{n-p} + Y_{n-p-1})$$

Si lo que se desea es alargar el tiempo de decaimiento, lo que debemos modificar es el lazo de realimentación dado por $H_a(z)$, de manera tal de que ahora sea $H_a(z) = (1 - s) + s \cdot z^{-1}$, donde $0 < s < 1$ es el factor de estiramiento. Con esto en cuenta, ahora la ganancia de lazo queda conformada:

$$G_a(z) = \sqrt{(1-s)^2 + s^2 + 2s(1-s) \cdot \cos(2\pi f \cdot T_s)}$$

1.5 Conclusiones

El Algoritmo de Karplus Strong es un algoritmo fácil y rápido que nos permite emular el sonido de una cuerda de guitarra con mucha precisión. Es un modelo versatil, puesto que con menores modificaciones no solo se pueden lograr sonidos de otros instrumentos como los de percusión, sino que también es posible de mejorar la calidad del sonido que se sintetiza.