

Interfaces gráficas para la síntesis de audio digital

Daniel Gómez
Instituto Tecnológico Metro-
politano
danielgomez@itm.edu.co

Resumen

Este artículo plantea una discusión en torno a los sintetizadores digitales de audio, analizando particularmente el desarrollo de sus interfaces y la interacción que estos instrumentos ofrecen al intérprete. Se usará un esquema comparativo entre un sintetizador y un instrumento musical tradicional para analizar diferentes aspectos de su funcionamiento, relacionados con las propiedades físicas y los diferentes procesos cognitivos involucrados en el aprendizaje de ambos. Este ejercicio evidenciará una serie de características especiales que condicionan el uso de los sintetizadores relacionados con su forma, el tipo de aprendizaje que plantean y la respuesta física a la interacción. Adicionalmente al ejercicio comparativo, se analiza el sintetizador a la luz de la creciente práctica infográfica. Esta exposición revela posibilidades prácticas para el diseño de interfaces gráficas para sintetizadores digitales.

Este panorama es abordado con diversas estrategias de diseño de interfaces digitales, que dadas las características a intervenir usan herramientas conceptuales, gráficas, matemáticas y estadísticas. Como resultado, surgen prototipos de interfaces que desde sus posibilidades de interacción, retroalimentación visual y manera de agrupar los elementos, suponen nuevas posibilidades de control y cognición de la síntesis de audio digital.

David Sánchez
Instituto Tecnológico Metro-
politano
davidsanchez@itm.edu.co

Recibido: Agosto 5 de 2010
Aprobado: Noviembre 16 de 2010

Palabras clave: Cognición,
diseño de nuevos disposi-
tivos sonoros, interfaz, sínte-
sis de audio, ,

Graphic interfaces for digital audio synthesis.

Abstract

This article puts forward a discussion around audio digital synthesizers, particularly analyzing the development of their interfaces and the interaction these instruments offer to the interpreter. A comparative scheme between a synthesizer and a traditional musical instrument will be used to analyze different aspects of its functioning, related with the physical properties and the different cognitive processes involved in the learning to use them both. This exercise will evidence a series of special characteristics that condition the use of synthesizers which are related to their shape, the type of learning presented and the physical response to the interaction. Additionally to the comparative exercise, the synthesizer is analyzed under the light of the growing info graphics practice. This presentation reveals practical possibilities for the design of graphic interfaces for digital synthesizers.

Key words: Audio synthesis, cognition, design of new sound dispositives, interface.

138

This view is approached with a diversity of digital interfaces design strategies which, given the characteristics to be intervened, use conceptual, graphic, mathematics, and statistics tools. As a result, prototypes of interfaces have come up which, from their possibilities of interaction, visual feedback and way to group elements, suppose new control and cognition possibilities of the audio digital synthesis.

Evolución e interacción de los sintetizadores

El timbre, junto con la duración y la altura, es una de las tres dimensiones fundamentales usadas para caracterizar un sonido (Copland, 2002). El timbre es a su vez multi-dimensional: por un lado, para describir un timbre cualquiera, usualmente se usa una gran cantidad de palabras derivadas de la percepción de quien califica el sonido (brillante, opaco, sucio, denso, fuerte, etc.), decimos entonces que podemos clasificar un sonido, en lenguaje perceptivo, usando múltiples variables; por otro lado, los instrumentos tecnológicos usados para controlar el timbre, son también máquinas que definen el timbre basadas en una gran cantidad de variables científicamente universales (Russ, 2004) (como frecuencia de corte, tipo de filtro, tipo de onda, amplitud de la onda, amplitud de modulación, ataque, etc.). En ambos casos, tanto en la descripción como en el control del timbre, ya sea con lenguaje perceptivo o con lenguaje científico, el timbre es complejo y multidimensional, lo que lo hace engorroso muchas veces para alguien que se acerque a él ya sea para describirlo o para transformarlo.

Las máquinas usadas para el control del timbre son llamadas sintetizadores, y por razones prácticas y de estandarización son controladas usando el lenguaje científico. La distancia que hay entre un sintetista y el uso correcto de su máquina, es proporcional a la distancia entre su lenguaje perceptivo y el lenguaje científico (Seago, 2004). Esta distancia es un tema importante de estudio para la psicoacústica (Risset & Wessel, 1999), la interacción hombre-máquina y es un capítulo obligado para los diseñadores de nuevos instrumentos. Históricamente para cubrir esta distancia ha sido el sintetista quien ha debido traducir su lenguaje perceptivo al lenguaje de la máquina, y no al revés. Durante el siglo XX muchos investigadores (Wessel, Grey, Arfib, Jordà, Freed, Smith, Moog, entre otros) han tratado de conciliar esta distancia entre la percepción y la ciencia del timbre usando diversos métodos; pero a pesar de sus esfuerzos, las máquinas de síntesis comerciales aún se producen

usando lenguaje altamente científico. Una de las razones para que esto suceda es la imposibilidad de generar un modelo psicoacústico unificado, donde las unidades de medición del timbre sean unidades estandarizadas como dB o Hz, para amplitud y frecuencia respectivamente.

Instrumentos acústicos y sintetizadores

Los procesos de transformación del timbre son procesos inherentes a los sintetizadores, pero aparecen también en los instrumentos acústicos (Russ, 2004): cada instrumento acústico es en sí un sintetizador de sonido, con reglas propias muy definidas para el control de la altura, y con un cierto margen de trabajo, mucho más reducido, para el timbre. Las interfaces de los instrumentos tradicionales acústicos no están codificadas en lenguaje científico sino en un lenguaje físico: para acceder a su control debemos hacer esfuerzos de motricidad fina con un gran despliegue y coordinación física. Tanto los sintetizadores como los instrumentos acústicos requieren de una traducción del lenguaje perceptivo a otro lenguaje: físico / mecánico para los instrumentos acústicos y científico para los sintetizadores. En ambos la percepción del timbre debe ser traducida a un segundo lenguaje y el proceso de aprendizaje de ambos es la construcción de ese mecanismo de traducción entre una idea perceptiva y una certeza mecánica o científica (ver esquema en la Figura 1).

140

Diferencias formales

En los instrumentos tradicionales se accede corporalmente al proceso de generación del sonido, es por esto que la relación gestual con dichos instrumentos se vive de manera muy orgánica, ya que permiten una interacción en la que se puede sentir (ataque, presión, ángulo, vibración) (háptica) y algunas veces visualizar (golpe, velocidad, movimiento) (retroalimentación visual) el

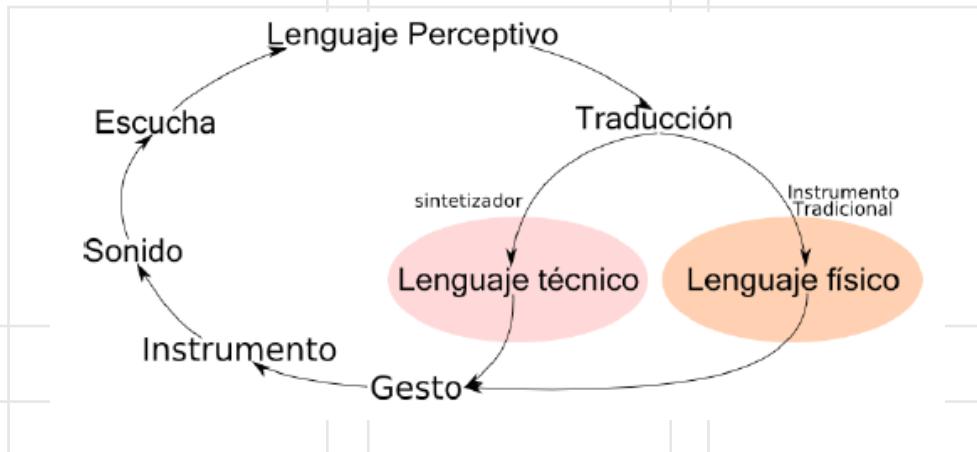


Figura 1. Retroalimentación cognitiva.

resultado sonoro. Los procesos cognitivos de aprendizaje consisten en la toma de conciencia de dichas relaciones (qué gesto me lleva a qué resultado) y de qué modificaciones o influencias en el sonido tienen las variaciones del gesto físico al tocar el instrumento (Goudeseune & Garnett, 1999).

En los sintetizadores físicos los procesos de generación del sonido no están basados en vibraciones mecánicas, sino en procesos electrónicos que solo se convierten en sonido al conectarse a un amplificador. Estos procesos se manipulan a través de una interfaz física que está dotada de potenciómetros y botones y codificada en un lenguaje científico. Por defecto en un sintetizador, el generador del sonido o motor de audio se encuentra separado de la interfaz física (Arfib, Couturier & Kessous, 2004). Los dispositivos que componen dicha interfaz se encargan de interpretar los gestos de un músico (giro, deslizamiento, presión, entre otros) y enviarlos al proceso electrónico. Los procesos cognitivos de aprendizaje de un sintetizador consisten entonces en hacer relaciones entre

el timbre del sonido percibido y el estado de las variables del sintetizador, principalmente en ausencia de guías hápticas o de retroalimentación visual, con muy baja exigencia motriz y sin referentes gestuales específicos.

Para aprender un instrumento acústico, el músico debe conocer a fondo la forma del instrumento y conectar las necesidades sonoras con gestos físicos. Sin embargo, la historia de la evolución de los sintetizadores dice que la forma de los instrumentos se basó en menor medida en la ergonomía y principalmente obedeciendo a necesidades mecánicas de la acústica de cada instrumento (Fletcher & Rossing, 1998). Por el lado de los sintetizadores, aunque existen formas estandarizadas y casi fetiches como la caja con potenciómetros y un teclado (Pinch & Trocco 2005), pueden tomar cualquier forma, ya que esta no tiene ninguna incidencia en el tipo de sonido que se genera y la exigencia física de sus controles es tan baja que podrían acoplarse a cualquier mecanismo.

Posibles roles de la infografía

Una consecuencia para el diseño de los sintetizadores generada por la separación entre forma y motor de audio es su mutación a sistemas digitales, en donde la aproximación del músico a estos es bastante diferente. Los sintetizadores digitales virtuales, representados por una imagen digital de los sintetizadores físicos, imponen una separación mayor entre el músico y la herramienta, pues plantean la interacción con dicha imagen, y solo pueden acceder a un control a la vez a través del ratón (Seago 2004). Aunque todas las posibilidades y la flexibilidad derivadas del avance en la computación gráfica están ahora en los computadores, estas aún no son lo suficientemente explotadas en los sintetizadores comerciales (Freed, 1995).

Tanto en los sintetizadores físicos como en los virtuales, la distribución de botones y potenciómetros en su superficie está relacionada con el flujo de voltaje interno entre los diferentes componentes del motor de audio. Las interfaces son entonces una especie de mapas con actuadores electrónicos que nos orientan a través del aparato. Estando el aparato conformado por esta gran cantidad de variables, y estando algunas conexiones no representadas en la interfaz, el mapa puede volverse confuso y superar el límite perceptivo del usuario, ya que estas interfaces pueden tener varias decenas de variables disponibles. Es muy probable que el uso de la infografía, que busca la optimización de la información en espacios bidimensionales (Tufte, 1990), pueda servir al diseño de las interfaces gráficas de los sintetizadores.

Otra metodología importante para la optimización de las interfaces de un sintetizador es la de crear un espacio tímbrico, usando técnicas de escalamiento multidimensional, que esté más cercano de un espacio perceptivo que de un espacio tecnológico o espacio-sintetizador (Grey, 1977). Esta estrategia, por un lado, busca definir variables perceptivas generales basados en experimentos de similaridad pero, por otro, ofrece una visión estructural y topográfica del timbre ubicado en un espacio perceptivo. Es importante ver que esta estructura multidimensional del timbre no es exclusiva de las variables perceptivas o del espacio tímbrico, sino que puede extrapolarse a espacios conformados por otras dimensiones.

Recopilación de la problemática

Para obtener guías para el diseño de sintetizadores virtuales que sobrepasen la frontera de la representación de los sintetizadores físicos, se han recopilado algunos decálogos y listas de deseo de investigadores en el área de la síntesis de sonido, donde se podrán encontrar diversos aspectos de esta problemática.

A finales de la década del 70 David Wessel manifestó, en relación a la síntesis aditiva, que era necesario buscar maneras de simplificar la manipulación de la cantidad de datos y variables trabajados, sin sacrificar la riqueza en el resultado tímbrico (Wessel, 1979). Goudeseune se refieren veinte años más tarde a este aspecto como el “problema de sobrecarga cognitiva” (Goudeseune & Garnett, 1999), sugiriendo como posible solución evitar mapeos muy simples, en los que el control de cada parámetro se haga uno a muchos (o muchos a uno), posibilitando que un gesto manipule varias variables. Jordà muestra cómo esto se encuentra en los instrumentos acústicos, donde el control de la altura siempre manipula timbre y amplitud a su vez (Jordà, 2005), y complementa exponiendo que los sintetizadores no deben mostrar todas sus posibilidades de síntesis al principio, sino que estas se deben ir percibiendo proporcionalmente al crecimiento del nivel de comprensión que el interprete tenga de los instrumentos; siempre en la búsqueda de mayor eficiencia con menor esfuerzo físico y conocimiento de ejecución, en contraste con los instrumentos tradicionales.

A mitad de la década del 90 Adrian Freed aparece con cinco sugerencias para el diseño de interfaces (Freed, 1995), las cuales, excepto la del procesamiento y síntesis de audio en tiempo real (área pertinente más a la tecnología que al diseño), siguen todavía muy poco exploradas. Las primeras dos apoyan la infografía digital y muestran cómo el uso del color y el modelado en 3D podrían ayudar mucho en el control, y retroalimentación visual de los procesos tímbricos, sin embargo a pesar de que Seago advierte que no hay ninguna representación del sonido psicológicamente significativa más que el sonido mismo (Seago, 2004), muestra cómo de los tres tipos de taxonomía de sintetizadores: arquitectura fija, especificación según la arquitectura y especificación directa; los últimos, que hacen representaciones de la onda, son los que permiten una manipulación más intuitiva y directa del procesamiento tímbrico. Por último, Seago realza la importancia de la escalabilidad multidimensional como un acercamiento a la especificación tímbrica.

El decálogo de Sergi Jordà de su tesis Digital Lutherie (Jordà, 2005) es quizá el más completo encontrado hasta ahora, sus recomendaciones se podrían dividir en cuatro clases: escalabilidad cognitiva, ejecución interactiva, amplitud y detalle en el control y comprensión del mecanismo interno. Las que tienen que ver con escalabilidad cognitiva fueron enunciadas en los párrafos anteriores en relación al problema de sobrecarga cognitiva expuesto por Goudeseune. Las que tienen que ver con ejecución interactiva sugieren delegar parte del proceso creativo al computador, de esta forma se aprovecharía su "inteligencia" y capacidad de intervención, lo que ayudaría a superar los límites de los instrumentos tradicionales en donde de una acción se deriva un resultado sonoro. Las que tienen que ver con amplitud y detalle en el control sugieren que los instrumentos deben permitir acceso a toda su complejidad, mas no dejar a un lado los cambios sutiles y precisos en su mecanismo; algo que le daría amplitud y riqueza a un instrumento sería la posibilidad de tener una plataforma multiusuario, en donde se podrían manipular más parámetros de forma más detallada sin causar sobrecarga en un solo usuario. Por último, las recomendaciones que tienen que ver con comprensión del mecanismo interno, sugieren aprovechar la capacidad de la pantalla para representar y monitorear procesos complejos, donde este soporte visual dé cuenta del estado y comportamiento de la plataforma operativa; en este punto habla también que es necesario evitar el ratón como controlador ya que, al igual que lo expresaba Freed en su decálogo de deseos (Freed, 1995), sus limitaciones no permiten manipular los instrumentos en toda su dimensión y amplitud.

Diseños

Con la información revisada anteriormente y el estudio de las problemáticas encontradas en diversas décadas, se ha iniciado un proceso de diseño que busca solucionar algunas de las necesidades existentes y escalarlas a un mundo virtual.

Interfaz Escalable

La dimensión de un sintetizador, en los inicios de su definición como instrumento (Moog, 1964) fue modular y escalable (era posible remover o aumentar algunos módulos, dependiendo de las necesidades de cada sonido). Aunque el montar y desmontar módulos físicos es poco práctico debido al tiempo y esfuerzo que debe ser invertido en ello, en los sintetizadores de software es posible retomar la idea de variar la dimensión de la interfaz dependiendo del uso en la conformación del sonido. La mayoría de sintetizadores software usa sin embargo una versión de interfaz digital más cercana al Minimoog físico en donde las conexiones entre los módulos son ocultas y la interfaz es rígida. Como estrategia para superar estas problemáticas, se ha propuesto el diseño y construcción de una interfaz escalable a la medida del usuario, en donde las variables se vayan mostrando o escondiendo a medida que el sonido pre programado se vuelve más complejo o que el usuario así lo desee.

Interfaz con variables perceptivas

Las interfaces físicas son en su gran mayoría interfaces diseñadas con lenguaje técnico. El lenguaje técnico es objetivo, pero no fácilmente interiorizable y para un músico no entrenado en síntesis de audio tan vago como sus propias descripciones de el sonido en términos perceptuales. Independiente de que exista un acuerdo o no entre el instrumento y el músico, la interfaz ya viene con unas dimensiones y unas conexiones propias del diseño del motor de audio. Las dimensiones de la interfaz podrían acercarse al usuario si éste simplemente las nombra con palabras de su propio lenguaje. Es decir, tener una interfaz funcional pero sin definiciones previas de las variables, para que sea el mismo usuario quien les asigne nombres de acuerdo a su funcionamiento. Obviamente sería un proceso de nombrar y corregir hasta que preceptivamente la interfaz tenga un lenguaje adecuado para cada usuario particular. Profundizando en esta

dirección, el diseño de interfaces adaptables al usuario en niveles superiores como la definición de un espacio tímbrico perceptivo de acuerdo a evaluaciones sencillas de calibración, en donde la herramienta le haga preguntas al usuario de tal manera que pueda calibrar automáticamente las dimensiones y los rangos de su interfaz. Esta familia de interfaces adaptables es otro de los prototipos a desarrollar.

Sintetizador Estadístico

La estadística se ha usado de diversas maneras en el control de nuevos instrumentos musicales. Herramientas estadísticas que sirven para auto completar discontinuidades en el control de un sintetizador o variables ínter relacionadas por métodos estadísticos (Loviscach, 2008), el uso de distribuciones estadísticas para asignar valores a variables que comparten un mismo plano (Momeni et al., 1995). Es posible también usar bases de datos de sonidos pre configurados en sintetizadores (presets) para navegar a través del espacio conformado por las variables de un sintetizador o espacio-sintetizador. Esta perspectiva genera la posibilidad de que el músico decida qué tan “amplia” es su interfaz. La amplitud de la interfaz puede ir desde cerrada, limitada por los puntos más comunes de toda la base de datos; máxima, ofreciendo todas las posibilidades independientes de su frecuencia de uso; hasta tener amplitud negativa, favoreciendo zonas del espacio-sintetizador que son estadísticamente desechadas.

Conclusiones

La multidimensionalidad del timbre y la falta de unidad de un modelo psicoacústico para su análisis, crean la necesidad de involucrar diferentes ramas del conocimiento para facilitar la tarea de traducción entre dimensiones del

lenguaje científico y perceptivo. Estas ramas deben aportar guías perceptivas adicionales al lenguaje científico para tratar de cerrar la brecha usuario-máquina estimulando la cognición del músico, usando anchos de banda disponibles en otros sentidos adicionales a la escucha. Los instrumentos acústicos hacen esto en mayor o menor grado.

Las limitaciones físicas de los instrumentos musicales tradicionales hacen que la motricidad fina de los músicos y la creatividad de los compositores, sean el recurso para expandir su sonoridad. Las diversas sonoridades y timbres derivados de variaciones de técnicas interpretativas, hacen parte de la paleta de colores que un instrumento acústico ofrece. Esta paleta es limitada. Un sintetizador ofrece muchas dimensiones de control y sus combinaciones prometen timbres y colores infinitos, pero la densidad en la cantidad de controles hace que un músico promedio se aleje de los instrumentos debido a la sobrecarga cognitiva. Es posible pensar que un instrumento musical a medio camino entre pocos parámetros y la sobrecarga cognitiva favorezca el acercamiento de nuevos usuarios. No necesariamente el instrumento con la mayor cantidad de variables y conexiones es el instrumento más adecuado para sintetistas promedio.

Algunas de las necesidades planteadas por investigadores en décadas pasadas, en relación con el uso de los gráficos para representar los instrumentos musicales, aún están vigentes. Éste es un campo de investigación que aún está en desarrollo. Tanto en la parte técnica como en la cognición y la respuesta al trabajo con diferentes tipos de interfaces gráficas. El desarrollo de herramientas gráficas con bajo nivel de procesamiento pero que estimule enlaces cognitivos entre la imagen y el lenguaje técnico del sintetizador (y entre la imagen y el sonido), es obligatorio para el diseño de mejores herramientas musicales.

Las mejoras en los nuevos instrumentos musicales pueden verse como optimizaciones del proceso cognitivo que acerquen el músico a su herramienta.

La mayoría de los esfuerzos se dan en mejorar la herramienta para que el músico la pueda entender mejor y en menor tiempo, pero hay un camino grande por recorrer en la optimización de la máquina para que entienda al usuario. Este camino va desde interfaces cuyos nombres sean personalizables por el usuario, hasta interfaces que generen espacios tímbricos multidimensionales basados en la percepción misma del usuario y que basados en entender su percepción ajusten sus controles de síntesis.

Muchos de los deseos y conclusiones de los investigadores para el diseño de nuevos instrumentos aun están por resolverse después de años de haber sido formulados. El trabajo presentado en este artículo no pretende dar soluciones definitivas a la problemática sino que plantea una metodología reflexiva para allanar el terreno, verificando literatura y proyectos de otros autores, analizando instrumentos acústicos y tratando de resolver problemas discretos o grupos de problemas en vez de enfrentarse a toda la problemática en un solo diseño.

Bibliografía

ARFIB, D., J. M. COUTURIER & L. KESSOUS. (2004). *Design and Use of Some New Digital Musical Instruments*. En *Gesture-Based Communication in HCI*. Berlín, Alemania: Springer.

COPLAND, A. (2002). *What to Listen for in Music*. First edition. USA: Signet Classics. p. 63.

FLETCHER, N. & T. ROSSING. (1998). *The Physics of Musical Instruments*. NY, USA: Springer.

FREED, A. (1995). *Improving Graphical User Interfaces For Computer Music Applications*. *Proceedings of the CHI*. *Computer Music Journal*, 19: 4-5.

GOUDESEUNE, C. & G. E. GARNETT. (1999). *Performance Factors in Control of High-Dimensional Spaces*. *Proceedings ICMC International Computer Music Conference*, Beijing, China.

GREY, J. M. (1977). *Multidimensional perceptual scaling of musical timbres*. *Journal Acoustical Society of America*, 61: 1270-1277.

JORDÀ, S. (2005). *Digital Lutherie*. Tesis Doctoral, Departamento de Tecnología, Universitat Pompeu Fabra, Barcelona, España.

LOVISCACH, J. (2008). *Programming a Music Synthesizer through Data Mining*. *Proceedings. New Interfaces for Musical Expression 2008*, Genova, Italia (207-210).

MOMENI, Ali & David WESSEL. (2003). *Characterizing and controlling musical material intuitively with geometric models*. Proceedings Conference on New Interfaces for Musical Expression NIME, Montreal, Quebec, Canada.

MOOG, Robert A. (1964). *Voltage-Controlled Electronic Music Modules*. AES Convention 16 (October 1964). p. 346.

PINCH, T. and F. TROCCO. (2005). *Analog Days: The Invention and Impact of the Moog Synthesizer*. Cambridge, MA, USA: Harvard University Press. p. 53-69.

RUSS, M. (2004). *Sound Synthesis and Sampling*. 2ª Edición. Cap. 2. Burlington, USA: Focal Press.

TUFTE, E. (1990). *Envisioning Information*. New York, USA: Graphics Press.

WESSEL, D. L. (1979). *Timbre space as a musical control structure*. Computer Music Journal, 3 (2): 45-52.