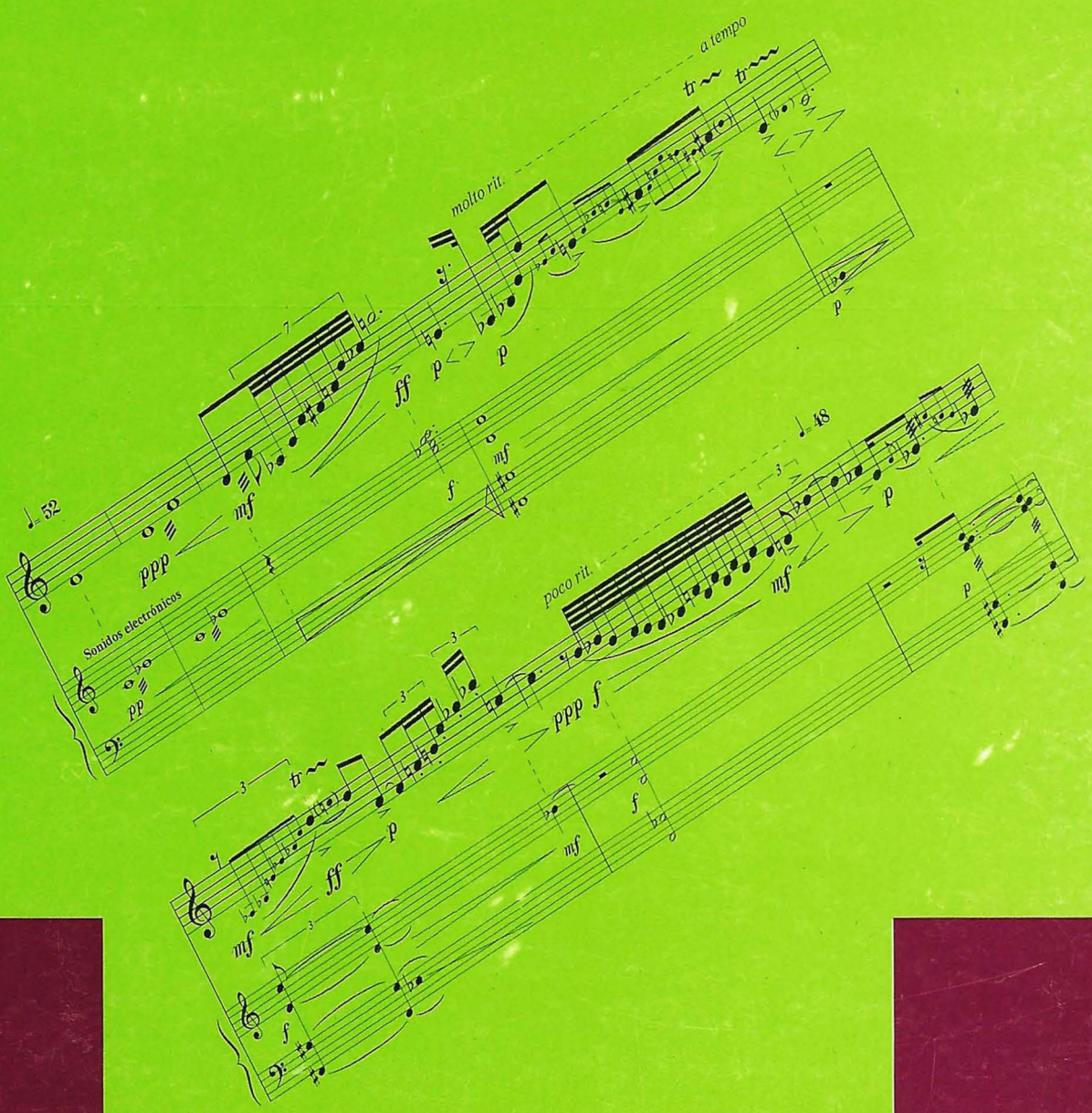
CUADERNO DE ESTUDIO Nº 5



ALTURA - TIMBRE - ESPACIO

PABLO CETTA • CARMELO SAITTA • EDUARDO MOGUILLANSKY • EDUARDO CHECCHI • OSCAR PABLO DI LISCIA • MATIAS GIULIANI • ROBERTO AZARETTO • GUSTAVO GARCIA NOVO

PONTIFICIA UNIVERSIDAD CATÓLICA ARGENTINA
"Santa María de los Buenos Aires"

FACULTAD DE ARTES Y CIENCIAS MUSICALES - INSTITUTO DE INVESTIGACIÓN MUSICOLÓGICA "CARLOS VEGA"

F A C U L T A D D E A R T E S Y C I E N C I A S M U S I C A L E S

CENTRO DE ESTUDIOS ELECTROACÚSTICOS

MODELOS DE LOCALIZACIÓN ESPACIAL DEL SONIDO Y SU IMPLEMENTACIÓN EN TIEMPO REAL

Pablo Cetta

INTRODUCCIÓN

El tratamiento espacial del sonido partiendo de la simulación de fuentes virtuales, y su desplazamiento en la sala de conciertos, es un recurso ampliamente utilizado en la composición, tanto de música mixta como electroacústica. A partir de un proyecto de trabajo con fines didácticos, originado en el Centro de Estudios Electroacústicos de la Facultad de Artes y Ciencias Musicales, se realizó la adaptación de diversos modelos de localización espacial del sonido al entorno gráfico de programación de audio digital Max-MSP. La traslación al mencionado entorno permite la operación de los programas en tiempo real, aumentando considerablemente sus posibilidades de utilización. En este artículo se tratan tres de los modelos considerados.

DESCRIPCIÓN DEL MODELO DE CHOWNING

El modelo de Chowning –ilustrado en la figura 1- se basa en un sistema cuadrafónico que reproduce el sonido directo (sin reflexiones) de la fuente virtual, reverberación local (diferenciada para cada uno de los parlantes) y reverberación global (común a los cuatro canales).

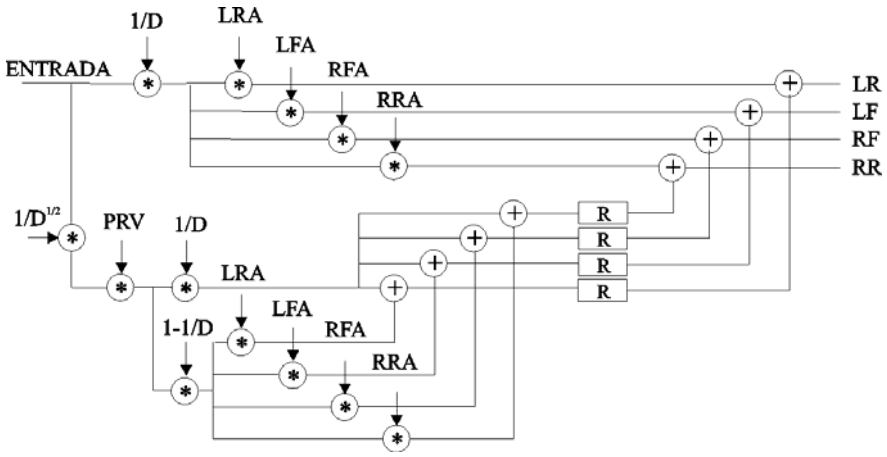


Figura 1

La amplitud de la señal de audio de entrada es atenuada en función de la distancia (1/D), y luego escalada por un coeficiente distinto para cada parlante (LRA, LFA, RFA, RRA) en relación al ángulo de posicionamiento de la fuente (0º-360º).

La ganancia de cada uno de los cuatro parlantes se obtiene mediante:

$$g_n(\theta) = \cos(\theta - \theta_n) \qquad si \quad |\theta - \theta_n| < 90° \quad \acute{o}\ bien \quad 0$$

donde g_n es la ganancia del parlante n, θ es el ángulo de posicionamiento de la fuente y θ_n es el ángulo de ubicación del parlante.

Para evitar la realización de los cálculos se implementa la función que se observa en la figura 2. Los valores de la función –cuya amplitud varía entre 0 y 1- determinan la amplitud relativa que debe poseer un parlante para generar un movimiento de 360º. Los cuatro canales leen la misma tabla, pero con fases iniciales diferentes. Vemos que se trata del hemiciclo de una sinusoide, y que la segunda mitad de la función permanece en el valor 0, que corresponde a la intensidad del parlante cuando la fuente sonora está ubicada del lado opuesto. La función posee 512 muestras, por lo que se vuelve necesaria una traslación del valor del ángulo a un número de muestra en la tabla que la aloja.

Figura 2

La parte inferior del esquema del modelo (figura 1) muestra la rama de tratamiento de la reverberación. La señal es también atenuada en función de la distancia, pero ahora por la raíz cuadrada de al inversa de la distancia, y posteriormente por un valor empírico (PRV) que representa al porcentaje de reverberación. La rama de reverberación es luego escalada por *1/D* a nivel global, mientras la local es atenuada por *1-1/D* y por los coeficientes en función del ángulo de posicionamiento de la fuente. El sistema utiliza cuatro unidades de reverberación diferenciadas, con el objeto de hacer más realista la simulación.

IMPLEMENTACIÓN EN MAX-MSP

En la figura 3 se muestra la unidad, denominada *Turenas* –por el título de la obra en la que Chowning utilizó su sistema de especialización-, donde se encapsuló el modelo. Los parámetros que recibe son –de izquierda a derecha- la señal de audio de entrada (en el ejemplo provista por un generador de ruido rosa) la distancia de la fuente respecto a la circunferencia imaginaria que rodea a los parlantes), el ángulo de posicionamiento, y el porcentaje de reverberación. La unidad *buffer* (parte superior del gráfico) almacena la función que aporta los coeficientes de amplitud en función del ángulo. Finalmente, las cuatro salidas son enviadas al conversor digital analógico para su audición.

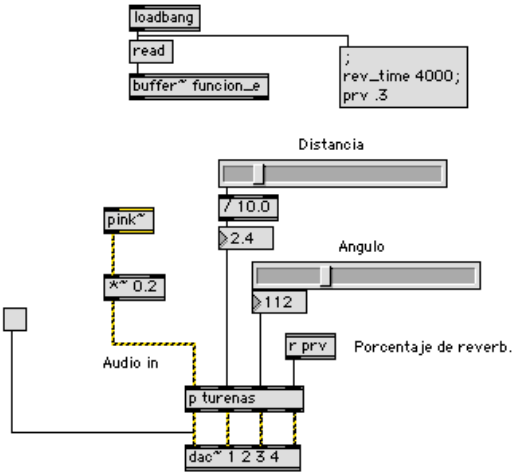


Figura 3

Observemos ahora, en la figura 4, el modelo propiamente dicho. Aquí encontramos parte del código nuevamente encapsulado: se trata de las operaciones destinadas a obtener cada coeficiente de amplitud en función del ángulo (ver figura 5). Según vimos antes, los cuatro canales leen la misma tabla (con la función *peek*) pero con ciertas diferencias de fase.

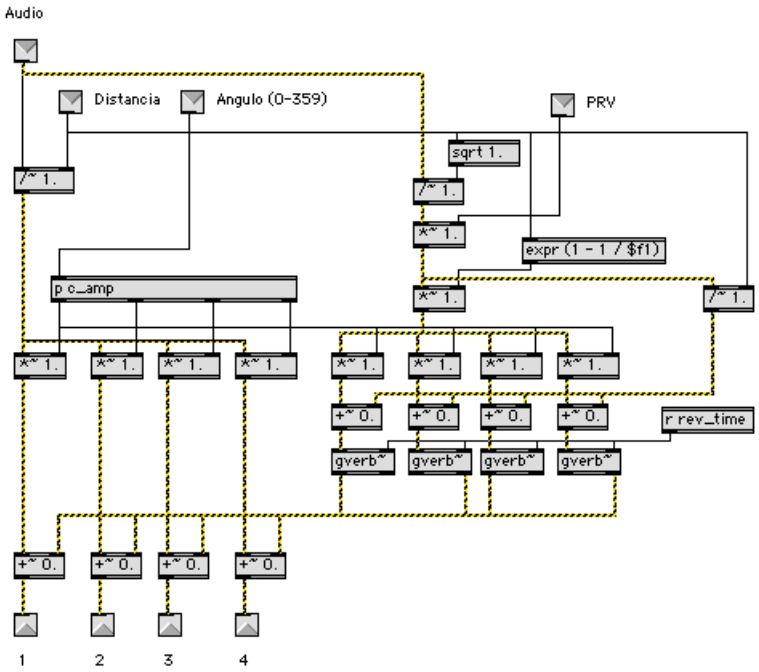


Figura 4

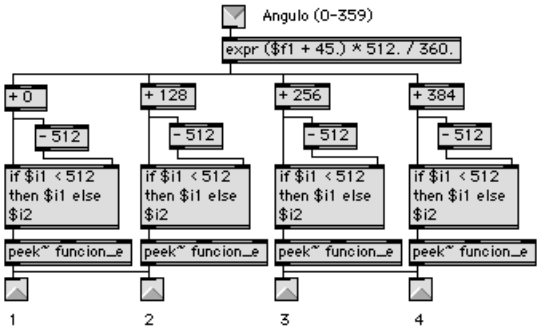


Figura 5

DESCRIPCIÓN DEL MODELO DE MOORE

El sistema de espacialización de F. Richard Moore se basa también en la reproducción cuadrafónica del sonido. Los cuatro parlantes se ubican en las esquinas del auditorio, y determinan los límites del "cuarto real", donde se ubican los oyentes, y donde se produce el fenómeno de simulación de fuentes aparentes. La posición de la fuente virtual se encuentra más allá de este espacio físico, dentro de lo que denominamos el "cuarto virtual" (ver figura 6).

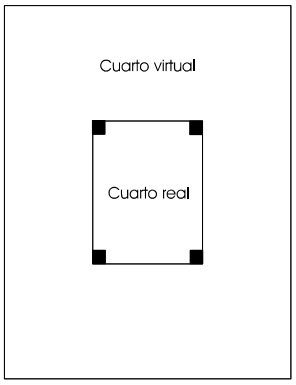


Figura 6

La ubicación de una fuente sonora entre la posición de un parlante y el oyente resulta prácticamente imposible. Para que esto suceda, es preciso simular las diferencias interaurales de intensidad, de tiempo y de espectro empleando reproducción transaural, acondicionar acústicamente la sala tratando de evitar las reflexiones en su interior, y aún así, la imagen espacial estará sumamente condicionada por la posición del oyente. El sistema de Moore no es ajeno a la deformación de la imagen en función de la posición del oyente, pero ésta resulta de un modo natural, creando una perspectiva sonora análoga a la perspectiva visual que cada individuo experimenta dentro de la sala.

El sonido que se produce en forma simulada en el cuarto virtual ingresa al cuarto real a través de cuatro orificios representados por los parlantes. Si imaginamos el movimiento de la fuente sonora en el exterior, obtendremos una idea de su ubicación por simple comparación: la fuente se encuentra cerca del parlante que suena más fuerte. Pero no sólo nos guiamos por la intensidad del sonido, sino también por su tiempo de arribo. Estamos simulando diferencias interaurales de intensidad y de tiempo, con una fuerte discriminación frente-atrás por tratarse de un sistema cuadrafónico.

Cuando experimentamos la sensación sonora en un ambiente cerrado llega primero a nuestros oídos el sonido directo, que es el que proviene directamente de la fuente. En segundo término las reflexiones de primer orden, que parten de la fuente sonora, chocan con un objeto (la pared, por ejemplo) y alcanzan nuestros oídos. Luego las de segundo orden (con dos reflexiones), y así hasta percibir una sensación difusa denominada reverberación. Las primeras reflexiones –en general seis, si sólo consideramos un cuarto con cuatro paredes, techo y piso- contribuyen con la determinación de la posición de la fuente, principalmente en el ataque del sonido. La reverberación, por su parte, nos brinda información sobre las características materiales de la sala, y sobre sus dimensiones. En este modelo se simula el sonido directo y las primeras reflexiones que parten de la fuente a cada uno de los orificios (parlantes), con sus niveles de intensidad y tiempos de llegada relativos, y también la reverberación global, que ayuda a determinar las características de la sala imaginaria. Sólo debemos limitarnos a reproducir, en cada parlante, copias del sonido original (directo y reflejado) con las intensidades y retardos calculados como si la fuente virtual realmente existiera. Vemos en la figura 7 una representación –por medio de rayos- del sonido directo y las primeras reflexiones que alcanzan a dos de los parlantes.

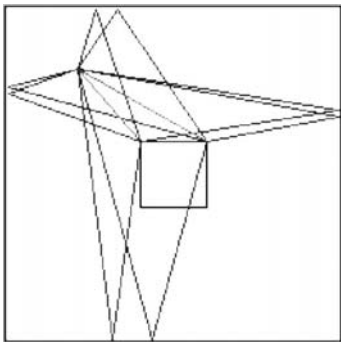


Figura 7

Un método para obtener la distancia recorrida por una reflexión consiste en rebatir la planta y trazar una línea entre la nueva posición de la fuente –fuente “fantasma”- y el punto de llegada –en nuestro caso el parlante (ver ejemplo 8). El cálculo de la distancia se basa en la determinación de las coordenadas de las fuentes fantasma y en operaciones simples.

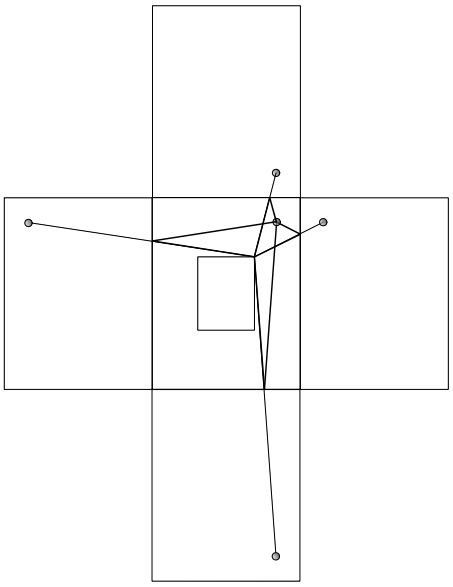


Figura 8

A partir de la obtención de la distancia de cada rayo (sonido directo y cuatro primeras reflexiones a cada uno de los cuatro parlantes, dando un total de 20) es posible calcular el retardo y el factor de atenuación de la amplitud.

La amplitud surge de la ley del cuadrado inverso:

$$A^2 \propto I \propto 1/D^2$$

En ocasiones es preferible, por razones preceptúales, cambiar el cuadrado por el cubo de la distancia, con lo cual nos queda:

$$A \propto \sqrt{I} \propto 1/D^{1.5}$$

Luego, los retardos también surgen a partir de las distancias:

$$r = D / c$$

donde *r* es el tiempo de retardo, *D* la distancia y *c* la velocidad del sonido.

La transformación dinámica del tiempo de retardo trae aparejado un cambio en la altura del sonido. Para entender esto, podemos imaginar un grabador a cinta con dos cabezas reproductoras, una fija y otra deslizable. La primera cabeza lee la información sonora, y la segunda lo hace con un cierto retraso. Si desplazamos la segunda cabeza mientras lee, cambiamos la velocidad de lectura: si la cabeza se mueve en el mismo sentido que la cinta, la velocidad baja, y por consiguiente la altura también. En cambio, si la desplazamos en sentido contrario la altura sube. Podemos reemplazar ahora al grabador por un archivo de sonido cuyas muestras se leen a frecuencia constante. Si el puntero de lectura se desplaza, se acorta o alarga el período de la forma de onda produciendo una variación perceptible de la altura. No obstante, la variación del retardo genera una transformación de la altura equivalente a la que produce el efecto Doppler, y nos brinda una nueva pista para la determinación del movimiento de la fuente virtual en el espacio.

IMPLEMENTACIÓN EN MAX-MSP

En la figura 9 se aprecia la sección de cálculo. Aquí se realizan las operaciones que dan por resultado las distancias, amplitudes y retardos relativos a cada vector en función de la posición de la fuente. Como vimos, se trata de 20 rayos ([sonido directo + 4 primeras reflexiones] * 4 parlantes). Dado que se trata de un número significativo de cálculos, pero todos basados en sólo tres ecuaciones, se optó por empaquetar a todas las variables en listas que son leídas por la unidad *vexpr*. Esta unidad lee tantas listas como variables tenga la expresión que ella contiene, y realiza las operaciones tomando los valores secuencialmente, por número de orden dentro de las listas. Los resultados se entregan empaquetados en una lista de salida. Las distancias, por ejemplo, se obtienen por la siguiente expresión, que tiene como variables a distintas coordenadas (fuentes, fuentes fantasma, posición de los parlantes):

$$D = \sqrt{(a-b)^2 + (c-d)^2}$$

Los retardos son calculados en función de la distancia, en milisegundos:

$$r_x = D_x / .34$$

y las amplitudes de cada rayo con la ley del cuadrado inverso. Para corregir perceptualmente la sensación de alejamiento, la potencia a la cual se eleva la distancia se establece como variable (f):

$$1 / (1 + D^f)$$

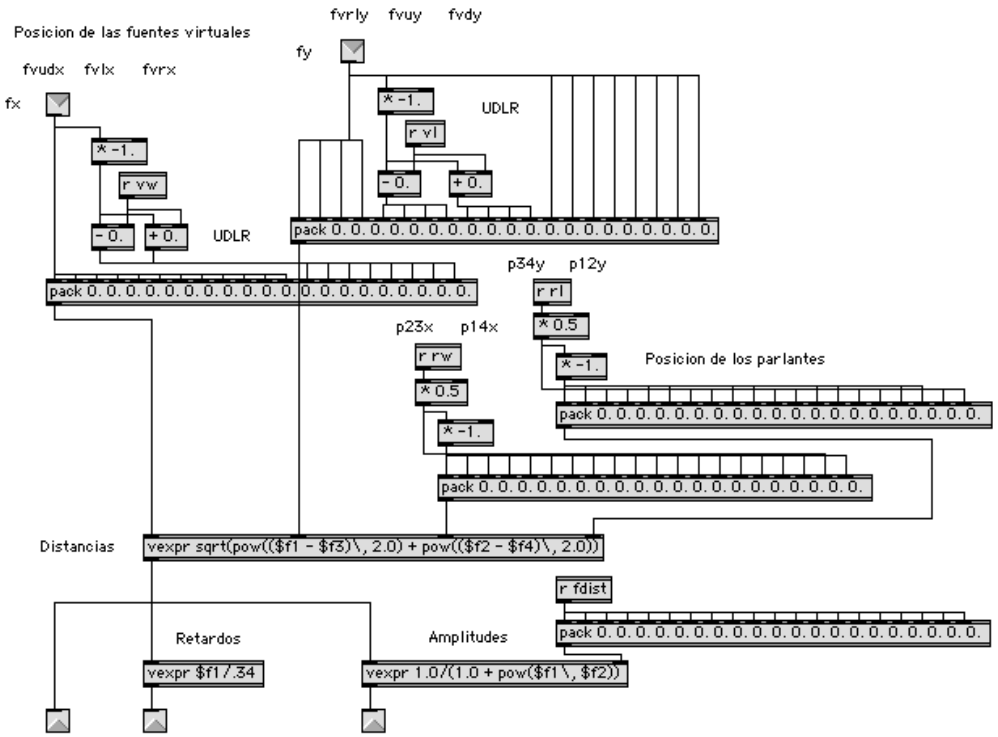


Figura 9

Obtenida esta información, se envía al procesador que se ilustra en la figura 10, que es el encargado de espacializar la señal de entrada. Pueden observarse las dos unidades *unpack*, que desempaquetan los veinte resultados de los cálculos de amplitud y de retardo, respectivamente. La señal original se bifurca en 20 ramas, en cada una de los cuales se produce un retardo (con *tapin* y *tapout*) y una modificación de la amplitud (multiplicando la señal original por los coeficientes de amplitud calculados). Los rayos relativos a las reflexiones se agrupan por canales (parlantes 1 a 4) y a cada canal se le aplica un filtro pasabajos, cuya frecuencia de corte es controlable, para simular la pérdida de energía en agudos propia de la reflexión. Finalmente, una versión sumada de los cuatro canales de las reflexiones (sin sonido directo) se envía a una unidad que simula la reverberación global.

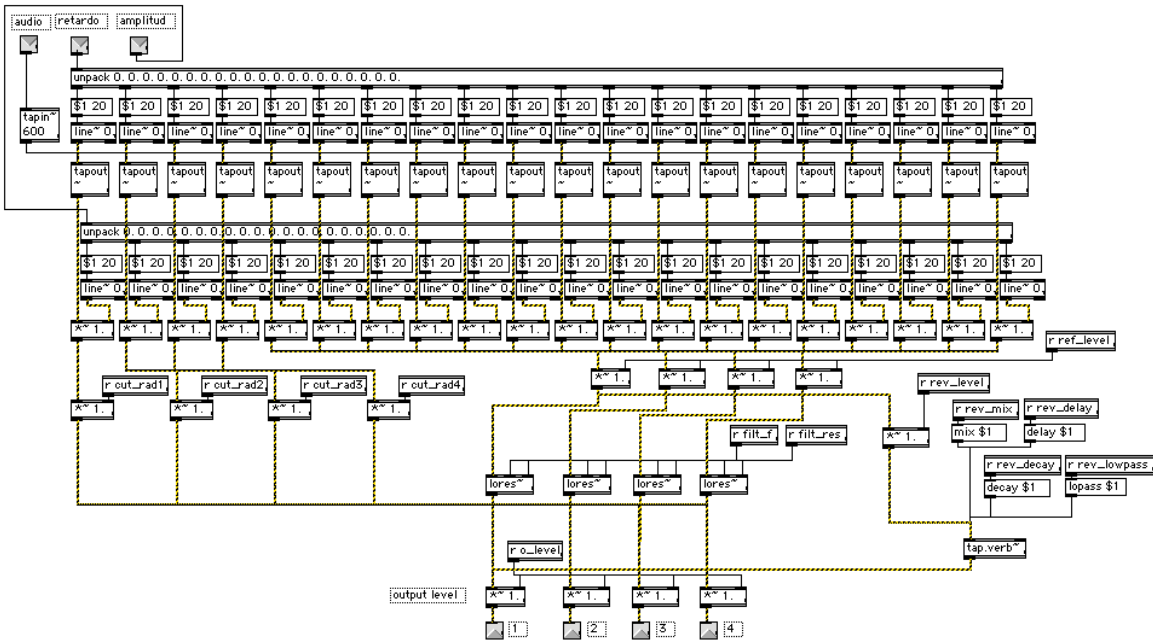


Figura 10

El modelo original de Moore incorpora, además, el control del ángulo de difusión de la fuente sonora y del esquema de radiación (forma cardioide, por ejemplo). Estas características no han sido todavía incorporadas en nuestra versión en tiempo real.

DESCRIPCIÓN DEL MODELO DE SÍNTESIS BINAURAL

Las grabaciones realizadas utilizando una cabeza artificial registran las señales que ingresan directamente a cada uno de los oídos, a través de dos micrófonos ubicados en la entrada de los canales auditivos. De este modo se capta la información acústica y espacial relativa a la ubicación de las fuentes sonoras, al ambiente y a la posición del supuesto oyente dentro de ese ambiente. Por medio de esta técnica –denominada *binaural*- se registran las diferencias interaurales de intensidad, de tiempo y de espectro, como así también las reflexiones que tienen lugar en las convoluciones de los pabellones auditivos, la cabeza y el torso. Todo esto brinda información al sistema perceptual para determinar la ubicación de las fuentes sonoras.

En música electroacústica suele emplearse la *síntesis binaural*, que consiste en el procesamiento de una señal sonora monofónica a la que se le atribuyen diferencias interaurales particulares, como si hubiera sido grabada con una cabeza artificial. Esto resulta sumamente útil cuando se trabaja con sonidos sintetizados. El resultado de la síntesis binaural es similar al producido al grabar un sonido cualquiera con una cabeza artificial, ubicando un parlante como fuente sonora en un lugar determinado del ambiente.

Se trata de una simulación basada en el proceso de convolución. Este proceso –que se logra en su forma más efectiva mediante la multiplicación de dos espectros- permite, entre muchos usos, la aplicación de reverberación natural a un sonido. Supongamos, a modo de ejemplo, que grabamos en un estudio a un instrumento, por ejemplo, un violín. El sonido es tomado directamente por un micrófono, y la grabación está prácticamente desprovista de características espaciales –direccionalidad, reflexiones del sonido sobre las paredes del recinto, reverberación del ambiente. Por otra parte, nos dirigimos a una sala de conciertos y grabamos un sonido extremadamente breve y fuerte –un impulso, que podría ser simulado con un petardo- que haga reaccionar a la sala con sus características propias –reflexiones, reverberación, respuesta a las

frecuencias. Si ahora realizamos una convolución entre ambos registros, el resultado a obtener será equivalente al que habríamos logrado grabando al violinista dentro de la sala de conciertos. La síntesis binaural consiste, entonces, en la convolución de las formas de onda de un sonido a espacializar y de una respuesta a impulso tomada con una cabeza artificial dentro de una cámara anecoica –es decir, sin reflexiones. Si la respuesta a impulso fue tomada a 90º de desplazamiento horizontal –considerando 0º a la posición frontal- y con una elevación de 30º, el resultado de la convolución nos permite escuchar al sonido original como si proviniera de esa posición. Cambiando dinámicamente las respuestas a impulso podemos describir trayectorias en un espacio de tres dimensiones.

IMPLEMENTACIÓN EN MAX-MSP

Para la realización de este modelo utilizamos respuestas a impulso grabadas en el MIT Media Lab por Bill Gardner y Keith Martin en 1994, empleando una cabeza artificial KEMAR (ver bibliografía). Los archivos –unos 367, tomados a diferentes ángulos de desplazamiento horizontal y elevación- constan cada uno de 128 muestras, tanto para el canal izquierdo como para el derecho. Para nuestro espacializador hemos colocado todas las funciones de transferencia en un único archivo, en orden sucesivo, habiendo agregado previamente 128 ceros a cada respuesta y calculado su espectro mediante la FFT (Transformada Rápida de Fourier). A cada respuesta se le agregan los ceros a fin de poder realizar correctamente el proceso de convolución aperiódica.

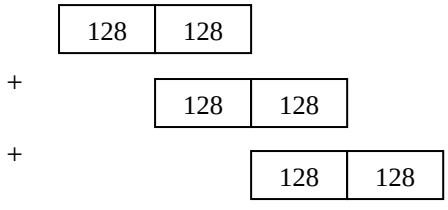
Debemos realizar la convolución en tiempo real entre alguna de estas respuestas a impulso y el sonido que ingresa al espacializador, multiplicando sus respectivos espectros y sumando parcialmente los bloques de resultados. Para aclarar esta técnica vamos a analizarlo por pasos:

1) Tomamos las primeras 128 muestras de sonido a espacializar y agregamos 128 ceros a continuación.

2) Realizamos la FFT de esas 256 muestras.

3) Elegimos la función de transferencia en función de los ángulos de posicionamiento buscados

4) Multiplicamos ambos espectros y realizamos la FFT inversa del resultado. Las primeras 128 muestras son enviadas a la salida, mientras que las 128 restantes se almacenan para sumarlas a las 128 próximas, y así sucesivamente.



Este método, sin embargo, presenta un problema de implementación en tiempo real. Cómo agregar 128 ceros luego de las 128 muestras recibidas sin que se pierda la información que continúa entrando al procesador. La solución consiste en determinar dos ramas de procesamiento en paralelo: recibidas las primeras 128 muestras en la primera rama, ésta

comienza a generar ceros, mientras la segunda rama lee las 128 muestras siguientes de la señal de entrada, y así, cíclicamente.

La figura 11 muestra las dos ramas de procesamiento, que derivan los datos a un objeto *pfft*, que realiza la FFT directa, la multiplicación de complejos, y luego la FFT inversa. A la salida, los resultados de ambas ramas de procesamiento se suman.

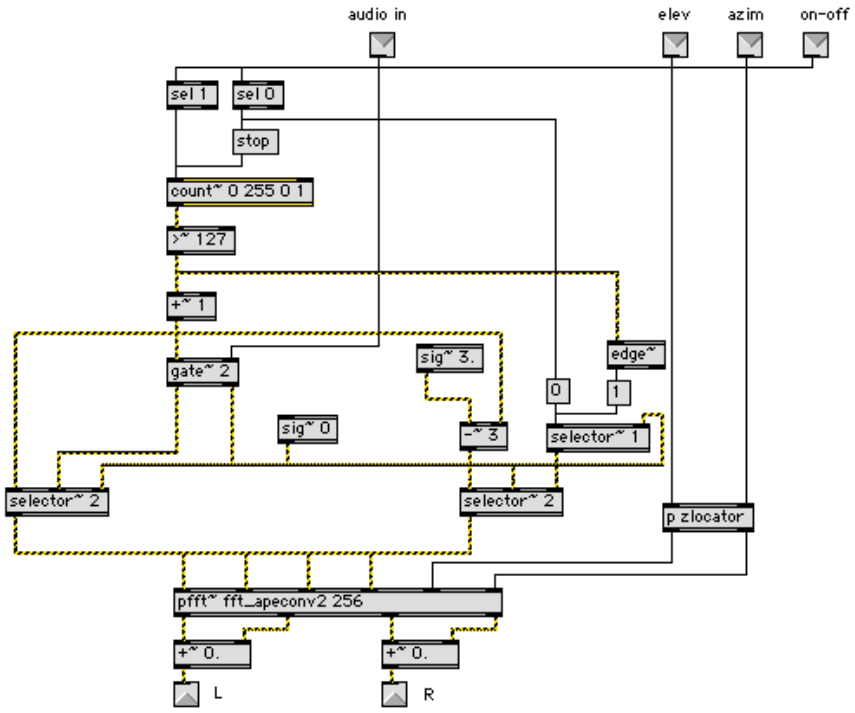


Figura 11

El *patcher* denominado *zlocator*, ubica a las funciones de transferencia dentro del archivo, en función de los ángulos de lateralización y elevación. Puede observarse en la figura siguiente.

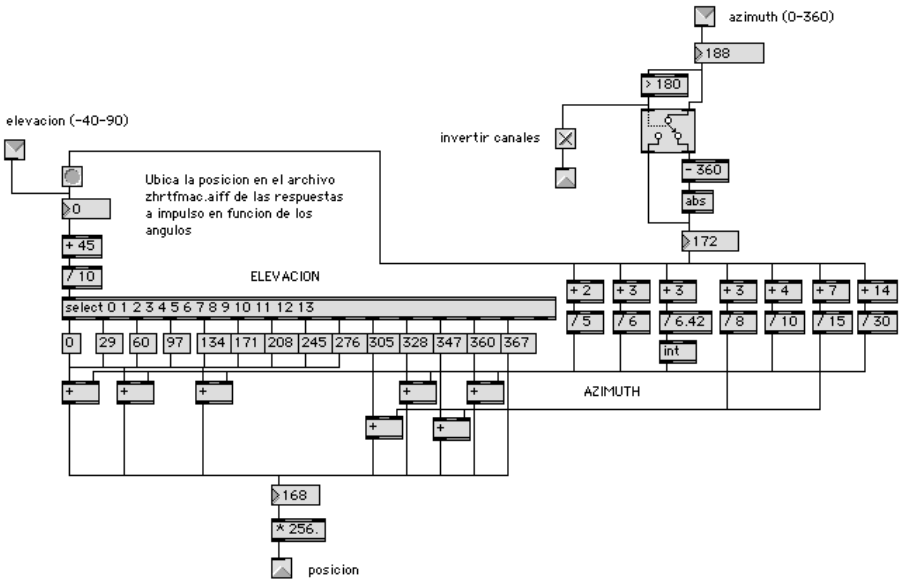


Figura 12

En la figura 13 observamos por dentro el *subpatch pfft*. Vemos cuatro unidades prácticamente iguales, un par (canales izquierdo y derecho) representa a una rama de procesamiento, y el otro par a la segunda rama. Un objeto *receive* (r) toma el valor de la variable *puntero*, que indica la posición de lectura dentro del archivo de funciones de transferencia. El valor de *puntero* es calculado por el objeto *zlocator*, mencionado anteriormente. En el *inlet* 6 de *pfft* se recibe un marcador que indica cuándo el ángulo de lateralización está fuera del rango comprendido entre 0 y 180º. Si la fuente virtual se encuentra a 270º, por ejemplo, las funciones de transferencia son las mismas que a 90º, pero con los canales invertidos. La inversión de derecho a izquierdo y viceversa se realiza con dos unidades *selector*, que operan de acuerdo al valor de la variable *hemi* (hemisferio).

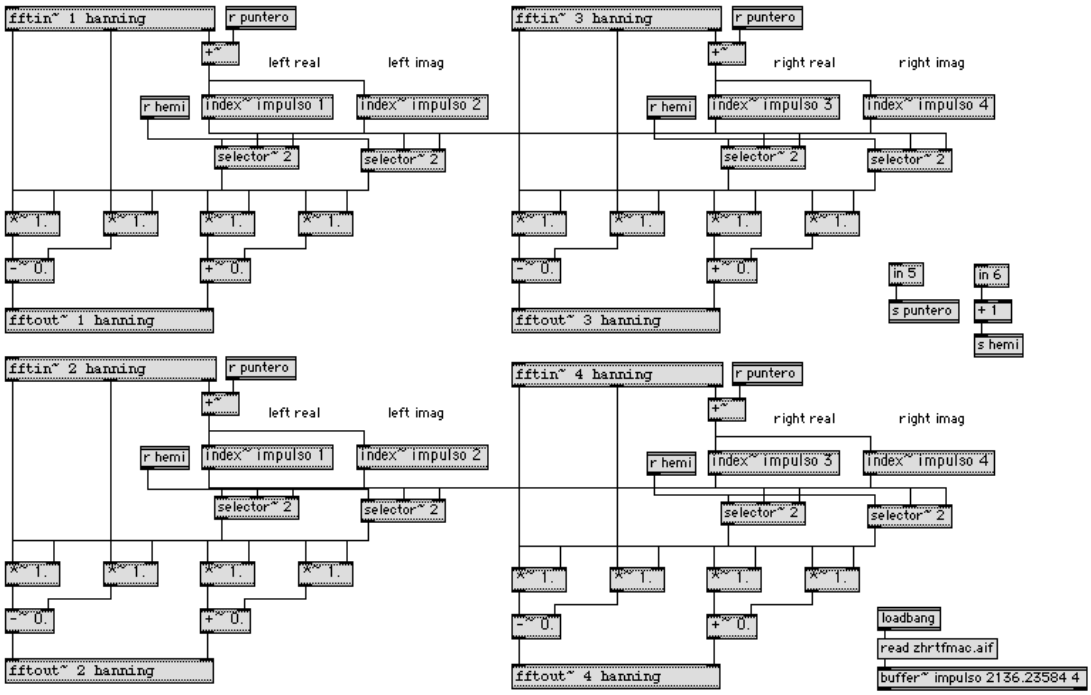


Figura 13

Por último, integramos todos los módulos vistos en un patcher (*zhrtf*) (figura 14). Al modelo se agrega reverberación y simulación de la distancia. Y para el posicionamiento de la fuente virtual, dos envolventes que describen los ángulos barridos por la trayectoria.

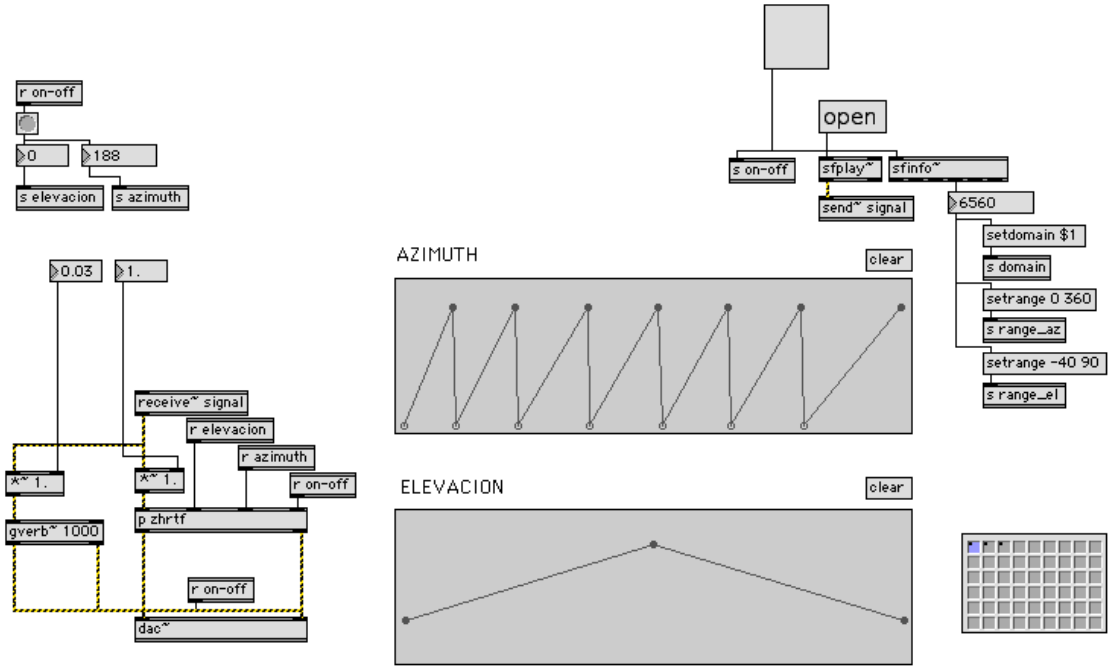


Figura 14

ELIMINACIÓN DEL CROSSTALK Y ECUALIZACIÓN

Para la reproducción de señales grabadas binauralmente, o tratadas con síntesis binaural, suelen emplearse auriculares, que aseguran que la señal destinada a cada uno de ambos oídos parte del mismo punto en el que fueron ubicados los micrófonos respecto a la cabeza artificial. La reproducción con parlantes es, en cambio, contaminada por las reflexiones naturales del ambiente en el que se reproduce la grabación, por la doble acción del oído externo (registrada al grabar y puesta nuevamente en acción al escuchar), por las reflexiones en torno a la cabeza y el torso, y fundamentalmente, por el cruzamiento de información que se produce entre los dos oídos. La señal destinada al oído izquierdo parte del parlante izquierdo e ingresa a ese oído, pero también al derecho, y viceversa. Esto trae como consecuencia una deformación significativa de la imagen espacial, debida a la percepción confusa de las diferencias interaurales. Este efecto de cruzamiento –denominado en inglés *crosstalk*, por analogía con ciertos fenómenos de la telefonía- puede ser disminuido significativamente empleando filtros. Básicamente, se busca que cada parlante emita la señal destinada al oído correspondiente, pero sumada a otra que anule el efecto no deseado respecto al oído contrario. A la reproducción binaural con parlantes se la denomina reproducción *transaural*.

La teoría de la cancelación del crosstalk fue tratada en principio por B. S. Atal y M. R. Schroeder, y luego reelaborada por D. H. Cooper y J. L. Bauck en diversos artículos. Dado que su puesta en práctica requiere del uso de filtros, vamos a detenernos brevemente sobre este punto para comentar algunas de sus particularidades.

La teoría de los filtros digitales explica que un filtro se caracteriza por su función de transferencia H(z), que es la representación en el dominio de la frecuencia –o en el dominio *z*- de h(n), la respuesta a impulso del filtro. Un caso particular de la transformada *z*, que nos permite acceder al dominio de la frecuencia, es la transformada de Fourier. Podemos decir, de forma simplificada, que la función de transferencia H(z) puede representarse como el espectro de la forma de onda h(n).

H(z) describe la relación entre la entrada y la salida del filtro.

$$H(z) = Y(z) / X(z)$$

Las figuras 15 y 16 representan al filtro visto desde el dominio del tiempo (formas de onda) y desde el dominio de la frecuencia (espectros), respectivamente.

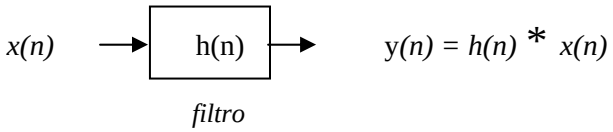


Figura 15

La señal filtrada *y(n)* se obtiene a través de la convolución $(*)$ entre la señal de entrada *x(n)* y la respuesta a impulso del filtro *h(n)*. Según vimos antes con un ejemplo, una sala de conciertos se asemeja a un filtro, en el sentido que transforma la señal proveniente de una fuente cualquiera –un instrumento musical, por ejemplo. Podíamos simular la reverberación natural de este mismo modo, es decir, mediante la convolución del sonido original de la fuente con la respuesta a impulso de la sala.

Si recordamos que la convolución de las formas de onda equivale a la multiplicación de los espectros, en la figura 16 observamos lo mismo que en la 15, pero el mismo resultado es ahora obtenido en el dominio de la frecuencia (espectros de las señales).

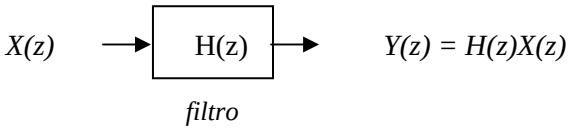


Figura 16

En su forma más general, la implementación del proceso de eliminación del *crosstalk* puede representarse a través de la siguiente matriz:

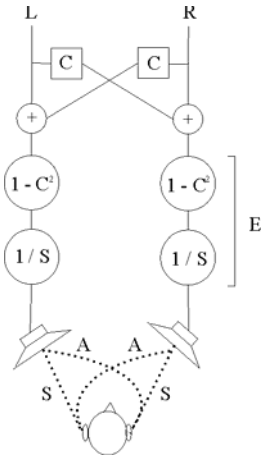


Figura 17

C(z) y E(z) son filtros de cancelación y ecualización, respectivamente.

Conocemos de antemano las respuestas a impulso *s(n)* y *a(n)*, que pueden obtenerse grabando con una cabeza artificial un impulso que parte del mismo ángulo en el que se ubican los parlantes (30º, por ejemplo). s(n) corresponde al oído más cercano y a(n) al más alejado. En el gráfico siguiente (figura 18) se observan estas respuestas, son 128 muestras tomadas a 44.1 kHz con la cabeza artificial KEMAR, sobre una fuente ubicada a 30º. Se observan claramente las diferencias interaurales de tiempo y de intensidad

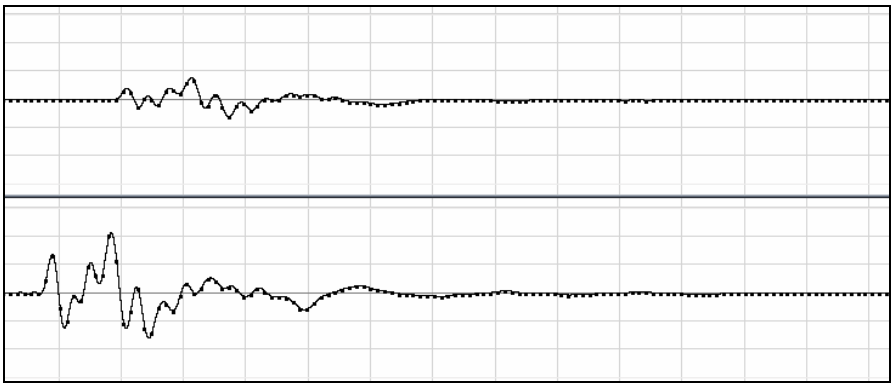


Figura 18

La figura 19 muestra el análisis espectral de las respuestas anteriores (ventana Blackman-Harris).

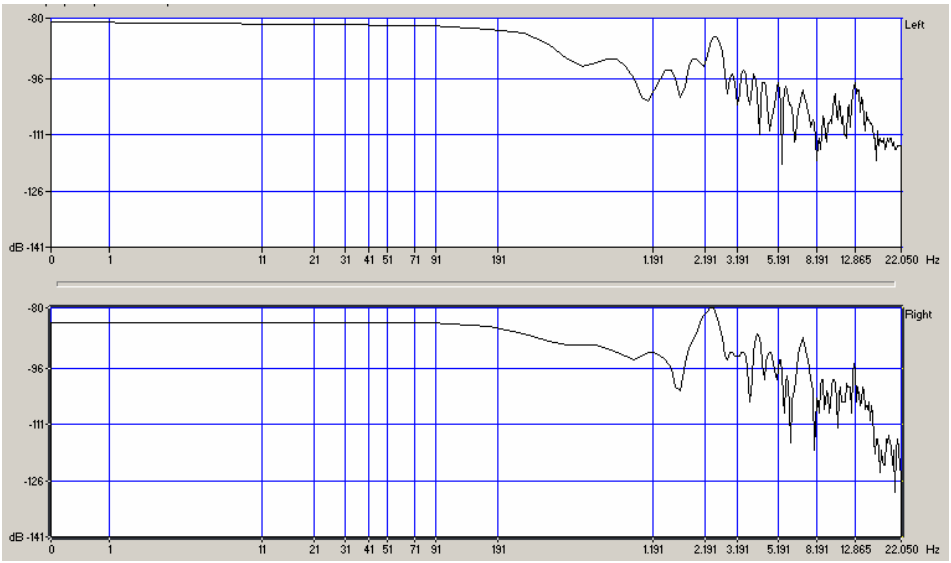


Figura 19

Vamos a tratar de determinar la función de transferencia de los filtros C y E en términos de *S(z)* y *A(z)*, las transformadas *z* de los impulsos de la figura 18.

De la observación de la matriz de la figura 17 deducimos lo siguiente:

$$R'(z) = R(z)\,[E(z)\,S(z) + C(z)\,E(z)\,A(z)] + L(z)\,[E(z)\,A(z) + C(z)\,E(z)\,S(z)]$$

$$R´(z) = [R(z)\,[S(z) + C(z)\,A(z)] + L(z)\,[A(z) + C(z)\,S(z)]]\,E(z)$$

Para que la eliminación del *crosstalk* se haga efectiva, es necesario que *R´* dependa sólo de *R* (señal destinada al oído derecho) y no de *L* (señal destinada al oído izquierdo). Por lo tanto, para que esto suceda, es preciso que lo que multiplica a *L(z)* sea igual a cero:

$$A(z) + C(z)\,S(z) = 0$$

Por lo tanto,

$$\boxed{C(z) = -\frac{A(z)}{S(z)}}$$

Por otra parte, se requiere que *R´(z)* sea igual a *R(z)*, para que no exista ninguna transformación de la señal luego de haberse eliminado el *crosstalk*. Por esto, y por el hecho de haber ubicado los parlantes en frente del oyente en lugar de ubicarlos contra sus oídos, es necesario emplear el filtro de ecualización *E*.

Para que *R´(z)* sea igual a *R(z)*,

$$[S(z) + C(z)\,A(z)]\,E(z) = 1$$

Reemplazando *C(z)*,

$$\left[S(z) + \left(-\frac{A(z)}{S(z)}\right)A(z)\right]E(z) = 1$$

$$E(z) = \frac{1}{S(z) + \left(-\frac{A(z)}{S(z)}\right)A(z)}$$

Multiplicando ambos términos por *S(z)* nos queda:

$$E(z) = \frac{S(z)}{S^2(z) - A^2(z)} = \frac{S(z)}{S^2(z)\left(1 - \frac{A^2(z)}{S^2(z)}\right)}$$

$$\boxed{E(z) = \frac{1}{S(z)}\frac{1}{\left(1 - C^2(z)\right)} = E_1.E_2}$$

Obtuvimos dos filtros de ecualización $E_1 = 1 / S(z)$, destinado a corregir la deformación introducida por la ubicación de los parlantes respecto al oyente y $E_2 = 1/(1 - C^2)$ que corrige la deformación generada a partir de la eliminación del *crosstalk*.

Implementarlo en Max-MSP será nuestro próximo paso...

BIBLIOGRAFÍA

Chowning, J. *"The simulation of moving sound sources"*. JAES 19: 2-6. 1971.

Moore, F. R. *"A general model for spatial processing of sound"*. Computer Music Journal 7(3): 6-15. 1983

Schroeder, M. y Atal, B.,*"Computer simulation of sound transmisión in rooms"*. IEEE Conv. Rec., pt 7, pp. 150-155. 1963.

Cooper H. y Bauck, J., *"Prospects for transaural recording"*. JAES. Vol. 37. pp. 3-19. 1989

Möller, H., *"Cancellation of crosstalk in artificial head recordings reproduced through loudspeakers"*. JAES, vol. 37, pp. 31-34. 1989

Gardner, W. G., y Martin, K. D. *HRTF measurements of a KEMAR dummy head microphone*. MIT Media Lab Perceptual Computing Technical Report #280. 1994. Incluído en el CD-ROM "Standards in Computer Generated Music", Goffredo Haus e Isabella Pighi, editores, publicado por IEEE CS Technical Committee on Computer Generated Music, 1996.

Gilkey, R. y Anderson, T. *Binaural and spatial hearing in real and virtual environments.* Lawrence Erlbaum Associates. New Jersey. 1997.

Artículo publicado en:
Cetta, P. (comp.). Altura-Timbre-Espacio. Cuaderno de Estudio Nº 5. IIMCV. Educa. 2004.