Capítulo 6

Diseño de espectros

Tal como se sugirió al comienzo del capítulo previo, una poderosa forma de sintetizar sonidos musicales es especificar –y luego realizar- trayectorias específicas de alturas de sonido (o más específicamente, frecuencias de parciales), a lo largo de las trayectorias de la envolvente espectral [Puc01]. La envolvente espectral es utilizada para determinar la amplitud de los parciales individuales, como una función de sus frecuencias, y se toma para controlar el timbre del sonido (que posiblemente varíe en el tiempo).

Un ejemplo simple podría ser imitar una cuerda que se pulsa, construyendo un sonido con parciales armónicamente espaciados, en el cual la envolvente espectral inicie rica pero luega muera exponencialmente con las frecuencias más altas decayendo más rápido que las bajas, de tal manera que el timbre suene melodioso en el tiempo. Los modelos de evolución espectral para varios instrumentos acústicos ya se han propuesto [GM77] [RM69]. Un ejemplo más complicado es la voz hablada o cantada, en la cual las vocales aparecen como envolventes espectrales, los diptongos y las consonantes aparecen como variaciones en el tiempo de las envolventes espectrales, y otras consonantes aparecen como ruido conformado espectralmente.

Las envolventes espectrales se pueden obtener del análisis de sonidos grabados (que se desarrolla en el capítulo 9) o de criterios puramente sintéticos. Especificar una envolvente espectral desde el inicio para cada frecuencia posible podría ser tedioso, y en la mayoría de los casos se la podría querer describir en términos de sus característica sobresalientes. La forma más popular de hacer esto es especificar el tamaño y la forma de los picos de la envolvente espectral llamados *formantes*. La figura 6.1 muestra una envolvente espectral con dos formantes. Aunque las formas de los dos picos en la envolvente espectral son diferentes, estos se pueden describir en términos generales por las coordenadas de cada ápice (que proporcionan el *centro de frecuencia* y su amplitud), y el *ancho de banda* de cada formante. Una medida típica del ancho de banda podría ser el ancho del pico a un nivel de 3 decibles por debajo de su cima. Note que si el pico está en (o cerca al) eje $f = 0$, esperamos que caiga a la izquierda de la misma manera a como lo hace (en realidad) a la derecha.

Suponga que queremos generar un sonido armónico con una colección específica de formantes. Independientemente de la frecuencia fundamental que se desee, queremos que el espectro tenga picos con centros de frecuencias, amplitudes y anchos de banda prescritos. Retornando a los espectros de modulación de fase mostrados en la figura 5.16, vemos que, con los índices de modulación más pequeños posibles, el resultado tiene un solo pico espectral bien definido. Podemos imaginar adicionar varios de estos <picos espectrales>, todos compartiendo una frecuencia fundamental (modulante) pero con frecuencias portadoras afinadas en diferentes armónicos para seleccionar los diversos centros de frecuencia <la frecuencia portadora = centro de frecuencia>, y con índices de modulación apropiados para dar los anchos de banda deseados. Esto lo exploró Chowning por primera vez, organizando formantes generados por la modulación de fase, para sintetizar voces cantadas. En este capítulo estableceremos un marco general para construir espectros armónicos con los formantes deseados, y con la posibilidad de variarlos en el tiempo.

6.1 El modelo portador/modulador

Vimos anteriormente cómo utilizar la modulación de anillo para modificar el espectro de una señal periódica, ubicando picos espectrales en diferentes lugares (ver la figura 5.4 en la página “125”). Para hacer esto debemos generar señales periódicas cuyos espectros tengan su máximo en DC y caigan monolíticamente con el decrecimiento de la frecuencia. Al fabricar una señal con un formante en la frecuencia cero -sin más formantes que este- podemos utilizar la modulación de anillo para desplazar el formante hacia cualquier armónico que queramos. Utilizando la conformación de ondas para generar el formante inicial, el producto de la modulación de anillo será de la forma

$$x[n] = \cos(\omega_c n) f(a \cos(\omega_m n))$$

donde ω_c (la *frecuencia portadora*) se ajusta para ser el centro de frecuencia del formante y $f(a\cos(\omega_m n))$ es una señal con frecuencia fundamental determinada por ω_m, que se produce utilizando una función de conformación de onda f y un índice a. Este segundo término es la señal a la cual le queremos dar un formante en DC con un ancho de banda controlable. Un diagrama de bloque para la síntesis de esta señal se muestra en la figura 6.2.

Mucho antes, en la sección 2.4 introdujimos la técnica de estiramiento del timbre, como parte de la discusión de la síntesis de tabla de ondas. Esta técnica, que genera timbres variables y complejos, se puede tomar dentro del mismo marco general. La salida para la tabla de onda envuelta para un ciclo es:

$$x(\phi) = T(c\phi)*W(a\phi),$$

donde ϕ, la fase, satisface $-\pi \leq \phi \leq \pi$ <ϕ no es frecuencia angular; estas son funciones del ángulo, no de la frecuencia angular>. Aquí T es la función almacenada en la tabla de ondas, W es la función de ventana, y c y a son el estiramiento y un índice de modulación para la función de ventana <esta función ventana es una envolvente con valores cero en sus extremos para evitar ruido entre el final y el inicio de la lectura de la tabla T>. La figura 6.3 muestra cómo realizar esto en la forma de un diagrama de bloques. Comparando esto con la figura 2.7, vemos que la única nueva característica significante es la adición del índice a.

En esta organización, como en la previa, el primer término especifica el emplazamiento de la energía en el espectro -en este caso con el parámetro c estirando el espectro de la tabla de ondas. Este es el papel que llevaba previamente <la función T almacenada en la tabla de onda> cuando se seleccionó como la frecuencia portadora en la modulación de anillo.

Ambas (la conformación de ondas por modulación de anillo y la síntesis de tabla de ondas estirada) pueden ser consideradas casos particulares de la aproximación a un caso más general que computa funciones de la forma,

$$x[n] = c(\omega n) m_a(\omega n)$$

donde c es una función periódica que describe la señal portadora y m_a es una función de un modulador periódico que depende del índice a. Las funciones de modulación que nos han interesado usualmente <envolvente de tablas de onda> tomarán la forma de trenes de pulsos <ventana de Hann, curva campana, distribución de Cauchy>, y el índice a tomará el control del ancho del pulso; valores más altos de a <diente de sierra más parada, lo que propicia una lectura de tabla de onda de la envolvente más rápida> darán pulsos más estrechos <más “cortos”; el índice a además de estrechar los pulsos cuando crece, fortalece los armónicos más altos de la función>. En el caso de la tabla de onda, la función de modulación debe ser cero en los puntos de la fase de la envolvente, para suprimir cualquier discontinuidad en la función portadora cuando la fase hace la envolvente. La señal portadora dará lugar a un único pico espectral (un formante) en el caso de la conformación de onda con modulación de anillo; las tablas de onda pueden tener un espectro más complicado.

En la sección siguiente haremos un desarrollo de las dos formas de modulación de la señal presentadas aquí, y en la siguiente, miraremos más de cerca a la señal portadora.

6.2 Trenes de pulsos

Los trenes de pulsos se pueden generar ya sea utilizando la formulación de la conformación de onda o la del estiramiento de tabla de ondas. La formulación de la conformación de onda es más fácil de analizar y controlar, y la consideraremos en primer lugar.

6.2.1 Trenes de pulso vía conformación de onda

Cuando utilizamos conformación de onda la configuración del formante está determinada por un término de modulación

$$m_a(n) = f(a\cos(\omega_m n))$$

Para valores pequeños del índice a el término de la modulación varía sólo ligeramente con respecto al valor constante $f(0)$, de tal manera que la mayor parte de la energía está concentrada en DC. En tanto a se incrementa la energía se esparce progresivamente hacia los armónicos más elevados de la fundamental ω. Dependiendo de la función f, este proceso puede ser llevado a cabo ordenada o desordenadamente. Puede quererse un esparcimiento ordenado al principio y luego no, dependiendo de si nuestro objetivo es un espectro predecible o un amplio rango de espectros (que son de esta manera difíciles de predecir).

La función de conformación de onda $f(x) = e^x$, analizada en la página “140”, proporciona resultados que se comportan bien, simples y predecibles. Después de normalizar apropiadamente, obtenemos los espectros mostrados en la figura 5.13. Una ligera re-escritura del modulador de conformación de onda para esta opción de f (y tomando en cuenta la renormalización) da:

$$m_a(n) = e^{a(\cos(\omega n) - 1)}$$

<$a(\cos(\omega_m n) - 1) = -a(1 - \cos(\omega_m n)) = -2a(\operatorname{sen}^2(\omega_m n/2))$; identidad trigonométrica>

$$= e^{-[b\operatorname{sen}(\omega n/2)]^2}$$ <en el libro hace falta el término n en el exponente>

donde $b^2 = 2a$ de tal manera que b es proporcional al ancho de banda. Esto puede re-escribirse como

$$m_a(n) = g[b\operatorname{sen}(\omega n/2)]$$

con

$$g(x) = e^{-x^2}$$

Excepto por un factor de normalización perdido, esta es una distribución gaussiana, usualmente llamada “curva campana” <será norma que una función exponencial se convierta en una curva campana, cuando la variable x es de tipo sinusoidal>. Las amplitudes de los armónicos están dadas por las funciones de Bessel tipo “I” <es decir, se llama “tren de pulsos” porque la “curva campana” funciona como una envolvente de tabla de onda; casi cero en sus extremos>.

Otra opción es la distribución de Cauchy (también sin normalizar):

$$h(x) = 1/(1 + x^2)$$

la cual hace aparecer un espectro de armónicos cayendo exponencialmente:

$$h[b\,\mathrm{sen}(\omega n/2)] = G \cdot (1/2 + H\cos(\omega n) + H^2\cos(2\omega n) + ...)$$

donde G y H son funciones del índice b (fórmulas explícitas se dan en [Puc95a]).

Tanto en esta como en la gaussiana del caso anterior, el ancho de banda <del formante> (contado en picos, es decir, en unidades de ω) es aproximadamente proporcional al índice b, y la amplitud del término DC (la cima del espectro) es aproximadamente igual a $1/(1 + b)$ <a esta deducción se debe llegar después de un análisis de las funciones mucho más profundo; hay que recordar que b^2 es proporcional al parámetro a por lo menos para el caso de la curva campana>. Para cualquiera de las dos funciones de conformación de onda (g ó h) si b es más grande que aproximadamente 2, la conformación de onda de $m_a(\omega n)$ <la onda resultante> es aproximadamente una lectura (hacia adelante o hacia atrás) de la función de transferencia, de tal manera que la onda resultante aparece como pulsos cuyas anchuras decrecen mientras los anchos de banda especificados se incrementan.

6.2.2 Trenes de pulsos vía estiramiento de tabla de ondas

En la formulación de tabla de ondas, un tren de pulsos puede ser hecho por una tabla de onda estirada:

$$M_a(\phi) = W(a\phi),$$

donde $-\pi \le \phi \le \pi$ es la fase, es decir, al valor ωn se le aplica una envolvente para quedar entre $-\pi$ y π. La función W deberá ser cero en y más allá de los puntos $-\pi$ y π, y elevarse a su máximo en 0. Una opción posible para la función W es

$$W(\phi) = (\cos(\phi) + 1)$$

lo cual se grafica en la parte (a) de la figura 6.4. Esta se conoce como la *función de la ventana de Hann*; que vendrá de nuevo en el capítulo 9.

Al realizar repiticiones de esta como una onda, obtenemos una sucesión de copias de la función W (sampleada apropiadamente), cuyo ciclo de trabajo es $1/a$ (partes b y c de la figura). Si no desea que las copias se traslapen, el índice a debe ser como mínimo 1. Si usted quiere permitir el traslape, la estrategia más simple es duplicar el diagrama de bloque (figura 6.3), como se describió en la sección 2.4 y se efectuó en la sección 2.6 <este procedimiento de duplicación no se describió como tal, si no que se efectuó en la sección 2.6, con el ejemplo B13.sampler.overlap.pd>.

6.2.3 Espectros resultantes

Antes de considerar señales más complejas que vayan con las moduladoras tal como lo hemos visto, es instructivo ver qué ondas y espectros nos da la multiplicación por una sinusoide pura. La figura 6.5 muestra el resultado de multiplicar dos trenes de pulsos diferentes por una sinusoide en el sexto parcial:

$$\cos(6\omega n)M_a(\omega n)$$

donde el índice de modulación a es dos en ambos casos. En la parte (a) M_a es la función de ventana de Hann, estirada; la parte (b) muestra una conformación de onda vía la distribución de Cauchy no normalizada. Se muestra un período de cada onda.

En ambas situaciones vemos, en efecto, el sexto armónico (la señal portadora) envuelto dentro de un *paquete de onda* centrado en la mitad del ciclo, donde la fase de la sinusoide es cero <en la parte (a) se ven en primera instancia 5 ciclos de la portadora en su sexto parcial únicamente, pero hay dos medios ciclos a lado y lado cuyas amplitudes no se alcanzan a destacar por encima de la línea horizontal; el resto del ciclo queda en cero -a lado y lado-; en (b) se alcanzan a visualizar dos ciclos de la portadora, pues las amplitudes de la ventana de Cauchy permiten esto>. Al cambiar la frecuencia de la sinusoide, cambia el centro de frecuencia del formante; al cambiar el ancho del paquete (la parte de la onda durante la cual la sinusoide es fuerte) cambia el ancho de banda. Note que la función de ventana de Hann estirada es cero al comienzo y al final del período, a diferencia del paquete al que se le ha aplicado la conformación de onda.

La figura 6.6 muestra cómo la configuración del formante depende del método de producción. La forma de la tabla de onda estirada <ventana de Hann> (parte (a) de la figura) se comporta bien en las vecindades del pico, pero de manera extraña a partir de una distancia del parcial cuatro desde el pico, pasado lo cual vemos los llamados *lóbulos laterales*, picos extras falsos a una amplitud más baja que la del pico central. Tal como el análisis de la sección 2.4 predice, el formante completo, los lóbulos laterales y todo <el espectro!>, se estira y se contrae de manera inversa a como el tren de pulsos es contraído y estirado en el tiempo.

El primero y más fuerte de los lóbulos laterales en cualquiera de los lados es de aproximadamente 32 dB más bajo en amplitud que el pico principal. Los demás lóbulos laterales caen más lentamente cuando se expresan en decibeles; las amplitudes decrecen con el cuadrado de la distancia desde el pico central de tal manera que el sexto lóbulo lateral a la derecha, a una distancia tres veces mayor que la del primero desde la frecuencia central, es aproximadamente veinte decibeles más bajo. El efecto de estos lóbulos laterales es usualmente audible como un ligero zumbido en el sonido.

Esta configuración del formante puede ser hecha arbitrariamente gorda (es decir con un ancho de banda alto) <recordar que este parámetro es proporcional al parámetro b>, pero hay un límite en lo delgada que pueda ser, ya que el ciclo de trabajo de la onda no puede exceder el 100%. En su ciclo de trabajo máximo la fortaleza del formante cae a cero a la distancia de dos armónicos del pico central. Si se requiere un ancho de banda aún más bajo, las ondas se pueden traslapar, tal como se describió en la sección 2.6.

Las partes (b) y (c) de la figura muestran los formantes generados utilizando la conformación de onda con modulación de anillo, con funciones gaussianas y de Cauchy. El índice de modulación es dos en ambos casos (el mismo que para la ventana de Hann de la parte a), y el ancho de banda es comparable al del ejemplo de la ventana de Hann. En estos ejemplos no hay lóbulos laterales, y es más, el índice de modulación puede ser llevado a cero, resultando una sinusoide pura; no hay límite inferior en el ancho de banda. De otro lado, debido a que la onda no alcanza el cero al final del ciclo, este tipo de tren de pulso no puede ser utilizado para ser ventana de una tabla de onda arbitraria, tal como lo puede ser el tren de pulsos de Hann.

El ejemplo de Cauchy es particularmente útil para el diseño de espectros, ya que la configuración del formante es un triángulo isósceles, cuando se grafica en decibeles. De otro lado, el ejemplo gaussiano reúne más energía hacia el formante, y cae más rápidamente en las colas, de tal manera que tiene un sonido más limpio y ofrece mayor protección contra el sobre-doblado.

6.3 Modulación de anillo móvil

Retornamos ahora a la señal portadora, buscando maneras de hacerla más controlable. En particular nos gustaría poder deslizar la energía espectral continuamente arriba y abajo en la frecuencia <sin que se escuchen “choques” o

discontinuidades>. Llevar en rampa la frecuencia del oscilador portador no logrará esto, ya que los espectros no serán armónicos excepto cuando la portadora sea un múltiplo entero de la frecuencia fundamental.

Con la tabla de ondas estirada podemos lograr esto sencillamente haciendo la lectura de muestras de una sinusoide y transponiéndola a la “altura” deseada. La altura transpuesta no es escuchada como algo periódico ya que la tabla de ondas misma es leída periódicamente a la frecuencia fundamental <recordar que el objetivo del estiramiento es independizar el período al cual comienza la lectura de una tabla de onda y la longitud de esa tabla de onda; dos parámetros que usualmente están controlados únicamente por la frecuencia>. En lugar de esto, la sinusoide es transpuesta como una envolvente espectral.

La figura 6.7 muestra una señal portadora producida de esta manera, afinada para producir un formante a 1.5 veces la frecuencia fundamental. La señal no tiene discontinuidad en la frecuencia de la fase envolvente, pero tiene una discontinuidad en el traslape <en la “conexión” entre el final del uno y el inicio del otro>, la cual, si no se remueve aplicando una modulación apropiada de la señal, podría llegar a tener componentes muy audibles en frecuencias altas.

Utilizando esta idea podemos hacer una descripción completa de cómo utilizar el diagrama de bloque de la figura 6.3 para producir el formante deseado. La tabla a leer en el lado izquierdo podría contener una sinusoide (ubicada simétricamente, de tal manera que la fase es cero en el centro de la tabla de ondas). La tabla de ondas de la derecha deberá contener una ventana de Hann u otra función de ventana apropiada. Siendo la frecuencia fundamental ω, la frecuencia central del formante ω_c, y el ancho de banda ω_b, obtenemos el parámetro “estiramiento” para el *cociente de frecuencia central*, definido como ω_c/ω, y el índice de modulación como el *cociente ancho de banda*, ω_b/ω.

La señal de salida es simplemente una muestra de una onda coseno a la frecuencia central deseada, repetida en el período deseado (no relacionados en general) y a la que se aplica una ventana para tomar las discontinuidades en las fronteras del período.

Aunque no somos capaces de derivar este resultado todavía (necesitaremos el análisis de Fourier), este se convertirá en que, en el lóbulo principal del formante, las fases son todas cero en el centro de la onda (es decir, los componentes son todos coseno si consideramos que la fase debe ser cero en el centro de la onda). Esto significa que podemos superponer cualquier número de estos formantes para construir un espectro más complejo y las amplitudes de los parciales se combinarán por adición. (Los lóbulos laterales no se comportan así: ellos son de signo opuesto alternativamente y producirán patrones de cancelación; pero usualmente los podemos ignorar como una señal residual, pequeña e incontrolable).

Este método nos conduce a una interesante generalización, que es tomar una secuencia de tablas de ondas grabadas, alinear todas las fases de sus componentes con las de los cosenos, y utilizarlas en lugar de la función coseno como señal portadora. La alineación de la fase es necesaria para permitir la atenuación cruzada entre las muestras, de tal manera que la envolvente espectral pueda cambiar con suavidad. Si, por ejemplo usamos retazos de una muestra vocal como entrada, obtenemos un sorpendentemente efectivo vocoder; ver la sección 9.6.

Otra técnica para fabricar señales portadoras que se puedan deslizar continuamente arriba y abajo en la frecuencia mientras se mantiene la frecuencia fundamental es simplemente hacer atenuación cruzada entre los armónicos. La señal portadora es entonces:

$$c(\phi) = c(\omega n) = p\cos(k\omega n) + q\cos((k + 1)\omega n)$$

donde $p + q = 1$ y k es un entero, y todos los tres se escogen de tal manera que

$$(k + q)*\ \omega = \omega_c$$

de tal manera que el centro espectral de masa de los dos cosenos esté ubicado en ω_c. (Note que hacemos que la adición de las amplitudes de los dos cosenos sea uno, en lugar de ajustar la potencia total a uno; hacemos esto porque el modulador operará coherentemente con la fase sobre ellos.) Para lograr esto ajustamos k y q para que sean parte entera y parte fraccional, respectivamente, del cociente del centro de frecuencia ω_c/ω.

La forma más simple de hacer una interfaz de control para esta técnica de síntesis podría ser utilizar rampas para actualizar ω y ω_c, y luego computar q y k como señales de audio que van en rampa, variando suavemente ω y ω_c. Extrañamente sin embargo, a pesar del hecho de que k, p y q son funciones discontinuas de ω_c/ω, la portadora $c(\phi)$ se convierte para variar continuamente con ω_c/ω, y así si el centro de frecuencia ω_c va en rampa de valor en valor, el resultado es un cambio continuo en el centro de frecuencia. Sin embargo se requiere más trabajo si se necesitan cambio discontinuos en el centro de frecuencia. No es irracional querer esto pues es análogo a cambiar la frecuencia de un oscilador de manera discontinua.

Hay una buena manera de acomodarlo. El truco para actualizar k y q es notar que $c(\phi) = 1$ siempre que ϕ sea un múltiplo de 2π, sin importar cuáles valores de k, p y q escojamos, y siempre que $p + q = 1$. Por lo tanto podemos hacer cambios discontinuos en k, p y q una vez por período (cuando la fase es un múltiplo de 2π) sin hacer discontinuidades en la señal portadora.

En el caso específico de FM, si queremos podemos ir atrás y modificar la formulación original a:

$$p\cos(n\omega_2 t + r\cos(\omega_1 t)) + q\cos((n + 1)\omega_2 t + r\cos(\omega_1 t))$$

Esto nos permite adicionar glisandos (los cuales son escuchados como diptongos) a la técnica original de la síntesis vocal basada en la modulación de fase, de Chowning.

6.4 Generador de formantes alineados en fase (PAF)

Combinando las dos señales portadoras de dos cosenos con el generador de pulsos por conformación de onda, resulta un generador de *formantes alineados en fase*, usualmente llamado por su acrónimo PAF. (El PAF es tema de una patente de 1994 perteneciente al IRCAM.) La fórmula combinada es,

$$x[n] = g(a\,\mathrm{sen}(\omega n/2))[p\cos(k\omega n) + q\cos((k + 1)\omega n)]$$

<donde $g(a\,\mathrm{sen}(\omega n/2))$ es el modulador (generador de pulsos por conformación de onda) y $p\cos(k\omega n) + q\cos((k + 1)\omega n)$ son las señales portadoras>

Aquí la función g puede ser la función de conformación de onda gaussiana o de Cauchy, ω es la frecuencia fundamental, a es el índice de modulación que controla el ancho de banda y k, p y q controlan la frecuencia central del formante.

La figura 6.8 muestra el PAF como un diagrama de bloques, separado en una sección de generador de fase, una sección portadora y una sección moduladora. La sección de la generación de fase da a la salida una señal diente de sierra en la frecuencia fundamental. El modulador fue hecho por conformación de onda normal con una ligera modificación. La fórmula para las señales del modulador tienen una entrada sinusoidal a la mitad de la frecuencia fundamental, es decir sen(ω/2), y esto normalmente nos hace utilizar una diente de sierra a la mitad de la frecuencia fundamental. Sin embargo, debido a que la función de conformación de onda es par, podemos sustituír por el valor absoluto de la sinusoide:

$$|\mathrm{sen}(\omega/2)|$$

la cual se repite a la frecuencia ω (la primera mitad del ciclo es igual a la segunda mitad.) Podemos calcular esto simplemente utilizando una sinusoide de medio ciclo como una función de lectura de una tabla de onda (con la fase de -π/2 a π/2), y esta es la sinusoide rectificada que pasamos a la función de conformación de onda.

Aunque la función de la tabla de onda está dibujada en valores positivos y negativos (de -10 a 10), de hecho estamos utilizando únicamente el lado positivo para la lectura, llevando de 0 a b el índice de modulación. Si el índice de modulación excede el rango de la tabla (aquí lo ajustamos para que se detenga en 10 como un ejemplo), la lectura de la tabla será recortada. La tabla debe extenderse suficientemente lejos hacia la cola de la función de conformación de onda, de tal manera que el efecto del recorte resulte inaudible.

La señal portadora es una suma ponderada de dos cosenos, cuyas frecuencias se incrementan multiplicándolas (por k y k + 1 respectivamente) y aplicándoles envolventes. De esta manera todas las fases de la lectura están controladas por el mismo oscilador diente de sierra.

Las cantidades k, q y el índice de tabla de onda b se calculan como se muestra en la figura 6.9. Estas son funciones de la frecuencia fundamental dada, de la frecuencia central del formante y del ancho de banda, que son parámetros originales del algoritmo. La cantidad p que no se muestra en la figura es precisamente $1 - q$.

Tal como se describió en la sección previa, las cantidades k, p y q deberán cambiar únicamente en los puntos de envoltura de la fase, es decir en períodos de $2\pi/\omega$. Ya que el cálculo de k, etc depende del valor del parámetro ω, se sigue que ω mismo deberá ser actualizado cuando la fase sea un múltiplo de 2π; de otra manera, un cambio en ω podría enviar el centro de frecuencia $(k + q)/\omega$ a un valor incorrecto por una fracción (muy notable) del período. En efecto, todos los cálculos del parámetro deberán ser sincronizados a la fase del oscilador original.

Tener un control de la fase del oscilador que actualice su propia frecuencia es un ejemplo de *retroalimentación*, que significa en general el uso de la salida de un proceso como una de sus entradas. Cuando procesamos señales de audio digital a una velocidad fija de lectura de muestras (tal y como lo estamos haciendo), nunca es posible tener la salida *actual* del proceso como una entrada, debido a que el tiempo que podríamos necesitar aún no lo hemos calculado. Lo mejor que podemos esperar es utilizar la muestra previa de la salida -en efecto, adicionando una muestra de retraso. En ambientes de bloques (tales como Max, Pd y Csound) la situación se vuelve más complicada, pero dejaremos la discusión para el capítulo siguiente (y simplemente deseo esquivar el problema en los ejemplos y al final de este capítulo).

La amplitud del pico central en el espectro del generador PAF es aproximadamente $1/(1 + b)$; en otras palabras, cercana a la unidad cuando b es menor que uno, y cayendo de manera inversa con valores más grandes de b. Para los valores de b aproximándose a diez por debajo, la fortaleza de la señal puede no variar mucho, ya que la introducción de otros parciales, incluso a amplitudes más bajas, desajusta el decrecimiento de la amplitud del parcial central. Sin embargo, si se utiliza PAF para generar formantes con amplitudes pico específicas, la salida deberá multiplicarse por 1 + b (o incluso, si es necesario, por una aproximación del factor de corrección, cuyo valor exacto depende de la función de conformación de onda). Esta corrección de amplitud debe aplicarse en rampa y no por muestra-y-sostenimiento.

Debido a que la expansión de la señal de conformación de onda (moduladora) consiste de todos los términos coseno (es decir, debido a que todos tienen fase inicial cero), como lo son los dos componentes de la portadora, se sigue de la fórmula del producto de cosenos que los componentes del resultado son todos coseno también. Esto significa que cualquier número de generadores PAF, si están fabricados para compartir el mismo oscilador de generador de fase, estarán todos en fase y al combinarlos el resultado será la suma de los espectros individuales. De esta manera podemos hacer versiones de formantes múltiples como se muestra en la figura 6.10.

La figura 6.12 muestra una salida posible de un par de formantes generados de esta manera; el primer formante está centrado en la mitad de los parciales 3 y 4, y el segundo en parcial 12, con amplitud y ancho de banda menores. Se utilizó la función de conformación de onda de Cauchy, la cual produce espectros traslapados linealmente (visto en dB). Los dos se superponen aditivamente, de tal manera que la envolvente espectral es una curva suave entre un formante y el otro. El formante más bajo también adiciona su propia reflexión alrededor del eje vertical, de tal manera que aparece ligeramente curvado delante de allí.

El generador PAF puede alterarse si se desea fabricar espectros inarmónicos deslizando los parciales adelante y atrás en su frecuencia. Para hacer esto, adicionamos un segundo oscilador para la fase de los dos cosenos de la portadora, pero no para la fase de la porción moduladora del diagrama, ni para controlar la fase de las unidades de muestra-y-sostenimiento. En cambio, la estrategia de muestra-y-sostenimiento para suavizar la actualización de los parámetros funciona todavía; y aún más, los generadores PAF comparten la misma porción de generación de fase con lo que seguirán en fase uno con otro.

La técnica de superposición de espectros no trabaja de manera predecible para la modulación de fase como sí lo hace el generador PAF; la salida de los parciales de la modulación de fase tiene relaciones de fase complicadas y parecen difíciles de combinar de manera coherente. En general la modulación de fase dará patrones más complicados de evolución espectral, toda vez que el PAF es más fácil de predecir y convertir en efectos específicos deseados.

6.5 Ejemplos

Tren de pulsos de tabla de ondas

El ejemplo F01.pulse.pd (figura 6.13) genera un tren de pulsos de altura variable, utilizando la lectura de una tabla de onda estirada. La figura 6.14 muestra dos productos intermedios del parche y su salida. El parche lleva a cabo la labor de la manera más simple posible, ubicando el pulso en el ángulo (fase) π en lugar de 0; en los ejemplos posteriores esto se solucionará adicionando 0.5 a la fase y aplicando una envolvente.

La fase inicial se ajusta para que vaya de -0.5 a 0.5 y se escala luego por un multiplicador que es uno como mínimo, dando como resultado la señal de la figura 6.14 (parte a); esta corresponde a la salida del objeto *~, el quinto de abajo hacia arriba en el parche. La gráfica en la parte (b) muestra el resultado de recortar la onda diente de sierra al intervalo -0.5 y 0.5, usando el objeto clip~ <la formulación con la función de corte (clip~) es básica para obtener los diferentes ciclos de trabajo de una onda; se aplica clip~ a la función diente de sierra phasor~, con un modificador de amplitud para esta>. Si el multiplicador que escala es el mínimo (uno) <en el parche aparece como “index”, pero en la explicación también se le identifica como “bandwidth” (ancho de banda)>, la diente de sierra deberá ir desde -0.5 a 0.5 en cualquier caso y el recorte no tendrá efecto. Para cualquier valor del multiplicador mayor que uno, la salida recortada se sitúa en el valor -0.5, hace la rampa hasta 0.5, y luego se estaciona en 0.5. Mientras más alto sea el multiplicador, más rápidamente la onda hace la rampa y más tiempo demora el recorte inferior y superior.

El objeto cos~ convierte luego esta onda en un pulso. Las entradas -0.5 y 0.5 van a -1 (están en un ciclo aparte); así que en el punto medio de la onda la entrada es 0 y la salida es, por lo tanto, 1. La salida por lo tanto se sitúa en -1, dibuja un ciclo completo de la función coseno y luego regresa a -1 para quedar en silencio. La proporción de tiempo que la onda gasta trazando la función coseno es uno dividido por el multiplicador; así, es 100% para un multiplicador de uno, 50% para 2 y así sucesivamente. Finalmente, la salida del pulso se ajusta para que sus valores estén entre 0 y 1; es lo que se grafica en la parte (c) de la figura <estamos ante un pulso con forma coseno: ventana de Hann>.

Generador de formantes sencillo

Los siguientes tres ejemplos demuestran el sonido de un pulso con anchura variable, grafican su espectro, y hacen el contraste con el generador de pulsos por conformación de onda. Saltando al ejemplo F05.ring.modulation.pd (figura 6.15), mostramos la manera más simple de combinar el generador de pulsos con un oscilador de modulación de anillo para fabricar un formante. El tren de pulsos de los ejemplo previos está contenido en el subparche pd pulse-train. Este se multiplica por un oscilador cuya frecuencia está controlada como un múltiplo de la frecuencia fundamental. Si el múltiplo es un entero, se obtiene un sonido armónico. No se ha intentado controlar las fases relativas de los componentes del tren de pulsos ni de las sinusoides portadoras.

El ejemplo siguiente, F06.packets.pd (figura 6.16), muestra cómo combinar el tren de pulsos de la tabla de onda estirada con una muestra de sinusoide para realizar formantes móviles, como se describió en la sección 6.3. El generador de pulsos está como antes, pero ahora la señal portadora es una sinusoide quebrada. Dado que su fase es la fase fundamental multiplicada por el cociente del centro de frecuencia, el incremento de la fase muestra por muestra es el mismo que para la sinusoide de la frecuencia central. Sin embargo cuando se aplica una envolvente a la fase, la fase de la portadora salta a un sitio diferente en el ciclo, <pues la sinusoide de la portadora no es continua> tal como se ilustró en la figura 6.7. Aunque el cociente del ancho de banda ω_b/ω debe ser por lo menos uno, el cociente del centro de frecuencia ω_c/ω puede llegar hasta cero si se desea <en este parche se evidencia de nuevo que al recortar el ciclo de trabajo, también se recorta la onda, cosa que no se estipuló cuando se enunció la teoría del estiramiento de onda>.

Señal portadora de dos cosenos

El ejemplo F08.two.cosines.pd (figura 6.17) muestra cómo fabricar una señal portadora que realice un cruzamiento atenuado entre armónicos para hacer centro de frecuencias continuamente variables. El cociente del centro de frecuencia aparece como la salida de un objeto line~. Esta es separada en dos partes fraccionales (utilizando el objeto wrap~) y su parte entera (al sustraer la parte fraccional de la original). Estas están etiquetadas como q y k de acuerdo con el tratamiento de la sección 6.3.

La fase -una onda diente de sierra a la frecuencia fundamental- es multiplicada por k y $k + 1$ (el último adicionando la diente de sierra original a la primera), y se toman los cosenos de ambos; ellos están de esta manera a k y $k + 1$ veces la frecuencia fundamental y no tienen discontinuidades en los puntos de la envolvente de fase. Los siguientes pocos objetos en el parche computan la suma ponderada $pc_1 + qc_2$, donde c_1 y c_2 son las dos sinusoides y $p = 1 - q$, al evaluar una expresión equivalente, $c_1 = q(c_2 - c_1)$. Esto nos da la frecuencia móvil deseada de la señal portadora.

El ejemplo F09.declickit.pd (que no se muestra aquí), muestra cómo adicionando un objeto samphold~ después del objeto line~ controlando la frecuencia central, usted puede evitar discontinuidades en la señal de salida incluso si la frecuencia central deseada cambia discontinuamente. En el ejemplo el cociente de frecuencia central alterna entre 4 y 13.5. Al efectuar la rampa multiplicada por un valor por debajo de casi 20 milisegundos se presentan artefactos audibles cuando se utiliza el objeto line~ solo, el cual desaparece cuando se añade el objeto samphold~. (Una desventaja de hacer muestra-y-sostenimiento para el cociente de la frecuencia, es que para frecuencias fundamentales muy bajas, los cambios pueden escucharse como pasos discretos. Así en situaciones donde la frecuencia fundamental es baja y la frecuencia central no necesita cambiarse muy rápidamente, puede ser mejor omitir el paso de muestra-y-sostenimiento.)

Los siguientes dos ejemplos demuestran el uso de los osciladores portadores con cruzamiento atenuado como parte de la técnica clásica de modulación de fase de dos operadores. El mismo oscilador de modulación se suma de manera separada a las fases de los dos cosenos. Los espectros resultantes pueden ser hechos para viajar arriba y abajo en frecuencia, pero debido a las complicadas relaciones de fase entre los picos vecinos en el espectro de modulación de fase, sin importar cómo alinee los dos espectros, usted nunca puede evitar cancelaciones de fase cuando se traslapan.

El generador PAF

El ejemplo F12.paf.pd (figura 6.18) es una realización del generador PAF, descrito en la sección 6.4. El control de la entrada especifica la frecuencia fundamental, la frecuencia central, y el ancho de banda, todo en “unidades” MIDI <no hay unidades MIDI en ese parche; se está refiriendo al parche F13.paf.control.pd>. Los primeros pasos tomados en la realización son dividir la frecuencia central por la fundamental (para obtener el cociente de la frecuencia central) y el ancho de banda por la fundamental para obtener el índice de modulación para el mecanismo de conformación de onda. Al cociente de la frecuencia central se le aplica muestra-y-sostenimiento de tal manera que esta se actualiza únicamente en períodos de la fundamental.

El oscilador uno (el objeto phasor~) corre a la frecuencia fundamental. Este es utilizado para controlar el objeto samphold~ que sincroniza las actualizaciones con el cociente del centro de frecuencia (etiquetado “C.F. relative to fundamental” en la figura), y para computar fases para los dos objetos cos~ que operan como se mostró antes en la figura 6.17.

La porción de conformación de onda del parche utiliza una sinusoide de medio período como una función de lectura (para compensar el doblaje de frecuencia debido a la simetría de la función de lectura). Para obtener medio ciclo de la función seno, multiplicamos la fase por 0.5 y sustraemos 0.25, de tal manera que el ajuste de la fase corre de -0.25 a 0.25, una vez cada período. Esto lee la mitad positiva del ciclo definido por el objeto cos~.

La amplitud de la mitad de la sinusoide se ajusta luego por un índice de modulación (el cual es precisamente el cociente del ancho de banda ω_b/ω). La tabla (“bell-curve”) contiene una curva gaussiana no normalizada, en muestras, desde -4 a 4 sobre 200 puntos (25 puntos por unidad), de tal manera que el centro de la tabla, en el punto 100, corresponde al pico central de la curva campana. Por fuera del intervalo -4 a 4 los valores de la curva gaussiana son prácticamente nulos.

La figura 6.19 muestra cómo se prepara la tabla de ondas gaussiana. Se requirió de un nuevo objeto de control:

until~: Cuando la entrada izquierda, “start”, se dispara, da como salida disparos secuenciales (sin que ningún tiempo transcurra entre ellos) iterativamente, hasta que la entrada derecha, “stop”, sea disparada. El mensaje “bang” de parada se debe originar de alguna manera desde la salida de los objetos until; de otra manera la salida enviará mensajes por siempre, congelando cualquier otro objeto que pudiera romper el lazo.

Tal como se utilizó aquí, un lazo conducido por un objeto until cuenta de 0 a 199, incluído. El lazo de conteo se mantiene por los objetos “f” y “+ 1”, cada uno de los cuales alimenta al otro. Pero debido a que la salida del objeto “+ 1” va a la entrada derecha de “f”, su resultado (una unidad más grande) únicamente surgirá del “f” la siguiente vez que sea disparado por “until”. De esta manera, cada disparo desde “until” incrementa el valor en uno.

Es importante el orden con el cual inicia el lazo: el objeto “t b b” de la parte superior (iniciales de “trigger bang bang”) deben enviar primero un cero a “f”, con el fin de darle inicio, y luego se ajusta el objeto until enviando disparos, incrementando el valor, hasta detenerse. Para detenerlo cuando el valor alcance 199, un objeto select examina el valor, y cuando ve que coincide, dispara la entrada “stop” del objeto until.

Entre tanto, por cada número de 0 a 199 que sale del objeto “f” creamos mensajes de pares ordenados para el objeto “tabwrite”. Primero, a la derecha, va el índice mismo, de 0 a 199. Luego, para la entrada izquierda, el primer objeto expr ajusta el índice para que vaya de -4 a 4 (iba previamente de 0 a 199) y el segundo evalúa la función gaussiana.

En este parche no tenemos totalmente solucionado el punto de actualizar el cociente del centro de frecuencia en los tiempos apropiados. Siempre que la frecuencia portadora sea cambiada en el paso muestra-y-sostenimiento se retrasa apropiadamente la actualización del cociente. Pero si, en lugar de la suma, la fundamental misma cambia abruptamente entonces, por una fracción de un período, la frecuencia del objeto phasor~ y el cociente quedan por fuera de sincronía. Pd no permite que la salida de samphold~ se conecte a la entrada de phasor~ nuevamente sin la inclusión de un retraso explícito (ver el capítulo siguiente) y no hay una manera sencilla de modificar el parche para resolver este problema.

Asumiendo que de alguna manera hicimos que la entrada del objeto phasor~ quedara en sincronía con sus propios puntos de aplicación de envolvente, deberíamos hacer entonces lo mismo con el cociente ancho-de-banda/fundamental en el lado derecho del parche. En el escenario corriente, sin embargo, no hay problema en actualizar ese valor continuamente.

Una solución práctica para este problema de actualización podría ser simplemente re-escribir el parche entero en C como un elemento Pd; esto también significa la utilización de mucho menos tiempo de CPU que el del parche mostrado, y es una solución más práctica en general -siempre y cuando no quiera experimentar embelleciendo o haciendo otros cambios en el algoritmo. Tales embellecimientos podrían incluír: adicionar un cambio inarmónico arriba y abajo en los parciales; permitiendo cambiar entre las actualizaciones suaves y de muestra-y-sostenimiento de la frecuencia central; adicionar controles de ganancia separados para parciales pares e impares introduciendo gravilla por la modulación irregular de la fase; permitiendo mezclas de dos o más funciones de conformación de onda; o fabricar ataques marcados alineando la fase del oscilador con los tiempos de un generador de envolvente de amplitud.

Un detalle final acerca de la amplitud está al orden: debido a que la amplitud de los parciales más fuertes decrece aproximadamente como $1/(1 + b)$, donde b es el índice de modulación, es a veces (pero no siempre) deseable corregir la amplitud de la salida multiplicándola por $1 + b$. Esto es únicamente una opción si b se actualiza de manera suave (como en este ejemplo), pero no cuando se hace con muestra-y-sostenimiento. Una situación en la cual esto es apropiado es en la simulacion de cuerdas pulsadas (al ajustar el centro de frecuencia a la fundamental, comenzando con índice de modulación alto y haciéndolo caer exponencialmente); debería ser apropiado escuchar la fundamental cayendo, no elevándose en su amplitud, cuando la cuerda decae.

Tablas de onda estiradas

En lugar de utilizar conformación de onda, la síntesis de formantes también es posible utilizando el estiramiento de tabla de ondas, como se demuestra en el ejemplo F14.wave.packet.pd (que no se muestra aquí). La técnica es escencialmente la del ejemplo B10.sampler.overlap.pd (que se describió en la sección 2.6), con un lector de coseno en lugar de la tabla de onda más general, pero con la adición de un control para ajustar el ciclo de tarea de las envolventes de amplitud. Las unidades están ajustadas para ser compatibles con las del ejemplo previo.

Ejercicios

1. Un tren de pulsos consiste en ventanas de Hann (cosenos elevados), de extremo a extremo, sin ningún espacio entre ellas. Cuál es el espectro resultante?

2. Para sintetizar un formante con centro de frecuencia de 2000 Hertz y fundamental de 300 Hertz, qué valores deberían tomar k y q (según la terminología de la figura 6.8)?

3. Cómo podría modificar el diagrama de bloques de la figura 6.8 para producir únicamente armónicos impares?