

Perspectivas actuales en la síntesis digital de sonidos musicales

Xavier Serra

Cuando se genera un sonido musical electrónicamente, es importante tener un buen modelo con parámetros que permitan manipular el sonido de una forma intuitiva. El músico, usuario, ha de poder desarrollar una intuición musical que le permita experimentar con la técnica de síntesis que utilice. En este artículo se presenta una visión general de la síntesis digital, concentrándose en dos líneas de investigación que seguramente nos permitirán romper las limitaciones existentes con los métodos utilizados hasta ahora y conseguir, finalmente, dar respuesta a las promesas que la música con ordenadores hizo en los años sesenta.

1. Introducción

La generación de sonidos ha sido el campo musical donde la tecnología digital ha tenido un impacto más grande. Pero la evolución de las técnicas de síntesis digital ha sido más lenta de lo que en un principio se esperaba, y hasta los últimos años no se ha encontrado soluciones de futuro que pueden dar un nuevo impulso a esta evolución.

En un artículo histórico que apareció en la revista Science el año 1963, Max Mathews (Mathews, 1963), mostraba una visión muy optimista del ordenador como instrumento musical. El autor, un pionero de la música con ordenadores, decía que generar sonidos a partir de números era una forma completamente general de sintetizar sonidos, ya que el ámbito de frecuencias y dinámicas del oído están acotados y que por tanto cualquier sonido que podamos percibir puede ser generado de esta forma. La promesa de la música con ordenadores era que el ordenador es capaz de generar cualquier sonido que pueda salir por un altavoz. Unos años más tarde, el mismo Max Mathews, en su libro sobre la tecnología de la música con ordenadores (Mathews, 1969) escribía:

"Los dos problemas fundamentales en la síntesis del sonido son (1) la gran cantidad de datos necesarios para definir una onda acústica -por tanto la necesidad de un programa muy rápido- y (2) la necesidad de un lenguaje simple y potente con que describir una secuencia compleja de sonidos".

El problema (1) ya ha sido resuelto en gran medida por la evolución tecnológica, ya que la velocidad de los procesadores digitales ha ido incrementando de forma exponencial durante los últimos veinte años. El problema (2) sigue sin tener una solución satisfactoria por el hecho de que es imposible describir los sonidos si se ha de definir cada uno de los números que representan una onda acústica. Se ha de poder describir el sonido a partir de menos números o se ha de partir de sonidos grabados. Afortunadamente la mayor parte de ondas no son musicalmente interesantes y muchas ondas físicamente diferentes, son perceptualmente iguales. Por tanto, no es necesario generar todas las ondas posibles y el objetivo se centra en encontrar un grupo reducido de técnicas de síntesis y de control que nos permitan explorar todo el espacio tímbrico de interés musical.

La aproximación tradicional a este problema ha sido intentar conseguir sonidos combinando elementos de síntesis simples. Pero en los últimos años ha quedado claro que es difícil generar de esta manera

sonidos complejos de interés musical y que para poder competir con los instrumentos acústicos, en términos de complejidad sonora y control expresivo, se requiere un nuevo enfoque de la síntesis digital. El enfoque que aquí proponemos parte del estudio de la realidad sonora que nos rodea y busca formas de extender esta realidad a un mundo sonoro nuevo, virtual.

2. Tradición de la Síntesis Digital

Las técnicas de síntesis digital heredaron los conocimientos desarrollados para la síntesis con medios analógicos. El primer sistema de síntesis digital, Music V, (Mathews, 1969) desarrollado por Max Mathews durante los años sesenta, introdujo el concepto de generador unitario como versión digital de los módulos de los sintetizadores analógicos. Un generador unitario acepta entradas numéricas de control y genera una señal, también numérica, que puede servir como entrada a otro generador unitario o ser directamente un sonido. Ejemplos de generadores unitarios son: osciladores, filtros, sumadores, multiplicadores, generadores de envoltorios, i generadores de números aleatorios. A partir de combinar estos elementos (figura 1) se pueden crear sonidos sintéticos similares a los obtenidos con los módulos controlados por voltaje de los sintetizadores analógicos, pero con un control mucho más preciso. La mayor parte de sistemas desarrollados desde entonces han partido de estos conceptos. Los sintetizadores basados en el protocolo MIDI son un ejemplo.

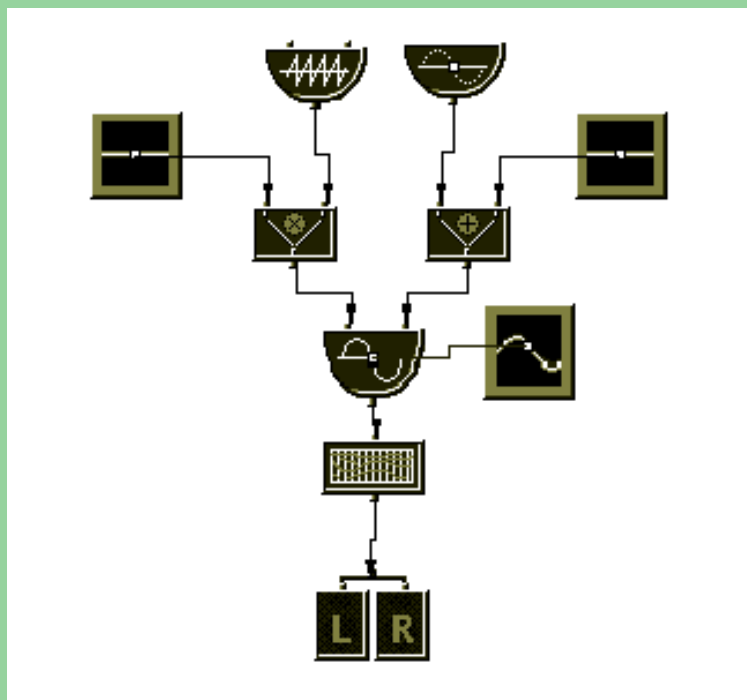


Fig.1: Algoritmo de síntesis a partir de la combinación de generadores unitarios

La implementación de los algoritmos de síntesis se ha hecho tanto a nivel de software como de hardware. La ventaja de las implementaciones en software es su flexibilidad y complejidad ilimitada ya que sólo están acotadas por el lenguaje de programación utilizado, pero por esta misma razón, difícilmente pueden funcionar a tiempo real. Por otro lado las implementaciones en hardware funcionan a tiempo real pero difícilmente pueden competir a nivel de flexibilidad y complejidad. Dadas las características del gran mercado, la mayor parte de sintetizadores comerciales utilizan implementaciones en hardware. De todos modos, gracias al constante aumento de la velocidad de procesamiento de los ordenadores, la utilización de sistemas basados en software es cada vez más eficiente y actualmente pueden funcionar a tiempo real un gran número de los algoritmos más conocidos.

Tradicionalmente las técnicas de síntesis digital se han clasificado en síntesis aditiva, síntesis por sustracción, y síntesis no-lineal. La síntesis aditiva está basada en la suma de sonidos elementales, cada uno de ellos generado por un oscilador. La síntesis por sustracción se basa en la idea complementaria,

es decir, filtrando energía de un sonido complejo. La síntesis no-lineal es un cajón de sastre donde se incluyen un gran número de técnicas basadas en ecuaciones matemáticas de comportamiento no-lineal.

Esta clasificación está hecha con una visión muy teórica que no nos lleva demasiado lejos en nuestro intento de encontrar nuevas soluciones de interés musical. Otra taxonomía más útil, propuesta por Julius Smith (Smith, 1991), organiza las técnicas de síntesis en : *procesamiento de grabaciones, modelos espectrales, modelos físicos, y algoritmos abstractos*.

Una forma de entender la clasificación de Smith es hablando de los métodos de síntesis como técnicas digitales que nos permiten obtener un continuo sonoro que va desde la reproducción de sonidos preexistentes (grabaciones) hasta la generación de sonidos a partir de una abstracción absoluta (sonidos imaginados), con todos los pasos intermedios. En este contexto, las técnicas basadas en procesar grabaciones parten del extremo de los sonidos preexistentes e intentan crear sonidos nuevos, imaginados, retocando directamente el sonido real (este es el caso de los aparatos llamados "samplers"). En el otro extremo de la clasificación están los algoritmos abstractos, que a partir de ecuaciones matemáticas generan sonidos sintéticos muy alejados de los sonidos "naturales", pero manipulando estas ecuaciones se intenta obtener sonidos que permitan una comunicación musical determinada (por ejemplo con los sintetizadores basados en la técnica de modulación de frecuencia). Los modelos espectrales y los físicos están en la zona intermedia entre estos dos extremos y partiendo de modelos, o abstracciones, que nos describen sonidos preexistentes y objetos generadores de sonidos, respectivamente, permiten explorar gran parte del espacio comprendido entre realidades sonoras concretas y realidades nuevas, virtuales. Más adelante profundizaremos en estos modelos.

3. Objetivos Musicales en la Síntesis Digital

Si preguntamos a los músicos que trabajan con medios informáticos qué querrían poder hacer a nivel de generación de sonidos, muy posiblemente, casi todos estarían de acuerdo con las siguientes respuestas: (1) poder crear cualquier sonido imaginable y (2) poder manipular cualquier sonido preexistente de cualquier manera que podamos concebir. Estos objetivos son evidentemente utópicos, no solamente por las limitaciones tecnológicas sino también, y más importante, por los límites prácticos de nuestra imaginación. Es difícil poder idear sonidos sin un referente del mundo sonoro que nos rodea y en realidad nuestra imaginación siempre parte de este referente. De aquí, el gran interés de las técnicas que tienen la realidad sonora como punto de partida.

Las limitaciones tecnológicas nos obligan a valorar una serie de compromisos que hemos de tener en cuenta a la hora de diseñar o utilizar una determinada técnica de síntesis. En concreto querríamos mencionar cuatro de estos compromisos:

1. Calidad del sonido. Por calidad del sonido entendemos la riqueza interna del sonido. Un sonido con mucha calidad podría ser un sonido natural y al otro extremo podríamos tener un sonido simple, sintetizado electrónicamente y sin ninguna microvariación al largo de su duración.
2. Flexibilidad. Con este término se describe la capacidad de una determinada técnica de síntesis para modificar el sonido a partir de una serie de parámetros de control. Con este criterio, un "sampler" sería un aparato muy poco flexible, y la síntesis por modulación de frecuencia sería muy flexible.
3. Generalidad. Por generalidad se entiende la posibilidad de que una misma técnica de síntesis pueda generar un gran número de timbres. La síntesis aditiva sería una técnica muy general y la grabación de un sonido sería muy específico.
4. Tiempo de cálculo. El tiempo de cálculo se refiere al número de instrucciones de ordenador que hacen falta para generar cada una de las muestras de sonido sintetizado. En este sentido la síntesis por

modulación de frecuencia es una técnica muy económica y la síntesis aditiva requiere mucho más tiempo de cálculo.

Idealmente querríamos maximizar calidad, flexibilidad, y generalidad, minimizando el tiempo de cálculo. Eso no es posible tecnológicamente y en cada caso concreto hemos de valorar qué es lo que nos interesa más, y escoger la técnica de acuerdo con estas consideraciones. Por ejemplo, si queremos máxima calidad seguramente habremos de renunciar en los aspectos de flexibilidad y generalidad y si queremos una gran flexibilidad no podremos tener mucha calidad. Dado que estamos mirando más de cara al futuro y es evidente que la velocidad de los procesadores digitales irá aumentando nos permitimos el lujo, en nuestra valoración de las técnicas de síntesis, de no preocuparnos demasiado del tiempo de cálculo. Nuestra prioridad será maximizar la calidad, flexibilidad y generalidad.

Además de valorar estos compromisos a la hora de escoger un sistema, es necesario considerar que para que una técnica de síntesis sea útil para un músico, su control ha de ser intuitivo y por tanto ha de partir de una realidad sonora y musical existente. Hay dos realidades sonoras que los músicos están acostumbrados a tratar y que son un buen punto de partida para crear nuevos sonidos. Una de estas realidades es la de los objetos físicos que generan sonido mecánicamente, por ejemplo, los instrumentos tradicionales. Los músicos tienen una intuición clara sobre la relación que hay entre los objetos físicos y el sonido que producen. Con este conocimiento es factible el diseño de nuevos objetos y por tanto nuevos sonidos. La síntesis por modelos físicos nos permite partir de esta realidad y crear objetos acústicos virtuales que van más allá de la realidad física que nos rodea.

La otra realidad sonora de interés para los músicos es la perceptual, es decir, el sonido que el oyente percibe. A partir de esta realidad el músico es capaz de describir sonidos e imaginar de nuevos. Con los modelos espectrales representamos las características perceptuales de un determinado sonido y a la vez podemos manipular estas características para obtener nuevas realidades perceptuales.

A continuación presentamos con más detalle las técnicas de síntesis digital basadas en los modelos físicos y espectrales. Estas técnicas dan respuesta a muchos de los problemas actuales de la síntesis sonora y permiten una elección muy favorable dentro de los compromisos antes mencionados, especialmente a nivel de calidad, flexibilidad y generalidad. Hemos de tener presente, de todos modos, que por el hecho de ser algoritmos bastante complejos difícilmente se obtienen tiempos de cálculo bajos.

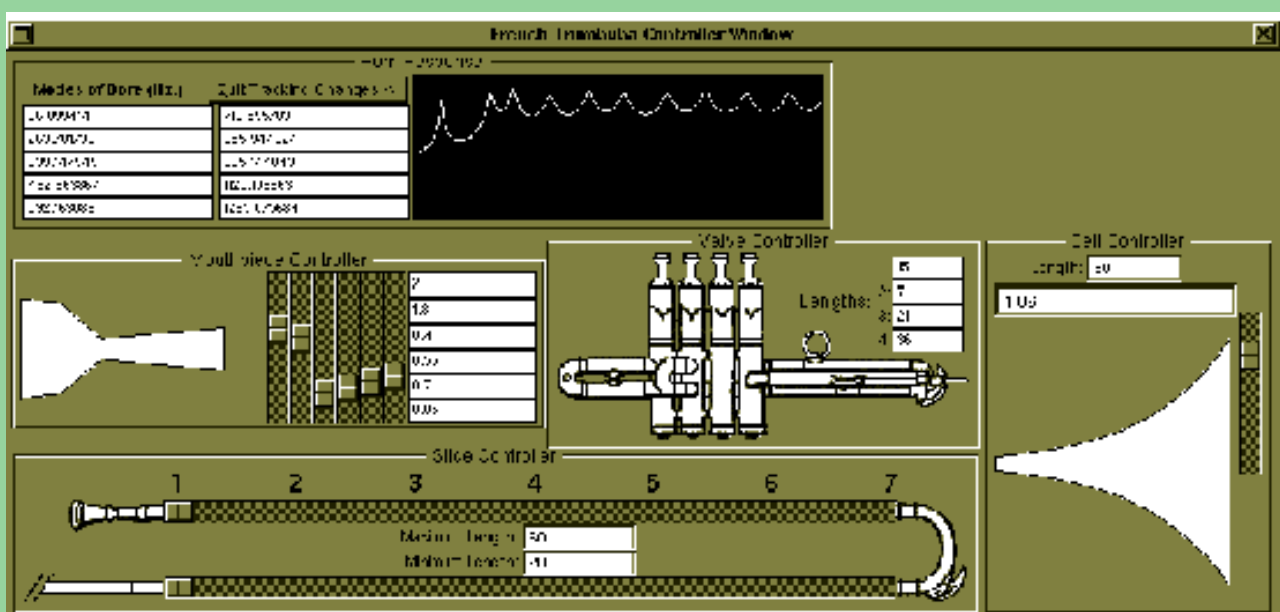


Fig. 2: Interficie gráfica para un modelo físico de un instrumento musical (por Perry Cook).

4. Síntesis con Modelos Físicos

Estos métodos de síntesis generan sonidos describiendo el comportamiento de los elementos que componen un instrumento musical, como cuerdas, lengüetas, labios, tubos, membranas y cavidades resonadoras. Todos estos elementos, excitados de una forma mecánica, vibran y producen perturbaciones, generalmente periódicas, en el aire que les rodea. Es esta perturbación la que nos llega a nuestro sistema auditivo y es percibida como sonido.

La acústica se ha preocupado de encontrar modelos, relaciones matemáticas, para describir estos sistemas vibratorios. Principalmente estos modelos se han utilizado para entender fenómenos físicos concretos, pero desde el trabajo pionero de Hiller y Ruíz (Hiller y Ruíz, 1971) también se han utilizado para la síntesis. El primer paso para la implementación de estos modelos consiste en definir y medir las características físicas del objeto a reproducir. Por ejemplo, para generar un sonido de cuerda hemos de considerar su longitud, grosor, densidad, etcétera... Una vez introducidas estas medidas físicas, el modelo nos permite reproducir numéricamente dentro del ordenador, el movimiento de cuerda y, a la vez, convertir este movimiento en sonido.

Generalmente todos los modelos instrumentales están compuestos por dos tipos de elementos: excitadores y resonadores (Borin et al., 1992). Los excitadores son los elementos que causan, y a la vez mantienen, el fenómeno vibratorio y los resonadores son donde las vibraciones de interés musical tienen lugar. En el caso de un violín, el arco funciona como excitador y la combinación de la cuerda con la cavidad de madera hace la función de resonador. Dentro del ordenador podemos tener modelos de los diferentes elementos, que combinándolos y detallando las características físicas de cada uno de ellos el usuario crea "instrumentos musicales". Con estos medios podemos diseñar y escuchar el sonido de un instrumento mecánico que sea físicamente imposible de construir. Por ejemplo, acoplando un modelo de doble lengüeta de caña para crear una excitación sonora, con un modelo de membrana de piel de ternero para hacer de resonador, generaremos sonido. Dudo que alguien pueda construir un instrumento como este fuera del ordenador.

El control de estos métodos de síntesis se hace a través de los controles mecánicos que tendría un instrumento acústico con las mismas características. Así, a un modelo de instrumento de viento de metal, hemos de dar, además de características de la excitación producida por los labios, los diferentes tamaños de los tubos y cavidades que componen el instrumento (Figura 2).

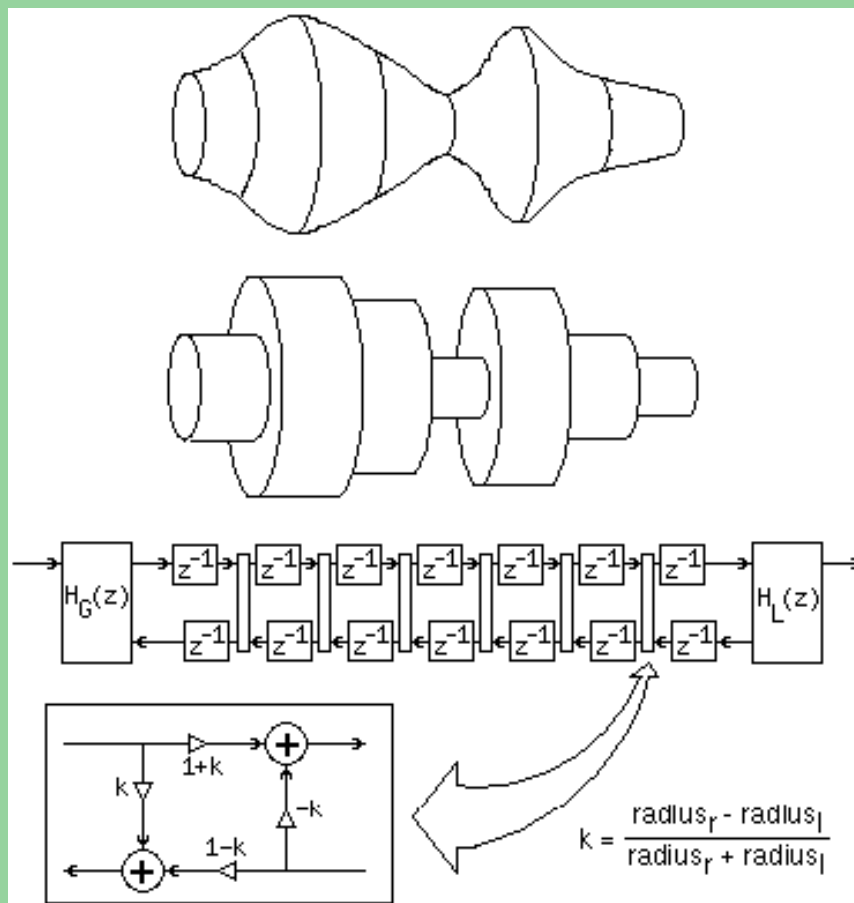


Fig. 3: Un tubo acústico (primer dibujo) puede mostrarse en el espacio de la misma forma que los sonidos (segundo dibujo). El tubo mostrado puede convertirse directamente en un filtro digital (tercer dibujo).

Históricamente los modelos físicos han sido implementados con algoritmos muy complejos que difícilmente pueden funcionar a tiempo real con la tecnología actual. Estas implementaciones se han basado en la integración numérica de la ecuación que describe la propagación de ondas en un fluido (Fletcher y Rossing, 1991). Recientemente se han encontrado soluciones más eficientes para este problema (Smith 1992) (Figura 3) y han comenzado a aparecer sistemas con interés para los músicos.

Estos modelos físicos pueden ser considerados como "generadores de realidad", no solamente en el sentido de imitar los instrumentos tradicionales sino también para ayudar a conceptualizar esta realidad y crear estructuras que no tienen por qué tener una interpretación física. En este caso la realidad física es utilizada como fuente de inspiración y no como una referencia de calidad del sonido producido.

5. Síntesis con Modelos Espectrales

Los modelos espectrales se basan en la descripción de las características sonoras que el oyente percibe. Para obtener un sonido de cuerda, en lugar de especificar las propiedades físicas, describiremos las características tímbricas, o espectrales, del sonido de cuerda. Entonces, la generación del sonido a partir de estos datos perceptuales se hace gracias a diversos procedimientos matemáticos desarrollados en los últimos años.

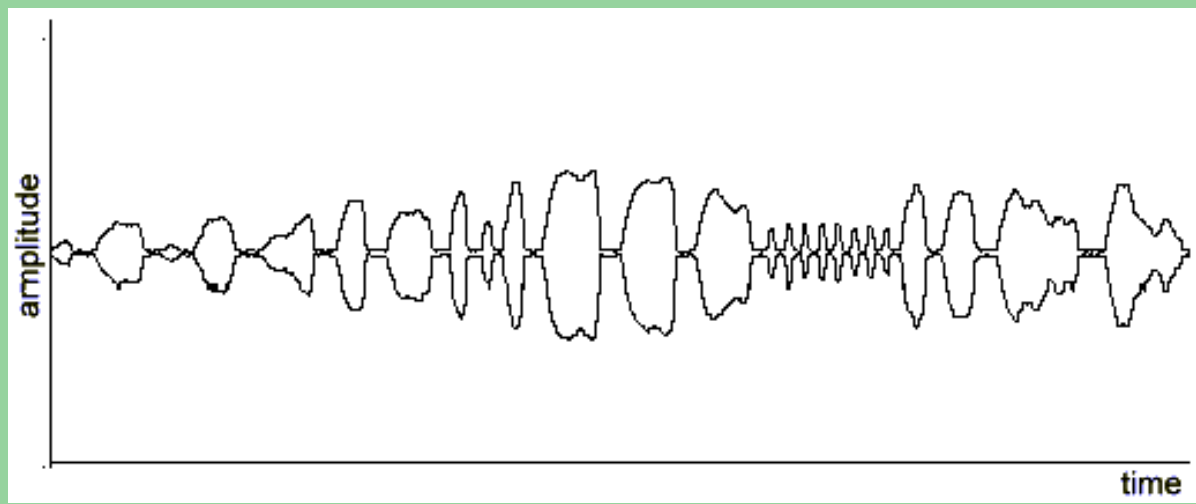


Fig. 4: Representación temporal del sonido de un pájaro.

Una ventaja de estos modelos es que existen técnicas para analizar sonidos y obtener los correspondientes parámetros perceptuales. Es decir, analizando un sonido determinado sacaremos sus parámetros de percepción. A partir del análisis es posible volver a sintetizar el sonido original y en el proceso, los parámetros pueden ser modificados de tal forma que el sonido resultante es nuevo pero conserva aspectos del sonido analizado.

Para un músico, es bastante intuitivo sintetizar sonidos con estos tipos de modelos. Por ejemplo, podemos comenzar con un sonido de clarinete, previamente analizado, e ir modificándolo con instrucciones como: hacer el sonido más inarmónico, hacer el ataque más percutivo, mezclar algunas características tímbricas de un sonido de voz, hacer el timbre más brillante, etcétera...El sonido se convierte en un material plástico que podemos ir modelando a nuestro gusto.

La transformada de Fourier es el primer paso hacia un modelaje perceptual del sonido. Con esta técnica un sonido (Figura 4) es descompuesto en sus armónicos, o componentes frecuenciales, de los cuales se puede estudiar su evolución en el tiempo (Figura 5). Un paso más respecto a la transformada de Fourier es el descomponer los sonidos en sinusoides (parciales) y ruido (componente residual), es decir, analizando sonidos con este modelo y generando de nuevos a partir de los datos de análisis (Serra, 1994; Serra y Smith, 1990). El análisis detecta los parciales estudiando las características espectrales de un sonido y los representa con sinusoides. Estos parciales se restan del ruido original y el "residuo" que queda se representa como sonido blanco filtrado. La parte de síntesis del sistema es una combinación de síntesis aditiva por la parte sinusoidal y síntesis por sustracción por la parte del ruido. Esta estrategia de análisis y síntesis puede ser utilizada tanto para generar sonidos (síntesis) como para transformar sonidos preexistentes (procesamiento). Para sintetizar sonidos, generalmente queremos modelar toda una familia tímbrica, por ejemplo, un instrumento, y eso se puede conseguir analizando notas aisladas y transiciones entre notas interpretadas con un instrumento, y construyendo una base de datos que caracterice todo un instrumento o cualquier familia tímbrica, a partir de la cual sintetizamos nuevos sonidos. En el caso de la aplicación del procesamiento sonoro, el objetivo es manipular cualquier sonido, es decir, no estar restringidos a notas aisladas ni requerir la obtención previa de una base de datos de análisis.

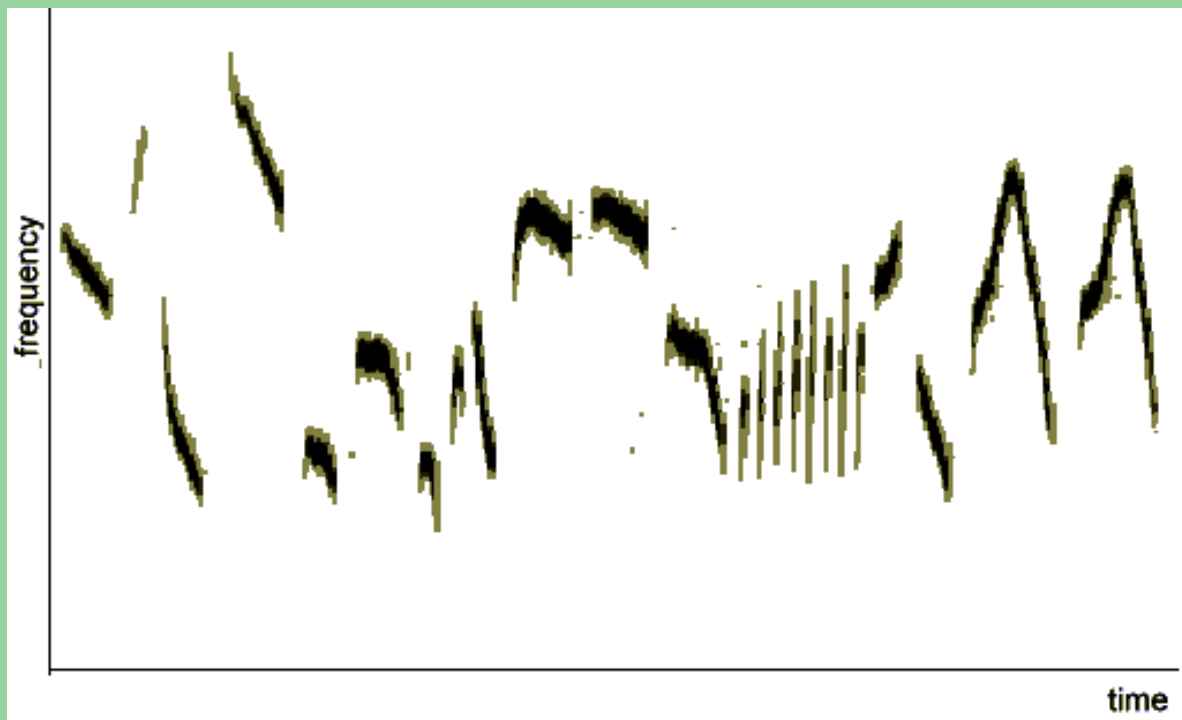


Fig. 5: Representación espectral del sonido de pájaro de la figura 4.

Los modelos espectrales pueden ser considerados como "transformadores de la realidad". Al igual que con los modelos físicos podemos reproducir una realidad ya existente, en este caso perceptual, y modificarla para obtener nuevos sonidos. La comparación entre estos dos tipos de modelos (Tabla 1) muestra una gran complementariedad entre ellos.

<i>modelos físicos</i>	<i>modelos espectrales</i>
basados en la realidad física	basados en la realidad perceptual
controlados por parámetros físicos	controlados por parámetros perceptuales
modelos específicos para cada instrumento	modelos generados para todos los sonidos
síntesis independientes de análisis	síntesis a partir de análisis
ideal para reproducir instrumentos tradicionales	menos ideal para reproducir instrumentos tradicionales
menos rápido para reproducir sonidos naturales	ideal para reproducir sonidos naturales
tiempo de cálculo alto	tiempo de cálculo alto
permite escoger entre calidad y flexibilidad	permite escoger entre calidad y flexibilidad
más cerca de la flexibilidad que de la calidad	más cerca de la calidad que de la flexibilidad

Tabla 1: Comparación de las características de los modelos físicos con los modelos espectrales

6. Conclusiones

La situación real de la síntesis digital todavía no está en estos niveles, los sintetizadores comerciales no permiten una aproximación a la generación sonora tal y como se ha presentado aquí. Pero nos estamos acercando. Concretamente no hace demasiado que salió el primer sintetizador comercial basado en modelos físicos (VL1 de Yamaha) y hay diversos sistemas informáticos, fuera de los canales comerciales tradicionales, que permiten experimentar tanto con modelos físicos como con modelos espectrales (por ejemplo el software distribuido por el IRCAM de París o el software de dominio público distribuido por la universidad de Stanford). Pero incluso estos programas todavía no están al nivel que querríamos. Esperamos que la investigación en este campo continúe y que no tardemos demasiado en disponer de este tipo de herramientas para la creación musical. Que finalmente el ordenador llegue a la madurez musical y cumpla las promesas que nos hicieron en los años sesenta.

7. Bibliografía

- BORIN, G.; DE POLI, G.; SARTI, A. (1992) "Algorithms and Structures for Using Physical Models". *Computer Music Journal*, 16(4): 30-42.
- MATHEWS, M. V. (1963) "The Digital Computer as a Musical Instrument". *Science*, 142(11): 553-557.
- MATHEWS, M. V., y otros (1969) *The Technology of Computer Music*. Cambridge (Mass.): MIT Press.
- FLETCHER, N. H.; ROSSING, T. D. (1991) *The Physics of Musical Instruments*. Nueva York: Springer-Verlag.
- SMITH, J. O. (1991) "Viewpoints on the History of Digital Synthesis". *Proceedings of the 1991 International Computer Music Conference*. San Francisco: International Computer Music Association.
- SMITH, J. O. (1992) "Physical Modeling Using Digital Waveguides". *Computer Music Journal* 16(4): 74-87
- HILLER, L.; RUIZ, P. (1971) "Synthesizing Musical Sounds by Solving the Wave Equation for Vibrating Objects: Parts I and II". *Journal of the Audio Engineering Society* 19 (6): 462-470 i 19(7): 542-551.
- SERRA, X. (1994) "Musical Sound Modeling with Sinusoids plus Noise". *Models and Representations of Musical Signals*. Cambridge (Mass) MIT Press. [llibre en preparació].
- SERRA, X. (1994) "Sound hybridization techniques based on a deterministic plus stochastic decomposition model". *Proceedings of the International Computer Music Conference*, 1994.
- SERRA, X.; Smith, J. (1990) "Spectral Modeling Synthesis: A Sound Analysis/Synthesis Based on a Deterministic plus Stochastic Decomposition". *Computer Music Journal*, 14(4): 12-24

