

Utilización alternativa de algunas técnicas de Codificación Matricial como herramientas para la deconstrucción, edición y recomposición de la imagen estereofónica.

por Daniel Schachter¹

1. Introducción

El presente trabajo forma parte del segundo Informe de Avance del Proyecto de Investigación *Producción del Discurso Sonoro para la distribución en formatos y soportes de sonido envolvente. Evolución tecnológica e influencia estética*, radicada en el CEPESA (Centro de Estudios y Producción Sonora y Audiovisual) del Departamento de Humanidades y Artes de la Universidad Nacional de Lanús.

Al cabo de dos años de trabajo en la investigación relativa a la influencia estética de los nuevos formatos y soportes para Sonido Envolvente, como así también al estudio de la proyección de dichos formatos y soportes en la determinación de criterios a seguir por parte de la industria discográfica, surge con claridad que el formato de audio estéreo mantiene su vigencia en el tiempo, dado que a pesar de que nos encontramos seguramente en un punto de inflexión histórico respecto de los formatos utilizados para la distribución masiva de música grabada, el nuevo escenario sobre el cual nos proyectamos muestra la coexistencia de formatos nuevos y viejos, así como de alta y baja calidad de audio. Por otra parte, el concepto mismo de estereofonía se refiere a las diversas técnicas de grabación y reproducción del sonido llevadas a cabo en dos o mas canales que toman como referencia e intentan reconstruir las condiciones de nuestra percepción auditiva.

¹ Profesor-Investigador en el Departamento de Humanidades y Artes, Licenciatura en Audiovisión, en la UNLa (Universidad Nacional de Lanús, Buenos Aires - Argentina). Investigador y Coordinador Académico en el CEPESA (Centro de Estudios en Producción Sonora y Audiovisual) de la UNLa, Co-Director del Festival Internacional Acusmático y Multimedial y del Ciclo Sonoimágenes. dans.ds@gmail.com

A partir de estas ideas, este artículo intenta explorar acerca de las razones de esta aparente resistencia al cambio, analizar las características propias del formato de audio estereofónico que pueden explicar su persistencia en el tiempo como referente paradigmático, como así también estudiar en que medida algunas de las técnicas de Codificación Matricial desarrolladas para el Sonido Envolvente pueden ser utilizadas total o parcialmente y en forma alternativa como herramientas eficientes para la deconstrucción², edición y recomposición de la imagen estéreo.

2. Vigencia de la estereofonía en los formatos de audio para la distribución masiva

En su trabajo *Revista y extensión del Pseudo-Stereo para Composiciones Electroacústicas Multicanal*³ Philippe Aubert Gauthier plantea que las primeras experiencias con tecnologías de grabación representaron una fuerte discontinuidad respecto de la experiencia auditiva misma, dado que esas grabaciones primitivas fueron monofónicas y aparecieron como un hecho perceptivo nuevo frente a la experiencia auditiva existente que era naturalmente espacial.

“... esta discontinuidad está caracterizada por la pérdida de la cualidad perceptiva relativa a la espacialidad... que era remarcable en la creencia de que el sonido grabado estaba fuertemente relacionado con aquello que era tecnológicamente disponible...”
(Gauthier, 2006).

El sonido estereofónico introducido por Blumlein a comienzos de la década de 1930 se basa en la idea del empleo de diferencias de amplitud entre la emisión

² Entendiendo por tal la descomposición de la imagen sonora estereofónica, no por el camino inverso a su construcción por medio de las diversas técnicas de mezcla, postproducción y masterización, sino mas bien a partir de la búsqueda de elementos implícitos en dicha mezcla pero no evidentes a través del proceso que llevó a su obtención.

³ Citado en Bibliografía de acuerdo a su título original en inglés. (Trad. D. Schachter).

de dos altoparlantes intentando recrear la percepción del oído humano. Afirma Gauthier:

“... las imágenes fantasmas⁴ (phantom images en el original inglés) utilizadas en sistemas de canales múltiples o de sonido envolvente emplean hoy en día la técnica del estéreo por diferencias de amplitud ...”. (Gauthier, 2006)

Las técnicas descritas por este autor para la conversión de una señal monoaural en otra pseudo-estereofónica toman en cuenta la percepción de nuestro aparato auditivo y buscan obtener como resultado una situación de escucha realista, considerando el concepto de sonido estereofónico como una unidad y no como una sumatoria de fuentes puntuales de ubicación identificable e inequívoca. De tal forma, por ejemplo, un sonido situado a nuestra derecha provendrá de un sector determinado dentro de un campo sonoro dado y no exacta y necesariamente del altoparlante allí colocado, como sí ocurre con el audio monoaural. Esta característica es propia e inseparable de nuestro sistema auditivo, la tecnología ha tratado de emularla a partir de la creación de la estereofonía y ha intentado consolidar este camino mediante nuevos desarrollos en múltiples canales. Sin embargo, la cuestión original sigue siendo la misma: crear sistemas de audio que permitan recrear la capacidad propia e inseparable de nuestro sistema auditivo: la orientación dentro de un escenario sonoro.

En efecto, contamos con dos oídos precisamente para podernos orientar en el espacio en función del sonido percibido, sin necesidad del apoyo visual. Tomado esto en cuenta, las técnicas desarrolladas para la captación del sonido intentan emular nuestra percepción mediante el uso en la grabación de diversos micrófonos que registran un cúmulo de pequeñas diferencias de tiempo, reflexiones diversas originadas en diferentes trayectorias que al fin y al cabo nos resultan especialmente útiles para ubicar el sonido en el espacio

⁴ En el sonido estéreo concebido sobre dos canales de audio, se denomina phantom (fantasma) al canal central físicamente inexistente pero auditivamente reconocible como resultante del efecto de la percepción simultánea de un canal izquierdo y un canal derecho.

circundante. Podemos afirmar que a los efectos de lograr el resultado esperado, las técnicas de grabación son cuanto menos tan importantes como el diseño de sistemas de amplificación y reproducción de dos o mas altoparlantes.

3. El carácter binaural de la audición humana

En un trabajo presentado ante el Segundo Congreso Internacional de Acústica realizado en Buenos Aires, Cardozo y Romo plantean la aplicación de técnicas para emular tridimensionalidad en tomas de sonido estereofónicas, con lo que ponen la percepción auditiva en el centro de la cuestión. La percepción de espacialidad, la posibilidad de predecir la ubicación de una fuente sonora y el grado de *credibilidad o sensación de presencia* que varios autores denominan *inmersión* en el escenario sonoro no dependerá tanto de la cantidad de altoparlantes puestos en juego como de las cualidades específicas del audio contenido en los mismos. En dicho trabajo, los autores describen las cualidades del *sonido binaural* propio de nuestro sistema auditivo, relacionado con las llamadas *diferencias interaurales* que corresponden a los valores de tiempo (ITD) y de intensidad (IID) que cada uno de nuestros oídos recibe a partir de la emisión de una fuente sonora dada.

El cerebro utiliza esa información para ubicar el sonido en el plano horizontal. Además de esta información, el cerebro interpreta el efecto producido por el pabellón auditivo conocido como *efecto pinna*, que actúa como un filtro *rechaza-banda* o *notch* para deducir la ubicación de las fuentes sonoras en el plano vertical. Los modelos que utilizan esta información para crear la situación perceptiva de tridimensionalidad, creando indicios de localización de fuentes sonoras virtuales son conocidos como *Modelos Binaurales* y utilizan para este fin la emulación de una situación de percepción auditiva real mediante la técnica de respuesta de impulso del pabellón auditivo. (Cardozo, Romo, 2010)

En su artículo incluido en el Primer Informe de Avance del presente proyecto, Gabriel Data hace referencia a las primeras experiencias como así también a la evolución posterior del audio estereofónico⁵.

El sonido estereofónico convencional en dos vías depende del correcto emplazamiento de los parlantes para una escucha aceptable. Conocemos como imagen estéreo al escenario sonoro que se produce delante del oyente distribuido aproximadamente en un arco imaginario que une ambas vías de sonido. El control de la percepción de la distancia y el direccionamiento del sonido son elementos fundamentales para que nuestro sistema auditivo pueda deducir la localización del sonido. Al respecto escribe Lund:

“...Mientras el estéreo de dos canales puede producir una imagen estereofónica que pierde consistencia mas allá del emplazamiento de altoparlantes en un ángulo de 30°, la ubicación lateral no es siquiera predecible aún considerando que el oyente se encuentra colocado precisamente en el centro de la sala...”. (Lund, 2000)

En este punto resulta importante aclarar la diferencia entre binaural y estereofónico. En su artículo *Principios básicos del Sonido Estéreo*, dice Snow:

“...un sistema binaural utiliza dos micrófonos para grabar, preferentemente colocados en una cabeza artificial, y un par de auriculares independientes para cada oído. El estereofónico es por su parte un sistema de dos o mas micrófonos colocados en una zona de captura, conectados a dos o mas altoparlantes distribuidos en una zona de escucha, lo que crea la ilusión de que los sonidos poseen direccionalidad y profundidad en la zona intermedia a los altoparlantes. Un sistema binaural efectivamente duplica en los oídos del oyente los sonidos que el mismo escucharía en los puntos de captura otorgando una sensación normal de direccionamiento en todas las direcciones. Un sistema

⁵ Data G. : Sonido Envolvente en formatos y soportes comerciales. Relevamiento de los formatos mas difundidos. Reseña histórica de la evolución del sonido multicanal. (pp. 88 – 92)

estereofónico produce un patrón de audición anormal en los oídos del oyente cuyo sentido auditivo interpreta como indicativo de direccionamiento dentro de un espacio limitado entre los altoparlantes. Se afirma correctamente que un sistema Binaural transporta al oyente al escenario original, mientras que un sistema Estereofónico transporta las fuentes sonoras a la sala de audición..” (Snow, 1955)

Para aclarar aun mas la diferencia entre estas dos situaciones perceptivas basta citar a Rumsey, quien se refiere al rol del ingeniero de mezcla de una producción estereofónica señalando que aún cuando éste pudiera decidir crear señales espaciales que son consecuentes con el sonido en ambientes naturales, en la mayoría de las ediciones comerciales no hay ningún ambiente natural para referir, y se trabaja sobre una creación artificial que no tiene ninguna referencia “natural” o ancla perceptual, donde el ambiente acústico diseñado por el ingeniero de grabación y productor es una forma de ficción o arte acústico. (Rumsey, 2002)

La escucha *Binaural* fue descubierta en 1839 por el físico prusiano Heinrich Wilhelm Dove quien fue el primero en plantear la diferencia perceptiva registrada por cada oído. Esta idea fue en realidad mas un antecedente histórico del sonido estereofónico que un punto de partida para futuras investigaciones relativas a la binauralidad. A mediados de la década de 1950 el ingeniero estadounidense Robert Monroe partiendo de esas ideas desarrolla la técnica *Hemi Sync* orientada a aplicaciones terapéuticas⁶.

Así, mientras el desarrollo de la estereofonía se relacionó fuertemente con la industria del disco y la producción de equipos para la reproducción de música grabada, permitiendo una situación satisfactoria para crear la sensación de espacialidad sin requerir técnicas especiales para la grabación ni el uso obligado de auriculares para la escucha, el audio *binaural* siguió en principio un camino principalmente experimental.

⁶ Hemi Sync: sincronización de los hemisferios cerebrales, aplicada a la relajación y en la terapia contra el insomnio.

El audio *binaural* requiere el empleo de dos micrófonos que deben ser colocados a ambos lados de una cabeza de un maniquí diseñada para emular no solo la ubicación de los oídos sino para imitar la función de la cabeza que actúa a modo de filtro, evitando que se produzcan interferencias entre el registro de uno y otro micrófono. Esa interferencia recibe el nombre de efecto *Diafonía* o *Crosstalk* y se produce habitualmente en el audio estereofónico emitido a través de un par de altoparlantes, cuando una parte del sonido proveniente de cada altoparlante es percibido por ambos oídos, y se refiere a las trayectorias acústicas que se producen entre canales que aparecen como interferencia de la información del canal izquierdo sobre el derecho y viceversa.

Cuando analizamos la escucha con auriculares, la situación es diferente porque no se produce interferencia acústica entre los oídos del oyente. El *Crosstalk* es sin dudas un aspecto normal y esperable del sonido estereofónico pero su presencia tiende a limitar la percepción de imágenes sonoras localizadas más allá del espacio mismo de los altoparlantes. Por esta razón, desde sus comienzos, el audio *Binaural* requiere la escucha a través de audífonos o auriculares para asegurar una correcta percepción sin interferencias por *Crosstalk*.

En años recientes se han desarrollado nuevas técnicas especialmente diseñadas para mejorar la percepción auditiva del audio *estereofónico* proveniente de un par de altavoces, que buscan eliminar la interferencia por efecto *Crosstalk*. Para lograr este objetivo, se han diseñado diversos dispositivos para ser conectados a la entrada de los altoparlantes que permitan la cancelación por contrafase de la señal de *Crosstalk*. Esta cancelación puede mejorar notablemente la percepción y convertir la experiencia en un equivalente a la escucha binaural, pero en la práctica este procedimiento dependerá en gran medida de la relación geométrica existente entre los altoparlantes y los oídos del espectador, sin descuidar el hecho de que las reflexiones del recinto también conspiran contra la correcta cancelación buscada. (Maher, Lindemann, Barish, 1996)

El éxito de estas investigaciones, en la medida que logren efectivamente anular la percepción de tales interferencias, significará sin dudas un aporte decisivo para lograr la expansión de la percepción espacial del audio proveniente de sistemas estereofónicos de dos canales, cuando el emplazamiento de los parlantes deje de ser un punto de referencia dado que el discurso sonoro liberado por completo de tales interferencias revelará a nuestra percepción diversas procedencias reconocibles. El desarrollo completo de las técnicas relativas al audio *binaural* excede el marco de este trabajo de investigación.

4. Enfatización de la imagen estereofónica.

Entendemos por percepción estereofónica a la posibilidad de ubicar en el espacio la procedencia de los sonidos reproducidos por un sistema de audio dado, compuesto por dos canales que responden a la Ley de Panorámico Seno-Coseno (Anderson, 2009). Toda estación de trabajo de audio digital incorpora este concepto para el posicionamiento de las señales monofónicas en un escenario estereofónico y la definición de la ubicación espacial del sonido es una cuestión central tanto para los compositores como para los técnicos especializados. En este sentido dice Anderson:

“...Cuando se trata de manejar las grabaciones resultantes, tanto el acusmático como el purista tienen un problema similar. El compositor necesita un conjunto de herramientas para manipular los atributos espaciales de los sonidos recolectados, si este es uno de los parámetros a operar para crear sentido –y para que como es de esperarse, sea reconocida la habilidad del artista. Para el ingeniero de grabación purista el asunto puede ser algo más correctivo. Particularmente para grabaciones hechas en conciertos, en las que cuestiones de puesta en escena y línea de visión pueden no siempre dar como resultado grabaciones bien balanceadas, centradas o de alguna manera bien representadas, la acción reparadora puede ser necesaria. El ingeniero de grabación

purista puede necesitar re-apuntar, re-balancear o re-guiar el resultado...” (Anderson, 2009)

Maher, Lindemann y Barish plantean que el sonido estereofónico supone la percepción de un ambiente natural donde sucede el evento sonoro en cuestión. Las diversas técnicas desarrolladas para la enfatización de la imagen estéreo tienen como objetivo reducir la dependencia del sonido reproducido respecto de las particularidades del recinto donde se escucha, tendiendo a independizar el escenario sonoro de los altoparlantes del sistema, yendo mas allá de los mismos, a través de la extensión virtual de la escena sonora, tanto en cuanto a la anchura como a la profundidad del campo perceptivo. Existen distintas alternativas para llevar a cabo la enfatización de la imagen estereofónica. La gama de soluciones disponibles abarca desde circuitos integrados de tipo analógico específicamente diseñados para ese fin, pasando por microprocesadores de uso habitual en aplicaciones de audio, llegando hasta micro chips programables para el procesamiento digital de la señal. Estos autores mencionan los métodos de codificación como alternativa para la enfatización de la imagen estereofónica al plantear dos posibles metodologías para llevarla a cabo:

- Enfatización de la imagen en un solo sentido⁷

Considerando en este caso dos etapas: una de Pre-procesamiento constituido por las tareas habituales de edición del sonido donde el realce de la imagen estéreo se encuentra incluida en las técnicas de producción de estudio convencionales y por lo tanto su efecto aparece incluido en la mezcla de audio destinada a la distribución; o bien como una etapa de Post-procesamiento realizado a través de equipamiento especialmente diseñado para tal fin, puesto en juego en oportunidad de la reproducción de la grabación. En este caso, la calidad del resultado dependerá del diseño del dispositivo en cuestión.

⁷ *Single-ended image enhancement* de acuerdo al original en idioma inglés.

- *Enfatización Complementaria de la imagen*⁸

Los autores plantean aquí la posibilidad de aplicar un proceso de dos sentidos, entendiendo por tal la codificación y decodificación en el cual tres o más canales de audio son combinados en sólo dos canales de audio, con el objetivo de la distribución discográfica, seguida de una operación en sentido contrario que permita descifrar y reconstruir los canales de audio originales. Al respecto indican los autores que mediante éste procedimiento: "...no es posible en general realizar una reconstrucción perfecta de los canales de audio originales en estas condiciones, pero en base a esto pueden emplearse diferentes métodos para realizar la separación aparente de los canales de salida..." (Maher, Lindemann, Barish, 1996).

Esta referencia a la consideración del empleo de la codificación para la enfatización de la imagen estereofónica es una referencia principal para este artículo, por lo que las diferentes alternativas planteadas en el punto 7 pueden considerarse como correspondientes a esta categoría.

5. El carácter inmersivo del sonido y la cantidad de altoparlantes en juego

El término *estereofonía* indica en realidad mucho más que una organización de equipamiento tecnológico de audio dotada de dos canales físicos. En realidad al hablar de *estereofonía* nos referimos a una situación particular de escucha en la cual la percepción del oyente tiende a desvincularse por completo de la cantidad de altoparlantes disponibles para percibir el escenario sonoro. Esa cualidad inmersiva propia de la escucha binaural es el elemento principal buscado por los desarrolladores de sistemas de audio estereofónicos. Tomando esto en cuenta, podemos afirmar que toda técnica que investigue a cerca de cómo aproximarse a la percepción de la audición humana dará un paso hacia la real percepción del panorama estéreo y cuanto más se acerque a

⁸ Complementary Image Enhancement de acuerdo al original en idioma inglés.

este resultado, mas se relacionará con la audición y menos con la cantidad de altoparlantes disponibles.

A partir de las primeras experiencias con estereofonía el diseño de los sistemas de audio buscó satisfacer la necesidad de percibir un escenario sonoro verosímil dentro de un panorama virtual ubicado en forma de arco detrás de la línea imaginaria que une ambos altoparlantes. En el primer informe de avance del presente proyecto hemos hecho referencia a la aparición y evolución del formato estereofónico a partir de la década de 1930 y su evolución posterior en búsqueda de resultados de un realismo cada vez mas acentuado. De la misma forma que un par de auriculares confortables pueden pasar desapercibidos al escuchar audio binaural, la cantidad de altoparlantes puede transformarse en un dato poco relevante para el oyente si está bien logrado el carácter inmersivo del audio que está escuchando, mas allá de la cantidad de altoparlantes involucrados. Este ha sido un objetivo principal de los desarrolladores de tecnologías para la eliminación de la interferencia por *crosstalk* en los sistemas estereofónicos de dos altoparlantes y precisamente resulta evidente que cuanto menos lograda sea la percepción de inmersión, se producirá en el oyente una natural tendencia a identificar a los altoparlantes con las fuentes sonoras, tal como ocurría con el audio monoaural. Así, la búsqueda de la inmersión dentro de un escenario sonoro tan inmersivo y expandido como sea posible es una característica propia de los sistemas multicanal, y constituye uno de los elementos cruciales para la percepción del discurso que por lo tanto dependerá tanto de las técnicas de grabación empleadas como de la calidad del equipamiento de audio, las características propias de la sala y el emplazamiento de los altoparlantes.

Si pensáramos un espacio octofónico como un doble cuarteto de imágenes estéreo cruzadas en una serie de puntos ubicados en el espacio interno delimitado por el emplazamiento de los altoparlantes, y aceptáramos que la búsqueda de la cualidad inmersiva de los formatos de sonido multicanal tiene su origen en la percepción del espacio sonoro proveniente del sonido estereofónico, entonces podríamos inferir que toda organización de canales que incluyese un número impar de altoparlantes, introduciría un elemento

distractor para nuestra percepción auditiva. Esto provocaría un desequilibrio perceptivo y el sonido proveniente de dicho altoparlante tendería a ser percibido como proveniente desde el mismo altoparlante por efecto de lo que podríamos denominar *carencia de juego inmersivo*.

La distribución mas habitual de sonido envolvente y comercialmente distribuida es la de 5.1 canales, con un par de canales frontales, otro par de canales traseros y un canal delantero central de gran utilidad para la reproducción de los diálogos en el sonido cinematográfico pero cuya presencia y relevancia en el discurso musical es cuanto menos discutible tomando en cuenta la idea de sonido inmersivo expuesta mas arriba, por lo que su utilización para la música debería ser evaluada con especial atención. Esta aparente debilidad estructural de las distribuciones de número impar de vías de sonido puede solucionarse omitiendo el uso del canal delantero central toda vez que el discurso sonoro sea musical.

Con respecto a la diferencia entre considerar el escenario estereofónico de dos canales en comparación con el sonido de cine que incluye el canal central frontal sostiene Lund que ese canal central tan importante para el cine será difícilmente aceptado por los músicos que no desean asignar las voces al canal central por las dificultades que se producen al momento de asignar efectos a ese canal. (Lund, 2000)

6. Sistemas de Codificación Matricial de audio

Una toma de audio estereofónica puede provenir tanto de una toma directa realizada mediante dos micrófonos, o bien a través de la mezcla de un número determinado de canales de audio, llevada a cabo con el empleo de una cantidad dada de micrófonos. En el primer caso, podríamos afirmar que los componentes de la toma serían los dos canales que le dieron origen, mientras que en el segundo caso, los componentes serían múltiples, tantos como los canales de audio que fueron mezclados, y en ambos casos el resultado arrojará una determinada imagen estereofónica.

Conocemos como Codificación Matricial a la técnica de conversión de la información contenida en un sistema de audio compuesto por una cantidad dada de canales en un número diferente de canales, consistente en una serie de relaciones matemáticas entre ellos incluyendo la suma, la resta y los corrimientos de fase.

Un sistema de Codificación Matricial se compone de dos unidades: un *Codificador* o *Encoder* y un *Decodificador* o *Decoder* los que determinan el sentido de la codificación. Para que el sistema lleve a cabo la tarea para la cual fue diseñado, es necesario respetar los roles del *Encoder* y el *Decoder*. Usualmente se hace referencia a la codificación mediante una simbología que representa la cantidad de canales de entrada respecto de la cantidad de canales de salida en el sentido del *Decodificador*. Por ejemplo una codificación **2:5** indica que dos canales discretos son decodificados en cinco canales y una codificación **4:6** indica que cuatro canales se decodificarán en seis.

Como se puede apreciar, no es imprescindible que uno de los formatos sea el estereofónico en dos canales, sin embargo algunas de las codificaciones mas utilizadas sirven para la conversión de la estereofonía de dos canales en algunos de los formatos mas difundidos de Sonido Envolverte. Cuando el Codificador toma como punto de partida una cantidad dada de canales del sistema en cuestión y obtiene como salida una señal de audio estereofónica que contiene todo el material originalmente asignado a esa organización múltiple de canales, decimos que el sentido de la codificación es el siguiente:

N canales → Lt Rt

Donde **Lt Rt**⁹ constituyen los dos canales resultantes. Por su parte el Decodificador parte de ese audio estereofónico producido por el Encoder y reconstruye o recupera la organización multicanal original:

Lt Rt → N canales

⁹ Lt Rt significan respectivamente Izquierdo total y Derecho total.

Los formatos de Codificación Matricial fueron descritos detalladamente por Gabriel Data en su artículo que integra el Primer Informe de Avance de este proyecto de investigación¹⁰. No será por lo tanto mi objetivo analizar los mismos en este artículo sino tomar algunos de ellos a modo de ejemplo y describir algunos casos que a mi juicio resultan especialmente relevantes.

7. Deconstrucción edición y recomposición de la imagen estereofónica por medio de la utilización alternativa de la Codificación Matricial

A partir de la idea de la búsqueda de una enfatización de la imagen estereofónica que permita incrementar la percepción de inmersión y profundidad del escenario sonoro, es posible utilizar en todo o en parte algunos de los distintos métodos de Codificación Matricial, no ya con la intención de codificar n canales en dos ni recuperar n canales a partir de dos sino con un sentido totalmente diferente como el de deconstruir o re-interpretar la imagen estéreo. Para eso diseñaremos caminos alternativos, buscando descomponer en partes esa imagen sonora para operar sobre la misma, intervenir sus partes y luego obtener como resultado una imagen estereofónica mas satisfactoria..

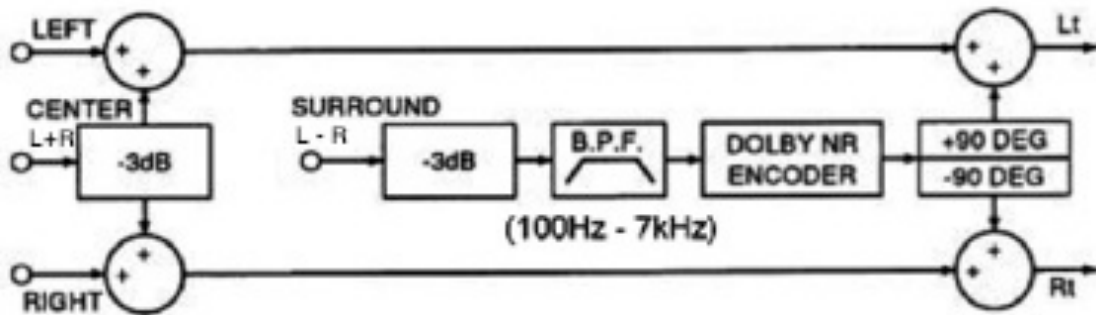
Para lograr este fin podremos por ejemplo prescindir del *Encoder* y aplicar directamente el *Decoder* sobre un archivo de audio estereofónico en dos canales, interpretando que en todos los casos un par **L R** lleva en sí mismo implícito un par **Lt Rt**. Esta especulación nos permitirá utilizar las herramientas de codificación, pero dado que nuestra intención será intervenir el archivo de audio y en ningún caso recuperar aquellos teóricos canales originales que dieron lugar al par **Lt Rt** podremos entonces entre otras cosas alterar los valores de los cálculos matriciales para alcanzar el fin buscado.

¹⁰ Op. cit. en Nota 4

A continuación se describen algunas de las posibles alternativas para llevar a cabo la enfatización de la imagen estéreo a partir de algunos de los sistemas de Codificación Matricial mas habituales.

7.1 Codificación Matricial Dolby Surround

Recordamos por un instante el esquema gráfico de la Codificación Matricial utilizada para Dolby Surround mencionada por Gabriel Data en su artículo ya mencionado mas arriba e incluido en el primer informe de avance del presente proyecto.



Es decir:

$$L_t = L + 0.707 * C + j * 0.707 * S$$

$$R_t = R + 0.707 * C - j * 0.707 * S$$

Aquí, dos canales se convierten en cuatro L C R S según la normativa Dolby Surround que considera un solo canal trasero S, los que luego de acuerdo a la normativa Dolby Pro Logic al aplicar a S el desfase de 90° hacia adelante y hacia atrás pasan a ser cinco canales L C R Ls Rs

Como hemos planteado, consideraremos que toda toma de sonido estereofónica L R lleva implícito un Lt Rt, y por lo tanto aplicando el Decoder podremos obtener aquellos teóricos cinco canales que tendría incluidos, para intervenir esa imagen estéreo y luego reconstruirla. De acuerdo a como fue diseñado, el Decoder toma el par Lt Rt y literalmente lo desarma para obtener los citados cinco canales. Observamos aquí que 0,707 equivale a la atenuación

en 3dB y el operador j indica el corrimiento de 90° en la fase, según sea su signo positivo o negativo. También C corresponde al canal delantero central que es igual a $L + R$ atenuado en 3 dB; y S corresponde al canal Surround que surge de $L - R$ filtrado en banda y luego pasado por el reductor de ruido de Dolby Labs. Inc. El corrimiento de fase en 90° positivo o negativo aplicado a S por aplicación de j o $-j$ dará como resultado los dos canales traseros L_s y R_s . Así, a partir de los dos originales obtendremos cinco canales cuyas denominaciones pueden ser por ejemplo:

$$L_f \quad R_f \quad C \quad L_t \quad R_t^{11}$$

Al llevar a cabo ese proceso decidiremos en que medida seguir total o parcialmente las premisas planteadas por la codificación, de acuerdo a las características propias de cada toma de sonido y a cual sea en cada caso el resultado esperado de nuestra intervención sobre la imagen estéreo. Por ejemplo, podremos entre otras cosas respetar en todo o en parte la atenuación, aplicar o no el filtro pasa banda establecido por la codificación o en su lugar definir valores de ecualización y/o compresión dinámica diferente para cada uno de las cinco componentes que obtuvimos a partir del archivo estéreo original, modificar el grado de desplazamiento de fase, etc. Con respecto al filtro reductor de ruido de Dolby, su utilización corresponde a la aplicación estricta de la codificación pero no se relaciona con la utilización alternativa de la misma para enfatizar la imagen estéreo. Luego, podremos mezclar los cinco canales y obtener como resultado una nueva imagen estereofónica.

7.2 Codificación Matricial para formatos Cuadrafónicos

Con respecto a los formatos de audio cuadrafónicos, existen diversas matrices que se diferencian entre sí por los coeficientes de atenuación aplicados a los canales frontales y traseros y por el corrimiento de fases. Cuando hablamos de

¹¹ Donde f significa Frontal y t significa trasero. Las denominaciones son similares pero no iguales a las que surgen de la Codificación Matricial, dado que el procedimiento seguido no sigue estrictamente dichas especificaciones.

cuadrafonía, las siglas utilizadas por las normativas para identificar los cuatro canales son las siguientes:

$$\text{LF} \quad \text{RF} \quad \text{LB} \quad \text{RB}^{12}$$

A modo de ejemplo, citaremos a continuación las matrices correspondientes a dos de los formatos cuadrafónicos mas habituales (Sessions, 1974).

a) La matriz QS conocida como Quadraphonic Sound desarrollado por la empresa Sansui indica lo siguiente:

$$\begin{aligned} L_t &= 0,92*LF + 0,38*RF + j*0.92*LB + j*0.38*RB \\ R_t &= 0,38*LF + 0,92*RF - j*0.38*LB - j*0.92*RB \end{aligned}$$

b) La matriz SQ básica desarrollada por la empresa CBS Inc. indica¹³:

$$\begin{aligned} L_t &= LF - j*0.707*LB + 0.707*RB \\ R_t &= RF - 0.707*LB + j*0.707*RB \end{aligned}$$

c) Existe una variante de la misma matriz SQ de CBS Inc. que indica¹⁴:

$$\begin{aligned} L_t &= LF + 0.707*LB - j*0.707*RB \\ R_t &= RF - j*0.707*LB + 0.707*RB \end{aligned}$$

Como se describió anteriormente, la idea será partir de un archivo de audio estéreo para obtener cuatro canales a partir de la modificación de algunas de las premisas de la codificación, cuyos valores pueden modificarse y aplicar además diversos procesos a cada uno de ellos para luego obtener una nueva mezcla estereofónica.

Para el caso de la tomar como punto de partida la matriz QS o Quadraphonic Sound podríamos obtener los siguientes cuatro canales:

$$+-j*0,92*L \quad +-j*0,92*R \quad +-j*0,38*L \quad +-j*0,38*R$$

¹² De acuerdo a lo normado, F significa Frontal y B significa Trasero (en inglés Back)

¹³ La matriz básica provee máxima expansión estéreo pero no compatibilidad con monoaural.

¹⁴ Versión alternativa para audio en vivo por su alta compatibilidad con el audio monoaural.

Para el caso de la tomar como punto de partida la matriz SQ podríamos obtener los siguientes cuatro canales:

$$+-j*L \quad +-j*R \quad +-j*0.707*L \quad +-j*0.707*R$$

En ambos casos, para cada uno de los cuatro canales obtenidos con este procedimiento podríamos decidir o no la aplicación de la inversión de fase j tanto en mas como en menos (por eso se la indica como $+-j$). Adicionalmente pueden aplicarse valores diferentes de ecualización, compresión dinámica, etc. a cada uno de los canales.

7.3 Codificación Matricial MS

La técnica **MS**, conocida como Mid/Side ha sido desarrollada para la grabación con el objeto de otorgar una mayor sensación de presencia. Se trata de separar la información ubicada al centro de la imagen de la información restante. El primero en formularla fue Alan Blumlein en la década del los años 30 y fue desarrollada luego por Michael Gerzon quien la denominó *stereo shuffling*. **MS** asume que todo par $L R$ surge como resultado de una codificación **MS** en un sentido similar al $L_t R_t$ expuesto para los casos anteriores. Así, **S** representa los elementos que L y R no comparten, o sea lo que entre ambos es diferente, mientras que **M** indica lo que ambos canales tienen en común. De acuerdo a esta técnica los valores de L y R se obtienen de la siguiente forma:

$$L = M + S \quad \text{y} \quad R = M - S$$

Sumando $L+R$ tendremos:

$$L + R = M + S + M - S = 2M$$

y por lo tanto

$$M = (L+R)/2 \quad \text{o bien} \quad M = 0,50 (L+R)$$

Restando L-R tendremos

$$L - R = M + S - (M - S) = M + S - M + S = 2S$$

y por lo tanto

$$S = (L-R)/2 \quad \text{o bien} \quad S = 0,50 (L+R)$$

Donde el signo negativo representa la inversión de fase en 180° y 0,50 implica una reducción de amplitud de 6 dB (Gerzon, 1986).

Así, a partir de un par estéreo podemos obtener cuatro canales:

L R M S

A partir de aquí podemos trabajar con estos elementos siguiendo las mismas pautas expuestas para los casos anteriores y manipularlos con el fin de obtener valores independientes de amplitud, ecualización, compresión dinámica y fase a cada uno de ellos de forma tal que nos permita obtener las componentes de una nueva mezcla estéreo cuyos protagonistas serán precisamente las resultantes de estos cálculos entre canales. Siguiendo estas pautas, cualquier sistema de codificación podría ser reinterpretado para la deconstrucción y edición de la imagen estereofónica, incluyendo matrices de codificación mucho mas complejas como por ejemplo la del sistema Ambisonics, que por su complejidad excede el presente trabajo y podría ser objeto de una investigación específica.

8. Conclusiones

En base a algunos indicios ya expuestos en el Primer Informe de Avance, hemos avanzado en esa línea con el objetivo de indagar a cerca de las razones por las que se aprecia una importante resistencia la cambio de formato, manteniéndose la vigencia de la producción de música en formato estéreo.

Tal como explica Philippe Gauthier, el paso de la experiencia perceptiva puramente acústica a las primeras experiencias del sonido grabado monofónico plantea una discontinuidad respecto de nuestra natural percepción

auditiva, por efecto del uso de la tecnología disponible en el momento (Gauthier, 2006).

El diseño de los sistemas de audio buscó desde sus comienzos satisfacer la necesidad de percibir una imagen estereofónica, entendiéndose por tal a un escenario sonoro verosímil dentro de un panorama virtual ubicado en forma de arco detrás de la línea imaginaria que une ambos altoparlantes. Esa concepción de la idea de estereofonía fuertemente relacionada con la percepción del espacio, dio lugar a la idea de inmersión

Tomando en cuenta el carácter esencialmente binaural de la audición humana podemos interpretar toda organización de altoparlantes como una extensión de la estereofonía de dos canales. Así por ejemplo un espacio octofónico funcionará como un doble cuarteto de imágenes estéreo cruzadas en diversos puntos del espacio interno delimitado por el emplazamiento de los altoparlantes. Si consideráramos correcta la idea de que la búsqueda de inmersividad que plantean los formatos de sonido multicanal tiene su origen en la percepción de la espacialidad proveniente del sonido estereofónico, entonces podríamos inferir que toda organización de canales dotada de un número impar de altoparlantes, introduciría un elemento distractor para nuestra percepción auditiva. Esto provocaría un desequilibrio perceptivo y el sonido proveniente de dicho altoparlante podría ser percibido como proveniente del interior del mismo altoparlante por efecto de su carencia de juego. Esta tal vez sea la razón principal de la resistencia de los músicos a adoptar los formatos 5.0 o 5.1 para la producción musical. Esta resistencia no es tan marcada en los compositores electroacústicos, sin embargo debería atenderse esta cuestión ya que de la pérdida de juego puesta de manifiesto por el canal central frontal de las disposiciones ideadas para el sonido envolvente presentaría las mismas carencias en todos los casos, mas allá de cual sea el estilo musical registrado. En el mismo sentido, Lund explica que una de los motivos de esa resistencia es el de la asignación de los efectos y no solo a la ubicación de las voces en el canal central. (Lund, 2000).

Las diferentes codificaciones matriciales de audio, pensadas para permitir la

compatibilidad del audio multicanal grabado en sistemas de sonido envolvente en todos los casos toman como punto de partida una cantidad dada de canales y llevan a cabo una serie de operaciones matemáticas a partir de sus valores, que permite obtener como resultado una salida de audio estereofónica que contiene todo el material originalmente asignado a dicha organización múltiple de vías de sonido.

Bibliografía:

CARDOZO P.M., ROMO M.W. (2010) *Empleo de técnicas de grabación para emulación de tridimensionalidad en una aplicación estereofónica*. Actas del 2º Congreso Internacional de Acústica UNTREF, Buenos Aires, Argentina

ANDERSON, J. (2009) *Transformadas clásicas de la imagen estéreo. Un análisis*, en *Música y espacio: ciencia, tecnología y estética* (pp 127-160) Compiladores: Basso, G., Di Liscia O., Pampin J. Bernal: Ed. Universidad Nacional de Quilmes.

CHRISTENSEN, K.; LUND, T. (1999) *Room Simulation for Multichannel Film and Music*. Pre-impresión Nr. 4993 - AES 107o Congreso - Nueva York, EE.UU.

GAUTHIER, P. (2006) *A Review and an Extension of Pseudo-Stereo for Multichannel electroacoustic Compositions: Simple DIY ideas*. En Revista Electrónica econtact 8.3 ed. CEC Comunidad Electroacústica Canadiense.
http://cec.sonus.ca/econtact/8_3/gauthier.html

GERZON, M. (1986) *Stereo Shuffling, New Approach - Old Technique*. En Revista Studio Sound, vol. 28, pp. 122-130. Link House Publications, Poole, Reino Unido.

LUND, T. (2000) *Enhanced Localization in 5.1 Production*. Pre-impresión Nr. 5243 - AES 109a Congreso - Los Angeles, EE.UU.

MAHER, R.; LINDEMANN, E.; BARISH J. (1996) *Old and New Techniques for Artificial Stereophonic Image Enhancement*. AES Audio Engineering. Society, 101^a Convention – Los Angeles, EE.UU.

NIELSEN, S. (2001) *Multichannel Signal Processing Tools - difference to multiple single channel processing*. AES Audio Engineering. Society, 19a Conferencia Internacional - Schloss Elmau, Alemania.

RUMSEY, F. (2002) *Spatial Quality Evaluation for Reproduced Sound: Terminology, Meaning and a Scene-Based Paradigm* en Revista de la AES Audio Engineering. Society, Vol. 50, No. 9, pp. 651-666

RUMSEY, F. (2001) *Spatial Audio*. Music Technology Series. Elsevier, Focal Press, Oxford, Reino Unido

SARROFF, A.; BELLO, J.P. (2008) *Measurements of Spaciousness for Stereophonic Music*. AES Audio Engineering. Society, 125^a Convention – San Francisco, EE.UU.

SARROFF, A.; BELLO, J.P. (2009) *Predicting the Perceived Spaciousness of Stereophonic Music Recordings*. Actas del Congreso SMC Sound And Music Computing Conference - Porto, Portugal

SESSIONS, K. (1974) *4 channel stereo from source to sound*. 2a Ed. Tab Books, ISBN: 0830637567; Blue Ridge Summit, PA. EE.UU

SNOW, W. (1955) , *Basic principles of stereophonic sound*, *IRE Transactions on Audio*, vol.3, no.2, pp. 42- 53, doi: 10.1109/TAU.1955.1165407

TOOLE, F. (2008) *Sound Reproduction - Loudspeakers and Rooms*. Elsevier Focal Press, ISBN 978-0-240-52009-4, Burlington MA EE.UU.

USHER, J. (2006) *A new upmixer for enhancement of reverberance imagery in multichannel loudspeaker audioscenes*. AES Audio Engineering. Society, 121^a International Convention – San Francisco, EE.UU.