

ESTUDIO COMPARATIVO DE LAS TÉCNICAS ESTÉREO Y LA TÉCNICA BINAURAL EN LA GRABACIÓN DE RUIDO INTERIOR DE VEHICULOS

REFERENCIA PACS: 43.38.Md

Ordoyo Ansorena, Fabiola;
Ingeniería La Salle, Universidad Ramon Llull
Paseo Bonanova 8
08022 Barcelona. España
Tel: 34 932 902 427
Fax: 34 932 902 416
E-Mail: fabiola@salleurl.edu

ABSTRACT

The representation of sound reality in our brains is done by a binaural take, so we can localize the origin of the sound source in the sphere that is surrounding us. But when one records this reality in order to analyze it and, among other experiments, to anticipate the measure of annoyance or satisfaction that her listening produces, we must study other microphone techniques to improve some sides of the binaural method. In order to come to good conclusions about this controversy, we have done an exhaustive research that compares objectively and subjectively all the stereophonic techniques and binaural method. An special account will be done in car applications.

RESUMEN

La representación de la realidad sonora en nuestro cerebro se realiza a través de una toma binaural, lo cual nos permite localizar el origen de la fuente de sonido en la esfera que nos envuelve. Pero cuando se trata de registrar esta realidad para poder analizarla y, entre otros experimentos, prever el grado de molestia o de satisfacción que produce su escucha, nos planteamos el uso de otras técnicas de microfonía que mejoren diferentes aspectos respecto la toma binaural. Para poder obtener conclusiones válidas en esta polémica, realizaremos un exhaustivo estudio comparativo, a nivel objetivo y subjetivo, de las diferentes técnicas estereofónicas y de la técnica binaural. Se hará un inciso en el resultado del estudio para aplicaciones en el mundo de la automoción.

INTRODUCCIÓN

En el campo de la automoción se ha demostrado que no son suficientes las medidas con sistemas convencionales que nos dan niveles en dBA y el análisis frecuencial en tercios de octava. Para mejorar la calidad del sonido en el interior de un vehículo es necesario realizar otras medidas que tengan en cuenta aspectos psicoacústicos. El desarrollo de los sistemas de medida psicoacústicos requiere cuantificar estadísticamente la impresión subjetiva que tienen diferentes personas. Posteriormente se podrá relacionar este primer resultado subjetivo con el resultado analítico.

Tenemos que escoger técnicas que nos den una buena localización de la fuente sonora. La fusión o formación de la imagen sonora que sintetiza la información de diferentes fuentes debe respetar la imagen original cuanto sea posible. Debe existir un equilibrio entre L y R, y las diferencias de retrasos y niveles nos darán la localización de la imagen y extensión estereofónica.

PARÁMETROS A TENER EN CUENTA EN UNA MEDIDA A TRAVÉS DE MICRÓFONOS

En función de la técnica utilizada nuestra medida se verá afectada a nivel objetivo y a nivel subjetivo. El cambio en las características del sonido grabado a través de los micrófonos es básicamente frecuencial y esto afecta tanto en el momento de analizarlo científicamente como auditivamente. Este fenómeno se llama comunmente coloración del sonido. En este último análisis es donde la subjetividad juega un papel importante. Parece magia cuando a través de un micrófono podemos escuchar sonidos que suceden en la realidad, pero que en una situación de escucha normal nuestro cerebro ignora casi por completo.

La respuesta frecuencial del sistema de micrófonos, ya sean uno, dos o más, afecta implícitamente a la forma de onda del sonido. La envolvente de esta forma de onda entonces también varia; podremos caracterizar el cambio introducido por el uso de una técnica microfónica en la curva ADSR (Attack Decay Sustain Release) de éste.

Respuesta frecuencial: Medida de la sensibilidad en todo el espectro. Una respuesta irregular introduce coloraciones en el sonido. Cuando hablamos de técnicas estereofónicas, debemos contemplar la "respuesta frecuencial" del conjunto, ya que nosotros interpretamos las diferencias de fase y de nivel que se dan en los dos señales grabados a través de dos micrófonos. Estas diferencias de fase y de nivel se pueden traducir en una respuesta global del sistema, que creemos puede ser proporcional al espectro cruzado de los dos señales obtenidos. Más adelante veremos como se ha realizado la medida.

Respuesta impulsional: Medida de la velocidad de respuesta a un transitorio. Una respuesta impulsional lenta se traduce en una acumulación del ataque de un sonido con el release del sonido anterior y el sonido es más sucio, no tan definido. Se pierde inteligibilidad secuencial.

En el caso de una buena respuesta impulsional, tendremos más riqueza en la alta frecuencia y más claridad en el seguimiento del ADSR del sonido.

Ruido equivalente: Se define como el valor de la presión incidente que daría la misma tensión provocada por el ruido del micrófono. Este ruido se llama térmico o eléctrico, y corresponde a la expresión $\sqrt{4KRT}$, donde K es la cte. de Boltzman, R la resistencia del conductor y T la temperatura. Dada esta expresión podemos deducir cuales serán los micrófonos más ruidosos; son los dinámicos, ya que tienen un conductor (bobina) que es el que da la resistencia.

Esto afectará directamente al margen dinámico del sistema, y dependerá del micrófono utilizado.

Sobrecarga o potencia máxima: La saturación de una membrana acústica implica que ésta deje de vibrar según la excitación del sonido y le sume distorsión armónica. Se deberá controlar también el nivel de entrada del amplificador, ya que éste tiene un nivel óptimo de trabajo.

TÉCNICAS ESTEREOFÓNICAS MÁS EMPLEADAS EN LA GRABACIÓN DE EVENTOS

La grabación estereofónica plasma el paso del sonido a través del espacio. Esta información incluye la posición de las diferentes fuentes en el espacio, la distribución de las primeras reflexiones y la reverberación. Para registrar en estéreo se necesitan como mínimo

dos micrófonos de medida. Lo más importante a la hora de establecer una técnica de grabación como óptima es ver si esta registra el campo sonoro de una manera simple y exacta.

Existe una primera distinción entre las tomas en función de la distancia a la fuente (toma distante, toma ambiente y toma cercana). También podemos clasificar en función del número de micrófonos utilizado y el método de medida. Nos encontramos con métodos de intensidad (MS, XY, mono, un micrófono principal más auxiliares,), métodos de retraso de propagación (AB, varios micrófonos,...) y métodos de mezcla (ORTF, OOS, técnica binaural, un micrófono principal más auxiliares,...).

La utilizada en la primera fase del estudio es la toma cercana, puesto que es la que se da en la grabación de sonido interior de vehículos. Haremos una breve explicación de las técnicas estereofónicas más utilizadas en grabaciones musicales, donde nos hemos basado para mejorar los sistemas de medida:

. Técnica X-Y: Es un sistema dependiente de la intensidad, funciona si existen diferencias de presión entre las señales de los micrófonos. La técnica es de par coincidente, y el ángulo entre micrófonos varía de 90 a 180°. Aun estando los micrófonos muy cerca, la imagen estéreo es excelente. El patrón polar utilizado es el cardiode, aunque el uso de bidireccionales (técnica Blumlein) también da muy buenos resultados. Si estos micrófonos están aparejados en fase, el resultado es óptimo.

. Técnica ORTF (Organisation Radio TV Française): Es el procedimiento ideal, donde se establece una separación de 15 a 20 cm., y el ángulo entre micrófonos es de unos 110°. Esta técnica de par casi coincidente tiene una buena localización, ya que estamos aumentando las diferencias de nivel y fase. A altas frecuencias se actúa básicamente por incremento de nivel y podremos separar mas los micrófonos.

. Técnica NOS: Técnica holandesa donde la separación de micrófonos es de 30 cm. y la fase relativa de 90°. Se obtienen unas diferencias temporales muy acusadas.

. Técnica OOS: Técnica con micrófonos omnidireccionales situados en forma de cruz y separados unos 16'5cm.. Se sitúa un disco de 30 cm. de diámetro y forrado de espuma lateralmente. Este sistema simula el comportamiento binaural. Para frecuencias menores a 200 Hz no hay imagen estéreo. Se podrían utilizar cardiodes, pero los omnidireccionales dan menos distorsión y mejor respuesta.

USO DE LA TÉCNICA BINAURAL

La acústica binaural estudia una estereofonía determinada, que es la que reproduce la audición humana. Se realiza con maniquis con torso, cabeza y orejas para simular un cuerpo humano. En la entrada del canal auditivo se sitúan micrófonos de medida para poder registrar los señales justamente antes de su paso por el canal. Estos señales han sufrido un modelado ("spectral shapping") debido a las reflexiones en el torso, cabeza y pabellón auditivo.

La ventaja de este método de medida es que las diferencias de nivel y de fase acústica (debido a los retrasos en la propagación) se contemplan y los cambios en el campo sonoro causados por la persona pueden ser reproducidos aproximadamente.

Teóricamente es una técnica que permite diferenciar entre las incidencias frontales y posteriores gracias a la no simetría de la cabeza y de los pabellones auditivos. En la práctica, esto ocurre con un margen de error bastante grande, ya que una cabeza artificial tiene unas dimensiones y unas características basadas en unos valores promedio y, por tanto, a la hora de la reproducción, se dan problemas de localización. En concreto, una de las zonas donde se producen mas errores es en el plano medio (hasta un 50%).

El mayor inconveniente en el uso de esta técnica es el no poder controlar los cambios a nivel frecuencial (coloración) que sufre la señal como resultado del modelado. Otro problema

añadido es el hecho que, incluso en un mismo fabricante de "dummy-heads", se producen variaciones en la coloración producida por la cabeza artificial.

En el caso de incidencia frontal en campo libre, las funciones de las cabezas artificiales variarán. Estas diferencias se pueden ajustar con funciones de corrección adecuadas, pero es muy difícil corregir el resto de las características direccionales con un filtro de este tipo. Esto implica un serio problema de compatibilidad entre distintos fabricantes, por lo que se hace imposible crear bases de datos y realizar comparaciones válidas de medidas.

Otro inconveniente de esta técnica es el sistema necesario para su reproducción mediante altavoces. Es necesario todo un complejo procesado de la señal estéreo a través de un cancelador de crosstalk para poder reproducir correctamente la señal binaural en el punto de escucha, punto que será muy limitado en cuanto a cobertura ("sweet spot"). La dificultad del proceso es alta, ya que invertir HRTFs (Head Related Transfer Functions) no es trivial debido a su naturaleza (Fig. 1).

Teniendo en cuenta que estas señales se deberán analizar, se deberán filtrar por las HRTFs inversas del maniquí para poder recuperar la señal original procedente del campo sonoro real, ya que ésta ha sido modificada a nivel frecuencial ("shapping") por el maniquí.

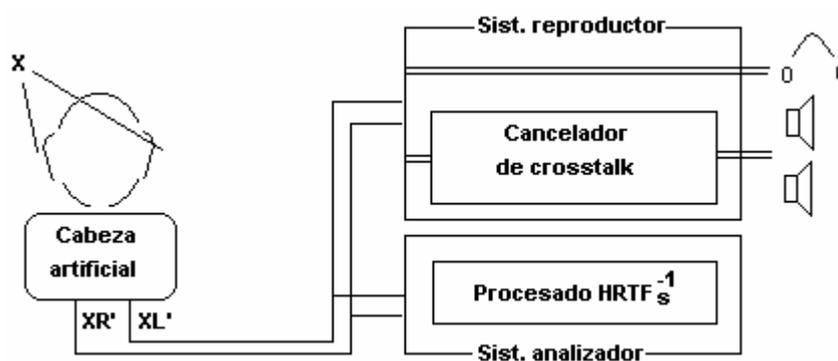


Figura 1. Sistema reproductor y analizador utilizando la técnica binaural

VARIACIONES EN LA SENSACIÓN AUDITIVA DE LAS DIFERENTES TÉCNICAS

El pitch es la sensación subjetiva percibida por el hombre de una determinada frecuencia. La manera en que se percibe el pitch es muy enigmática, e intervienen muchas variables. Como primera experiencia a comentar, en la escucha de varios tonos simultáneamente, se produce distorsión no lineal en la cóclea que provoca la audición de otras frecuencias, las cuales son combinaciones de las anteriores. Los productos más simples de la distorsión producida por el sistema auditivo son los armónicos aurales. Como su nombre indica, son productos de la distorsión que contienen frecuencias que son múltiplos de las frecuencias del estímulo.

Si la diferencia entre dos tonos es pequeña (por ejemplo 3 Hz), entonces los patrones de excitación en la cóclea se solapan considerablemente, así que los dos estímulos no se podrán distinguir. La diferencia entre los dos tonos causará que estos estén en fase y en contra fase cíclicamente, de manera que se produzcan batidos a esta frecuencia diferencia. La percepción de los "beats" aurales refleja la poca habilidad del oído a la hora de resolver frecuencias.

La percepción del pitch también se basa en factores temporales (periodicidad). La percepción del pitch basada en una periodicidad en el estímulo se llama descriptivamente *periodicity pitch* (pitch de periodicidad). Los armónicos más importantes a tener en cuenta en

la percepción de tonos complejos son, como resultado de algunos experimentos, hasta los terceros y quintos armónicos.

El "missing fundamental" (frecuencia diferencia de dos tonos simultáneos) se percibe incluso en situación de enmascaramiento a esa frecuencia. Esto demuestra que es un resultado de las distorsiones producidas en la cóclea. En otras palabras, tonos de baja frecuencia y frecuencias diferencia reales si pueden enmascarse entre ellas, pero el "missing fundamental" no. Es evidente que este resultado se basa en una percepción neuronal para el pitch de periodicidad.

La base neuronal de todas estas experiencias, es decir, el hecho que el sistema nervioso central auditivo participe en la percepción de las variaciones de pitch queda demostrado en el siguiente estudio, donde se utilizan estímulos dicóticos, diferentes estímulos presentados en cada oído. Houtsma y Goldstein propusieron estimular con una frecuencia cada oído independientemente. Los sujetos identificaban melodías basadas en la percepción de la señal diferencia (el "missing fundamental"). Si esta señal solo fuera producto de la distorsión en la cóclea, estos individuos no podrían oírlo, ya que en cada cóclea solamente existe un tono. Si esta frecuencia fantasma aparece incluso cuando los armónicos se presentan separadamente en los oídos, el fenómeno tiene que ocurrir en el sistema nervioso auditivo central, ya que ésta es la única región donde los armónicos se representan simultáneamente. Podríamos estar hablando de un posible procesador central de pitch.

Este fenómeno explica en parte las diferencias audibles que se presentan al registrar diferentes sonidos con distintas técnicas estereofónicas. Podemos deducir que lo que está haciendo el sistema central auditivo es similar a correlacionar las dos informaciones que le llegan desde los dos oídos. En función de la distancia entre los micrófonos en el momento de la grabación, y juntamente con la respuesta en frecuencia de los transductores en función de la directividad, sobretodo a alta frecuencia, obtendremos un espectro cruzado de las dos señales u otro. Esto es importante tenerlo muy en cuenta, para no realizar diagnósticos equivocados respecto a la molestia o al agrado que pueden producir sonidos grabados en el interior de un coche, ya que la sensación que produce la grabación será distinta, en mayor o en menor medida.

MEDIDAS REALIZADAS Y CONCLUSIONES

Experimentando con distintas técnicas podemos obtener la respuesta de la técnica a nivel subjetivo: qué técnica representa mejor la realidad, tanto en la localización como en las características audibles; y la respuesta a nivel objetivo, el espectro cruzado que nos muestra las diferencias a nivel frecuencial que percibiremos en función de la técnica (ejemplos: figuras 2 y 3).

El método de la medida subjetiva debe ser estadístico; evidentemente, cada individuo puede presentar una respuesta diferente en la localización de las fuentes de ruido. Se realizará en el entorno real, comparando con las grabaciones realizadas en el mismo lugar.

Los micrófonos utilizados son de medida, de respuesta plana y poca variación de la directividad a altas frecuencias. Las medidas las hemos realizado con una fuente de ruido rosa reproducida a través de un solo altavoz de directividad suficientemente constante en la zona de campo cercano, donde se han realizado las medidas. El lugar de estudio ha sido en cámara anecoica, para minimizar así el número de variables influyentes en la medida. No olvidemos que nuestro primer propósito era caracterizar mediante una respuesta en frecuencia del conjunto cada técnica estereofónica que nos interesara. Para el test a nivel subjetivo se grabarán sonidos en el campo real. Evidentemente, no podemos aspirar a localizar la fuente (altavoz emisor de ruido rosa) en un entorno como el que hemos comentado, ya que no hay información de reflexiones, reverberación, etc. de la sala, que es la que facilita la localización de las fuentes sonoras.

Hemos calculado el espectro cruzado, así como la correlación temporal entre las dos señales, y hemos visto las diferencias audibles entre las distintas técnicas plasmadas en el análisis. Estas diferencias son atribuidas mayoritariamente al procesado interno de las señales

en nuestro sistema nervioso auditivo. Otro factor que influye en la obtención de resultados diferentes es la variación de la directividad a altas frecuencias en los micrófonos, aunque los micrófonos utilizados sean lo menos variables posible. Esta dependencia se nota escuchando o analizando directamente un mismo canal en dos técnicas diferentes.

Como es de esperar, la técnica binaural es la que permite una mayor localización, pero como ya hemos comentado, modifica en exceso el señal a analizar y la sensación de realismo disminuye. La técnica XY con ángulo 180° es la que introduce mas pérdidas a altas frecuencias, notándose un sonido mas compacto y grave en comparación con otras técnicas.

Para poder comparar análisis así como poder ser capaz de crear bases de datos compatibles de grabaciones de sonido interior de coches, es necesario definir un estándar común para grabaciones con cabeza artificial.

Las ventajas del uso de técnicas microfónicas estereofónicas simples son una facilidad en el tratamiento de datos, un menor coste, con diferencia, que los sistemas con dummy-heads, el uso de micrófonos ordinario, que todo el camino que recorre la señal está calibrado, los prototipos se pueden describir matemáticamente fácilmente y una normativa estándar se puede obtener de forma sencilla.

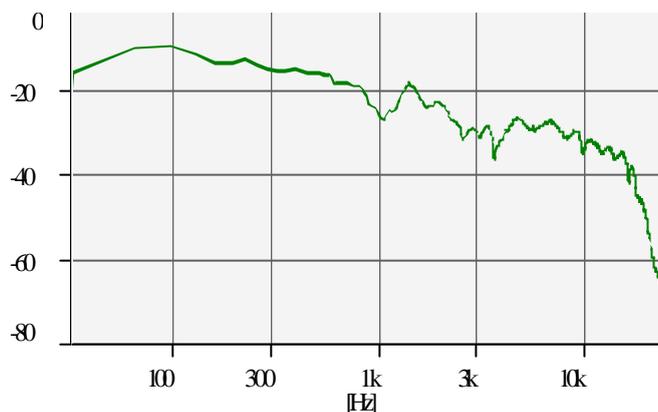


Figura 2. Cross-spectrum técnica ORTF.

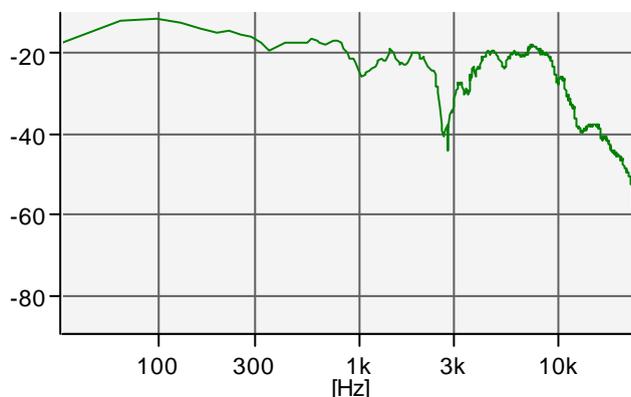


Figura 3. Cross-spectrum técnica binaural.

REFERENCIAS BIBLIOGRÁFICAS

- [1] Gelfand, Stanley A., "Hearing - An introduction to psychological and physiological acoustics", Marcel Dekker, Inc., Butterworths, 1981.
- [2] Pflüger, M., "Evaluation of recording methods for vehicle interior sound - Different artificial heads and comparative systems", AVL - Sound Engineering Conference, 1999.
- [3] Ordoyo, Fabiola, " Processament de senyal binaural per reproduir so 3-D mitjançant altaveus ", Dept. d'Acústica d'Enginyeria La Salle, Barcelona, 1998.
- [4] Begault, Durand R., " 3D Sound for virtual reality and multimedia ", AP Professional, Academic Press, Inc., 1994.
- [5] Ballou, Glen, " Handbook for Sound Engineers ", SAMS, 1991.
- [6] Ordoyo, Fabiola; Trayter, Joan, " Àudio digital ", Apunts i edicions La Salle, 1998.