

Comparación de algoritmos para compresión de voz aplicados a dispositivos de asistencia auditiva

Alejandro J. Uriz¹, Pablo D. Agüero², Jorge Castiñeira Moreira¹, Juan C. Tulli², Esteban González² y Roberto M. Hidalgo²

¹CONICET – Laboratorio de Comunicaciones, Facultad de Ingeniería, Universidad Nacional de Mar del Plata, Buenos Aires, Argentina

Facultad de Ingeniería, Universidad Nacional de Mar del Plata, Buenos Aires, Argentina

Resumen— Las deficiencias auditivas afectan a un gran porcentaje de la Sociedad. Aunque se han desarrollado un gran número de dispositivos destinados a mejorar la calidad de vida a los individuos que padecen de este tipo de dolencias, estos suelen tener precios de venta muy altos y en algunos casos no cubren por completo las necesidades del usuario. Este trabajo se enfoca en el estudio e implementación de diversos algoritmos para compresión espectral de la voz, los cuales tienen como objetivo trasladar espectralmente las componentes de los sonidos de forma tal que una persona con una deficiencia auditiva específica pueda mejorar su percepción de los mismos. Los algoritmos fueron implementados en un dispositivo de asistencia auditiva de bajo costo desarrollado por el Grupo en base al DSP dsPIC33EP512MU810 de Microchip. Se realizaron experimentos subjetivos con el fin de validar el funcionamiento de los algoritmos bajo estudio. Por último, un conjunto de mediciones experimentales fueron llevadas a cabo para determinar el rendimiento del dispositivo de asistencia desarrollado. Los resultados obtenidos demuestran la ventaja de aplicar el algoritmo de compresión SPINC frente a otros métodos bajo estudio. De la misma forma, se demostró el correcto funcionamiento del dispositivo implementado.

Palabras clave— deficiencias auditivas, transposición de frecuencias, procesamiento digital de señales, dsPIC, dispositivo de asistencia auditiva.

Abstract— Hearing impairments are conditions that affect a high percentage of the Society. Although there are devices that improve the quality of life for people with these problems, they tend to be expensive and not to fulfill the user requirements. This work focuses on a study and implementation of a spectral compression algorithm, whose goal is to spectrally map the sounds to let a person with a specific hearing condition perceive them. The algorithms are implemented on a low-cost hearing aid device based on the dsPIC33EP512MU810 of Microchip. Subjective experiments using a hearing impairment simulator are done in order to validate the study. Finally, experimental measurements are carried out to analyze the performance of the implemented hearing aid device.

Obtained results demonstrate the advantage of the SPINC algorithm over the other methods under study. The frequency transposition algorithm is successfully implemented in the developed hearing aid device.

Keywords— hearing impairments, frequency transposition, digital signal processing, dsPIC, hearing aid

I. INTRODUCCIÓN

Las Deficiencias auditivas son una condición que aqueja a un gran número de personas. En Argentina, de acuerdo al censo de 2010, hay más de 300.000 personas que padecen alguna deficiencia auditiva. Además, este número está en aumento debido al uso inapropiado de reproductores de música y a la exposición de los individuos a entornos ruidosos.

Con el fin de mejorar la calidad de vida de las personas que padecen de este tipo de trastornos, se ha desarrollado una amplia variedad de técnicas y dispositivos. En esta línea, existen compañías que ofrecen distintos tipos de

dispositivos de asistencia auditiva, los cuales van desde simples amplificadores analógicos, hasta complejos dispositivos digitales implementados en base a microcontroladores [1, 2, 3, 4]. Estos últimos, son sistemas muy complejos y disponen de gran cantidad de funcionalidades, pero tienen la desventaja de que su precio de venta es elevado respecto a los del primer grupo.

La característica más importante de un dispositivo de asistencia auditiva es el extensor de audibilidad [1]. Esta funcionalidad permite al usuario percibir sonidos que sin el dispositivo de asistencia auditiva no podría percibir. Esto se realiza desplazando la información desde la región del espectro donde el individuo padece una dolencia, hacia una región donde el individuo tenga un resto auditivo.

Algunos trabajos [3, 5] plantean un método para mejorar la percepción de componentes de alta frecuencia en individuos con deficiencias auditivas severas por medio de

Dirección de contacto:

Alejandro José Uriz, Juan B. Justo 4302 (CP 7600), Mar del Plata, Argentina, +54-0223-4816600, ajuriz@fi.mdp.edu.ar.

la transposición de componentes de alta frecuencia, hacia frecuencias inferiores. Esto lo realizan por medio de un circuito multiplicador de frecuencias, el cual es sintonizado de acuerdo al valor requerido de desplazamiento espectral. Por otro lado, hay métodos de transposición espectral que realizan un desplazamiento siguiendo una ley lineal [6, 7], esto se realiza desplazando las componentes de la representación espectral de la señal por un factor fijo (ya sea uniforme o por tramos). Si bien es una técnica ampliamente utilizada, tiene la desventaja que puede realizar un mapeo inapropiado de los formantes, por lo que el usuario podría percibir los sonidos con una reducción de inteligibilidad, lo cual en señales de voz se debe al desplazamiento en las componentes más importantes de la misma. Existe otro grupo de técnicas que tienen en cuenta la naturaleza de la señal de voz a la hora de realizar el mapeo de la información, una de ellas es la basada en la función SPINC (Terhardt *et al* [8]). Este método mapea las componentes de espectrales inferiores a 700 Hz utilizando un factor casi unitario, mientras que las componentes de mas alta frecuencia las comprime utilizando una ley logarítmica. De esta forma, los formantes más importantes de la voz casi no son desplazados, mientras que los superiores son comprimidos utilizando un factor elevado. De esta forma, la inteligibilidad de la voz es elevada respecto al caso de utilizar técnicas de compresión lineal. Cabe destacar, que en condiciones auditivas más específicas, se debería utilizar un método de transposición de frecuencias específico a las necesidades del individuo.

En el mercado existen pocos dispositivos de asistencia auditiva que permiten al especialista configurarlo de acuerdo a las necesidades del usuario. Además, hay que tener en cuenta que los mismos son costosos, por lo que muchos individuos que requieren de este tipo de soluciones, quedan marginados del acceso a las mismas.

El principal objetivo de este trabajo es realizar un estudio entre diversos algoritmos de transposición de frecuencias e implementarlos en un dispositivo de asistencia auditiva de bajo costo desarrollado por el Grupo. Por ello, en primer lugar se describen los métodos bajo estudio y se los implementa en MATLAB. El rendimiento de los mismos es evaluado utilizando un simulador de deficiencias auditivas [9] desarrollado para este propósito. Luego, los algoritmos son implementados en el dispositivo de asistencia auditiva, el cual está basado en un procesador digital de señales de bajo costo de la firma Microchip [10], y finalmente se realizan mediciones experimentales para analizar el funcionamiento del mismo.

El trabajo está organizado de la siguiente manera. En la Sección 2 un conjunto de métodos para transposición de frecuencias es presentado y estudiado. A lo largo de la Sección 3 se describe el dispositivo de asistencia auditiva desarrollado. La Sección 4 presenta los experimentos y mediciones llevadas a cabo. Mientras que, por último, en la Sección 5 se presentan las conclusiones del trabajo y las líneas de investigación a seguir en el futuro.

II. ALGORITMOS PARA TRANSPOSICIÓN DE FRECUENCIAS

El objetivo de los algoritmos de transposición espectral es desplazar y escalar las componentes espectrales de una señal. De esta forma, para cada índice de una componente espectral se obtiene un nuevo índice de salida. El

desplazamiento a aplicar depende de las necesidades del individuo que utilizará el dispositivo de asistencia en el cual se implementará el algoritmo.

Como se dijo previamente, la principal función de los métodos de transposición de frecuencias en un dispositivo de asistencia es mejorar la percepción de sonidos que sin uso de los mismos no podrían ser percibidos. De esta forma, los métodos de transposición espectral mejoran la interacción social del individuo, ya que permiten mejorar la inteligibilidad de la voz. Por ello, es necesario que los algoritmos utilizados para implementar esta funcionalidad, tengan en cuenta la naturaleza de la voz, y que de ser posible no alteren los formantes mas importantes de la misma, los cuales están ubicados entre 0 Hz y 3200 Hz.

Por lo antes descrito, se debe aplicar un factor bajo de compresión en las frecuencias mas bajas, mientras que para frecuencias mas elevadas se puede aplicar un factor mas elevado. En estos casos debe tenerse en cuenta que la información requerida para reconocer a un determinado usuario, además de estar en bajas frecuencias, está contenida entre 1400 Hz y 3000 Hz [11]. Por ello, los algoritmos basados en esta filosofía permiten percibir correctamente el mensaje pronunciado por una persona, pero no permiten identificar claramente su identidad. Algunos de los algoritmos de transposición espectral mas utilizados se describen en las siguientes subsecciones.

A. Algoritmos implementados en dispositivos analógicos

Los primeros dispositivos de asistencia que realizaban transposición de frecuencias eran analógicos, y realizaban el escalamiento de frecuencias mediante multiplicadores.

En la Fig. 1 se muestra un esquema de un sistema de este tipo, el cual se describe en el trabajo de Johansson *et al* [3]. En este modelo, se divide el espectro en dos bandas: 0 Hz – 3000 Hz y 3000 Hz a 6000 Hz, una vez realizada esta separación, la banda superior es desplazada espectralmente 3000 Hz y la señal resultante se suma con la banda inferior. De esta forma, se obtiene una señal limitada a la banda de 0 Hz a 3000 Hz. En Velmans *et al* [5] se propone un enfoque similar, pero en dicho caso se utilizan las bandas 0Hz a 4000 Hz y 4000Hz a 8000 Hz, la transposición de frecuencias en este caso se realiza mediante un oscilador de $f=4000$ Hz.

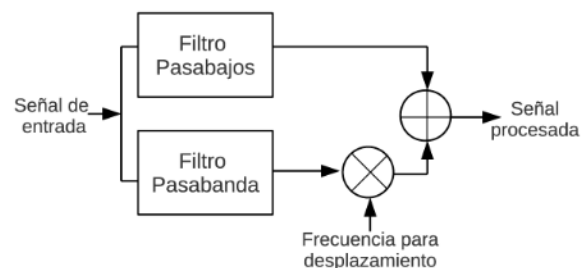


Fig. 1: Sistema transponedor de frecuencias analógico.

Estas aproximaciones permiten mejorar la percepción de sonidos tales como /s/ y /ch/. Si bien estos métodos desplazan las componentes espectrales a un rango limitado de frecuencias, existen casos donde debido al enmascaramiento espectral [15], las componentes de bajas frecuencias pueden enmascarar a las de alta frecuencia, reduciendo la percepción de las mismas.

Otra limitación de este tipo de sistemas es que generalmente sus especificaciones son fijadas en fábrica y no pueden ser reconfiguradas en campo. Por ello, con el surgimiento de procesadores digitales de señales más potentes se desarrollaron dispositivos de asistencia digital, los cuales pueden ser programados de acuerdo a las necesidades de cada usuario. Algunas de las funciones de transposición de frecuencia utilizadas en este tipo de dispositivos se describen en la siguiente subsección.

B. Algoritmos implementados en dispositivos de asistencia digitales

Como se describió previamente para mejorar la percepción de sonidos en personas con deficiencias auditivas es necesario aplicar compresión espectral al menos a las componentes ubicadas entre 0 Hz y 8000 Hz. Además, debe tenerse en cuenta que no se debe alterar significativamente a los formantes más importantes de la voz. Esto se debe a que un cambio en estos formantes de la voz produce un deterioro en la inteligibilidad de los sonidos. De esta forma, se deben implementar algoritmos de compresión espectral más complejos, los cuales están basados en técnicas de procesamiento digital de señales, los cuales son implementados en microcontroladores. Permitiendo, de esta forma, implementar funciones de transposición espectral más complejas. Tal como se estableció previamente, para obtener un método de compresión espectral que se ajuste a las necesidades de los individuos, se debe conservar la posición de los formantes más importantes de la voz. Por ello, la única situación donde el algoritmo de transposición debería afectar a estos formantes, se da en el caso donde la discapacidad auditiva afecta a la región del espectro donde se ubican. En esta situación se deben encarar métodos más complejos, en muchos casos específicos para cada condición.

Existen muchos algoritmos para realizar la transposición espectral. Uno de los más utilizados es el que desplaza las componentes espectrales mediante una función lineal, en este caso el índice de cada componente es afectado por un factor fijo. Este método, presenta la desventaja que comprime espectralmente las componentes de baja frecuencia, motivo por el cual deteriora la inteligibilidad de algunos sonidos. Por otro lado, si se adopta como factor de compresión $k = 1.3$, el rango de los sonidos audibles (0 Hz – 20 KHz), se comprime a un ancho de banda de 15 KHz. En la Ec. 1 se presenta la función utilizada en este caso. Tal como puede verse, la relación entre la frecuencia de salida $\Phi(f)$ y la de entrada (f) es la constante k .

$$\Phi(f) = (f/k) \quad (1)$$

Con el fin de mejorar la inteligibilidad del método anterior, se han desarrollado métodos de compresión lineal por tramos [6, 7]. Muchos dispositivos de asistencia auditiva implementan funciones de este tipo, ya que permiten aplicar un índice de compresión cercano a la unidad a los formantes más importantes, mientras que para las altas frecuencias se utiliza un factor más elevado. Aunque esta función tiene mayor rendimiento que las descritas previamente, el ancho de banda requerido para comprimir el espectro audible sigue siendo muy grande.

Con el fin de mejorar las prestaciones de las funciones previamente descritas, se han desarrollado funciones de compresión más complejas. Una de ellas, está basada en la función arcotangente y ha sido desarrollada teniendo en cuenta la naturaleza de la voz. Esta función es llamada SPINC [8, 12] y en la Ec. 2 se aprecia la expresión que la caracteriza. En este caso, la frecuencia de salida $\Phi(f)$, es calculada a partir de la frecuencia de entrada (f) a partir de aplicar el arcotangente y un factor de escala.

$$\Phi(f) = 1414 \arctg(f/1414) \quad (2)$$

En la Fig. 2, en trazo sólido se presenta la función SPINC y dos funciones de compresión lineal: una con un factor de compresión fijo de $k = 1.4$ (trazo rayado), y otra también lineal, pero con un factor de compresión definido por tramos (trazo punteado). Por último, con el fin de establecer una referencia, se presenta una función en la que no se realiza transposición, la cual se representa mediante una función lineal con $k = 1$, (trazo rayado y punteado).

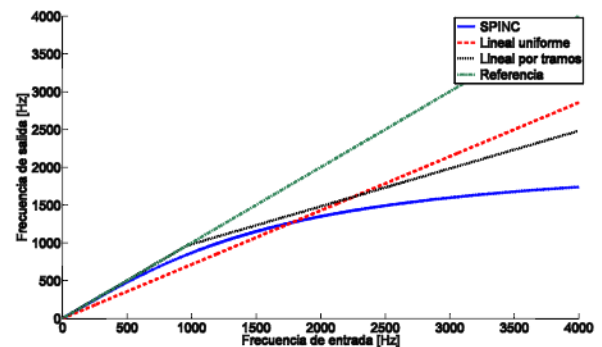


Fig. 2: Funciones para transposición de frecuencias bajo estudio.

La función de compresión SPINC tiene muchas ventajas respecto a la compresión lineal. En primer lugar, este algoritmo no comprime las componentes espectrales entre 0 Hz y 500 Hz, y de esta forma no afecta en forma considerable la inteligibilidad de la voz. Además, el espectro audible es completamente trasladado al rango de frecuencias entre 0 Hz y 2000 Hz, lo que significa una reducción de ancho de banda considerable respecto a otros métodos. Debe tenerse en cuenta que si bien esto deteriorará el reconocimiento de la identidad del locutor, el objetivo de un dispositivo de asistencia auditiva es percibir correctamente el mensaje, mientras que la identidad de quién emitió el mensaje es secundaria. Otra ventaja de este algoritmo, es que permite a personas afectadas por una discapacidad auditiva debida al envejecimiento, percibir sonidos de alarma que de otra forma no podrían oír.

En este trabajo, se aplica el algoritmo SPINC siguiendo lo establecido en Timms et al [12], donde se realiza un análisis de la señal utilizando ventanas de Hanning solapadas un 50%. Luego, a cada una de ellas se les aplica la Transformada Rápida de Fourier (FFT). Una vez que la información se encuentra en el dominio espectral, se le aplica el algoritmo de transposición espectral SPINC, reubicando de esta forma las componentes de la señal. Luego se aplica la Transformada Rápida de Fourier Inversa (IFFT) y se sintetiza la señal utilizando el método de la suma solapada en el dominio del tiempo (TD-OLA).

La carga computacional de estos métodos depende del número de muestras que se utilicen para las ventanas a procesar y del número de bins utilizados para la FFT. Por ello, se puede decir que dicha carga depende de la resolución espectral adoptada.

En la siguiente Sección se describen las características más relevantes del dispositivo de asistencia utilizado para implementar los algoritmos.

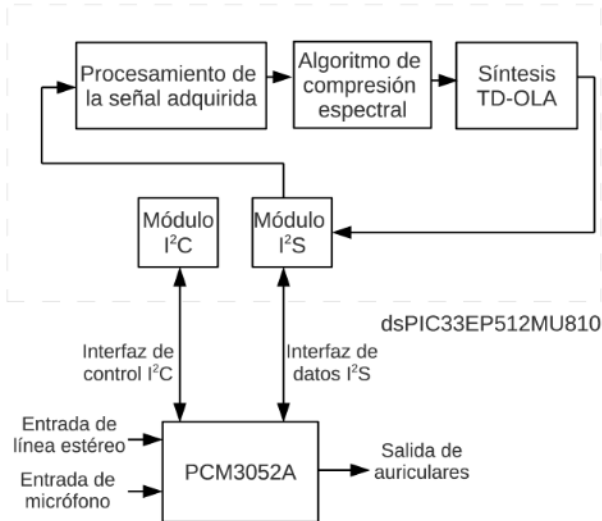


Fig. 3: Diagrama de bloques del sistema construido.

III. DISPOSITIVO IMPLEMENTADO

Tal como se describió en previamente, los dispositivos de asistencia auditiva comerciales tienen un elevado precio de venta, por ello uno de los objetivos del Grupo de investigación es desarrollar un dispositivo de asistencia auditiva comercial de bajo costo. Por ello, a continuación se describen las características relevantes de los componentes utilizados para la construcción del mismo.

Este equipo desarrollado está compuesto por dos componentes principales: un códec integrado y un microcontrolador con un procesador digital de señales integrado. En las siguientes subsecciones se describen las características más importantes de los mismos.

A. Códec de audio

Con el fin de lograr el nivel de integración y repetibilidad del equipo desarrollado, se optó por utilizar un códec de audio para las etapas analógicas del sistema. Este tipo de circuitos integrados permiten integrar en un solo componente etapas tanto de entrada como de salida que de otra forma tendrían que ser implementadas por separado. Se seleccionó el códec PCM3052A [13] de Texas Instruments®. Este es un circuito integrado de bajo costo y reúne un conjunto de características que lo hacen apropiado para esta implementación, algunas de dichas características se listan a continuación.

- Conversor Analógico a Digital integrado (ADC) con filtro antialiasing incluido (SNR=101dB).
- Conversor Digital a Analógico integrado (DAC) con amplificador de salida incluido.
- Entrada de micrófono con preamplificador, entrada de línea estéreo y salida estéreo.
- Múltiples funciones programables a través de una interfaz I²C (*Inter Integrated Chip*).

- Interfaz de datos utilizando un enlace I²S (*Integrated Interchip Sound*).

B. Procesador digital de señales

El procesador digital de señales (DSP) elegido para la implementación es el dsPIC33EP512MU810 [10] de Microchip®. Este se caracteriza por reunir un conjunto de funcionalidades que lo hacen ideal para el desarrollo de un dispositivo de asistencia auditiva reprogramable. Una de las grandes ventajas de este DSP es que puede ser programado utilizando librerías de alto nivel basadas en el lenguaje de programación C. Esto facilita las tareas de desarrollo e implementación de funciones, ya que pueden utilizarse funciones similares a las existentes en otros entornos de programación como MATLAB®. Otras características relevantes del DSP seleccionado se presentan a continuación:

- **52KB de memoria de datos.** Esta cantidad de memoria es superior a la existente en otros DSP de bajo costo.
- **512KB de memoria de programa.** Tal como se mencionó previamente, el dispositivo dispone de memoria suficiente para utilizar compiladores cruzados.
- **70 Millones de Instrucciones por Segundo (MIPS):** Esta característica es fundamental para esta aplicación, donde el algoritmo debe operar en tiempo real.
- **Acceso Directo a Memoria (DMA):** Este dispositivo dispone de hasta 4KB de memoria para ser utilizada por los periféricos para acceso directo a la misma, de esta forma se pueden realizar tareas sin afectar el flujo de funcionamiento del programa principal.
- **Interface I²S (*Interface Interchip Sound*):** Esta interface es ampliamente utilizada para transferir datos entre dispositivos que procesen señales de audio como por ejemplo conversores analógico-digital, digital-analógico y códecs de audio, entre otros.

Una de las mayores ventajas del dsPIC33EP512MU810 es que está capacitado para ejecutar múltiples tareas en simultáneo. En este caso, por un lado se lo ha programado para ejecutar la recepción y transmisión de los datos desde y hacia el códec a través de la interfaz I²S. Por otro lado, en paralelo se ejecuta el procesamiento de los datos almacenados. De esta forma, es posible estar adquiriendo los datos correspondientes al segmento actual, mientras en paralelo se procesan los datos del segmento previo. Esto se realiza debido a que ha configurado a la interfaz I²S para que opere utilizando Acceso Directo a Memoria (DMA) [10], lo que incrementa de forma notable el rendimiento del sistema, ya que las fuentes de interrupción del programa principal se reducen notablemente. En este caso en particular, se ha programado al módulo DMA para que interrumpa al programa cada vez que un buffer de 256 datos esté disponible para ser utilizado.

En el sistema implementado, cuyo diagrama de bloques se presenta en la Fig. 3, los sonidos son adquiridos utilizando el conversor analógico a digital del códec. Entonces, los datos son enviados al DSP utilizando la interfaz I²S. Tal como se mencionó previamente, los datos son agrupados en segmentos de 256 muestras, las cuales son enventanadas utilizando una ventana de Hanning. Luego, a cada segmento se le aplica la transformada rápida de Fourier (FFT), obteniéndose de esta forma una representación en el dominio espectral de la señal adquirida. Entonces, se aplica el algoritmo de transposición espectral

que consiste en desplazar las componentes espectrales de la señal de acuerdo a las necesidades específicas del usuario. Una vez realizada dicha operación, se aplica al segmento transformado la transformada rápida de Fourier inversa (IFFT), y se obtiene una representación temporal de la misma. Luego, dicho segmento es sintetizado utilizando el método de la suma solapada en el dominio del tiempo ó TD-OLA (*Time Domain Overlap and Add*). Por último, el segmento sintetizado se envía nuevamente al códec de audio a través de la interfaz I²S y el mismo se encarga de realizar la conversión digital-analógica y la amplificación necesaria para que la señal excite los auriculares.

En este caso se ha utilizado una frecuencia de muestreo $f_s = 16384 \text{ Hz}$, y los parámetros de configuración del códec fueron enviados a través de la interfaz I²C.

Debido a que el dsPIC fue configurado para operar con DMA para la recepción y transmisión de los datos, se puede decir que mientras el sistema adquiere un segmento de $N=256$ muestras, está procesando el adquirido previamente. Por ello, se puede decir que las tareas de adquisición y procesamiento se realizan en paralelo. Debido a esto, para establecer el tiempo de operación del sistema, se debe considerar el mayor de ambos, siendo un requisito necesario para afirmar que el sistema opera en tiempo real, que el tiempo de procesamiento sea menor al de adquisición de los datos.

Con el fin de realizar un diseño robusto desde el punto de vista electromagnético, se tuvieron en cuenta un conjunto de reglas de diseño [14, 15] las cuales fueron consideradas para el diseño del circuito impreso. Cabe destacar que además, el mismo fue optimizado utilizando una versión de evaluación de la aplicación EMISStream®, la cual tiene como objetivo optimizar un circuito impreso de forma tal de reducir tanto las emisiones como la susceptibilidad del mismo a fuentes externas de ruido electromagnético.

IV. EXPERIMENTOS

Con el fin de analizar el rendimiento de los métodos bajo estudio, un conjunto de experimentos subjetivos y mediciones fueron llevados a cabo.

A. Experimentos subjetivos

En primer lugar, los algoritmos fueron implementados en MATLAB® y un experimento subjetivo se llevó a cabo utilizando un simulador de deficiencias auditivas también desarrollado por el grupo. En este experimento, un conjunto de sonidos extraídos de la base de datos del Dr. Tato [16] fueron procesados utilizando los algoritmos de compresión, los audios obtenidos fueron procesados utilizando el simulador, configurado en cuatro condiciones auditivas específicas. Se eligió para el experimento esta base de datos debido a que es la que es utilizada en la Argentina por los fonoaudiólogos para evaluar este tipo de dolencias. La base de datos contiene 300 palabras fonéticamente balanceadas, las cuales fueron grabadas utilizando una frecuencia de muestreo de 16 KHz y su duración promedio fue de 1 segundo.

Un experimento subjetivo fue realizado para evaluar la **inteligibilidad** y la **calidad** de los sonidos procesados. Para ello, un conjunto de 15 voluntarios (quienes aceptaron expresamente un consentimiento informado) escucharon un conjunto de los sonidos procesados en orden aleatorio. Luego, debieron evaluar la inteligibilidad con puntajes del 1

(la palabra no es inteligible) al 5 (la palabra es totalmente inteligible). Por otro lado, la calidad fue evaluada con puntajes del 1 (pésima) al 5 (excelente). El experimento estuvo compuesto por audios comprimidos utilizando los siguientes métodos:

1. Sin transposición de frecuencias (referencia).
2. Johansson et al (presentado en la Subsección 2.1).
3. Compresión lineal uniforme ($k = 1.3$) (presentado en la Subsección 2.2).
4. Compresión lineal por tramos. Se utiliza la función presentada en la Fig. 2).
5. Compresión SPINC (Terhard et al).

Debido a que, como se describió previamente, el principal objetivo de los algoritmos bajo estudio es mejorar la percepción de sonidos de personas con alguna discapacidad auditiva, el rendimiento de los mismos debe ser analizado en personas con algún tipo de dolencia. Por ello, mediante el uso de un simulador de deficiencias auditivas específicas [8], se generaron condiciones auditivas específicas. El simulador está basado en el estándar MPEG-I Audio Layer 3 [17, 18]. Gracias al uso de este simulador es posible que un voluntario sin ningún tipo de dolencia evalúe los sonidos tal como lo haría una persona con alguna dolencia específica. De esta forma, es posible evaluar el rendimiento del sistema utilizando un mayor número de evaluadores para cada dolencia que se desee analizar. Por otro lado, los evaluadores con dolencias auditivas se utilizarán en un estadio más avanzado de las pruebas. Las deficiencias auditivas simuladas para el experimento se describen a continuación.

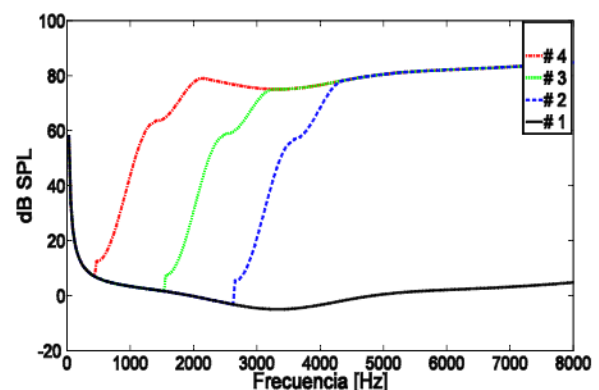


Fig. 4: Umbrales de enmascaramiento de las condiciones simuladas.

En la Fig. 4 se muestra el umbral absoluto de audición ó ATH por sus siglas en inglés (*Absolute Threshold of Hearing*) de cada condición auditiva utilizada para la simulación. El umbral ATH representa la mínima energía en función de la frecuencia que necesita una componente espectral para poder ser percibida en un ambiente sin ruido.

En el experimento, en primer lugar, en trazo continuo, se considera como referencia (**condición #1**), el caso dónde no existe ningún tipo de dolencia auditiva. En otros colores se presentan los umbrales ATH para diversas condiciones auditivas, estas a su vez representan el umbral ATH asociado al deterioro de un oído causado por envejecimiento a lo largo del tiempo. En estos casos, los colores representan dicho deterioro del sistema auditivo, dónde el individuo pierde progresivamente la capacidad de percibir sonidos agudos. De esta forma, para el

experimento se simularon tres estadios de la deficiencia auditiva que actualmente mayor porcentaje de personas afecta: la debida al envejecimiento. Estas curvas se presentan en otros tipos de trazos. En línea de trazo rayado (**condición #2**), el umbral ATH de este caso se puede representar también con un rechazo que aparece alrededor de 2200 Hz. En segundo lugar, en línea de trazo punteado (**condición #3**) se presenta un umbral ATH asociado a un rechazo equivalente a partir de alrededor de 1500 Hz. Finalmente, en línea de trazo rayado y punteado (**condición #4**) se grafica el peor de los casos simulados, donde el umbral ATH posee un rechazo adicional a partir de alrededor 700 Hz. En la figura se puede apreciar claramente el incremento de los valores de la curva ATH debido al deterioro de la salud auditiva del individuo. Un esquema del experimento desarrollado se presenta en la Fig. 5.

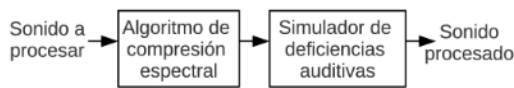


Fig. 5: Representación del experimento realizado.

TABLA I

RESULTADOS DEL EXPERIMENTO SUBJETIVO DE INTELIGIBILIDAD.

Método/ Condición auditiva	1	2	3	4
1 - Sin compresión	4.7	3.8	3	2.7
2- Johansson et al	2.7	2.7	1.9	2.2
3- Lineal uniforme	3.1	3.4	2.6	3.1
4- Lineal por tramos	3.4	3.3	3.5	3
5 - SPINC	4.2	4	3.6	3.9

Los resultados del experimento subjetivo de inteligibilidad se presentan en la Tabla 1. En la misma, se puede ver que tal como era de esperarse, en el caso donde no se aplica ningún tipo de método de transposición espectral, en el caso donde se simula un oído saludable, se obtiene un puntaje cercano al máximo, mientras que en el resto de los casos, la inteligibilidad se deteriora junto con la salud auditiva del individuo. En las restantes filas se aprecian los resultados de los casos donde se aplican métodos de transposición espectral. Puede verse que en la última fila se presentan los resultados del método SPINC, el cual obtiene los mejores resultados. Esto se debe a que como se mencionó previamente, este método de transposición considera a la naturaleza de la voz a la hora de aplicar el procesamiento matemático de la señal. El segundo mejor conjunto de resultados es obtenido por el algoritmo de compresión lineal por tramos, mientras que el peor conjunto de resultados es obtenido por el método de compresión lineal uniforme. Estos resultados son consistentes con lo descrito en la Subsección 2.B, donde se estableció que en el caso de los algoritmos de compresión lineal uniforme, se realiza un desplazamiento de los formantes de mayor energía y la inteligibilidad de la voz se altera. La implementación del método propuesto en Johansson et al obtuvo resultados de inteligibilidad intermedios, esto en principio nos permite establecer que si bien mejora la percepción de algunos sonidos en sorderas leves, al aumentar la complejidad de la deficiencia, el algoritmo no mejora la inteligibilidad de los sonidos.

Por otro lado, el experimento subjetivo que evaluaba la calidad de los audios procesados, obtuvo valores MOS

superiores a 4 para todos los casos, de esta forma la calidad de los audios obtenidos es aceptable para esta aplicación.

Del experimento subjetivo se puede concluir que el método que presenta un mejor rendimiento utilizando la base de datos del Dr. Tato es el SPINC. Además, debe tenerse en cuenta que dicho método, debido al alto factor de compresión que presenta para sonidos con componentes por encima de los 1400 Hz, permite percibir sonidos de alarma, que el individuo utilizando métodos de transposición basados en otros algoritmos no podría percibir. De esta forma, este método favorece la independencia del usuario, lo cual es uno de los principios de la bioética.

En la siguiente subsección se describirán las mediciones experimentales realizadas para verificar en funcionamiento de los métodos en el dispositivo de asistencia desarrollado por el Grupo.

B. Mediciones experimentales

Una vez determinado el rendimiento de los métodos bajo estudio, los algoritmos fueron implementados en un dsPIC33EP512MU810 de Microchip®. Debido a que la única diferencia entre los algoritmos estudiados radica en la forma en que desplazan las componentes espectrales, parámetros como el tiempo de procesamiento y el uso de memoria son iguales para cada implementación. Por ello, en la Tabla 2 solo se describe el uso de recursos del algoritmo SPINC, el cual obtuvo los mejores resultados en los experimentos subjetivos. En primer lugar, se puede ver que existen gran cantidad de memoria de datos y de programa remanente. En segundo lugar, la memoria de datos asignada a DMA está totalmente utilizada. En tercer lugar, debido a que la suma de los tiempos de procesamiento es ampliamente inferior al de adquisición, se puede decir que el sistema opera en tiempo real. Por último, el retardo del sistema es de 15.63ms, lo cual es permisible para aplicaciones de este tipo.

TABLA II

USO DE RECURSOS DE LA IMPLEMENTACIÓN.

Memoria de datos	35%
Memoria de programa	7%
Uso de DMA	100%
Tiempo de adquisición	15.63 ms
Tiempo de procesamiento	3.3 ms

Por último, con el fin de analizar el rendimiento del dispositivo de asistencia, se repitió el experimento subjetivo previo. Pero, en este caso, en lugar de utilizar sonidos procesados en MATLAB®, se utilizó para procesarlos el dispositivo de asistencia auditiva. De esta forma, los voluntarios evaluaron un conjunto de audios con dos etapas de procesamiento: en la primera se aplicaron los algoritmos de transposición espectral utilizando el dispositivo de asistencia, mientras que en segundo lugar, la salida del equipo fue utilizada para excitar el simulador de deficiencias auditivas. De esta forma, se simuló el funcionamiento del equipo desarrollado. Un esquema del banco de pruebas utilizado se presenta en la Fig. 6.

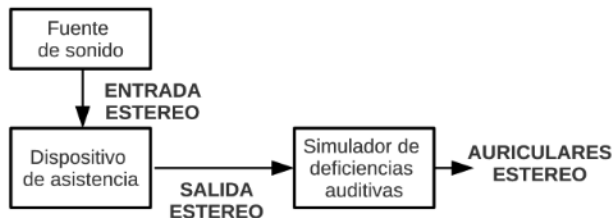


Fig. 6: Diagrama de bloques del banco de pruebas utilizado en el experimento.

Para evaluar los resultados de este experimento se utilizó el mismo criterio que el realizado previamente. Los resultados obtenidos son similares a los obtenidos en la Tabla 1, por ello el sistema funciona de acuerdo a lo esperado. Por último, tal como puede verse en la Tabla 2, para la implementación llevada a cabo existe un remanente de memoria de datos y de programa. Además, como el tiempo de procesamiento es inferior al de adquisición, se podrían implementar otras funcionalidades existentes en otros equipos comerciales. Algunas de estas características son: localización de locutor [19], algoritmos para reducción de ruido [20] y tratamiento de acúfenos [21], entre otros.

V. CONCLUSIONES

En este trabajo se desarrolló una comparación entre algoritmos de transposición de frecuencias, los cuales fueron utilizados para comprimir espectralmente sonidos en un dispositivo de asistencia auditiva. Con el fin de analizar los métodos bajo estudio, se realizaron experimentos utilizando un dispositivo de asistencia auditiva. El algoritmo SPINC obtuvo el mejor resultado en el test de inteligibilidad. Por otro lado, en el test de calidad todos los métodos obtuvieron resultados similares.

También, se demostró que los algoritmos implementados operan en tiempo real en un DSP de bajo costo de la firma Microchip®. Debido a que aún existe memoria de datos y de programa remanente, es posible implementar simultáneamente otras funcionalidades existentes en dispositivos comerciales.

En el futuro, se estudiarán otros algoritmos de transposición espectral con el fin de mejorar las prestaciones del equipo desarrollado. Finalmente, debido a los resultados obtenidos, se desea comenzar a realizar experimentos con individuos que padezcan dolencias auditivas.

AGRADECIMIENTOS

Deseamos agradecer al CONICET y a la Universidad Nacional de Mar del Plata por el financiamiento fruto del cual se desprende este trabajo. El mismo ha sido obtenido a través de los proyectos "Análisis de señales destinados a mejorar la calidad de vida" (15G/380) y "Secuencias caóticas digitales en procesamiento y encriptado de señales" (PIP 112 - 200801 - 01420).

Además, deseamos agradecer a la Lic. Graciela Moscardi, la Fonoaudióloga Lila Rovedo, al Dr. Ing. Enrique Spinelli y al Ing. Francisco Denk por su colaboración en este trabajo.

REFERENCIAS

- [1] Widex Inc. <http://www.widex.com/>, Lyngø, Denmark, 2009.
- [2] Siemens Hearing Instruments, Introduction to Tinnitus Management, <http://hearing.siemens.com>, 2008.
- [3] B. Johansson, A new coding amplifier system for the severely hard of hearing. Stuttgart. pp. 655-657.
- [4] Phonak Hearing Systems, <http://www.phonak.com/>. Zurich, Switzerland, 2010.
- [5] M. Velmans, Aids for deaf persons. U.K. patent 13401105, 1971.
- [6] F. Kuk, H. Peeters, D. Keenan, A. Jessen y H. Andersen, Linear frequency transposition: Extending the audibility of high-frequency information. *Hearing Review*, 13, pp.44-46, 2006.
- [7] F. Kuk, H. Peeters, P. Koorhonenen y C.C. Lau, Efficacy of linear frequency transposition on consonant identification in quiet and in noise. *Journal of the American Academy of Audiology*, 20(8), pp.465-479, 2009.
- [8] E. Terhardt, The SPINC Function for scaling of frequency in Auditory Models. *Journal of Acoustic*. Vol.77 pp.40-42, 1992.
- [9] A. J. Uriz, P. D. Agüero, J. Castiñeira Moreira, J. C. Tulli y E. L. González, Simulador de Deficiencias Auditivas, *Revista Iberoamericana de Automática e Informática e Industrial*, Vol. 8, No.2, pp. 52-62, 2011.
- [10] Microchip Inc., dsPIC33EP512MU810 Data Sheet, High Performance 16-bit Digital Signal Controllers. <http://www.microchip.com/>, 2011.
- [11] O. Turk y L.M. Arslan, Subband based voice conversion, *Proceedings of ISCLP 2002*. pp. 289-292, 2002.
- [12] O. Timms, Speech Processing strategies based on the sinusoidal speech model for the profoundly Hearing Impaired, Ph.D. Thesis, 2003.
- [13] Texas Instruments, PCM3052A Datasheet, <http://www.ti.com/>, 2005.
- [14] B. R. Archambeault y J. Drewniak, *PCB Design for Real-World EMI Control*, Springer, 2002.
- [15] NEC Corporation, EMISream Website, <http://www.nec.com/en/global/prod/emistream/>, 2011.
- [16] G. De Sebastian, J.J. Badaroco y D.G. Postan, *Audiología Práctica*, Ed. Médica Panamericana, 1992.
- [17] International Organization of Standardization, ISO/IEC JTC1/SC29/WG111MPEG IS13813(1994): Generic Coding of Moving Pictures and Associated Audio, International Organization of Standardization. Ginebra, Suiza, 1994.
- [18] H. Fletcher, Auditory Patterns, *Modern Physics* Vol. 12, No. 1, pp. 4765, 1940.
- [19] A. K. Tellakula, Acoustic Source Localization Using Time Delay Estimation, Degree Thesis. Bangalore, India: Supercomputer Education and Research Centre Indian Institute of Science, 2007.
- [20] V. Balakrishnan, N. Borgesa y L. Parchment, Wavelet denoising and speech enhancement, 2006.
- [21] J. A. Henry, T. L. Zaugg, P. J. Myers y M. A. Schechter, Using Therapeutic Sound With Progressive Audiologic Tinnitus Management, *Trends on amplification*, Vol. 12, No. 3, pp.188-209, 2008.