

# Implementación y evaluación de cancelación de crosstalk aéreo mediante regularización de matrices

Franco Jofré\*, Georgina Lizaso\*\*, Jorge Petrosino\*\*\*

Universidad Nacional de Tres de Febrero, Caseros, Argentina

francoj.jofre@gmail.com / glizaso@unla.edu.ar / jpetrosino@untref.edu.ar

## RESUMEN

En el presente trabajo se mostrará la aplicación de una técnica de cancelación de crosstalk aéreo que es posible implementar con muy pocos recursos tecnológicos. En un sistema de configuración estéreo la señal del canal derecho llegará en forma más directa al oído derecho, pero parte de su emisión también llegará al oído izquierdo luego de rodear la cabeza del oyente. Lo mismo sucederá con la señal del canal izquierdo. Esto significa que las señales que llegan a cada oído contendrán una determinada mezcla de lo que fue emitido por ambos canales. Los sistemas de cancelación de crosstalk aéreo procesan el audio para que cada canal sea percibido por el oído ipsilateral cancelando la contribución en el contralateral, logrando así la localización espacial de eventos sonoros. Para conseguirlo se implementará analíticamente la técnica de regularización de matrices descrita por Choueiri (2018) y se evaluará su implementación en condiciones de escucha no ideales. Las pruebas experimentales fueron realizadas utilizando muestras binaurales (sonido 3D) con un parlante portátil bluetooth en un espacio no anecoico, arrojando resultados satisfactorios tanto a nivel objetivo como perceptivo.

**Palabras clave**— Audio Transaural, Regularización de Matrices, Sistemas de Cancelación de Crosstalk.

**Abstract**— This paper shows the application of crosstalk cancellation techniques for binaural audio through loudspeakers (XTC-BAL) which can be implemented with very few technological resources. In a stereo setup, the right channel signal will reach the right ear more directly, but some of its output will also reach the left ear after surrounding the listener's head. The same will happen with the left channel signal. This means that the signals reaching each ear will contain a certain mix of what was emitted by both channels. XTC-BAL systems process the audio so that each channel is perceived by the ipsilateral ear, canceling the contribution in the contralateral ear, thus achieving spatial localization of sound events. To achieve this, the matrix regularization technique described by Choueiri (2018) will be analytically implemented and its implemen-

tation under non-ideal listening conditions will be evaluated. The experimental tests were carried out using binaural samples (3D sound) with a portable Bluetooth speaker in a non-anechoic space, yielding satisfactory results for measurements and subjective listening.

**Keywords**— Transaural Audio, Crosstalk Cancellation Systems, Matrix Regularization.

## I. INTRODUCCIÓN

Cuando se utilizan altavoces para emitir eventos sonoros en un espacio físico ya sea en condiciones de campo libre o dentro de un recinto, se generarán variaciones de presión de aire que crearán un campo sonoro. El cómo este campo sonoro sea percibido por quienes estén en situación de escucha dependerá de múltiples factores, entre ellos las características acústicas del ambiente y las características físicas del cuerpo humano como la forma de la cabeza, torso y pabellón auricular del oyente. Estas características físicas modifican el espectro de la señal sonora en función del ángulo de incidencia en relación con el eje medio de la cabeza y es lo que se conoce como función de transferencia de la cabeza (HRTF por Head Related Transfer Function).

Si bien en cuanto a la reproducción esto podría resolverse con el uso de auriculares, en donde a cada oído llegue una señal distinta que no esté alterada por el HRTF ni por el ambiente, para ciertas aplicaciones esto no representa una solución. A la hora de reproducir campos sonoros que incluyan espacialización o características tridimensionales las alteraciones producidas por el ambiente y el HRTF se vuelven determinantes. Es por ello que para este caso las señales enviadas a cada oído deberán incluir procesos que imiten las condiciones físicas del HRTF y el ambiente, o bien simulen indicios que promuevan la localización (señales binaurales). Sin embargo, para lograr una correcta binauralidad es necesario compensar la función de transferencia entre la señal emitida por el auricular y la detectada por el tímpano. Para lograrlo resulta necesario aplicar una ecualización precisa y diferente para cada individuo (Vorlander, 2000). En la mayoría de los casos esta compensación no se realiza por la dificultad de obtener la función de transferencia específica para cada oyente y además existen aplicaciones donde el uso de auriculares puede resultar indeseable.

\* F. Jofré estudia en la Universidad Nacional de Tres de Febrero, Valentín Gómez 4772, B1678ABH Caseros, Buenos Aires, Argentina. (email: francoj.jofre@gmail.com).

\*\* G. Lizaso trabaja en la Universidad Nacional de Lanús y en CONICET (email: glizaso@unla.edu.ar)

\*\*\* J. Petrosino trabaja en la Universidad Nacional de Lanús y en la Universidad Nacional de Tres de Febrero (e-mail: jpetrosino@untref.edu.ar).

Las técnicas de audio transaural intentan reproducir señales binaurales por medio de altavoces en donde cada oído recibe una señal distinta. Como se dijo anteriormente, el campo sonoro percibido por los oídos del oyente contendrá las modificaciones por el HRTF por lo que las señales a emitir deben incluir procesos que cancelen la función de transferencia de la cabeza debido a la propagación al momento de la escucha, es decir que la señal binaural deberá incluir una HRTF que podrá ser la del oyente o una genérica.

Los desafíos técnicos principales que implica el audio transaural se presentan ya en el caso más básico que corresponde a la utilización de dos altavoces. En la Fig. 1 la señal del canal derecho llega en forma más directa al oído derecho, pero parte de su emisión también alcanza al oído izquierdo luego de rodear la cabeza del oyente. Lo mismo ocurre con la señal emitida por el canal izquierdo. Esto implica que lo detectado por cada oído será una mezcla de lo emitido por cada altavoz. Los sistemas de audio transaural tienen que procesar la señal de información de ambos canales de modo que al emitirlos la contribución cruzada tienda a cancelarse para que cada oído reciba la información original de cada canal. Por este motivo también se los denomina sistemas de cancelación de crosstalk aéreo (sistemas XTC).

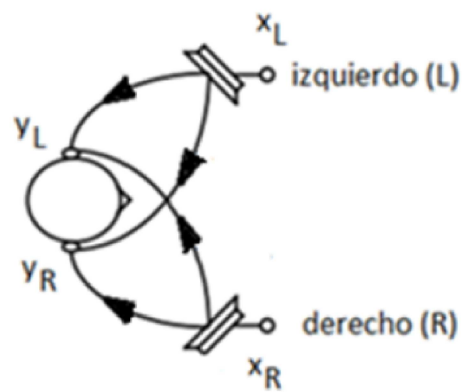


Fig. 1.

Recorrido de las señales emitidas por un par de altavoces.

La primera patente de un sistema XTC fue presentada en 1966 por Atal y Schroeder con el fin de obtener grabaciones binaurales para su auralización por medio de altavoces. Donde el desafío central a resolver fue poder transmitir dos señales independientes desde un conjunto de altavoces a los dos oídos de un oyente mediante la aplicación de filtros analógicos que compensen el crosstalk producido cuando un oído recibe información del altavoz ubicado del lado opuesto. Sin embargo los sistemas XTC a priori funcionan para un único oyente en una posición y orientación predeterminadas (poseen un área de funcionamiento denominado sweet spot muy limitado) en donde un sistema razonablemente bien diseñado tolerará pequeñas variaciones en la posición de la cabeza, pero los grandes cambios hacen que el filtrado no entregue la señal binaural adecuada al oyente (Gardner, 1995). Por otra parte, pueden mencionarse tres problemas asociados a estos sistemas: lograr una separación de canales para poder conseguir el nivel XTC deseado, mantener dentro de niveles aceptables la exigencia de rango dinámico en los altavoces y la coloración tonal en relación al tamaño del sweet spot. Para lograr sistemas XTC perfectos deberían darse las condi-

ciones de que los altavoces y toda la cadena electroacústica se encuentren apareados, un recinto anecoico y que la posición de la cabeza del oyente no varíe. A lo largo de los años se han propuesto variadas técnicas que intentan dar solución a las dificultades mencionadas. En este trabajo se presenta la técnica de regulación de matrices de crosstalk (Choueri, 2018), técnica que no requiere la compensación por HRTF ya que calcula los filtros a utilizar suponiendo dos puntos de escucha que representen la separación entre oídos y sin presencia de obstáculos (en este caso la representación de la cabeza) logrando buenos resultados tanto a nivel analítico como perceptivo. Para evaluar la implementación de esta técnica se realizaron mediciones en condiciones no anecoicas.

## II. CANCELACIÓN DE CROSSTALK AÉREO

### A. Técnica de regularización de matrices

En la Fig. 2 se muestra el diagrama en bloques de un sistema de cancelación de crosstalk, donde se busca que al oído  $Y_L$  llegue la señal  $S_L$  y al  $Y_R$  llegue  $S_R$ . Lo que se emita por el altavoz  $X_L$  va a llegar tanto al oído  $Y_L$  como al  $Y_R$ . Por lo cual el altavoz  $X_R$  deberá emitir una señal que logre cancelar esta contribución. Esta nueva señal también llegará al oído  $Y_L$ , por lo cual, se debe diseñar un conjunto de filtros  $C$  que logre cancelar ambas contribuciones simultáneamente.

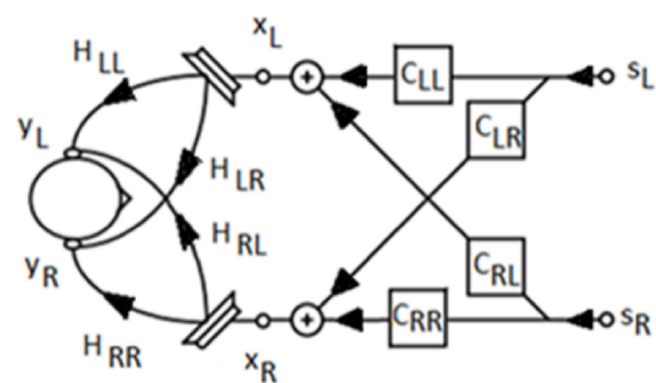


Fig. 2.

Diagrama en bloques de un sistema de cancelación de crosstalk.

Esto implica resolver un sistema de ecuaciones que puede representarse en forma matricial. Este sistema de ecuaciones tiene la desventaja de que frente a pequeñas variaciones en las condiciones iniciales del sistema, se producen grandes variaciones en el resultado. Pequeños movimientos de la cabeza generan una caída del nivel de cancelación y grandes coloraciones. Por otro lado, las condiciones no ideales del ambiente y los altavoces utilizados también afectan negativamente. Para reducir este problema se aplica la técnica conocida como regularización de matrices (1) mediante la cual los resultados teóricos desmejoran pero el conjunto de filtros se vuelve más estable ante las variaciones del sistema.

$$C_{\beta} = [H^{TC} \cdot H + \beta \cdot I]^{-1} \cdot H^{TC} \cdot K \cdot e^{-j\omega\Delta t} \quad (1)$$

En la ecuación (1)  $C$  es la matriz que contiene los cuatro filtros de la Fig. 2,  $H$  representa los caminos de propagación de sonido al aire,  $\beta$  es el parámetro de regularización

(1) [github.com/JofreFranco/AES-LAC-2022.git](https://github.com/JofreFranco/AES-LAC-2022.git)

(una constante mayor que cero),  $I$  es la matriz identidad y el supraíndice TC significa transpuesto conjugado.

### B. Medición de crosstalk

Para evaluar la técnica implementada se deben realizar mediciones del nivel de cancelación de crosstalk conseguido. Las mismas fueron realizadas en la sala del estudio de grabación Enrique Santos Discépolo de la Universidad Nacional de Lanús. El filtro inverso fue calculado para ser utilizado con un parlante bluetooth (JBL Flip 4) a una distancia aproximada de un brazo semiextendido del oyente. Esto se realizó en base a un trabajo de Choueiri (2022), disponible en la página web de la Princeton University que contiene audios binaurales procesados como ejemplo para su descarga. Estos fueron calculados para ser utilizados con un dispositivo portátil similar al mencionado. Para la medición se utilizó un micrófono Earthworks M-30 y una placa de audio externa Komplete Audio 6.

El micrófono se colocó en la posición donde estaría el oído derecho de un potencial oyente con una distancia entre el plano del parlante y el del micrófono de 35 cm. Se consideró que el oído se encuentra a una distancia de 6 cm del plano de simetría de los parlantes. No se utilizaron obstáculos físicos que representen la cabeza del oyente (propagación libre), siguiendo los lineamientos descritos por Choueiri (2018). Posteriormente, se repitió la medición colocando el micrófono en la posición del oído izquierdo. En la Fig. 3 se puede observar la disposición de los instrumentos en la medición.



**Fig. 3.**

*Distribución de los instrumentos de medición.*

Como señal para la medición se utilizó una secuencia de dos barridos exponenciales (sine sweep). El primero fue localizado totalmente a la derecha y el segundo a la izquierda. Al procesar los audios utilizando el filtro diseñado se obtiene una señal que pretende reproducir el barrido en primer lugar en el oído izquierdo, cancelándolo en el derecho y posteriormente en el oído derecho, cancelando la señal en el oído izquierdo. En la Fig. 4 se puede ver un esquema de las señales enviadas al altoparlante y el registro en cada oído.

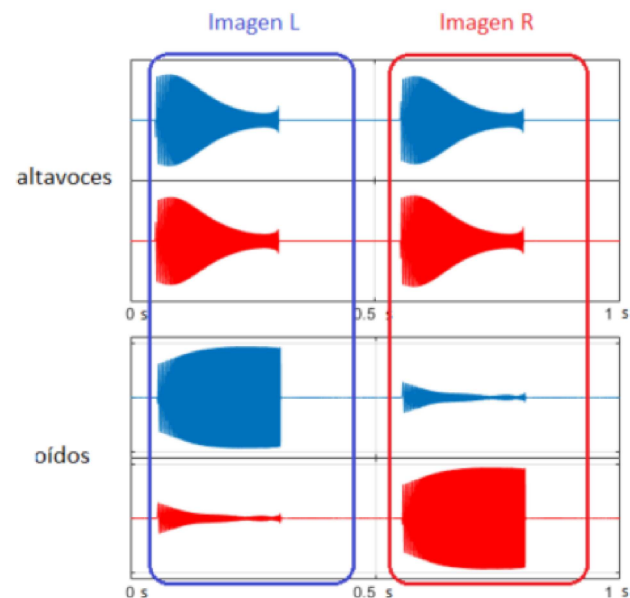


**Fig. 4.**

*Esquema de señales. a) Señales procesadas para los altavoces; b) Señales registradas en los oídos.*

Con esta metodología puede obtenerse el nivel de cancelación de crosstalk siguiendo dos métodos. La primera posibilidad es comparar los registros entre los oídos, es decir, comparar el nivel del oído derecho con el del izquierdo (las señales registradas en el oído como imagen L o R en la Fig. 4). La segunda posibilidad es analizar los canales por separado, donde uno de los barridos debe estar presente y el otro tiende a cancelarse.

Adicionalmente, se realizaron mediciones del ruido de fondo del estudio para asegurar que sea lo suficientemente bajo para que no afecte las mediciones.

Se pone a disposición el link a un repositorio de GitHub con audios binaurales que han sido grabados con dummy head y procesados por los autores con esta técnica con el fin de que puedan ser evaluados por los interesados.

### III. RESULTADOS

En primer lugar se muestran los resultados de la medición de ruido de fondo en la Fig. 5. Dado que el parlante utilizado no es capaz de reproducir correctamente frecuencias bajas, se toma como frecuencia de corte inferior 200 Hz. Como se puede ver, el ruido de fondo es aceptable para este rango.

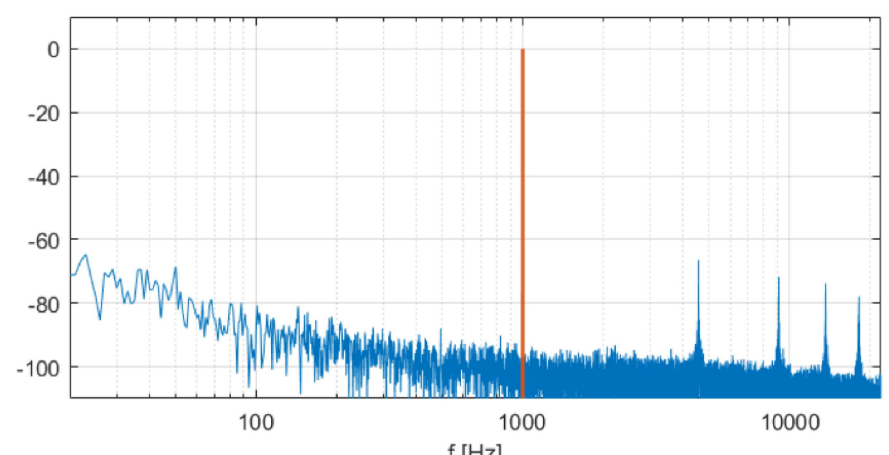


**Fig. 5.**

*Relación señal ruido de los registros (relativo a una senoidal tomada como referencia en rojo).*

Los picos presentes después de los 4000 Hz son provocados por la placa de audio utilizada, los de frecuencia inferior a 10000 Hz fueron eliminados utilizando filtros notch con un  $Q = 20$  centrado en cada una de las frecuencias de interés, el resto fueron eliminados con un filtro pasabajos en 10 kHz y esto representa el límite superior de la medición.

Para determinar el crosstalk se registraron tres secuencias correspondientes a la señal imagen R y tres para la señal imagen L. Se calculó la potencia por tercios de octava de cada registro y se promediaron los resultados. En la Fig. 6 se pueden observar los niveles de señal para el oído ipsilateral y para el contralateral.

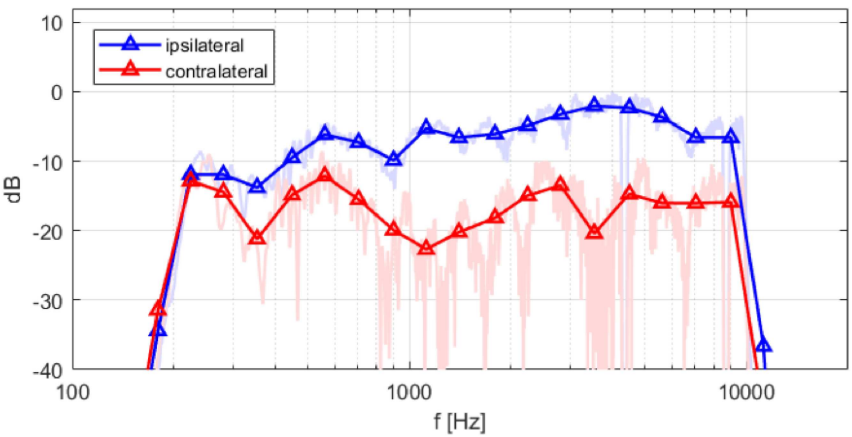


**Fig. 6.**  
Niveles detectados en el oído ipsilateral y contralateral por tercios de octava.

Restando estos valores se obtiene el nivel de cancelación de crosstalk para el sistema en la configuración utilizada para la medición. En la Fig. 7 se muestran los valores de cancelación por frecuencia.

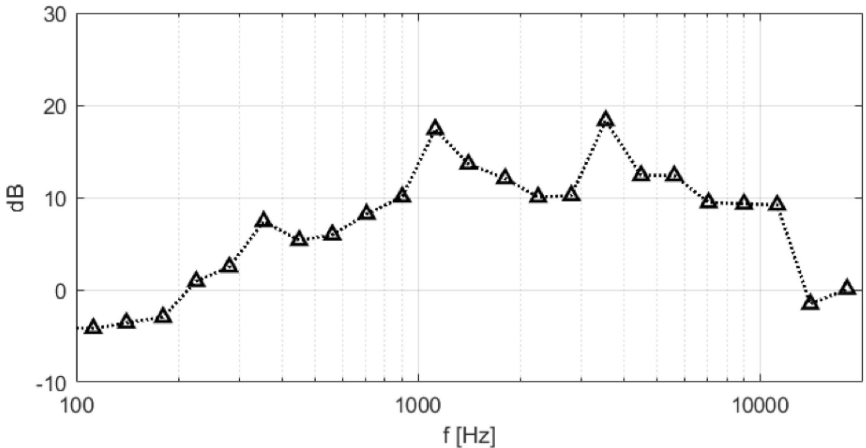


**Fig. 7.**  
Nivel de cancelación de crosstalk por tercios de octava.

#### IV. CONCLUSIONES

Las pruebas realizadas muestran que es posible diseñar filtros inversos que logren cancelación de crosstalk aéreo utilizando solo dos altavoces en configuración estéreo y equipamiento no profesional de bajo costo. Es importante mencionar que el estado del arte sugiere que para lograr sensación de binauralidad se requieren al menos 20 dB de cancelación de crosstalk, sin embargo Choueiri (2008) afirma que en muchas situaciones prácticas este grado de cancelación de 20 dB no es necesario

ni alcanzable. Los resultados obtenidos luego de realizar el presente trabajo son coincidentes con esta afirmación dada la interesante percepción de la distribución espacial de las fuentes sonoras. Incluso para casos donde las fuentes se perciben por detrás de la cabeza del oyente, algo difícil de lograr aún utilizando auriculares. Como trabajo futuro se espera formalizar las pruebas subjetivas para evaluar la técnica y realizar mediciones de HRTF personalizadas para evaluar el nivel de mejoras que pudieran significar.

#### V. REFERENCIAS

- Atal, B. S., & Schroeder, M. R. (1966). U.S. Patent No. 3,236,949. Washington, DC: U.S. Patent and Trademark Office.
- Choueiri, E. (2018) Binaural Audio Through Loudspeakers. En Roginska, A., Geluso, P. (Ed) Immersive Sound. The Art and Science of Binaural and Multi-Channel Audio. (pp. 124-179). Taylor & Francis.
- Choueiri, E. (20 de julio de 2022). Bacch demos for the jambox. The 3D Audio and Applied Acoustics (3D3A) Laboratory at Princeton University. <http://www.princeton.edu/3D3A/Jambox+BACCHDemos.html>
- Gardner, W. G. (1995). Transaural 3-D audio. Perceptual Computing Section, Media Laboratory, Massachusetts Institute of Technology.
- Vorländer, M. (2020). Auralization. Berlin/Heidelberg, Germany: Springer International Publishing.