

SLD184 ALGORITMOS PARA PROTESIS AUDITIVAS UTILIZANDO LA TRANSFORMADA WAVELET PACKETS

SLD184 ALGORITHMS FOR DIGITAL HEARING AIDS BASED ON WAVELET PACKETS TRANSFORM

Ing. Beatriz Peón Pérez*. Ing. Roberto Jiménez Yera**. Dr.C.T. Julián Cárdenas Barrera***.

* Centro Provincial de Ingeniería Clínica y Electromedicina. Villa Clara.

Correo electrónico: beatrizpp@capiro.vcl.sld.cu

** Centro Provincial de Ingeniería Clínica y Electromedicina. Villa Clara.

*** Centro de Estudios de la Electrónica y la Tecnología de la Información
Facultad de Ing. Eléctrica. UCLV.

RESUMEN: *La habilidad de escuchar todos los sonidos del mundo real depende del correcto funcionamiento de las estructuras que componen el Sistema Auditivo. Una discapacidad auditiva común conocida como hipoacusia, pérdida auditiva sensorineural de origen coclear, afecta a una de cada diez personas en el mundo según datos de OMS. El tratamiento de estas patologías depende de la correcta selección y funcionamiento de las prótesis auditivas. Varios métodos basados en wavelet han sido usados en el procesamiento digital de señales de las prótesis auditivas para compensar la discapacidad. Los métodos más completos utilizan los mismos coeficientes wavelet para las etapas de reducción de ruido y compensación de la hipoacusia. Este trabajo propone dos algoritmos basados en la Transformada Wavelet Packet, uno para la etapa de reducción de ruido y el otro es un algoritmo de control automático de ganancia, ambos para procesamiento digital de señales en prótesis auditivas digitales.*

Palabras claves: Hipoacusias, Prótesis Auditivas, reducción de ruido, umbral, Transformada Wavelet Packets

ABSTRACT: *The ability of listen to all sounds of the real world depends on the correct performance of the structures that make up the Additive System. A common hearing impairment known as recruitment of loudness, a sensorineural hearing loss of cochlear origin affect one out to ten people worldwide according to OMS data. The treatment of these pathologies depends on the correct choice and performance of hearing aids. Several wavelet-based methods have been used in the digital signal processing of hearing aids to compensate hearing impairment. The more complete method performs both denoising and amplitude compression using the same wavelet coefficients for both stages. This work proposes two algorithms based on Wavelet Packet Transform, one is a noise reduction algorithm and the other is an automatic gain control algorithm, both for digital signal processing in digital hearing aids.*

Keywords: Hipoacusia, additive prótesis, Wavelet packets

1. INTRODUCCIÓN

Para el hombre, el sistema auditivo es uno de los sistemas sensoriales más importantes debido a la gran cantidad de información ambiental que procesa. Además, es uno de los sistemas involucrados en la adquisición, desarrollo y mantenimiento del lenguaje hablado, por lo que su funcionamiento tiene una gran influencia en el papel desempeñado por cada individuo en la sociedad [1].

Las pérdidas auditivas más comunes en los pacientes son las hipoacusias; éstas constituyen afecciones del sistema auditivo de manera parcial, y afectan tanto a niños como a adultos, sin importar el sexo o nivel social.

A principios del siglo XX aparecieron los primeros aparatos eléctricos destinados a corregir los problemas relacionados con las hipoacusias y se describen como las primeras prótesis auditivas que se desarrollaron. Su funcionamiento se basó en el teléfono; donde el sonido emitido ante un micrófono de granulos de carbón convierte la onda sonora en señal eléctrica que es transmitida hasta un auricular y en este tiene lugar el proceso inverso. El promotor de estos ingenios fue el estadounidense Millar Reese Hutchinson, cuyos esfuerzos tuvieron una brillante repercusión internacional en 1901 con motivo de la coronación de la reina Alejandra, esposa de Eduardo VI de Inglaterra: la reina era sorda y gracias al invento de Hutchinson, una especie de voluminoso teléfono portátil, pudo seguir la ceremonia y atender a las preguntas de ritual con entera corrección.

En la actualidad ha existido un desarrollo vertiginoso de este tipo de dispositivos y las prótesis auditivas digitales han ocupado un espacio importante en el diseño de equipos electrónicos con resultados muy alentadores para el tratamiento de las hipoacusias en todas sus modalidades. Estas transforman el sonido y lo procesan antes de amplificar la señal. Una prótesis auditiva digital puede ser programada y ajustada individualmente para satisfacer cada deficiencia de manera personalizada [2]. Con la introducción del circuito digital, el sonido se puede amplificar o atenuar según se requiera sin el temor de sobrepasar los límites tolerables de comodidad, con la ayuda de controles de ganancia que mantienen un nivel estable de intensidad sonora.

Los programas que utiliza este tipo de dispositivo pueden ser más flexibles según necesidades específicas, realizando una amplificación determinada de ciertas frecuencias, de modo que el paciente pueda oír claramente las voces de sus

interlocutores, incluso cuando se encuentre en ambientes ruidosos.

La descomposición en bandas de frecuencia de la señal es una de las etapas más importante en la corrección del sonido para la estimulación correcta del oído en pacientes con hipoacusias específicas y constituye la base de la implementación de prótesis auditivas digitales de tipo personalizadas, las cuales se ajustan a los requerimientos individuales del paciente basados en el tipo de pérdida individual detectada por las audiometrías tonales de diagnóstico [2].

La adquisición de estos dispositivos en el mercado internacional para su utilización en nuestro país constituye uno de los problemas a los cuales se debe enfrentar el Estado, debido a las políticas internacionales y su alta cotización monetaria (aproximadamente 3000 USD). Es por eso que el desarrollo de una prótesis auditiva y su producción nacional resulta de vital importancia para el tratamiento de estas afectaciones en Cuba.

La aplicación de la Transformada Wavelet (TW) a una señal es equivalente a la implementación de un banco de filtros no uniforme que usa una ventana de tamaño variable para diferentes componentes de frecuencia [3]. Permite obtener intervalos largos de tiempo para bajas frecuencias y cortos para frecuencias altas de esta forma se logra una mejor representación de la señal en bandas de frecuencia, este tipo de análisis de señales se conoce como *análisis multiresolución* y es fundamental en el estudio de señales no estacionarias como las señales de voz.

El cálculo de la Transformada Wavelet Discreta (TWD) consiste en la obtención de los Coeficientes Wavelets (CW) a partir del cálculo de la correlación entre la señal original $x(n)$ y funciones prototipo llamadas función escala $\varphi_{j,k}(t)$: ecuación 1 y función wavelet $\psi_{j,k}(t)$: ecuación 2.

$$\varphi_{j,k}(t) = 2^{j/2}\varphi(2^j t - k), \quad j, k \in \mathbf{Z}, \quad 1$$

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k), \quad j, k \in \mathbf{Z}, \quad 2$$

Donde: j es el factor de compresión o dilatación y k es el factor de traslación en el tiempo. La forma de onda de las dos funciones varía según el tipo de wavelet que se utilice [4].

Los CW son clasificados en dos tipos: coeficientes de aproximación obtenidos al hallar la correlación de $x(n)$ con $\varphi_{j,k}(t)$ y coeficientes de detalle que se obtienen con el cálculo de la correlación de $x(n)$ con

$\psi_{j,k}(t)$. Estos coeficientes se caracterizan por concentrar en ellos distintos componentes de frecuencia de la señal. De manera tal que los coeficientes de aproximación contienen los componentes de baja frecuencia y los coeficientes de detalle contienen los componentes de alta frecuencia de $x(n)$.

Esta técnica para el cálculo de la TW hace que el proceso sea análogo a un filtrado paso bajo y paso alto siendo $h_0(n)$ y $h_1(n)$ los filtros de descomposición (fig. 1). El proceso puede ser iterativo mediante la descomposición sucesiva de las aproximaciones, siendo N el nivel de descomposición. Así la señal queda distribuida en bandas de frecuencia de manera similar a un banco de filtros. Luego del filtrado se realiza un diezmado por dos como muestra la figura 1.

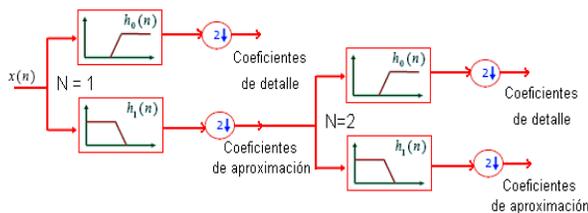


Figura 1. Proceso de descomposición (dos niveles) de la señal $x(n)$ mediante la TWD.

El método de Wavelet-Packets (WP) es una generalización del proceso de descomposición de la Transformada Wavelet que ofrece un mayor rango de probabilidades para el análisis de la señal. En el análisis con wavelet la señal se distribuye en coeficientes de aproximación y detalle. Los coeficientes de aproximación son transformados igualmente, para un segundo nivel de descomposición, en nuevos coeficientes de aproximación y detalle y así sucesivamente.

Las técnicas que utilizan WP tienen muy buena aceptación en este tipo de problemas pues se pueden lograr bandas de frecuencia tan finas como se desee.

La descomposición es análoga a un filtrado sucesivo cuya frecuencia de corte es siempre la mitad de la mayor frecuencia de la señal. Esta forma de descomposición de la señal se asemeja en gran medida a la descomposición en frecuencias del sonido realizada por la cóclea del sistema auditivo humano. Esto sugiere que la aplicación de la Transformada Wavelet y TWP en este tipo de procesamiento sea adecuada [4].

Los filtros de análisis y reconstrucción son conocidos como "Filtros de Reconstrucción

Perfecta" o "Filtros de Espejo en Cuadratura" por su respuesta de frecuencia característica, que permite la reconstrucción perfecta de la señal [5]. Esto evita la introducción de errores en el proceso de síntesis a partir de los coeficientes modificados.

Teniendo en cuenta lo antes expuesto este trabajo se basa en la aplicación de la Transformada Wavelet Packets (TWP) al procesamiento digital de señales sonoras para su utilización en la reducción de ruido y la compensación de hipoacusias en prótesis auditivas digitales. Además se realiza una evaluación del desempeño de ambos algoritmos.

2. CONTENIDO

MATERIALES Y MÉTODOS

– Algoritmo de reducción de ruido basado en TWP.

El ruido de fondo presente en las señales de voz es un problema que necesita ser solucionado para cualquier técnica de procesamiento que involucre a estas señales. Este tipo de ruido puede causar pérdida de la inteligibilidad del habla y distorsión en la información presente en la señal.

Sin embargo, en la actualidad, una de las dificultades que se presentan al diseñar técnicas que empleen procesamiento de voz, como la comunicación telefónica y las prótesis auditivas, es el logro efectivo de la reducción de ruido presente en las señales del habla. En gran medida esto se debe a las características aleatorias y no estacionarias del ruido y a la similitud con él de algunos componentes de la voz como señal. Los efectos de una pérdida auditiva sensorineural incrementan el problema de la escucha tratando de entender el habla en un ambiente ruidoso, lo que evidencia la necesidad del empleo de un buen sistema de reducción de ruido antes de cualquier otro tipo de procesamiento para la compensación de la hipoacusia [6].

Las técnicas de reducción de ruido basadas en la TWP se sustentan en las diferencias espectrales del ruido y la señal del habla. Los componentes de ruido se encuentran distribuidos en coeficientes de valor absoluto cercano a cero en las descomposiciones mediante TWP, lo que posibilita una mejor eliminación del mismo sin afectar la información útil de la señal de voz. De manera general el proceso consiste en la eliminación de todos los coeficientes wavelet de la señal que son menores que un umbral predeterminado. Este método de eliminar los componentes de ruido mediante la aplicación de umbrales a los

coeficientes wavelet está basado en el hecho de que para muchas señales, como el habla, la energía se concentra en un pequeño número de dimensiones wavelets [7] [4].

Existen algunas dificultades en los métodos de utilización de umbrales basados en wavelet cuando son aplicados a señales corrompidas por ruido, sobre todo si es ruido real. En primer lugar los métodos básicos asumen que el espectro del ruido es blanco, sin embargo el ruido real presente en este tipo de señales es ruido coloreado. De esta forma los métodos basados en wavelet no resultan en una buena calidad de voz y no pueden extraer de manera eficiente el ruido coloreado [7].

Se propone utilizar un algoritmo que modifique la aplicación de umbrales a los coeficientes wavelet así como la forma del cálculo de estos umbrales, de tal forma que sea capaz de adaptarse a cambios en las características, tanto de la señal de voz, como del ruido de fondo. El algoritmo elegido se basa en los descritos por Ghanbari y cols. 2005, [8] haciendo modificaciones en algunos aspectos de su funcionamiento con el objetivo de mejorar su desempeño. En la figura 2 se muestra un diagrama en bloques que representa las etapas de dicho algoritmo, donde *s* corresponde a la señal de original y *sm* la señal modificada por el algoritmo.



Figura 2: Diagrama en bloques del algoritmo de reducción de ruido basado en TWP.

El proceso de eliminación de ruido se basa en la estimación y posterior actualización de un patrón de ruido a partir de la señal contaminada. En primer lugar se hace un proceso de división en tramas de la señal original, sin solapamiento y se asume como punto de partida que la primera trama de la señal presenta ausencia de voz (silencio), lo que implica que se tome esta trama de la señal como ruido. Inicialmente se calculan los coeficientes wavelets de la trama de la señal a partir de la descomposición mediante WP (DWP por sus siglas en inglés).

Luego se realiza un proceso de detección de pausa basado en la energía de los coeficientes. Si la trama de la señal es detectada como pausa se actualizan los valores del patrón de ruido y el umbral y energía de ruido en cada banda de frecuencia. Posteriormente se obtiene la relación señal a ruido, también por banda entre la trama de

la señal no detectada como silencio y el ruido estimado. A partir de esta relación y del umbral actualizado se calcula un nuevo umbral.

El bloque nombrado como Clasificador de actividad de voz se encarga de etiquetar los segmentos como sordos "unvoiced" o sonoros "voiced". Para prevenir degradación en las regiones sordas, primero se detectan los segmentos sordos de los sonoros. En este sistema la detección sorda-sonora se hace basada en la TWP [8]. A partir de este bloque se modifican los valores de umbral obtenidos luego de la detección de pausa.

Posterior al cálculo del umbral, este se le aplica a los coeficientes de señal para obtener los coeficientes de la señal filtrada.

El último bloque del diagrama (RWP) corresponde al proceso de recomposición de la señal mediante TWP inversa. De esta forma se obtiene la señal modificada con el ruido sustraído.

- Estimación del ruido y parámetros relacionados con este.

Al descomponer la señal mediante TWP se generan 2J bandas de frecuencia, siendo J los niveles de descomposición de la misma. En análisis de señales de voz lo más común es utilizar 8 o 16 bandas de frecuencia. En este caso se decidió realizar una descomposición en tres niveles lo que provocaría un total de 8 bandas de frecuencia.

Cuando una trama de la señal es detectada como pausa o en ausencia de voz se toma el patrón de ruido como la trama en sí y los coeficientes wavelet del ruido (CWR) como los coeficientes que resultan de la descomposición de la trama, en todas las bandas de frecuencia. Para todas las tramas de la señal se actualiza la energía estimada del ruido (EEruido), en cada banda de frecuencia, a partir de los CWR de la siguiente forma:

$$EEruido_{k,n} = \alpha EEruido_{k,n-1} (1 - \alpha) \sum_{i=1}^{L_k} CWR_k^2(i), \quad 3$$

Se denota cada banda de frecuencia con el índice *k* y cada trama de la señal con *n*, α es un factor establecido por optimización, $0 < \alpha < 1$. La cantidad de coeficientes de la *k*-ésima banda de frecuencia se identifica por *L_k*.

Para la primera trama de la señal la energía del ruido se calcula de diferente manera, como indica la ecuación 4.

$$EEruido_{k,1} = \sum_{i=1}^{L_k} CWR_k^2(i), \quad 4$$

Los valores de la E_{ruido} se utilizan para la detección de segmentos de pausa, comparándolos con los valores de la energía de los coeficientes wavelet (CW) luego de las modificaciones realizadas por el proceso de reducción de ruido.

A partir de la actualización del patrón de ruido se calcula la Relación Señal a Ruido por banda (RSRb) según la ecuación 3 mediante la cual se logra determinar las correspondencias entre las tramas de señal y el ruido.

$$RSRb_{k,n} = 10 \log_{10} \frac{\sum_{i=1}^{L_k} SW_{k,n}^2(i)}{\sum_{i=1}^{L_k} CWR_{k,n}^2(i)}, \quad 5$$

Donde $SW_{k,n}$ y $CWR_{k,n}$ son los coeficientes wavelet de la k -ésima banda de frecuencia de las n -ésimas tramas de señal ruidosa y ruido estimado respectivamente.

La RSRb se utiliza para modificar el umbral adaptándolo en correspondencia con su valor. Así el umbral logra una correspondencia con las variaciones temporales del ruido.

- Determinación y aplicación del umbral adaptativo.

Existen varios métodos para la obtención de umbrales con el objetivo de eliminar ruido en señales de voz, el umbral básico para uso en algoritmos basados en Wavelet Packets se muestra en la ecuación 6 y es el utilizado en este algoritmo.

$$U = \hat{\sigma} \sqrt{2 \log(L \log_2 L)}, \quad 6$$

Siendo L la longitud de la señal ruidosa s y $\hat{\sigma}$ la desviación estándar del ruido que puede ser estimada como:

$$\hat{\sigma} = \frac{mad(CWR)}{0.6745}, \quad 7$$

El valor de este umbral no se emplea de forma invariable, sino que es adaptado teniendo en cuenta la RSRb, entre otros factores, obteniéndose un nuevo valor para cada segmento de la señal y cada banda de frecuencia.

Cuando una trama de la señal es detectada como silencio se actualiza el umbral según indica la ecuación 8. El umbral correspondiente a la primera trama de la señal se calcula a partir de la ecuación 6. Para los demás segmentos se modifica su valor de acuerdo al umbral obtenido en la trama anterior.

$$U_{k,n} = \alpha U_{k,n-1} + (1-\alpha) \hat{\sigma}_{k,n} \sqrt{2 \log(L_k)}, \quad 8$$

La desviación estándar del ruido se calcula mediante la ecuación 7 y α es el mismo factor calculado experimentalmente de la ecuación 8. Para ambas ecuaciones se utilizó un valor de α igual a 0,9.

Se realiza el cálculo de un umbral según la ecuación 9 que posibilita el ajuste a los cambios en las características inherentes a la señal y el ruido. Se obtiene a partir del umbral calculado en la ecuación anterior y depende de la RSRb.

$$U_{k,n} = \begin{cases} U_{k,n} \left(1 + e^{-\frac{RSRb_{k,n}}{t}} \right) RSRb \geq 0 \\ 2U_{k,n} RSRb < 0 \end{cases}, \quad 9$$

El factor t , utilizado en la expresión anterior, se calcula de manera experimental. Para mejores resultados su valor debe estar entre 2 y 3, en este caso se utilizó un valor de t igual a 2,2.

La aplicación del umbral adaptativo $U_{k,n}$ a los coeficientes wavelet de la señal (CWS) se hace en cada trama y para todas las bandas de frecuencia; a partir de una modificación del umbral duro convencional. Cuando el valor absoluto del coeficiente está por debajo del valor del umbral, este no se elimina completamente sino que se reduce de manera sustancial según la ecuación 10, donde el factor d se obtiene de manera experimental. En el algoritmo se utilizó un valor de d igual a 4, este hace una eliminación lo suficientemente ruda de los coeficientes, sin perder el objetivo de la ecuación 10 de modificar el método clásico de aplicación del umbral.

$$CWSM_{k,n} = \begin{cases} CWS_{k,n} | CWS_{k,n} | \geq U_{k,n} \\ \text{signo}(CWS_{k,n}) \frac{|CWS_{k,n}|^d}{U_{k,n}^{d-1}} | CWS_{k,n} | < U_{k,n} \end{cases}, \quad 10$$

De esta forma se obtienen los coeficientes wavelets de la señal modificada de la k -ésima banda de frecuencia y la n -ésima trama de la señal. A partir de estos coeficientes se obtiene la señal filtrada mediante la recomposición por WP.

- Detección de los segmentos de pausa.

La detección de pausas en la actividad de voz es una etapa importante en los algoritmos de reducción de ruido basados en la estimación del ruido presente en la señal. Permite establecer las diferencias entre silencio y voz en presencia de un ambiente ruidoso.

La clasificación de un segmento de señal del habla como silencio se puede realizar por varias vías, mediante el cálculo de cruces por cero, mediante la determinación de la energía de la señal, etc. En este caso se utiliza una comparación entre la energía de los coeficientes de la señal modificada y la energía del ruido estimado.

- Clasificación de la actividad de voz.

La clasificación de las tramas de la señal en sordas o sonoras se hace basada en la Transformada Wavelet Packets. En este método para cada banda de frecuencia se calcula la energía promedio. Acumulando la energía promedio de las bandas por debajo de 2kHz se obtiene EL (energía de baja frecuencia) y similarmente, mediante la acumulación la energía promedio de las bandas por encima de 2kHz tenemos EH (energía de alta frecuencia) [8].

Cuando se detecta una pausa se calcula la razón entre EL y EH como EL_p/EH_p para esa trama. Luego tomando este valor como límite para la clasificación tenemos que si para la trama actual EL/EH es mayor que EL_p/EH_p , esa trama se clasifica como sonora, de lo contrario se considera como segmento sordo [8].

Luego de la clasificación se modifica el valor de umbral que le será aplicado a los coeficientes, incrementando el mismo para las bandas altas de los segmentos etiquetados como sonoros y para las bandas bajas de frecuencia de los segmentos considerados como sordos. El umbral se incrementa teniendo en cuenta un porcentaje del mismo, 25% en este caso. Las bandas por debajo de 2kHz son las nombradas como bandas bajas y las que se encuentran por encima de este valor como bandas altas.

- Algoritmo de compensación de pérdidas auditivas mediante TWP.

Los pacientes con sorderas sensorineurales generalmente experimentan pérdidas de la información de la señal, disminuyendo cuantiosamente el desarrollo del individuo en la sociedad. Teniendo como resultado un deterioro en la inteligibilidad de la voz debido a la reducción del rango dinámico de la audición. Esta característica hace necesaria la transformación de la señal para que los niveles de intensidad de la misma se sitúen dentro del rango dinámico propio del individuo discapacitado.

Las técnicas utilizadas para la compensación de pérdidas auditivas deben garantizar en primer lugar la amplificación de los niveles de intensidad de la señal sonora de manera tal que proporcione una compensación de la pérdida auditiva. Luego del

procesamiento, la intensidad del sonido escuchado por el individuo debe estar por encima del umbral de audición propio de éste, y al amplificar la intensidad de la señal, no debe sobrepasar el umbral de inconformidad o del dolor. Se debe concentrar la intensidad de la señal sonora dentro del rango dinámico de audición propio del sujeto [9]. Las técnicas de compensación lineales aplican una ganancia directamente a la amplitud de la señal, por lo que solo son eficientes en sorderas conductivas. En el caso de sorderas sensorineurales con importantes restricciones en el rango dinámico, las técnicas lineales no son aplicables, es necesaria la utilización de técnicas no lineales que varíen la ganancia aplicada a la amplitud de la señal en función de la frecuencia. Estas técnicas tienen un comportamiento similar a los ecualizadores de frecuencia [2].

- Control automático de ganancia mediante TWP.

El análisis de señales mediante TWP posibilita una forma de descomposición en frecuencia útil en este tipo de aplicaciones. A partir del umbral de audición de discapacidad y del rango dinámico audible es posible lograr una modificación acertada de los coeficientes proporcionados en la descomposición mediante WP con el objetivo de compensar la pérdida auditiva. Además la compresión de la amplitud mediante WP brinda un ajuste rápido de la ganancia de la señal del habla para asegurar la amplificación de los componentes de bajo nivel de intensidad sin que los de mayor amplitud sobrepasen el umbral del dolor [10].

Con el propósito de alcanzar una modificación enfocada a la compensación de este tipo de pérdidas auditivas; se propone un algoritmo que utiliza técnicas de compresión de amplitud con TWP basado en el diseño utilizado por Trenas, M. A, y cols, 1999; [10] y descrito en [11].

El objetivo de esta modificación de la señal es aumentar la intensidad por encima del umbral de audición del individuo, posibilitando una compensación para su hipoacusia. El incremento de la intensidad se hace por bandas de frecuencia teniendo en cuenta las características de la pérdida del sujeto en particular.

La modificación realizada a los CW de la señal en cada banda de frecuencia se corresponde con el umbral de audición propio del individuo, además de la razón existente entre el rango dinámico de discapacidad del sujeto y el rango dinámico normal de audición. Luego de obtener la equivalencia en dB de los coeficientes de la señal ($CWM_{db_{k,n}}$) se obtienen los coeficientes modificados también en dB según la ecuación 11 [11].

$$CWMdb_{k,n} = Ud + (CWdb_{k,n} - Un) * Rrd, \quad 11$$

La relación entre los coeficientes de la señal y su valor en dB es de la forma:

$$CWdb = 20 \log Cw, \quad 12$$

Los valores de umbral de discapacidad, umbral normal y umbral del dolor de la k-esima banda de frecuencia se denotan como Udk, Unk y Udolork, respectivamente y la razón del rango dinámico (Rrd) se calcula para cada banda como:

$$Rrd = \frac{Udolor - Ud}{Udolor - Un}, \quad 13$$

3. RESULTADOS Y DISCUSIÓN

Los dos algoritmos explicados anteriormente fueron implementados sobre la plataforma de MATLAB 7.9 [12], ambos desarrollados utilizando la wavelet madre db8. Esta decisión se tomó luego de un experimento donde se evaluaron distintos tipos de wavelet y su comportamiento ante este tipo de señales; además, db8, es de las más utilizadas en la literatura para aplicaciones de este tipo [8] [7].

- Desempeño del algoritmo de reducción de ruido.

El algoritmo de reducción de ruido propuesto fue comparado con otros dos en cuanto a la posibilidad de reducir el ruido presente en la señal con la mínima distorsión posible de la voz. Se seleccionaron para la comparación los algoritmos siguientes:

- Método de sustracción espectral multibanda (SEM) utilizando banco de filtros gammatone [13].
- Algoritmo de sustracción espectral geométrico (SEG) [14].

Se comparó con estos algoritmos pues ambos son fáciles en su implementación y por consiguiente no mantienen un elevado costo computacional [15]; característica importante a tener en cuenta a la hora de seleccionar algoritmos para su implementación en prótesis auditivas digitales y que los métodos basados en TWP comparten.

Las comparaciones se realizaron teniendo en cuenta las medidas objetivas compuestas como distorsión de la señal (Cs), distorsión del ruido (Cn)

y calidad general de la voz (Cg) desarrolladas por Hu y Loizou, 2006 [16]. Se tomaron tres señales para la comparación de los algoritmos procedentes de la base de datos española AHUMDA [17], muestreadas a 8000Hz, todas pertenecientes a voces masculinas.

Las señales fueron contaminadas artificialmente con dos tipos de ruido: murmullo, "babble" y ruido presente en autos; con dos relaciones señales a ruido (RSR) 5dB y 10dB. Los ruidos fueron tomados de la base de datos AURORA [18] con una frecuencia de muestreo de 19980Hz.

En la Figura 3 se muestran las formas de onda de una de las señales utilizadas para la comparación de los algoritmos: señal limpia en la parte superior, la señal contaminada con ruido murmullo con una RSR de 5dB en el centro y la señal filtrada por el método de reducción de ruido basado en TWP propuesto. Se observa que este algoritmo realiza una reducción del ruido de fondo, sin embargo, no lo elimina completamente.

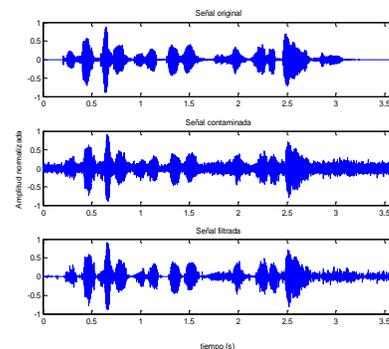


Figura 3. Formas de onda de la señal limpia (arriba), contaminada (centro) y filtrada (debajo) por el algoritmo basado en WP

También se muestran en la Figura 4 los espectrogramas correspondientes a las señales representadas en la Figura 3 que reafirma lo anteriormente concluido.

En la Tabla 1 se muestran los valores resultantes del cálculo de las medidas objetivas compuestas mencionadas anteriormente para tres señales de voz contaminadas con ruido murmullo a 5 y 10dB. La nomenclatura que se utiliza en las tablas indica el número de la señal y la RSR a la que se le aplicó el ruido de la siguiente manera: (1_10) en el caso de la señal 1 contaminada con RSR de 10dB.

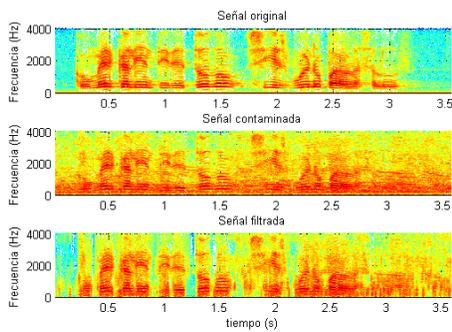


Figura 4. Espectrogramas de la señal limpia (arriba), contaminada (centro) y filtrada (debajo) por el algoritmo basado en WP

Al analizar los valores de las medidas contempladas resultado del experimento realizado se deduce claramente que el algoritmo propuesto de reducción de ruido mediante TWP es superior al de sustracción espectral multibanda. Mantiene mejores resultados de todas las medidas calculadas, tanto a las referidas a la reducción de ruido como a la conservación de la calidad de voz. Sin embargo tiene un peor comportamiento que el algoritmo de sustracción espectral geométrico, cuyas medidas se sitúan próximas a sus valores óptimos. A pesar de los resultados arrojados por los valores de todas las medidas objetivas utilizadas en la comparación, desde el punto de vista perceptual, el algoritmo de reducción de ruido basado en WP no realiza una reducción sustancial del ruido de fondo. Para la total comprobación del desempeño de este algoritmo sería necesario realizar evaluaciones del mismo mediante medidas subjetivas según normas de la ITU-T P. 835 [19].

Tabla 1: Valores de medidas objetivas compuestas aplicadas a los algoritmos a comparar para tres señales distintas contaminadas con ruido murmullo a una RSR de 5 y 10dB.

Señal	Algoritmo basado en WP			Algoritmo SEG			Algoritmo SEM		
	Cs	Cn	Cg	Cs	Cn	Cg	Cs	Cn	Cg
1_5	2,84	2,06	2,53	3,01	2,31	2,58	2,08	1,72	2,00
1_10	3,21	2,30	2,92	3,44	2,73	3,01	2,56	1,78	2,40
2_5	2,69	2,01	2,41	3,03	2,34	2,61	1,74	1,37	1,72
2_10	3,03	2,20	2,75	3,47	2,75	3,03	2,33	1,68	2,23
3_5	2,35	1,76	2,04	2,86	2,25	2,41	1,56	1,18	1,49
3_10	2,85	2,01	2,50	3,37	2,69	2,87	2,25	1,47	2,06

- Desempeño del algoritmo para la compensación de la pérdida auditiva.

Para la validación del algoritmo de compensación se utilizaron señales de voz que tenían aplicada una simulación de pérdidas auditivas; de manera tal que tienen una variación específica de la potencia por bandas de frecuencia. Estas señales se obtuvieron en el sitio web "HearingLossSamples" [20] con una frecuencia de muestreo de 44KHz y para utilizarlas fueron submuestreadas a 16kHz. Para la comprobación del desempeño del algoritmo de reducción de ruido se utilizaron señales con pérdidas inducidas en las frecuencias medias y altas.

Se tomó como umbral de discapacidad los valores de potencia por banda de las señales con la simulación de pérdidas y como umbral normal de audición 15dB hasta 1KHz y 0dB en las bandas restantes, están documentados estos valores como los umbrales normales de las personas normo-oyentes [11]. Luego de la aplicación del algoritmo de compensación se calcularon los valores de potencia por banda de las señales modificadas. Se utilizaron tres tipos de pérdidas simuladas, dos de ellas provocadas por exposición al ruido, teniendo como característica principal afectaciones en las altas frecuencias y la tercera que se utilizó presenta pérdidas en las frecuencias medias. En las figuras 5 y 6 se muestran representaciones de las potencias por banda de la señal original, la señal con la simulación de pérdidas y la señal modificada mediante el algoritmo.

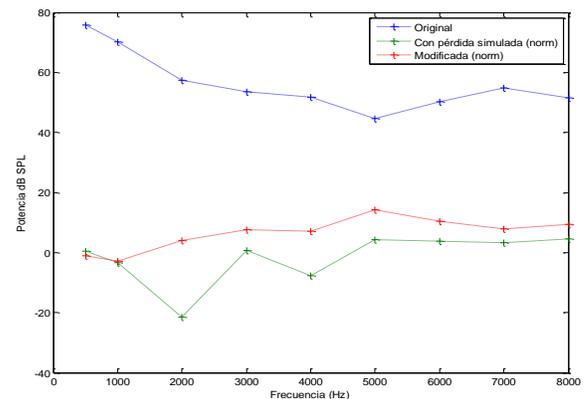


Figura 5. Representación de la potencia en dB SPL de la señal original (azul), potencia normalizada de señal con pérdida simulada en las frecuencias medias (verde) y de la señal modificada por el algoritmo basado en TWP (rojo).

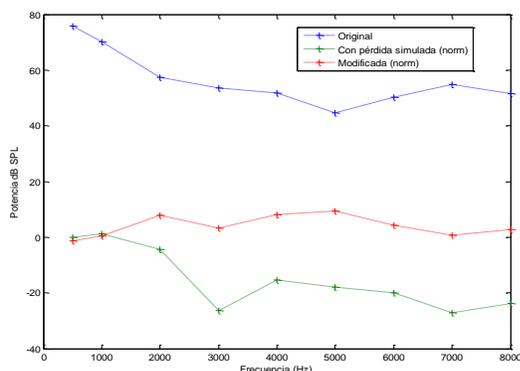


Figura. 6. Representación de la potencia en dB SPL de la señal original (azul), potencia normalizada de señal con pérdida simulada en las altas frecuencias (verde) y de la señal modificada por el algoritmo basado en WP (rojo).

En el caso de las señales con pérdida simulada y modificada se realizó una normalización a sus potencias con respecto a la potencia de la señal original para su representación. Se demuestra que el algoritmo es capaz de compensar la pérdida auditiva manteniendo la potencia normalizada de la señal modificada alrededor de 0dB, lo que evidencia una correcta amplificación de las bandas de frecuencia afectadas por las pérdidas. De igual manera logra corregir las pérdidas auditivas simuladas en frecuencias altas y bajas.

– Desempeño de la prótesis.

Con el objetivo de verificar el comportamiento de la prótesis en general se verificó el desempeño de los dos algoritmos propuestos, primero se aplicó el de reducción de ruido y posteriormente el de compensación de pérdida auditiva. Las señales utilizadas son versiones contaminadas con ruido de las empleadas en la validación del algoritmo de compensación, disponibles también en el sitio web Hearing Loss Samples [20]. Están contaminadas con ruido murmullo con una RSR de 0dB.

En este caso la señal que se modificó fue la resultante del proceso de reducción de ruido. Se utilizaron solo dos tipos de pérdidas simuladas, (pérdidas en las frecuencias medias y pérdidas moderadas en las altas frecuencias). Igualmente se espera que la potencia normalizada en cada banda de frecuencia de la señal con la pérdida compensada se mantenga alrededor de 0dB.

Para realizar la evaluación del desempeño del algoritmo de reducción de ruido ante la señal contaminada con ruido murmullo con una RSR de 0dB, se compararon los valores de las medidas objetivas compuestas al filtrar la señal con el

algoritmo propuesto y el algoritmo de SEG. En la Tabla 2 se muestran los valores de estas medidas.

Tabla 2: Valores de medidas objetivas compuestas aplicadas a las señales procesadas por el algoritmo basado en WP y el algoritmo de SEG. Señal ruidosa contaminada con ruido murmullo con una RSR de 0dB.

Señal contaminada	Algoritmo basado en WP			Algoritmo de SEG		
	Cs	Cn	Cg	Cs	Cn	Cg
Con pérdidas simuladas en las frecuencias altas.	1,74	1,75	1,81	1,15	1,80	1,28
Con pérdidas simuladas en las frecuencias medias.	2,34	1,75	2,14	2,00	1,88	1,76

Se puede observar en la Tabla 2 para el caso de RSR peores, (0dB) el algoritmo basado en WP tiene mejor comportamiento que el de SEG. Vale destacar que las medidas obtenidas para el algoritmo basado en WP tienen un mejor comportamiento para RSR de 0dB que para RSR mejores como es el caso de la Tabla 1, es decir que se comporta mejor en ambientes más ruidosos.

Luego de realizar la reducción de ruido se procedió a efectuar la compensación de las pérdidas a las señales filtradas con los umbrales utilizados en la compensación de la señal limpia anteriormente correspondiente a cada caso.

En la figura 7 se muestra la distribución de la potencia por bandas de la señal con la simulación de la pérdida auditiva, la señal filtrada mediante el algoritmo de WP y la señal filtrada luego de la compensación de la pérdida simulada para simulación de pérdidas en las frecuencias medias.

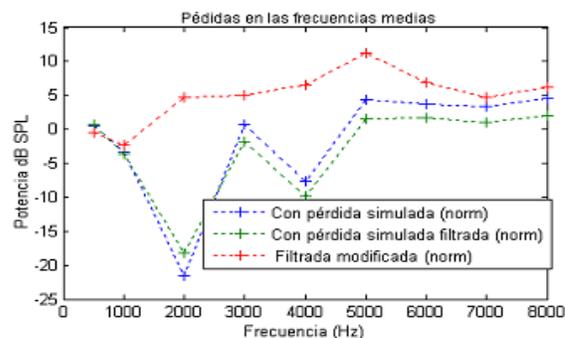


Figura 7. Representación de la potencia por bandas de frecuencia de la señal con simulación de pérdida auditiva, la señal con pérdida contaminada con ruido y posteriormente filtrada por el algoritmo basado en WP y de la señal filtrada con la pérdida compensada.

Como muestra la figura 7 el algoritmo es eficiente en la compensación de la pérdida auditiva, aún luego de un proceso de reducción de ruido se logran niveles válidos de la potencia de las señales

luego del control automático de ganancia. Aunque en algunas bandas la potencia está hasta 10dB por encima del nivel esperado de 0dB, esto no constituye una deficiencia significativa del algoritmo. Un aumento de 10dB es prácticamente imperceptible para el Sistema Auditivo humano.

4. CONCLUSIONES

En este trabajo se aplicó la Transformada Wavelet Packets al procesamiento digital de señales sonoras para su utilización en la reducción de ruido y la compensación de hipoacusias en prótesis auditivas digitales. Se determinaron los algoritmos de reducción de ruido y compensación de pérdidas auditivas basados en la Transformada Wavelet Packets y se le realizaron modificaciones a los mismos con el fin de lograr mejores resultados.

Se evaluó el desempeño de los algoritmos seleccionados basados en la Transformada Wavelet Packets, obteniéndose resultados que brindan ventajas en comparación con algoritmos basados en métodos de sustracción espectral en el caso del algoritmo de reducción de ruido y en el algoritmo de compensación de pérdidas auditivas se demostró su aplicabilidad en el tratamiento hipoacusias.

A pesar de que el algoritmo de reducción de ruido basado en WP no brinda un resultado óptimo, se logran resultados aceptables, mejorando los mismos para contaminaciones de ruido más severas. Además, el hecho de poder utilizar los coeficientes para ambas etapas reduce sustancialmente el costo computacional a la hora de implementar métodos de procesamiento digital de señales para prótesis auditivas lo que hace que los métodos basados en wavelet sean alternativas viables para estas aplicaciones.

5. REFERENCIAS BIBLIOGRÁFICAS

- 1 Arthur C. Guyton, "Sistema nervioso. B: Los sentidos especiales," in *Tratado de fisiología médica.*, 2001, ch. X, pp. 723-740.
- 2 V. Hamacher and J. Chalupper, "Signal Processing in High-End Hearing Aids: State of the Art, Challenges, and Future Trends," *EURASIP Journal on Applied Signal Processing*, vol. 18, pp. 2915–2929, September 2005.
- 3 Truong Q. Nguyen, "A Tutorial on Filter Banks and Wavelets," in *International Conference on Digital Signal Processing*, Cypress, 1995.
- 4 C. Sidney Burrus, Ramesh A. Gopinath, and Haitao Guo, *Filter Banks and the Discrete Wavelet Transform*, Primera ed., Horton Marcia, Ed. New Jersey, USA: Prentice Hall, 1998.
- 5 Jhon G. Proakis and Dimitris G. Manolakis, "Procesado digital de tasa múltiple," in *Tratamiento digital de señales. Principios, algoritmos y aplicaciones*, Tercera edición ed., Isabel Capella, Ed. Madrid: Prentice Hall, 1998, ch. 10, pp. 840-849.
- 6 McGraw-Hill Companies, *Principles of Neural Science*, Cuarta edición ed., Erik R. Kandel, James H. Schwartz, and Thomas M. Jessell, Eds. New York, United States of America: McGraw-Hill Companies, 2000.
- 7 Y. Ghambari, S. Y. Mortazavi, and M. R. Karami, "A New Speech Enhancement System Based on the Adaptive Thresholding of Wavelet Packets," *13th ICEE2005*, vol. 1, May 2005.
- 8 Yasser Ghanbari and Mohammad Reza Karami, "A Modified Speech Enhancement System Based on the Thresholding of the Wavelet Packets," *13th ICEE2005*, vol. 3, May 2005.
- 9 O. Khalifa, M. H. Makthar, and M. S. Baharom, "Hearing Aids System for Impaired Peoples," *International Journal of Computing & Information Sciences*, vol. 2, no. 1, April 2004.
- 10 María A. Trenas, Janet C. Rutledge, and Nathaniel A. Whitmall, "Wavelet-based speech enhancement for hearing aids," *European Medical & Biomedical Engineering Conf(EMBE'99)*, November 1999.
- 11 Janet C. Rutledge, "Speech Enhancement for Hearing Aids," in *Time Frequency and Wavelets in Biomedical Signal Processing*, Metin Akay, Ed. New York, USA: IEEE Press, 1998, ch. 13, pp. 347-366.
- 12 MathWorks Inc. (2009) Matlab. V7.9.
- 13 Sunil D. Kamath and Philipos C. Loizou, "A Multi-band Spectral Subtraction Method for Enhancing Speech Corrupted by Colored Noise," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 2002.
- 14 Yang Lu and Philipos Loizou, "A geometrical approach to spectral subtraction," *Speech Communication*, vol. 50, pp. 453-466, January 2008.
- 15 Philipos C. Loizou, *Speech Enhancement. Theory and Practice.*: Taylor and Francis Group, 2007.
- 16 Yi Hu and Philipos C. Loizou, "Evaluation of Objective Measures of Speech Enhancement," *Proc. Interspeech*, pp. 1447-1450, 2006.
- 17 Javier Ortega, Joaquin Gonzalez, and Victoria Marrero-Aguiar, "AHUMADA: A large speech corpus in Spanish for speaker characterization and identification," *Elsevier Science*, vol. 31, pp. 255-264, 2000.
- 18 H. Hirsch and D. Pearce, "The aurora experimental framework for the performance evaluation of speech recognition systems under noisy conditions.," *ISCA ITRW ASR2000*, September 2000.
- 19 Yi Hu and Philipos C. Loizou, "Subjective comparison and evaluation of speech enhancement algorithms," *Speech Communication*, vol. 49, pp. 588-601, 2007.

20 Scott Bradley. (2007, Dec.) HEARING LOSS SAMPLER. [Online].
<http://facstaff.uww.edu/bradleys/radio/hlsimulation/>

6. SINTESIS CURRICULAR AUTOR PRINCIPAL

Beatriz Peón Pérez recibió el título de Ingeniera Biomédica con distinción de Título de Oro en la Universidad Central de Las Villas, Santa Clara, Cuba, en el año 2010. En estos momentos se encuentra trabajando para obtener el título de Máster en Ciencias en Señales y Sistemas en esta universidad. Se encuentra trabajando como Especialista A en el Centro de Ingeniería Clínica y Electromedicina de la provincia de Villa Clara, Cuba. Sus investigaciones se vinculan al procesamiento de señales sobre todo en la reducción de ruido en señales de audio y procesamiento para prótesis auditivas digitales.

