Capítulo 9

Análisis y resíntesis de Fourier

Entre las aplicaciones de los filtros discutidos en el capítulo 8, vimos cómo utilizar el heterodino, combinado con un filtro pasa-bajo, para encontrar la amplitud y la fase de un componente sinusoidal de una señal (pág.”257”). En este capítulo refinaremos esta técnica dentro de lo que se ha llamado el *análisis de Fourier*. En su forma más simple el análisis de Fourier, toma camo su entrada cualquier señal periódica (de período N) y sus salidas son amplitudes de valor complejo de sus N posibles componentes sinusoidales. Estas N amplitudes complejas pueden ser utilizadas, en teoría, para reconstruír exactamente la señal original. Esta reconstrucción es llamada *resíntesis de Fourier*.

En este capítulo comenzaremos por desarrollar la teoría del análisis y la resíntesis de Fourier de señales periódicas en muestras. Luego mostraremos cómo aplicar la misma técnica para señales arbitrarias, periódicas o no. Fnalmente desarrollaremos algunas aplicaciones comunes tales como el vocoder de fase.

9.1 Análisis de Fourier para señales periódicas

Suponga que $X[n]$ es una señal de valor complejo que se repite cada N muestras. (Continuamos con el uso de las señales de valor complejo en lugar de valor real para para simplificar las matemáticas.) Debido a su período N, los valores de $X[n]$ para $n = 0,...,N - 1$ determinan a $X[n]$ para todos los valores enteros de n.

Suponga, además, que $X[n]$ puede escribirse como una suma de sinusoides complejas de frecuencia 0, $2\pi/N$, $4\pi/N,...,2(N -1)\pi/N$. Estos son los parciales, comenzando con el cero, para una señal de período N. Nos detenemos en el término N-ésimo debido a que el siguiente deberá tener frecuencia 2π, equivalente a la frecuencia 0, que ya está en el listado.

Dados los valores de X, deseamos encontrar las amplitudes complejas de los parciales. Suponga que queremos el k-ésimo parcial, donde $0 \leq k < N$. La frecuencia de este parcial es $2\pi k/N$. Podemos encontrar su amplitud compleja modulando hacia abajo $2\pi k/N$ radianes de frecuencia por muestra, de tal manera que el k-ésimo parcial esté modulado a la frecuencia cero. Pasamos luego la señal por un filtro pasa-bajos con una frecuencia de corte tal que nada, excepto el parcial de frecuencia cero, permanezca. Podemos hacer esto en efecto promediando sobre todas las muestras; pero debido a que la señal se repite cada N muestras, este promedio complejo, es el mismo que el promedio de las primeras N muestras. En resumen, para medir un componente sinusoidal de una señal periódica, modúlela a DC y promedie luego sobre un período.

Sea $\omega = 2\pi/N$ la frecuencia fundamental para el período N y sea U el número complejo de magnitud unitaria con argumento ω:

$$U = \cos(\omega) + i\operatorname{sen}(\omega)$$

El k-ésimo parcial de la señal $X[n]$ es de la forma:

$$P_k[n] = A_k[U^k]^n$$

donde A_k es la amplitud compleja del parcial, y la frecuencia del parcial es:

$$\angle(U^k) = k\angle(U) = k\omega$$

Asumimos por el momento que la señal $X[n]$ puede ser escrita como la suma de n parciales, o en otras palabras:

$$X[n] = A_0[U^0] + A_1[U^1] + ... + A_{N-1}[U^{N-1}]^n$$

Por el argumento del filtrado heterodino anterior, esperamos ser capaces de medir cada A_k multiplicando por la sinusoide de frecuencia $-k\omega$ y promediando sobre la mitad de un período:

$$A_k = ([U^{-k}]^0 X[0] + [U^{-k}]^1 X[1] + ... + [U^{-k}]^{N-1} X[N -1])/N$$

Es esta una fórmula tan útil que tiene su propia notación. La transformada de Fourier de una señal $X[n]$, sobre N muestras está definida como:

$$\mathcal{FT}\{X[n]\}(k) = V^0 X[0] + V^1 X[1] + ... + V^{N-1} X[N -1]$$

donde $V = U^{-k}$. La transformada de Fourier es una función de la variable k, que es igual a N veces la amplitud de la entrada del k-ésimo parcial. Hasta aquí hemos tomado a k como un enetero pero la fórmula tiene sentido para cualquier valor de k si definimos V más generalmente como:

$$V = \cos(-k\omega) + i\,\mathrm{sen}(-k\omega)$$

donde como antes, $\omega = 2\pi/N$ es la frecuencia fundamental (angular) asociada con el período N.

9.1.1 Periodicidad de la transformada de Fourier

Si $X[n]$ es, como se dijo, una señal que se repite cada N muestras, la transformada de Fourier de $X[n]$ también se repite cada N unidadaes de frecuencia, esto es,

$$\mathcal{FT}\{X[n]\}(k + N) = \mathcal{FT}\{X[n]\}(k)$$

para todos los valores reales de k. Esto se sigue inmediatamente de la definición de la transformada de Fourier, ya que el factor

$$V = \cos(-k\omega) + i\,\mathrm{sen}(-k\omega)$$

no cambia cuando añadimos N (o múltiplos de N) a k.

9.1.2 Transformada de Fourier como síntesis aditiva

Consideremos ahora una señal arbitraria $X[n]$ que se repite cada N muestras. (Previamente asumimos que $X[n]$ se puede obtener como una suma de sinusoides y aún no hemos encontrado las $X[n]$ periódicas que pueden ser obtenidas de esa manera). Sea $Y[k]$ la transformada de Fourier de X para $k = 0,..., N - 1$:

$$Y[k] = \mathcal{FT}\{X[n]\}(k)$$

$$= [U^{-k}]^0 X[0] + [U^{-k}]^1 X[1] + ... + [U^{-k}]^{N-1} X[N -1]$$

$$= [U^0]^{-k} X[0] + [U^1]^{-k} X[1] + ... + [U^{-(N-1)}]^k X[N -1]$$

En la segunda versión reacomodamos los exponentes para mostrar que $Y[k]$ es una suma de sinusoides complejas, con amplitudes complejas $X[m]$ y frecuencias $-m\omega$ para $m = 0,..., N - 1$. En otras palabras, $Y[k]$ puede ser considerada como una serie de Fourier por derecho propio, cuyos m-ésimos componentes tienen fortaleza $X[-m]$. (La expresión $X[m]$ tiene sentido ya que X es una señal periódica. También podemos exprear la amplitud de los parciales de $Y[k]$ en términos de su transformada de Fourier. Igualando ambas obtenemos:

$$(1/N)\mathcal{FT}\{Y[k]\}(m) = X[-m]$$

Esto significa, en cambio, que $X[-m]$ se puede obtener por la suma de sinusoides con amplitudes $Y[k]/N$. Haciendo $n = -m$ tenemos:

$$X[n] = (1/N)\mathcal{FT}\{Y[k]\}(-n)$$

$$= [U^0]^n X[0] + [U^1]^n X[1] + ... + [U^{N-1}]^n X[N -1]$$

Esto indica que cualquier $X[n]$ periódica puede ser obtenida de esta manera como una suma de sinusoides. Es más, la fórmula muestra explícitamente cómo reconstruir $X[n]$ de su transformada de Fourier $Y[k]$, si conocemos su valor para los enteros $k = 0,..., N - 1$.

9.2 Propiedades de las transformadas de Fourier

En esta sección investigaremos qué sucede cuando tomamos la transformada de Fourier de una sinusoide (compleja). La más simple de éstas es "DC", la sinusoide especial de frecuencia cero. Luego de que derivemos la transformada de Fourier de esta, desarrollaremos algunas propiedades de las transformadas de Fourier que nos permitirán aplicar el resultado para cualquier otra sinusoide.

9.2.1 La transformada de Fourier de DC

Sea $X[n] = 1$ para todo n (que se repite para cualquier período entero que queramos $N > 1$). De la discusión precedente esperamos encontrar que

$$\mathcal{FT}\{X[n]\}(k) = \begin{cases} N & k = 0 \\ 0 & k = 0,..., N-1 \end{cases}$$

Sin embargo, usualmente necesitaremos conocer la respuesta para los valores no enteros de k, y para esto no se puede hace nada mejor que calcular el valor directamente:

$$\mathcal{FT}\{X[n]\}(k) = V^0 X[0] + V^1 X[1] + ... + V^{N-1} X[N -1]$$

donde V es, como antes, un número complejo de magnitud unitario con argumento $-k\omega$. Esto es una serie geométrica; siempre que $V \neq 1$, tenemos:

$$\mathcal{FT}\{X[n]\}(k) = (V^N - 1)/(V - 1)$$

Simetrizamos ahora la parate superior y la parte inferior de la misma manera que lo hicimos antes en la sección 7.3. Para esto, hacemos:

$$\xi = \cos(\pi k/N) - i\,\mathrm{sen}(\pi k/N)$$

de tal manera que $\xi^2 = V$. Luego factorizando las potencias de ξ apropiadas del numerador y el denominador, tenemos:

$$\mathcal{FT}\{X[n]\}(k) = (\xi^{N-1})[(\xi^N - \xi^{-N})/(\xi - \xi^{-1})]$$

Es fácil simplificar ahora el numerador:

$$\xi^{N} - \xi^{-N} = (\cos(\pi k) - i\mathrm{sen}(\pi k)) - (\cos(\pi k) + i\mathrm{sen}(\pi k)) = -2i\mathrm{sen}(\pi k)$$

y de manera similar para el denominador, tenemos:

$$\mathcal{FT}\{X[n]\}(k) = \big(\cos(\pi k(N-1)/N) - i\mathrm{sen}(\pi k(N-1)/N)\big)(\mathrm{sen}(\pi k)/(\mathrm{sen}(\pi k/N)$$

Sea que $V = 1$ o no, tenemos

$$\mathcal{FT}\{X[n]\}(k) = \big(\cos(\pi k(N-1)/N) - i\mathrm{sen}(\pi k(N-1)/N)\big)D_N(k)$$

donde $D_N(k)$, conocido como el *núcleo de Dirichlet*, está definido como

$$D_N(k) = \begin{cases} N & k = 0 \\ (\mathrm{sen}(\pi k)/(\mathrm{sen}(\pi k/N) & k \neq 0,\ -N < k < N \end{cases}$$

La figura 9.1 muestra la transformada de Fourier de $X[n]$, con $N = 100$. La transformada se repita cada 100 muestras, con un pico a $k = 0$, otro a $k = 100$, y así sucesivamente. La figura se esfuerza en mostrar el comportamiento tanto de la fase como de la magnitud utilizando una gráfica tridimensional proyectada sobre el papel. El término de la fase

$$\cos(\pi k(N-1)/N) - i\mathrm{sen}(\pi k(N-1)/N)$$

tiene la función de hacer dar vueltas a los valores de $\mathcal{FT}\{X[n]\}(k)$ alrededor

del eje k con un período de aproximadamente dos. El núcleo Dirichlet $D_N(k)$, mostrado en la figura 9.2, controla la magnitud de $\mathcal{FT}\{X[n]\}(k)$. Tiene un pico

de dos unidades de amplitud, alrededor de $k = 0$. Este está rodeado de *lóbulos laterales* de una unidad de ancho, alternando en signos y decreciendo gradualmente en magnitud mientras k se incrementa o disminuye partiendo del cero. El término de la fase rota casi π radianes cada vez que el núcleo de Dirichlet cambia de signo, de tal manera que el producto de los dos permanece aproximadamente en la misma mitad del plano complejo para $k > 1$ (y en la mitad opuesta para $k < -1$). La fase rota por casi 2π radianes sobre el pico desde $k = -1$ a $k = 1$.

9.2.2 Movimientos y cambios de fase

La sección 7.2 mostró cómo los cambios de tiempo de una señal cambia las fases de sus componentes sinusoidales, y la sección 8.4.3 mostró cómo al multiplicar una señal por una sinusoide compleja se cambian las frecuencias que la componen. Estos dos efectos tienen identidades correspondientes dentro de la transformada de Fourier.

Consideremos primero un cambio de tiempo. Si $X[n]$, como es usual, es una señal de valor complejo que se repite cada N muestras, hagamos que $Y[n]$ sea $X[n]$ retrasada d muestras:

$$Y[n] = X[n - d]$$

que se repite también cada N muestras según lo hace X. Podemos rducir la transformada de Fourier de $Y[n]$ de esta manera:

$$\mathcal{FT}\{Y[n]\}(k) = V^0Y[0] + V^1Y[1] + ... + V^{N-1}Y[N-1]$$

$$= V^0X[-d] + V^1X[-d+1] + ... + V^{N-1}X[-d+N-1]$$

$$= V^dX[-d] + V^{d+1}X[1] + ... + V^{d+N-1}X[N-1]$$

$$= V^d(V^0X[0] + V^1X[1] + ... + V^{N-1}X[N-1])$$

$$= V^d\mathcal{FT}\{X[n]\}(k)$$

(La tercera línea es sólo la segunda con los términos sumados en un orden diferente). De esta manera obtenemos la Fórmula del Cambio de Tiempo para las Transformadas de Fourier:

$$\mathcal{FT}\{X[n-d]\}(k) = (\cos(-dk\omega) + i\,\mathrm{sen}(-dk\omega))\mathcal{FT}\{X[n]\}(k)$$

La transformada de Fourier de $X[n-d]$ es un término de fases que multiplica la transformada de Fourier de $X[n]$. La fase se cambió por $-dk\omega$, una función lineal de la frecuencia k. Suponga ahora que cambiamos nuestra señal inicial $X[n]$ multiplicándola por una exponencial compleja Z^n con frecuencia angular α:

$$Y[n] = Z^nX[n]$$

$$Z = \cos(\alpha) + i\,\mathrm{sen}(\alpha)$$

La transformada de Fourier es:

$$\mathcal{FT}\{Y[n]\}(k) = V^0Y[0] + V^1Y[1] + ... + V^{N-1}Y[N-1]$$

$$= V^0X[0] + V^1ZX[1] + ... + V^{N-1}Z^{N-1}X[N-1]$$

$$= (VZ)^0X[0] + (VZ)^1X[1] + ... + (VZ)^{N-1}X[N-1]$$

$$= \mathcal{FT}\{X[n]\}(k - \alpha/\omega)$$

Obtenemos así la Fórmula del Cambio de Fase para las Transformadas de Fourier:

$$\mathcal{FT}\{(\cos(\alpha) + i\,\mathrm{sen}(\alpha))X[n]\}(k) = \mathcal{FT}\{X[n]\}(k - \alpha N/2\pi)$$

9.2.3 Transformada de Fourier de una sinusoide

Podemos usar la fórmula del cambio de fase anterior para encontrar la transformada de Fourier de cualquier sinusoide compleja Z^n con frecuencia α, haciendo simplemente $X[n] = 1$ en la fórmula y utilizando la transformada de Fourier para DC:

$$\mathcal{FT}\{Z^n\}(k) = \mathcal{FT}\{1\}(k - \alpha/\omega)$$

$$= [\cos(\Phi(k)) + i\,\mathrm{sen}(\Phi(k))]D_N(k - \alpha/\omega)$$

donde D_N es el núcleo de Dirichlet y Φ es un término de fase feo:

$$\Phi(k) = -\pi \bullet (k - \alpha/\omega) \bullet (N - 1)/N$$

Si la frecuencia de la sinusoide α es un entero múltiplo de la frecuencia fundamental ω, el núcleo de Dirichlet es cambiado a la izaquierda o a la derecha por un entero. En este caso los ceros del núcleo de Dirichlet están alineados con valores enteros de k, de tal manera que únicamente un parcial no es cero. Esto se ilustra en la figura 9.3 (parte a).

La parte b muestra el resultado cuando la frecuencia α cae en el medio de los dos enteros. Los parciales tienen amplitudes que caen aproximadamente en $1/k$ en ambas direcciones, medidas desde la frecuencia actual α. Que la energía se deba esparcir entre muchos parciales cuando después de todo iniciamos con una sola sinusoide, podría parecer sorprendente al principio. Sin embargo, como se muestra en la figura 9.4, la señal se repite en el período N que no concuerda con la frecuencia de la sinusoide. Como resultado, hay aquí una discontinuidad al comienzo de cada período y la energía se arroja sobre un rango amplio de frecuencias.

9.3 El análisis de Fourier de señales no periódicas

La mayoría de las señales no son periódicas, e incluso una periódica podría tener un período desconocido. De esta manera debería estar preparado para hacer el análisis de Fourier en señales sin hacer la confortable suposición de que la señal a analizar se repita a un período fijo N. Por supuesto podemos simplemente tomar N muestras de la señal y *hacerla* periódica; esto es escencialmente lo que hicimos en la sección previa, en la cual una sinusoide pura nos dio la complicada transformada de Fourier de la figura 9.3 (parte b).

Sin embargo, debería ser mejor para tener un resultado en el cual la respuesta a una sinusoide pura estuviera mejor localizada alrededor del valor correspondiente de k. Podemos lograr esto usando la técnica de envolvente que introdujimos en la figura 2.7 (página “38”). Aplicando esta técnica para el análisis de Fourier no solamente mejorará nuestro análisis, si no que también arroja nueva luz en el muestreador de lazo con envolvente del capítulo 2.

Dada una señal $X[n]$, periódica o no, definida en los puntos 0 a $N-1$, la técnica es envolver la señal antes del análisis de Fourier. La forma de la envolvente es conocida como la *función ventana*. Dada una función ventana $\omega[n]$, la *transformada de Fourier de ventana* es:

$$\mathcal{FT}\{\omega[n]X[n]\}(k)$$

Mucha tinta se ha derramado acerca de las funciones de ventana apropiadas para situaciones particulares, pero consideraremos aquí la más sencilla, llamada la función ventana de *Hann* (el nombre a veces se corrompe a “Hanning” en los círculos del DSP). La ventana de Hann es:

$$\omega[n] = 1/2 - (1/2)\cos(2\pi n/N)$$

Es fácil analizar el efecto de multiplicar una señal por la ventana de Hann antes de tomar la transformada de Fourier, debido a que la ventana de Hann puede escribirse como la suma de tres exponenciales complejas:

$$\omega[n] = 1/2 - (1/4)U^{n} - (1/4)U^{-n}$$

donde como antes, U es el número complejo de magnitud unitaria con argumento $2\pi/N$. Podemos ahora calcular la transformada de Fourier de ventana de una sinusoide Z^n con frecuencia angular α como antes. Las fases viene confusas y las reemplazaremos con aproximaciones simplificadas:

$$\mathcal{FT}\{\omega[n]Z^{n}\}(k)$$

$= \mathcal{FT}\{(1/2)Z^n - (1/4)(UZ)^n - (1/4)(U^{-1}Z)^n\}(k)$

$\approx [\cos(\Phi(k)) + i\,\mathrm{sen}(\Phi(k))]M(k - \alpha/\omega)$

donde el término fase es (aproximadamente):

$\Phi(k) = -\pi \bullet (k - \alpha/\omega)$

y la magnitud de la función es:

$M(k) = [(1/2)D_N(k) + (1/4)D_N(k + 1) + (1/4)D_N(k - 1)]$

La magnitud de la función $M(k)$ es graficada en la figura 9.5. Los tres componentes del núcleo de Dirichlet se muestran también separadamente.

El lóbulo principal de $M(k)$ es de cuatro armónicos de ancho, dos veces el ancho del lóbulo principal del núcleo de Dirichlet. Los lóbulos laterales, de otro lado, tienen una magnitud mucho más pequeña. Cada lóbulo lateral de $M(k)$ es una suma de tres lóbulos laterales de $D_N(k)$, uno atenuado por 1/2 y los otros de signo opuesto, atenuados por 1/4. No se cancelan perfectamente, pero lo hacen aproximadamente bien.

Los lóbulos laterales alcanzan su amplitud máxima carca a sus puntos medios, y podemos estimar sus amplitudes allí, usando la aproximación:

$D_N(k) = [N\mathrm{sen}(\pi k)]/(\pi k)$

Haciendo k = 3/2,5/2,... se obtienen las amplitudes de los lóbulos laterales, relativos a la altura de pico N de:

$2/(3\pi) \approx -13\mathrm{dB}$, $2/(5\pi) \approx -18\mathrm{dB}$, $2/(7\pi) \approx -21\mathrm{dB}$, $2/(9\pi) \approx -23\mathrm{dB}$,...

Los lóbulos laterales caen progresivamente más lento de tal manera que el décimo es atenuado únicamente en cerca de 30 dB y el 32º en cerca de -40 dB. De otro lado los lóbulos laterales de la ventana de Hann son atenuados en:

$2/(5\pi) - (1/2)[2/(3\pi) + 2/(7\pi)] \approx -32.30\mathrm{dB}$

y -42, -49, -54 y -59 dB para los siguientes cuatro lóbulos laterales.

Esto muestra que aplicar la ventana de Hann antes de tomar la transformada e Fourier nos permitirá aislar mejor los componentes sinusoidales. Si una señal tiene muchos componentes sinusoidales, los lóbulos laterales engendrados por cada uno, interferirán con el lóbulo principal de los demás. Reducir la amplitud de los lóbulos laterales reducirá esta interferencia.

La figura 9.6 muestra un análisis de Fourier en ventana de Hann de una señal con componentes sinusoidales. Ambos están separados por cerca de 5 veces la frecuencia fundamental ω, y para cada uno vemos claramente la forma de la transformada de Fourier de venana de Hann. Cuatro puntos del análisis de Fourier caen dentro del lóbulo principal $M(k)$ correspondientes a cada sinusoide. La amplitud y la fase de cada sinusoide individual están reflejadas en las de los picos (de cuatro puntos de ancho). Los cuatro puntos dentro de un pico los cuales caen en valores enteros de k están sucesivamente fuera de fase en casi medio ciclo.

Para resolver por completo los parciales de una señal, debemos escoger el tamaño N del análisis suficientemente grande de tal manera que $\omega = 2\pi/N$ no sea mayor que un cuarto de la frecuencia de separación que hay entre los parciales vecinos. Para una señal periódica por ejemplo los parciales están separados por la frecuencia fundamental. En el análisis para resolver totalmente los parciales, el período de análisis N debe ser por lo menos cuatro períodos de la señal.

En alguna aplicaciones esto funciona de tal forma que permite el traslape de los picos en tanto el centro de frecuencia de cada pico es aislado de todos los demás picos; en este caso la regla de los cuatro períodos puede relajarse a tres o incluso un poco menos.

9.4 Análisis de Fourier y reconstrucción de señales de audo

El análisis de Fourier puede utilizarse a veces para resolver los componentes sinusoidales de una señal de audio. Incluso cuando este no puede ir así de lelos, se puede separar una señal de audo en dos regiones de frecuencias, en el sentido de que para cada k el k-ésimo punto de la transformada de Fourier deberá estar afectado por únicamente por los componentes cercanos a la frecuencia nominal $k\omega$. Esto sugiere muchas operaciones interesantes que podemos ejecutar sobre una señal tomando su transformada de Fourier, transformando el resultado y reconstruyendo luego una señal nueva, transformada, partiendo de la transformación modificada.

La figura 9.7 muestra cómo llevar a cabo un análisis de Fourier, la moficación y la reconstrucción de una señal de audio. El primer paso es dividir la señal en *ventanas*, las cuales son segmentos de la señal, de N muestras cada una, usualmente con algún traslape. Cada ventana es luego conformada multiplicándola por una función ventana (Hann, por ejemplo). Luego la transformada de Fourier es calculada para los N puntos $k = 0,1,..., N-1$. (A veces es deseable calcular la transformada de Fourier para más puntos que estos, pero estos N puntos serán suficientes aquí.)

Los análisis de Fourier nos dan un arreglo bi-dimensional de número complejos. Sea H el tamaño del salto, el número de muestras que avanza cada ventana pasada la ventana previa. Luego para cada $m = ...,0,1,..$ la m-ésima ventana consiste de N puntos iniciando en el punto mH. El n-ésimo punto de la m-ésima ventana es $mH + n$. La transformada de Fourier de ventana es así igual a:

$$S[m, k] = \mathcal{FT}\{\omega[n]X[n - mH]\}(k)$$

Esta es una función del tiempo (m en unidades de H muestras) y de la frecuencia (k, como un múltiplo de la frecuencia fundamental ω. Fijando el número marco m y buscando la transformada de Fourier de ventana como una función de k:

$$S[k] = S[m, k]$$

Nos da una medida del espectro momentáneo de la señal $X[n]$. De otro lado, fijarndo un frecuencia k podemos buscar esta como el k-ésimo canal de una señal de N canales:

$$C[m] = S[m, k]$$

Desde este punto de vista la transformada de Fourier de ventana separa la señal original $X[n]$ en N regiones de frecuencia estrechas, llamadas *bandas*.

Habiendo computado la transformada de Fourier de ventana, aplicamos a continuación la modificación que queramos. En la figura la modificación es simplemente reemplazar la mitad superior del espectro por cero, el cual nos da una filtro pasa-bajos altamente selectivo. (Otras dos posibles modificaciones son la compansión de banda estrecha y el vocoder, que son descritas en las secciones siguientes.)

Finalmente reconstruimos la señal de salida. Para hacer esto aplicamos la inversa de la transformada de Fourier (etiquetada "iFT" en la figura). Como se mostró en la sección 9.1.2 esto puede ser hecho tomando otra transformada de Fourier, normalizando y moviendo de un tirón el resultado hacia atrás. En el caso de que la ventana reconstruida no vaya a cero suavemente en sus dos extremos, aplicamos la función ventana de Hannpor segunda vez. Al hacer esto a cada ventana sucesiva de la entrada, adicionamos entonces las salidas, usando el mismo traslape del análisis.

Si usamos la ventana de Hann y un traslape de cuatro (esto es, escoger N como un múltiplo de cuatro y cada espacio de ventana $H = N/4$ pasado el previo), podemos reconstruír la señal original fielmente omitiendo el paso de la "modificación". Esto debido a que iFT deshace el trabajo de FT , y de esta manera estamos multiplicando cada ventana por la función de Hann al cuadrado. La salida es así la entrada, multiplicada por la ventana de Hann elevada al cuadrado, y traslape adicionado por cuatro. Un examen fácil muestra que este se convierte en la constante 3/2, de tal manera que la salida es igual a la entrada multiplicada por un factor constante.

La habilidad para reconstruir la señal de entrada exactamente es útil porque algunos tipos de modificaciones pueden ser hechas por grados, y así la salida se puede hacer variar suavemente entre la entrada y alguna modificación de esta.

9.4.1 Compansión de banda estrecha

Un *compansor* es una herramienta que amplifica una señal con una ganancia variable, dependiendo de la medida de amplitud de la señal. El término es una contracción de "compresor" y "expansor". La ganancia de un compresor decrece mientras el nivel de entrada se incrementa, de tal manera que el *rango dinámico*, esto es, la variación general del nivel de la señal, se reduce. Un expansor hace lo inverso, incrementando el rango dinámico. Frecuentemente la ganancia depende no sólo del nivel de señal inmediato, si no de su historia; por ejemplo, la velocidad de cambio podría estar limitada o debería ser un retraso de tiempo.

Utilizando el análisis de Fourier y la resíntesis, podemos hacer compansión individualmente en canales de banda estrecha. Si $C[m]$ es tal banda, le aplicamos una ganancia $g[m]$, para tener $g[m]C[m]$. Aunque $C[m]$ es un número complejo, la ganancia es número real no negativo. En general la ganancia debería ser una función no solamente de $C[m]$ si no también de cualquiera o de todos las muestras previas en el canal: $C[m - 1]$, $C[m - 2]$, y así sucesivamente. Consideraremos aquí la situación más sencilla donde la ganancia es simplemente una función de la magnitud de la muestra corriente $|C[m - 1]|$.

El parche diagramado en la figura 9.8 muestra una aplicación muy útil de la compansión, llamada *compuerta de ruido*. Aquí la ganancia $g[m]$ depende de la amplitud del canal $C[m]$ y del ruido de fondo que es una función f del número de canal k. Para claridad aplicaremos la frecuencia k suscrita a la ganancia, escrita ahora como $g[m, k]$, y para la transformada de Fourier de ventana $S[m, k] = C[m]$. la ganancia está dada por:

$$g[m, k] = \begin{cases} 1 - f[k]/|S[m, k]| & |S[m, k]| > f[k] \\ 0 & \text{en otros casos} \end{cases}$$

Siempre que la magnitud $S[m, k]$ sea menor que el umbral $f[k]$ la ganancia es cero y así la amplitud $S[m, k]$ se reemplaza por cero. De otra manera, multiplicar la amplitud por $g[m, k]$ reduce la magnitud hacia abajo para $|S[m, k]| - f[k]$. Ya que la ganancia es un número real no negativo, la fase se preserva.

En la figura la ganancia se computa como una función de umbral de la razón $x = f[k]/|S[m, k]|$ de la magnitud de la señal para el ruido de fondo; el umbral es $g(x) = 1\ 1/x$ cuando $x < 1$ y cero en los demás casos, aunque otras funciones de umbral pueden ser sustituidas fácilmente.

Esta técnica es útil para remover el ruido de un sonido grabado. Medimos o suponemos valores del ruido de fondo $f[k]$. Debido al diseño de la función de ganancia $g[m, k]$, solamente las amplitudes que están por encima del ruido de fondo alcanzan la salida. Ya que esto es hecho sobre bandas de frecuencias estrechas, a veces es posible eliminar la mayor parte del ruido incluso mientras la señal misma, en los rangos de frecuencia donde es más fuerte que el ruido de fondo, es preservada en su mayor parte.

La técnica también es útil como preparación para aplicar una operación no lineal, tal como la distorsión, a un sonido. Es usualmente mejor distorsionar sólo las frecuencias más sobresalientes del sonido. Al sustraer el sonido que se ha extraído por la compuerta de ruido, del original, se obtiene una señal residual que puede pasar sin distorsión.

9.4.2 Estampado de timbre (vocoder clásico)

Una segunda aplicación del análisis de Fourier y la resíntesis es un filtro variable en el tiempo capaza de hacer que un sonido evolucione espectralmente sobre la envolvente de otro. Esto es ampliamente conocido en los círculos de la música electrónica como un vocoder, llamado así, no con mucha exactitud, después del dispositivo de análisis/síntesis vocal original de los Laboratorios Bell. La técnica descrita aquí es llamada más acertadamente *estampado de timbre*. Se utilizan dos señales de entrada, una para ser filtrada, y la otra para controlar el filtro vía su envolvente espectral variante en el tiempo. La transformada de Fourier de ventana es utilizada tanto en la entrada de la señal de control para estimar su envolvente espectral, como sobre la entrada del filtro, para aplicarlo.

Un diagrama de bloque para el estampado de timbre es mostrado en la figura 9.9. Como en el ejemplo previo, el estampado del timbre actúa multiplicando la transformada de Fourier de ventana de valor complejo del filtro de entrada, por números reales no negativos, cambiando por lo tanto sus magnitudes pero dejando sus fases intactas. El giro aquí es que simplemente reemplazamos las magnitudes del $|S[m, k]|$ original, con las magnitudes otenidas de la entrada de control (llamémosla $|T[m, k]|$). La ganancia necesaria deberá ser,

$$g[m, k]| = |T[m, k]|/|S[m, k]|$$

En la práctica es mejor limitar la ganancia a algún valor máximo (que podría depender de la frecuencia) ya que de otra manera los canales no contienen más que ruido, lóbulos laterales, o incluso el error de truncado podría volverse audible. De tal amnera que una función de límite apropiada se aplica a la ganancia antes de utilizarla.

9.5 Fase

Hasta aquí hemos operado las señales alterando las magnitudes de sus transformadas de Fourier de ventana, pero dejando las fases intactas. Las magnitudes codifican la envolvente espectral del sonido. Las fases de otro lado, codifican frecuencia y tiempo, en el sentido de que el cambio de fase desde una ventana a otra diferente se acumula, en el tiempo, de acuerdo a la frecuencia. Para hacer una transformación que permita el control independiente sobre la frecuencia y el tiempo, se requiere analizar y reconstrír la fase.

En los ejemplos de análisis/síntesis de la sección previa, la fase de la salida es copiada directamente de la fase de la entrada. Esto es apropiado cuando la señal de salida se corresponde con la señal de entrada. A veces se requieren modificaciones del tiempo, por ejemplo para hacer estiramiento o contracción dl tiempo. Alternativamente, la fase de la salida podría depender de más de una entrada, por ejemplo para modificar un sonido y obtener otro.

La figra 9.10 muestra cómo la fase de la transformada de Fourier cambia de ventana a ventana, dando una sinusoide compleja como entrada. La frecuencia de la sinusoide es $\alpha = 3\omega$, de tal amnera que el pico en la transformada de Fourier está a $k = 3$. Si la fase inicial es ϕ, entonces las fases vecinas se pueden llenar así:

$$\begin{array}{lll} \angle S[0,2] = \phi + \pi & \angle S[0,3] = \phi & \angle S[0,4] = \phi + \pi \\ \angle S[1,2] = \phi + H\alpha + \pi & \angle S[1,3] = \phi + H\alpha & \angle S[1,4] = \phi + H\alpha + \pi \\ \angle S[2,2] = \phi + 2H\alpha + \pi & \angle S[2,3] = \phi + 2H\alpha & \angle S[2,4] = \phi + 2H\alpha + \pi \end{array}$$

Esto da una forma excelente de estimar la frecuencia α: seleccionar cualquier canal cuya amplitud esté dominada por la sinusoide y sustraiga dos fases sucesivas para obtener $H\alpha$:

$$H\alpha = \angle S[1,3] - \angle S[0,3]$$

$$\alpha = (\angle S[1,3] - \angle S[0,3] + 2p\pi)/H$$

donde p es un entero. Hay H posibles frecuencias, espaciadas por $2\pi/H$. Si estamos utilizando un traslape de 4, esto es $H = N/4$, las frecuencias están espaciadas por $8\pi/N = 4\omega$. Felizmente, este es el ancho del lóbulo principal para la ventana de Hann, de tal manera que no hay más que un valor posible de α que pueda explicar cualquier diferencia de fase medida dentro del lóbulo principal, de un pico. El valor correcto de p para escoger es aquel que da la frecuencia más cercana a la frecuencia nominal del canal, $k\omega$.

Cuando computamos fases para sintetizar una señal nueva o modificada, queremos mantener las relaciones de fase apropiadas entre las sucesivas ventanas de resíntesis y también, simultáneamente, entre los canales adyacentes. Sin embargo estos dos conjuntos de relaciones no son siempre compatibles. Haremos de nuestra primera obligación honrar las relaciones entre las ventanas sucesivas de resíntesis, y nos ocuparemos de las relaciones de fase entre los canales, después.

Suponga que queremos construir el m-ésimo espectro $S[m, k]$ para resíntesis (habiendo ya construido el previo, $m - 1$). SUponga que queremos que la las relaciones de fase entre las ventanas $m - 1$ y m sean aquellas de la señal $x[n]$, pero que las fases de la ventana $m - 1$ podrían haber venido de algún otro sitio y no podemos asumir que estén alineadas como lo queremos.

Para encontrar qué tanto difiere la fase de cada canal, del canal previo, realizamos dos análisis de la señal $x[n]$, separadas por el mismo salto de tamaño H que estamos utilizando para la resíntesis:

$$T[k] = \mathcal{FT}(W[n]X[n])(k)$$

$$T'[k] = \mathcal{FT}(W[n]X[n + H])(k)$$

La figura 9.11 muestra el proceso de acumulación de fase, en el cual cada fase a la salida depende de la fase de salida previa, y de la diferencia de la fase para dos análisis de ventana a la entrada. La figura 9.12 ilustra la relación de fase en el plano complejo. La fase de la nueva salida $S[m, k]$ debe ser la de la previa, más la diferencia entre las fases de los dos análisis:

$$\angle S[m, k] = \angle S[m - 1, k] + (\angle T'[k] - \angle T[k])$$

$$= \angle(S[m - 1, k]T'[k]/T[k])$$

Aquí utilizamos el hecho de que al multiplicar o dividir dos números complejos obtenemos la suma o diferencia de sus argumentos.

Si la magnitud deseada es un número real a, debemos ajustar entonces $S[m, k]$ para:

$$S[m, k] = a \bullet |S[m - 1, k]T'[k]/T[k]|^{-1} \bullet S[m - 1, k]T'[k]/T[k]$$

Las magnitudes del segundo y del tercer término se cancelan, y así la magnitud de $S[m, k]$ se reduce a a; el primero de los dos términos son números reales de tal manera que su argumento está controlado por el último término.

Si queremos finalizar con la magnitud del espectro T de la misma manera, podemos hacer $a = |T'[k]|$ y simplificar:

$$S[m, k] = |S[m - 1, k]/T[k]|^{-1} \bullet S[m - 1, k]T'[k]/T[k]$$

9.5.1 Relaciones de fase entre los canales

En el esquema anterior, la fase de cada $S[m, k]$ depende únicamente del valor previo del mismo canal. Las relaciones de fase entre los canales vecinos son dejadas al azar. Estas a veces trabajan bien, pero a veces la incoherencia de los canales vecinos da lugar a un efecto de chorus no buscado. Podríamos querer idealmente que $S[m, k]$ y $S[m, k + 1]$ tuvieran la misma relación de fase que hay entre $T'[k]$ y $T[k]$.

Estas $2N$ ecuaciones para N fases en general no tendrá solución, pero podemos alterar la ecuación anterior de $S[m, k]$ de tal manera que siempre que ocurra una solución en el sobre-apretado sistema de ecuaciones, el algoritmo de reconstrucción se aloje en la solución. Este enfoque es llamado *aseguramiento de fase* [Puc95b], y tiene la virtud de la simplicidad, aunque están disponibles técnicas más sofisticadas [DL97].

La relación de fase de salida deseada, en el cuadro m -1, es:

$$\angle T[k + 1] - \angle T[k] = \angle S[m - 1, k + 1] - \angle S[m - 1, k]$$

o re-ordenando:

$$\angle\{S[m - 1, k + 1]/T[k + 1]\} = \angle\{S[m - 1, k]/T[k]\}$$

En otras palabras, la fase del cociente S/T no debería depender de k. Con esto en mente podemos re-escribir la fórmula de recursión para $S[m, k]$:

$$S[m, k] = |R[k]|^{-1} \bullet R'[k]T'[k]$$

con

$$R[k] = (\overline{T[k]} \bullet S[m - 1, k])/|S[m - 1, k]|$$

y según la ecuación previa, los $R[k]$ deberán estar todos en fase. El truco es ahora reemplazar $R[k]$ para cada k con la suma de tres vecinos. El cómputo es entonces:

$$S[m, k] = |R'[k]|^{-1} \cdot R'[k]T'[k]$$

con

$$R'[k] = R[k + 1] + R[k] + R[k - 1]$$

Si los canales están ya en la relación de fase correcta, esto no tiene efecto (la fase resultante será la misma que si se utilizara únicamente $R[k]$.) Pero en general, la suma compartirá dos términos en común con su vecino en $k + 1$:

$$R'[k + 1] = R[k + 2] + R[k + 1] + R[k]$$

de tal manera que R' tenderá a apuntar más en la misma dirección que R. Aplicando esto iterativamente, eventualmente todo se alineará R' a la misma fase, en tanto las relaciones de fase entre los espectros medidos de T y T' lo permitan.

9.6 Bashing de fase

En la sección 2.3 en el muestreo con envolvente vimos cómo fabricar una onda periódica desde un sonido grabado, y así prestar el timbre del sonido original pero ejecutandolo a una altura específica. Si a la ventana del sonido grabado se le aplica precesión en el tiempo, el timbre resultante varía en la imitación del sonido grabado.

Aparece un problema importante, y es que si tomamos formas de onda de diferentes ventanas de una muestra (ode diferentes muestras), no hay garantía de que las dos fases coincidan. Si no lo hacen el resultado es puede ser feo, ya que los cambios aleatorios de fase son escuchados como fluctuaciones de la frecuencia. Esto puede ser corregido utilizando análisis de Fourier y resíntesis [Puc05].

La figura 9.13 muestra una manera sencilla de utilizar el análisis de Fourier para alinear las fases en una serie de ventanas de una grabación. Simplemente tomamos el FFT de la ventana y luego ajustamos cada fase a cero para valores pares de k y a π cuando son impares. La fase en el centro de la ventana es, de esta manera, cero, tanto para los valores pares como para los valores impares de k. Para ajustar las fases (los argumentos de las amplitudes complejas del espectro) en la forma deseada, primero encontramos la magnitud, la cual puede ser considerada un número complejo con argumento cero. Luego, al multiplicar por $(-1)^k$ se ajusta la amplitud de tal forma que esta es positiva y negativa de manera alternada. A continuación tomamos la transformada inversa de Fourier sin siquiera molestarse en tomar la ventana de nuevo en el camino de regreso; probablemente querremos aplicar una envolvente de ventana más tarde, de todas maneras, tal como se mostró en la figura 2.7. Los resultados pueden combinarse con las técnicas de modulación del capítulo 6 produciendo herramientas poderosas para voces y otras síntesis imitativas.

9.7 Ejemplos

Análisis de Fourier y resíntesis en Pd

El ejemplo I01.Fourier.analysis.pd (figura 9.14, parte a) demuestra el cómputo de la transformada de Fourier de una señal de audio utilizando el objeto fft~:

fft~: Transformada de Fourier Rápida. Las dos entradas toman señales de audio que representan la parte real y la parte imaginaria de una señal de valor complejo. El tamaño N de ventana está dado por el tamaño de bloque de Pd. Una transformada de Fourier está hecha para cada bloque.

La transformada de Fourier Rápida [SI03] reduce el costo computacional del análisis de Fourier en Pd a sólo de 5 a 15 objetos osc~ en configuraciones típicas. El algoritmo en su forma más simple toma N como una potencia de dos, lo cual es también (normalmente) una restricción en el tamaño de los bloques en Pd.

El ejemplo I02.Hann.window.pd (figura 9.14, partes b y c) muestra cómo controlar el tamaño del bloque utilizando un obejeto block~, cómo aplicar una ventana de Hann, y una versión diferente de la transformada de Fourier. La parte (b) muestra la invocación de una sub-ventana la cual se muestra, a su vez, en la parte (c). Los nuevos objetos son:

rfft~: transformada de Fourier Rápida real. Se asume que la parte imaginaria de la entrada es cero. Ünicamente los dos primeros $N/2 + 1$ canales de salida se llenan (los otros son determinados por simetría). Este toma la mitad del tiempo de cómputo del objeto fft~, más general.

tabreceive~: da salidas repetidas del contenido de la tabla de ondas. Cada bloque de cómputo da a la salida las mismas primeras N muestras de la tabla.

En este ejemplo, la tabla “$0-hann” contiene una función ventana de Hann de longitud 512, de acuerdo con el tamaño especificado para los bloques. La señal que va a ser analizada aparece (desde el parche padre) vía el objeto inlet~. Las amplitudes de los canales (la salida del objeto rfft~) son reducidas a magnitudes de valor real: las partes real e imaginaria son elevadas al cuadrado y sumadas, y el resultado se hace pasar por el objeto sqrt~. Finalmente la magnitud es escrita (controlada por una conexión que no se muestra en la figura) vía tabwrite~ a otra tabla, “$0-magnitude”, para graficar.

El ejemplo I03.resynthesis.pd (figura 9.15) muestra cómo analizar y resintetizar una señal de audio siguiendo la estrategia de la figura 9.7. Como antes, hay una sub-ventana para realizar el trabajo con un tamaño de bloque apropiado para la tarea; la figura muesra únicamente la sub-ventana. Necesitamos un objeto nuevo para la transformada inversa de Fourier:

rifft~: transformada inversa de Fourier Rápida, real. Utilizando $N/2 + 1$ puntos de su entrada (tomados como un par real/imaginario), y asumiendo los valores apropiados para los otros canales por simetría, se reconstrye una salida de valor real. No se ha hecho ninguna normalización, de tal manera que un par rfft~/rifft~ junto, resulta con una ganancia de N. El objeto ifft~ está disponible igualmente, y computa un inverso no normalizado para el objeto fft~, reconstruyendo una salida de valor complejo.

El objeto block~, en la sub-ventana, es invocado con un segundo argumento que especifica un traslape de 4. Esto hará que la sub-ventana corra cuatro veces cada $N = 512$ muestras, en intervalos regulares de 128 muestras. El objeto inlet~ realiza el búfer necesario y la reorganización de las muestras de tal manera que su salida siempre da las últimas 512 muestras de la entrada, en orden. En la otra dirección, el objeto outlet~ adiciona segmentos de sus cuatro entradas previas para llevarlas al esquema de traslape sumado en la figura 9.7.

Los bloques de 512 muestras están multiplicados por la ventana de Hann tanto en la entrada como en la salida. Si los objetos rfft~ y rifft~ fueron conectados entre ellos sin ninguna modificación, la salida deberá reconstruír la entrada fielmente.

Sin embargo, una modificación es aplicada: cada canal está multiplicado por una ganacia (positiva y de valor real). La amplitud de valor complejo para cada canal es escalada separadamente, multiplicando las partes real e imaginaria por la ganancia. La ganancia (que depende del canal) viene de otra tabla, llamada “$0-gain”. El resultado es un filtro de ecualización gráfica; aplicando el ratón en la ventana gráfica de esta tabla, puede diseñar curvas de ganancia de frecuencia.

Hay un retraso inherente introducido al utilizar block~ para incrementar el tamaño del bloque (pero no lo hay si es utilizado, como se muestra en el capítulo 7, para reducir el tamaño del bloque relativo a la ventana padre.) El retraso puede medirse desde la entrada hasta la salida del sub-parche, y es igual a la diferencia de los dos tamaños de los bloques. En este ejemplo el retraso del búfer es de 512-64=448 muestras. El retraso del bloque no depende del traslape, sólo de los tamaños del bloque.

Compansión de banda estrecha: reducción de ruido

El ejemplo I04.noisegate.pd (figura 9.16) muestra un ejemplo de compansión de banda estrechautilizando análisis/resíntesis de Fourier. (Esta es una realización del diagrama de bloques de la figura 9.8.) La parte (a) de la figura muestra una configuración de filtro similar al del ejemplo previo, excepto que la ganancia para cada canal es ahora una función de la magnitud del canal.

Para cada k, si $s[k]$ denota la potencia en el canal k y $m[k]$ es el nivel de máscara (un nivel presumiblemente algo mayor que la potencia del ruido para el canal k), entonces la ganancia en el canal k está dada por

$$\begin{cases} \sqrt{(s[k]-m[k])/s[k]} & s[k] > m[k] \\ 0 & \text{para los demás valores} \end{cases}$$

La potencia en el k-ésimo canal es reducida de esta manera por $m[k]$ si es posible, y de otra manera, es reemplazada por cero.

La máscara misma es el producto del promedio de ruido medido en cada canal, el cual está contenido en la tabla “$0-mask”, multiplicado por un valor llamado “mask-level”. El ruido promedio es medido en un sub-parche (pd calculate-mask), cuyos contenidos se muestran en la parte (b) de la figura. Para computar la máscara estamos utilizando dos objetos nuevos:

bang~: envía un disparo con anticipación a cada bloque de cómputo. El disparo aparece en el tiempo lógico de la primera muestra en cada bloque (el primer tiempo lógico cuyo cómputo de control afecta a ese bloque y no al previo), siguiendo el esquema mostrado en la figura 3.2.

tabsend~: el objeto complementario de tabreceive~, copia repetidamente su entrada a los contenidos de una tabla, afectando las primeras N muestras de la tabla.

El proceso de promediar la potencia es iniciado al enviar un tiempo de duración en milisegundos para “hacer la máscara”. El parche computa el número equivalente de bloques b y genera una secuencia de pesos: 1, 1/2, 1/3,..., 1/b por los cuales cada una de las potencias de los bloques que siguen a continuación es promediada en lo que contenga la tabla de la máscara en el bloque previo. Al final de los b bloques la tabla contiene los promedios de igual peso de todas las medidas de la potencia b. Después de eso, el peso del promedio de las nuevas medidas de potencias es cero, de tal manera que el promedio medido detiene su evolución.

El uso de este parche para la supresión clásica de ruido requiere por lo menos unos pocos segundos de ruido grabado sin que “la señal” esté presente. Esta es reproducida en el parche, y su duración es enviada a “make-mask”, de tal manera que la tabla “$0-mask” contiene el promedio de la potencia medida del ruido para cada canal. Luego, asumiendo que la parte ruidosa de la señal raramente excede 10 veces su potencia promedio (por ejemplo), “mask-level” es enviada a 10, y la señal a la que se le suprimirá el ruido es enviada a través de la parte (a) del parche. El ruido se habrá ido en su mayor parte, pero aquellos canales en los cuales la señal excede 20 veces la potencia del ruido serán atenuados únicamente en 3 dB, y los canales con mayor potencia serán atenuados progresivamente menos. (Por supuesto, la supresión actual del ruido podría no ser la aplicación más interesante del parche, uno podría tratar de enmascarar una señal cualquiera, de otra.)

Estampado de timbre (“vocoder”)

El ejemplo I05.compressor.pd (figura 9.17) es otro compansor de canal presentado como preparación para el ejemplo I06.timbre.stamp.pd, que examinaremos a continuación. Este es una realización del estampado de timbre de la figura 9.9, ligeramente modificado.

Hay dos entradas, una a la izquierda, para ser filtrada (y cuya transformada de Fourier es utilizada para resíntesis después de modificar las magnitudes), y otra a la derecha que actúa como una fuente de control. En términos generales, si las dos magnitudes son $f[k]$ para la entrada de filtro y $c[k]$ para la fuente de control, sólo “blanquearemos” la entrada de filtro, multiplicándola por $1/f[k]$, y luego estampamos las magnitudes de control sobre el resultado, al multiplicarlas por $c[k]$. En la práctica, debemos limitar la ganancia a un valor máximo razonable. En el parche esto se hace limitando el factor de blanqueamiento $1/f[k]$ a un valor máximo especificado utilizando el objeto clip~. El límite está controlado por el parámetro “squelch”, que es elevado al cuadrado y dividido por 100 para mapear los valores de 0 a 100 a un rango útil.

Otro esquema posible es limitar la ganancia después de la formación del cociente $c[k]/f[k]$. La limitación de la ganancia puede ser en cualquier caso dependiente de la frecuencia. A veces es conveniente elevar la ganancia a una potencia p entre 0 y 1; si es 1 se trata de un estampado de timbre y si es 0, pasa la entrada del timbre sin cambios; los valores intermedios dan una suave interpolación entre ambos.

Ondulación de tiempo de vocoder de fase

El vocoder de fase usualmente se refiere a la técnica general de pasar de un canal de amplitudes (complejas) a pares consistentes en magnitudes (reales) y velocidades de precesión de fase (“frecuencias”), y al contrario, como se describió en la figura 9.11 (sección 9.5). En el ejemplo I07.phase.vocoder.pd (figura 9.18) utilizamos esta técnica con el ánimo específico de aplicar estiramiento o contracción del tiempo de un sonido grabado, con control de tiempo real. Es decir, controlamos en cualquier momento en el tiempo real, la localización del sonido grabado que estamos escuchando. Son utilizados dos nuevos objetos:

lrshift~: cambia un bloque de izquierdo o derecho (según el argumento de creación). Si el argumento es positivo, cada bloque de la salida es la entrada cambiada en ese número de espacios a la derecha, llenando con ceros, según se requiera, a la izquierda. Un argumento negativo cambia a la izquierda, llenando con ceros a la derecha.

q8_rsqrt~: raíz cuadrada rápida y aproximadamente recíproca. La salida del recíproco de la raíz cuadrada es su entrada, bien para una aproximación de una parte en 256, utilizando mucho menos cómputo que la de una raíz cuadrada de total precisión y recíproca.

El proceso comienza con un sub-parche, pd read-windows, que tiene como salida dos bloques de ventana de Hann de un sonido grabado, uno “atrás” y el otro “adelante”, con 1/4 de ventana adelantada en la grabación. La ventana mostrada utiliza las dos salidas del sub-parche para guiar el cambio de amplitud y fase de cada canal en su propia salida.

Los dos objetos tabreceive~ de la parte superior llaman al bloque previo de amplitudes complejas enviado al objeto rifft~ de la parte inferior, correspondiente al $S[m-1, k]$ de la discusión de la sección 9.5. El parche como un todo computa $S[m, k]$ y luego hace TF inversa de ventana de Hann, para la salida.

Después de normalizar $S[m-1, k]$, su complejo conjugado (el inverso normalizado) es multiplicado por la transformada de Fourier de la ventana de “atrás” $T[k]$, dando el producto $R[k]$ de la página “283”. A continuación, dependiendo del valor del parámetro “lock”, el valor computado de $R[k]$ es reemplazado condicionalmente con la versión de aseguramiento de fase $R'[k]$. Esto se hace utilizando los objetos lrshift~, cuyas salidas se adicionan a $R[k]$, si “lock” es ajustado a uno, o de la otra manera, si es cero. El resultado es luego normalizado y multiplicado por la transformada de Fourier de ventana de Hann de la ventana “delantera” ($T'[k]$), para dar $S[m, k]$.

Otras tres aplicaciones de síntesis/resíntesis de Fourie, que no se ilustran aquí, están provistas en los ejemplos de Pd. La primera, el ejemplo I08.pvoc.reverb.pd muestra cómo fabricar un vocoder de fase cuya salida recircula como en un reverberador, excepto que los canales individuales se han reemplazado por la entrada cuando esta es más poderosa de lo que es cuando ya está recirculando. El resultado es un efecto de revereración de sonido más coherente, que el que se logra con la forma clásica utilizando líneas de retrasos.

El ejemplo I09.sheep.from.goats.pd demuestra la técnica (imperfecta) para separar señales afinadas de las señales ruidosas, canal por canal, basado en la coherencia de fase que deberíamos esperar de una sinusoide en ventana de Hann. Si los tres canales adyacentes están desfasados aproximadamente en π radianes uno del otro, son juzgados para belong un pico sinusoidal. Los canales belonging los picos sinusoidales son reemplazados con ceros para extraer la porción ruidosa de la señal, o todos los demás son reemplazados por cero para obtener la porción sinusoidal.

El ejemplo I10.phase.bash.pd nos regresa al muestreador de lazo de tabla de ondas de la figura 2.7, y muestra cómo alinear las fases de la muestra de tal amnera que todos los componentes de la señal tengan fase cero en los puntos 0, N, $2N$, y sucesivamente. De esta manera dos copias de un muestreador de lazo ubicadas N muestras aparte pueden hacer cruzamiento atenuado coherentemente. Se puede fabricar una versión sintética, afinada del archivo de sonido original, utilizando cruzamiento atenuado en cadena daisy.

Ejercicios

1. Una señal $x[n]$ es 1 para $n = 0$ y cero para cualquier otro valor (un impulso). Cuál es su transformada de Fourier (N puntos) como una función de k?

2. Asumiendo además que N es un número par, cómo es la transformada de Fourier si $x[n]$ es 1 en $n = N/2$ en lugar de serlo en $n = 0$?

3. Para qué valores enteros de $k = 0$ es no cero la transformada de Fourier de N puntos de ventana de Hann?

4. Con el fin de realizar el análisis de Fourier de un sonido periódico de 100 Hertz (a una velocidad de muestras de 44100 Hertz) utilizando una ventana de Hann, qué valor de N se necesitará para resolver completamente los parciales

del sonido (en el sentido de tener picos que no se traslapen en el espectro)?

5. Suponga que una transformada de Fourier de N puntos está hecha sobre una sinusoide compleja de frecuencia de 2.5ω donde $\omega = 2\pi/N$ es la frecuencia fundamental. Qué porcentaje de la energía de la señal cae sobre el lóbulo principal, en los canales $k = 2$ y $k = 3$? Si a la señal se le aplica ventana de Hann, qué porcentaje de la energía está en el lóbulo principal (el cual está entonces en los canales 1 a 4)?