



# Estudio comparativo de técnicas digitales de reducción de ruido para su posible uso en auxiliares auditivos

Solís Estrella, H.\*  
Gutiérrez Aldana, A.\*

\* Centro de Investigación en Computación, IPN, U. Prof. Adolfo López Mateos, Av. Juan de Dios Bátiz s/n, 07738, México, DF.  
agutierr@cic.ipn.mx, Tel. 5720 6000 Ext. 56616

Correspondencia:  
Alfonso Gutiérrez Aldana  
agutierr@cic.ipn.mx

Artículo recibido: 20/septiembre/2006  
Artículo aceptado: 8/diciembre/2006

## RESUMEN

Se presenta un estudio comparativo de cuatro técnicas de procesamiento digital de señales que se enfocan a solventar el problema del ruido ambiental en auxiliares auditivos, con el objetivo de restaurar en el mayor grado posible la percepción del lenguaje oral. Las técnicas evaluadas fueron el filtro adaptable RACE, la sustracción espectral, la reducción de ruido por descomposición en valores singulares y el análisis de componentes independientes. Se efectuó la comparación calculando la mejoría de la relación señal a ruido en diferentes entornos y aplicando pruebas de significancia estadística sobre los resultados. La técnica que obtuvo mejor desempeño fue el análisis de componentes independientes, cuyos resultados en entorno de simulación muestran que es posible realzar señales de voz, aun cuando su nivel de energía se encuentre 15dB por debajo del nivel del ruido ambiental.

## Palabras clave:

Auxiliares auditivos, reducción de ruido, filtros adaptables, análisis de componentes independientes.

## ABSTRACT

In this paper we present a comparative study of four signal processing techniques focused to solve the environmental noise problem in hearing aids, aiming to restore speech communication to the maximum possible degree. The evaluated techniques are the RACE adaptive filter, the Spectral Subtraction, the Singular Value Decomposition based noise reduction and the Independent Component Analysis. The techniques were evaluated computing the improvement to the signal-to-noise ratio in different environments and performing statistical significance tests on the results. The Independent Component Analysis obtained the best results, which in a simulation environment is able to restore the speech signal even when its energy level is 15dB under that of the noise.

## Key Words:

Hearing aids, Noise reduction, Adaptive filters, Independent component analysis.

## INTRODUCCIÓN

Una de las características más distintivas de los humanos es la capacidad de lenguaje oral articulado, lo cual ha permitido la creación de esquemas sociales muy complejos. Cuando una persona sufre una pérdida auditiva neurosensorial, se ve afectada en sus capacidades de comunicación con su grupo social en mayor o menor medida.

Cualquier afección que imposibilite seriamente las funciones sensoriales o cognitivas tiende a provocar profundos efectos psicológicos no sólo en la persona que está afectada sino también en la familia, amigos, compañeros de trabajo o personas que les brindan cuidados<sup>1</sup>. En los casos donde una discapacidad auditiva conlleva a la percepción de aislamiento por parte del individuo discapacitado, dicho aislamiento puede tener efectos devastadores en él. Entre algunos de los efectos psicosociales que pueden sufrir las personas con discapacidades auditivas se encuentran: dificultad para iniciar y mantener amistades, indicios de baja autoestima, aplazamiento en el desarrollo social, rechazo por sus pares y depresión<sup>2</sup>.

De acuerdo con estadísticas del XII Censo General de Población y Vivienda, las discapacidades auditivas desde temprana edad limitan el nivel educativo de aquellos que las sufren, teniendo éstos un promedio de escolaridad de 4 años<sup>3</sup>. Asimismo, de acuerdo a estudios realizados en adultos mayores, con una mayor severidad de la discapacidad existe una tendencia general a incrementar las probabilidades en restricciones de las actividades instrumentales de la vida diaria y se reducen las probabilidades para el control de toma de decisiones y bienestar emocional<sup>4</sup>.

Los auxiliares auditivos son dispositivos diseñados para compensar ciertos tipos de pérdida auditiva. Un aspecto en común en el diseño de auxiliares auditivos es que constan de un transductor fonocaptor para convertir ondas sonoras en una señal eléctrica, posteriormente la señal pasa por algún tipo de manipulación de manera que compense la pérdida auditiva, para finalmente convertir la señal en sonido con un transductor electroacústico.

Durante las últimas dos décadas se han desarrollado algoritmos de procesamiento digital de señales extremadamente sofisticados, y el hardware digital para tiempo real avanza a paso acelerado, sin embargo hoy en día son relativamente pocos los auxiliares auditivos disponibles comercialmente que realicen procesamiento digital de señales. Muchos de los auxiliares "programables" en el mercado sim-

plemente permiten el control digital de los circuitos analógicos que procesan las señales<sup>5,6</sup>.

Los auxiliares más sencillos utilizan un amplificador para aumentar la intensidad de la señal, sin ser necesariamente de amplificación selectiva; sin embargo, se ha encontrado que el oído sufre un proceso degenerativo cuando esa pérdida de sensibilidad se ha querido compensar únicamente aumentando los niveles de la intensidad sonora<sup>5,7</sup>. Otros métodos que se han intentado para ayudar a las personas con pérdida auditiva están basados en el procesamiento de las frecuencias, compensando aquéllas donde se sufre la pérdida auditiva.

En la actualidad se pueden identificar diversos tipos de técnicas de procesamiento digital de señales cuya implementación es necesaria en los auxiliares auditivos. Un primer grupo de técnicas está conformado por la ecualización por bandas y la compresión de rango dinámico de la señal, se podría considerar como una versión digital de la amplificación selectiva y el control automático de ganancia de los auxiliares auditivos convencionales<sup>8-10</sup>. Una segunda temática es la eliminación del ruido ambiental de las conversaciones; tal vez es la que tiene más rango de acción, dado que el problema se puede atacar con diversas aproximaciones de complejidad variable<sup>11-14</sup>.

Es bien sabido que el ruido de fondo reduce la inteligibilidad de la voz. Es posible entender la voz en un ambiente moderadamente ruidoso porque la voz es una señal altamente redundante. La situación es distinta para una persona con pérdida auditiva; el ruido que enmascara aun una pequeña porción de la empobrecida señal de voz remanente, degradará significativamente la inteligibilidad pues existe menos redundancia disponible para compensar los efectos del ruido. Como consecuencia, las personas con pérdidas auditivas tienen una mayor dificultad para comprender la voz con ruido que las personas con audición normal<sup>11</sup>. Asimismo, si la señal amplificada está compuesta principalmente por voz, es posible utilizar menor ganancia en la amplificación y así se reduce el riesgo de dañar células sanas en el oído.

Un estudio<sup>16</sup> realizado entre usuarios de auxiliares auditivos concluye que los equipos actuales no son satisfactorios, en especial para aquellas personas en el rango de edad de la población económicamente activa. En dicho estudio se listan las áreas en las cuales los usuarios esperarían mejoras y un 95% de los usuarios piensa que es altamente deseable un mejor desempeño en ambientes ruido-

sos, seguido por otros aspectos de la calidad del sonido, mientras que sólo un pequeño porcentaje le dio importancia a cuestiones estéticas.

Considerando lo anteriormente planteado, se desarrolló este estudio para determinar si alguna técnica de procesamiento digital de señales, eliminaría significativamente el ruido recibido en un auxiliar auditivo, permitiendo así, mejorar su desempeño.

## DESARROLLO

Para diseñar técnicas de reducción de ruido se pueden tomar en cuenta ciertos aspectos como las características espectrales de la voz y del ruido o los umbrales de inteligibilidad para evaluar la interferencia con el lenguaje y así estimar ciertos parámetros<sup>11</sup>.

De tal forma surgen principios generales de reducción de ruido, uno de ellos el filtrado fijo en el cual se eliminan las bandas no críticas para la voz, esto es muy básico pues ni la voz ni tipos comunes de ruido son estacionarios o siempre tienen el mismo contenido en frecuencia.

Una evolución natural son los filtros adaptables, los cuales modifican su respuesta en función de las características de las señales. Asimismo, otro principio que se puede usar es el filtrado espacial aprovechando "observaciones" de dos o más transductores colocados de manera específica para obtener más información sobre las señales de voz y ruido.

Para el desarrollo del presente trabajo se evaluaron cuatro técnicas de reducción de ruido ambiental: el filtro adaptable RACE (realizado adaptable por correlación), la sustracción espectral, el filtrado por medio de GSVD (descomposición en valores singulares generalizada) y el análisis de componentes independientes. Las primeras dos son técnicas clásicas de reducción de ruido implementadas en prototipos de auxiliares auditivos, el filtrado por GSVD es una propuesta reciente la cual está implementada en tiempo real en una computadora. El análisis de componentes independientes es una técnica con un enfoque distinto al del filtrado adaptable (en el que se basan de manera general las otras tres técnicas), forma parte de una familia de técnicas utilizadas para la separación a ciegas de fuentes y no ha sido implementado en auxiliares auditivos pero ha mostrado resultados interesantes en otras áreas<sup>17</sup>. A continuación se examinarán diversas consideraciones teóricas de estas técnicas.

## FILTRO ADAPTABLE RACE

Magotra et al proponen una solución implementada en un DSP (procesador digital de señales) a la que llaman realzado adaptable por correlación en tiempo real o RACE, por sus siglas en inglés<sup>12,18</sup>. RACE es básicamente un filtro adaptable de respuesta al impulso finita (FIR) que actualiza sus coeficientes de acuerdo a la estimación recursiva de la correlación de la señal de entrada según la expresión<sup>1</sup>.

$$\hat{R}_{xx}(n, l) = \beta R_{xx}(n-1, l) + (1-\beta)x(n)x(n+l) \quad (1)$$

Donde  $\hat{R}_{xx}(n, l)$  y  $R_{xx}(n-1, l)$  son los estimados de la autocorrelación de la entrada  $x(n)$  con el valor de retraso  $l$  en los instantes de tiempo  $n$  y  $n-1$  respectivamente;  $\beta$  es una constante de suavizado la cual toma su valor en el intervalo  $(0, 1)$ . Si la constante de tiempo  $\tau$  se define como el tiempo que le toma a la amplitud caer a  $1/e$  o 37% de su valor inicial  $(1-\beta)$ , entonces

$$\tau = \frac{1}{(1-\beta) f_s} \quad (2)$$

Los coeficientes de la correlación son estimados para valores de retraso  $l$  que varían entre  $-L$  y  $+L$  resultando en un filtro de  $2L+1$  coeficientes; típicamente se elige un valor de  $L$  entre 4 y 7. Los valores de  $\beta$  y  $L$  deben ser elegidos para no exceder la suposición de estacionariedad a corto plazo de la voz. Para los resultados que se muestran en el presente trabajo, los parámetros  $\beta$  y  $L$  fueron de  $\beta=0.99$  y  $L=7$ .

## SUSTRACCIÓN ESPECTRAL

La idea principal de la sustracción espectral es estimar el ruido ambiental y posteriormente sustraer dicho estimado espectral de la señal ruidosa para así mejorar la componente de voz<sup>13,19,20</sup>. Las características del ruido se actualizan durante segmentos donde no hay voz, por lo tanto, se requiere un detector de actividad de voz<sup>21</sup> (VAD por sus siglas en inglés) para determinar secuencias de voz con ruido y aquellas compuestas únicamente por ruido.

Supóngase que una señal de voz  $s(n)$  se contamina con ruido aditivo  $v(n)$ . Se asume que el ruido no está correlacionado con la voz y que no es estacionario. Si se aplica la transformación de Fourier (por segmentos), entonces:

$$Y(\omega_k) = S(\omega_k) + V(\omega_k) \tag{3}$$

Como  $s(n)$  y  $v(n)$  no están correlacionadas, el espectro de potencia está dado por la expresión:

$$|Y(\omega_k)|^2 = |S(\omega_k)|^2 + |V(\omega_k)|^2 \tag{4}$$

El algoritmo se basa en obtener un filtro de Wiener a partir del estimado del ruido  $\hat{V}(\omega_k)$ , el cuadrado de la función de transferencia del filtro está dada por la expresión<sup>19</sup>:

$$H^2(\omega_k) = \frac{|S(\omega_k)|^2 - |\hat{V}(\omega_k)|^2}{|S(\omega_k)|^2} \tag{5}$$

Al estimar  $\hat{V}(\omega_k)$  se pueden aplicar ponderaciones para acentuar la reducción de ruido en bandas no críticas para la voz.

El método utilizado para realizar la reducción de ruido por sustracción espectral es el siguiente:

1. Obtener la FFT por segmentos de la señal contaminada.
2. Estimar el espectro de potencia del ruido en periodos de inactividad de voz y ponderarlo.
3. Sustraer el estimado de ruido del espectro de potencia de la señal contaminada.
4. Realizar la rectificación de media onda: las partes del espectro que resulten negativas después de la sustracción son reemplazadas por una versión muy atenuada de la señal o con cero.
5. La fase de la señal ruidosa se combina con el resultado de la rectificación de media onda y se aplica la IFFT

**FILTRADO POR GSVD**

Maj et al proponen un esquema de reducción de ruido basado en la descomposición en valores singulares generalizada (GSVD) para auxiliares auditivos con micrófonos duales<sup>14</sup>.

Se plantea la recuperación de una señal  $\mathbf{s}_k$  de una mezcla ruidosa  $\mathbf{y}_k = \mathbf{s}_k + \mathbf{v}_k$  (siendo  $\mathbf{v}_k$  la componente de ruido) mediante la construcción de un filtro  $\mathbf{W}_{WF}$  de tal forma que  $\mathbf{s}_k = \mathbf{W}_{WF}^T \mathbf{y}_k$  en el instante  $k$ . Usando el criterio de error de mínimos cuadrados promedio (MSE) se tiene que el filtro se expresa como<sup>22</sup>:

$$\mathbf{W}_{WF} = E \{ \mathbf{y}_k \cdot \mathbf{y}_k^T \}^{-1} \cdot (E \{ \mathbf{y}_k \cdot \mathbf{y}_k^T \} - E \{ \mathbf{v}_k \cdot \mathbf{v}_k^T \}) \tag{6}$$

$\mathbf{W}_{WF}$  se deriva de la descomposición en valores singulares de las matrices  $\mathbf{Y}_k$  y  $\mathbf{V}_k$ . Por lo tanto las matrices de datos de periodos de voz con ruido y de ruido únicamente se definen como:

$$\mathbf{Y}[k] = \begin{bmatrix} \mathbf{y}_k^T \\ \mathbf{y}_{k+1}^T \\ \vdots \\ \mathbf{y}_{k+p-1}^T \end{bmatrix} \quad \mathbf{V}[k] = \begin{bmatrix} \mathbf{v}_k^T \\ \mathbf{v}_{k+1}^T \\ \vdots \\ \mathbf{v}_{k+p-1}^T \end{bmatrix} \tag{7}$$

Donde

$$\mathbf{y}_k = [\mathbf{y}_{1k}^T \quad \mathbf{y}_{2k}^T]^T \tag{8}$$

$$\mathbf{y}_{jk} = [y_j(k) \quad y_j(k-1) \quad \dots \quad y_j(k-N+1)]^T \tag{9}$$

La GSVD de las matrices  $\mathbf{Y}[k]$  y  $\mathbf{V}[k]$  resulta en

$$\begin{cases} \mathbf{Y}[k] = \mathbf{U}_Y \Sigma_Y \mathbf{Q}^T \\ \mathbf{V}[k] = \mathbf{U}_V \Sigma_V \mathbf{Q}^T \end{cases} \tag{10}$$

Donde  $\Sigma_Y = \text{diag} \{ \sigma_i \}$ ,  $\Sigma_V = \text{diag} \{ \eta_i \}$ ,  $\mathbf{U}_Y$  y  $\mathbf{U}_V$  son matrices ortogonales,  $\mathbf{Q}$  es una matriz invertible pero no necesariamente ortogonal, y los valores singulares generalizados están dados por  $\frac{\sigma_i}{\eta_i}$ . De tal forma se obtiene la expresión para el estimado de la matriz de filtros<sup>22</sup>:

$$\mathbf{W}_{WF} \cong \mathbf{Q}^{-T} \text{diag} \left\{ 1 - \frac{p \sigma_i^2}{q \eta_i^2} \right\} \mathbf{Q}^T \tag{11}$$

Para los resultados obtenidos con esta técnica de reducción de ruido, se utilizaron filtros de 15 coeficientes con lo que se obtiene una matriz  $\mathbf{W}_{WF}$  con 30 estimadores. Para una implementación en tiempo real la complejidad del cálculo de la matriz completa es muy alta, y sólo se requeriría una columna de ésta, así pues, se utilizó la columna central de  $\mathbf{W}_{WF}$  para todas las pruebas y así poder dar resultados que serían más cercanos a la implementación real.

## Análisis de componentes independientes

Supóngase que se tienen dos transductores fonocaptadores que registran dos señales en función del tiempo, denotadas por  $x_1(t)$  y  $x_2(t)$ . Bajo ciertas condiciones, cada uno de esos registros se puede considerar como una suma ponderada de las señales emitidas por las fuentes de sonido (la voz y el ruido). Esto se puede expresar con las ecuaciones lineales.

$$x_1(t) = a_{11}s_1(t) + a_{12}s_2(t) \quad (12a)$$

$$x_2(t) = a_{21}s_1(t) + a_{22}s_2(t) \quad (12b)$$

Donde  $a_{ij}$  son parámetros que dependen de las distancias de los transductores a las fuentes. Usando una notación matricial, el modelo de mezcla anterior se puede escribir como:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (13)$$

Así pues, se define el análisis de componentes independientes (ICA por sus siglas en inglés) de un vector aleatorio  $\mathbf{x}$  como la transformación invertible  $\mathbf{s} = \mathbf{W}\mathbf{x}$ , donde la matriz  $\mathbf{W}$  se determina de tal forma que sea minimizada la información mutua de las componentes transformadas  $s_i$ <sup>23,24</sup>.

Para el cálculo de ICA se requiere cierto preprocesamiento de la señal<sup>17</sup>. El primer paso es centrar a  $\mathbf{x}$ , esto es, sustraer la media  $\mathbf{m} = E\{\mathbf{x}\}$  tal que se haga a  $\mathbf{x}$  una variable de media cero. Una vez encontradas las componentes se puede añadir de nuevo la media.

Otra estrategia de preprocesamiento en ICA es blanquear previamente las variables observadas. Esto significa que se transforma linealmente a  $\mathbf{x}$  tal que se obtenga una  $\tilde{\mathbf{x}}$ , cuya matriz de covarianza equivale a la matriz identidad

$$E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T\} = \mathbf{I} \quad (14)$$

La transformación de blanqueado siempre es posible. Un método popular de blanqueado es usar la descomposición en eigenvalores de la matriz de covarianza  $E\{\mathbf{x}\mathbf{x}^T\} = \mathbf{E}\mathbf{D}\mathbf{E}^T$ , donde  $\mathbf{E}$  es la matriz ortogonal de eigenvectores de  $E\{\mathbf{x}\mathbf{x}^T\}$  y  $\mathbf{D}$  es la matriz diagonal de sus eigenvalores. El blanqueado entonces se realiza con

$$\tilde{\mathbf{x}} = \mathbf{E}\mathbf{D}^{-\frac{1}{2}}\mathbf{E}^T\mathbf{x} \quad (15)$$

Donde  $\mathbf{D}^{-\frac{1}{2}}$  se calcula realizando la operación por elemento.

Finalmente, para encontrar las columnas de la matriz  $\mathbf{W}$  de la transformación ICA, se aplica el siguiente algoritmo de punto fijo, propuesto por Hyvärinen y Oja<sup>23,24</sup>:

$$\begin{aligned} \mathbf{w}^+ &= E\{\mathbf{x}g(\mathbf{w}^T\mathbf{x})\} - E\{g'(\mathbf{w}^T\mathbf{x})\}\mathbf{w} \\ \mathbf{w}^* &= \mathbf{w}^+ / \|\mathbf{w}^+\| \end{aligned} \quad (16)$$

Donde la función  $g$ , llamada función de contraste, es prácticamente cualquier función no cuadrática, y  $g'$  es su derivada. En la práctica se puede elegir entre las funciones  $\tanh(u)$ ,  $\exp(u)$ , y  $u^3$ , que satisfacen criterios de simplicidad, consistencia y robustez<sup>23,24</sup>.

Puesto que en ocasiones la convergencia del método puede ser incierta, se puede introducir un parámetro de incremento  $\mu$ , obteniendo la versión estabilizada del algoritmo<sup>23,24</sup>.

$$\begin{aligned} \mathbf{w}^+ &= \mathbf{w} - \mu \frac{E\{\mathbf{x}g(\mathbf{w}^T\mathbf{x})\} - \beta\mathbf{w}}{E\{g'(\mathbf{w}^T\mathbf{x})\} - \beta} \\ \mathbf{w}^* &= \mathbf{w}^+ / \|\mathbf{w}^+\| \end{aligned} \quad (17)$$

Para prevenir que dos vectores columna de la matriz  $\mathbf{W}$  converjan al mismo máximo, se deben descorrelacionar las salidas  $\mathbf{w}_n^T\mathbf{x}$ . Para tal efecto se sustraen las proyecciones de vectores previamente estimados después de cada iteración del algoritmo, esto es:

$$\mathbf{w}_{p+1} = \mathbf{w}_{p+1} - \sum_{j=1}^p \mathbf{w}_{p+1}^T \mathbf{w}_j \mathbf{w}_j \quad (18)$$

## Conjunto de datos experimentales

El conjunto de registros de voz que se utilizó para realizar la evaluación consta de treinta grabaciones de voz de siete locutores diferentes. Quince de las grabaciones provienen de un locutor de género masculino y las quince restantes de un locutor de género femenino. Son archivos en formato WAV con codificación PCM lineal, una frecuencia de muestreo de 11025 Hz y 16 bits por muestra. Los archivos consisten de una oración inteligible en lenguaje español.



Se aplicaron las pruebas de normalidad de Lilliefors<sup>25</sup> y de Jarque-Bera<sup>26</sup> a los niveles de energía de los archivos y a la duración; en ambos casos, las pruebas arrojaron como resultado que no se puede rechazar la hipótesis de normalidad, esto indica que los elementos del conjunto de datos tiene duración y nivel de energía comparables. Se aplicó la prueba de Kolmogorov-Smirnov<sup>25</sup> para comparar las distribuciones entre los treinta registros y siempre que se compare con un registro distinto, se puede rechazar la hipótesis de que provienen de una misma distribución, esto indica que las evaluaciones no se hicieron sobre un mismo tipo de señal y por lo tanto que resultados similares al evaluar las técnicas no serán consecuencia de haber usado datos muy parecidos.

Para las evaluaciones se utilizaron cinco tipos de ruido, dos de ellos artificiales y tres grabados en situaciones que se podrían encontrar en la vida cotidiana de una persona.

Los ruidos artificiales son ruido blanco y ruido no estacionario de banda limitada, el ruido blanco fue generado con la función randn de MATLAB. Para generar el ruido no estacionario de banda limitada, se comienza generando ruido blanco, posteriormente se creó un filtro pasabandas de 400 a 4000 Hz y se filtró el ruido blanco; para hacerlo no estacionario con características similares a la voz, se tomó una grabación de voz ajena al conjunto de datos experimentales, se aplicó la transformación de Hilbert y su valor absoluto, que representa la amplitud instantánea de la señal y, se multiplicó por el ruido filtrado.

Los ruidos grabados son ruido de oficina, ruido de conversaciones indistintas en un lugar concurrido y ruido de un medio de transporte. El ruido de oficina consta del sonido de equipo de cómputo funcionando y de escritura en teclados y no incluye voces de personas. El ruido de conversaciones indistintas fue grabado en la sala 20 del Aeropuerto Internacional de la Ciudad de México. El ruido de un medio de transporte fue grabado en un vuelo comercial de una aeronave tipo MD-80.

### Detección de actividad de voz

Debido a que los detectores de actividad de voz más comunes no son confiables en condiciones donde la SNR es cercana a 0 dB<sup>21</sup>, y el objetivo del presente trabajo es comparar el desempeño de las técnicas de reducción de ruido, no del VAD, para las técnicas que requieren de un detector de actividad de voz, la señal de actividad de voz fue calculada a partir de los registros limpios de voz. La

señal de actividad de voz fue estimada utilizando un umbral de energía por segmentos de la señal; si en un segmento de 150 ms la energía de la señal caía 10 dB por debajo del nivel calculado de toda la señal, se consideró como un periodo de silencio.

### Ajuste de nivel en decibeles

Para ajustar una señal digital  $x(n)$  a cierto nivel deseado  $L$  en decibeles con referencia a la escala completa (dB FS) se utilizó la siguiente expresión para encontrar el escalar  $g$  por el cual se debe multiplicar la señal para que tenga el nivel deseado.

$$g = \left( \frac{N}{\sum x^2(n)} \cdot 10^{0.1L} \right)^{1/2} \quad (19)$$

### Relación señal a ruido

Para calcular la relación señal a ruido, se estimó la energía de las señales digitales por medio de la varianza, por lo tanto se calculó con la expresión

$$SNR = 10 \log_{10} \left( \frac{\sigma_S^2}{\sigma_R^2} \right) = 10 \log_{10} \left( \frac{\frac{1}{N_S} \sum_{n=1}^{N_S} s^2(n)}{\frac{1}{N_R} \sum_{n=1}^{N_R} r^2(n)} \right) \quad (20)$$

### Ponderación en escala A

Para los resultados que fueron ponderados en escala A, se utilizó el filtro con la siguiente función de transferencia:

$$H_A(s) = \frac{7.39705 \times 10^9 \cdot s^4}{(s+129.4)^2(s+676.7)(s+4636)(s+76655)^2} \quad (21)$$

### Metodología de evaluación en entorno de simulación

Para evaluar el desempeño de las técnicas de reducción de ruido se utilizó la siguiente metodología:

1. Se carga de un archivo la señal limpia de voz y se calcula el nivel en dB FS.
2. Se detecta la actividad de voz.
3. Se genera o se carga de un archivo con la señal de ruido.
4. Se ajusta el nivel del ruido con la expresión (19) de forma que se obtenga una SNR de -9 dB,

- 6 dB, -3 dB, 0dB, 3 dB y 6 dB en referencia al nivel de la señal de voz calculado previamente.
- Se hacen las mezclas aditivas de voz con ruido con las SNR previamente mencionadas, la mezcla se hace de la forma 0.5 (voz + ruido). En la mezcla se toma la media aritmética en vez de la suma para evitar que la señal salga del rango  $[-1, 1]$ . En el caso del filtrado por GSVD y de ICA, que requieren dos canales, el segundo canal está compuesto por una mezcla en la cual el ruido tiene 1 dB más que en el paso anterior y la voz se atenúa 1 dB, simulando que un micrófono está más orientado hacia la voz y otro más hacia el ruido.
  - Se obtiene una señal mejorada con las técnicas utilizando los archivos de mezcla de voz con ruido.
  - Se calcula la SNR de la señal mejorada de acuerdo a la expresión (20) tomando en cuenta que si se asume que en la señal mejorada se recuperó la componente de voz de la mezcla, se debe ajustar el nivel de la energía al nivel de ésta. Así pues la señal  $s$  en (20) es la componente de voz de la mezcla, esto es,  $0.5 \cdot \text{voz}$  y la señal de ruido residual  $r$  es por lo tanto la diferencia entre la señal mejorada y la componente de voz en la mezcla.
  - Se calcula la SNR ponderada en escala A de manera similar al paso anterior, con la única diferencia que las señales  $s$  y  $r$  son filtradas previamente. El filtrado se realiza multiplicando en el

dominio de la frecuencia la FFT de las señales con la función de transferencia denotada en la expresión (21).

- La mejora a la SNR que proporciona la técnica se obtiene como la diferencia entre la SNR de la señal mejorada y la SNR de la mezcla de voz con ruido.

Una vez que se calcularon las mejoras a la SNR de todos los registros, se realizan las pruebas de normalidad de Lilliefors y Jarque-Bera a los resultados. Las pruebas se aplican debido a que si los resultados forman una distribución normal, se puede afirmar con el mismo grado de confiabilidad de la prueba estadística que si se hacen experimentos con más registros de voz bajo las mismas condiciones, en un 66% la mejora será de  $\mu \pm \sigma$  y en un 99% de  $\mu \pm 3\sigma$ , donde  $\mu$  es la media de la distribución y  $\sigma$  su desviación estándar.

## RESULTADOS

En los cuadros 1 a 10 se pueden comparar los resultados obtenidos después de evaluar las técnicas de reducción de ruido en distintas situaciones. Con el propósito de condensar la información y que sea más sencillo estudiarla, en los cuadros sólo se muestran la media y desviación estándar de las mejoras obtenidas en cada experimento.

Para los resultados de ICA que se muestran en los cuadros 1 y 2, se omitieron los resultados de un

**Cuadro 1.** Estadísticas de las mejoras a la SNR obtenidas para los registros de voz contaminados con ruido blanco (dB).

	SNR -9dB		SNR -6dB		SNR -3dB		SNR 0dB		SNR +3dB		SNR +6dB	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
RACE	9.440	0.317	8.557	0.371	7.804	0.468	6.685	0.637	5.056	0.855	2.965	1.086
S. Esp.	4.863	0.170	4.677	0.176	4.438	0.184	4.143	0.198	3.786	0.218	3.357	0.258
GSVD	10.058	0.357	8.615	0.433	6.989	0.523	5.085	0.629	2.911	0.735	0.494	0.826
ICA	64.437	9.701	61.439	15.438	58.437	9.702	55.437	9.700	52.437	9.701	49.436	9.700

**Cuadro 2.** Estadísticas de las mejoras a la SNR ponderadas en la escala A obtenidas para los registros de voz contaminados con ruido blanco (dBA).

	SNR -9dB		SNR -6dB		SNR -3dB		SNR 0dB		SNR +3dB		SNR +6dB	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
RACE	6.098	0.627	5.516	0.439	5.188	0.339	4.222	0.518	2.396	0.747	-0.022	0.913
S. Esp.	0.174	1.034	0.066	1.013	-0.084	0.987	-0.286	0.956	-0.543	0.924	-0.863	0.890
GSVD	7.305	0.692	6.186	0.628	4.813	0.565	3.077	0.541	1.016	0.548	-1.294	0.571
ICA	59.667	10.006	56.669	10.012	53.667	10.007	50.667	10.006	47.667	10.006	44.666	10.005

registro de voz pues éstos eran muy altos (aproximadamente el doble del promedio) y se consideró como una anomalía estadística. Asimismo, las mejoras a la SNR de -9 dB, -6 dB y -3 dB ponderadas en escala A utilizando la técnica de filtrado por GSVD no pasaron la prueba de normalidad.

En el caso del ruido de oficina (Cuadro 3) los resultados para el filtrado por GSVD no pasaron la prueba de normalidad de Lilliefors con una SNR de -6 dB y -3dB y la de Jarque-Bera con una SNR de -9 dB y -6dB, en ambos casos al ponderar en escala A se puede asumir la distribución normal. Para la técnica de ICA se omitieron los resultados de otro registro, pues eran muy altos y con ellos no pasaban algunas pruebas de normalidad.

## DISCUSIÓN

Antes de obtener conclusiones sobre los resultados, es conveniente mencionar la complejidad computacional de las técnicas. Es un factor importante a considerar para tomar decisiones de diseño dado que sería de poca utilidad una técnica con un desempeño excepcional si no es factible una implementación en tiempo real con algún procesador disponible en la actualidad.

La complejidad computacional del filtro RACE es de  $O(n)$ . Cada actualización del filtro con  $T$  coeficientes requiere aproximadamente  $3T$  multiplicaciones y  $T$  adiciones, y la operación de filtrado requiere  $T$  multiplicaciones y  $T$  adiciones por cada elemento del vector de entrada, es decir, un número cercano a  $6TN$  operaciones.

La complejidad computacional de la sustracción espectral es de  $O(n \log n)$ . Si se toman segmentos cuya longitud  $N$  es una potencia de 2, la FFT puede ser calculada en la práctica con  $2N \log_2 N$  multiplicaciones de escalares y  $3N \log_2 N$  sumas de escalares, se requerirían aproximadamente  $16N$  corrimientos y  $16N$  adiciones para estimar el espectro de potencia y fase utilizando el algoritmo CORDIC<sup>27</sup>,  $N/2$  multiplicaciones para ponderar el estimado de

ruido,  $N/2$  sustracciones para filtrar el ruido, aproximadamente  $16N$  corrimientos y  $16N$  adiciones para convertir el espectro de potencia y fase a coordenadas rectangulares (una vez más utilizando el algoritmo CORDIC) y finalmente  $2N \log_2 N$  multiplicaciones y  $3N \log_3 N$  sumas para calcular la IFFT.

La complejidad computacional del filtrado por GSVD es de  $O(n^3)$ , y en casos especiales,  $O(n^2)$ . El cálculo de la GSVD de dos matrices de  $p \times M$  requiere aproximadamente  $17M^3 + 3pM^2$  operaciones<sup>28</sup>. Utilizando algoritmos de actualización de la GSVD el número se reduce a  $23.5M^2$  y con versiones que no utilizan raíces cuadradas, se puede reducir a  $17.5M^2$ . El cálculo de una sola columna de la matriz  $\mathbf{W}_{WF}$  tiene una complejidad computacional de orden  $4M^2$ .

El algoritmo de ICA tiene una complejidad de  $O(n)$ . Para un vector de datos de tamaño  $2 \times N$ , durante el blanqueado de los datos, el cálculo de la matriz de covarianza requiere  $3N$  sumas y  $3N$  multiplicaciones, al resultar una matriz de solamente  $2 \times 2$ , su descomposición en eigenvalores requiere aproximadamente 200 operaciones y el producto de los datos con la matriz de blanqueado requiere  $2N$  multiplicaciones y  $2N$  sumas. Cada iteración del algoritmo de ICA requiere dos multiplicaciones de la matriz de datos blanqueados por un vector de  $2 \times 1$ , es decir aproximadamente  $4N$  sumas y  $4N$  multiplicaciones; se estima un tope de 100 iteraciones para que el algoritmo converja, aunque para los experimentos realizados, el algoritmo convergía en menos de 15 iteraciones, esto se debe a la propiedad de convergencia cúbica del algoritmo, demostrada en (23).

Como se aprecia al comparar los Cuadros 1 y 2, con el ruido blanco las mejoras disminuyen al ponderarlos en escala A. Probablemente esto se deba a que el ruido blanco, ocupa toda la banda de frecuencias con la que se trabaja, y al ponderarlo en escala A ya no se toma en cuenta que se haya eliminado el ruido en una banda que no era crítica.

**Cuadro 3.** Estadísticas de las mejoras a la SNR obtenidas para los registros de voz contaminados con ruido no estacionario de banda limitada (dB).

	SNR -9dB		SNR -6dB		SNR -3dB		SNR 0dB		SNR +3dB		SNR +6dB	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
RACE	8.735	0.467	7.419	0.706	6.583	0.863	5.701	0.858	4.217	0.911	2.026	1.127
S. Esp.	2.596	0.673	2.534	0.652	2.445	0.623	2.323	0.588	2.150	0.546	1.910	0.495
GSVD	7.517	0.419	5.193	0.573	2.976	0.726	0.808	0.877	-1.489	1.054	-3.982	1.261
ICA	64.049	8.988	62.150	11.821	56.574	11.173	56.118	13.230	52.830	12.769	50.830	12.766



También se puede apreciar que la sustracción espectral no tiene un buen desempeño para este tipo de ruido, de hecho, al ponderarlo en escala A empeora la señal en vez de mejorarla. Tal vez la resolución ofrecida por ventanas cortas limite la capacidad de estimar adecuadamente el espectro del ruido blanco, pues éste, al ser estacionario no debería ser afectado por la detección de actividad de voz.

Cabe destacar que con ICA se recupera la señal a la misma calidad, sin importar la SNR de la mezcla con ruido, si se observan las columnas de la mejora promedio, decrecen 3 dB al igual que las mezclas. Esto también puede tener consecuencias negativas pues da indicios que es un enfoque "todo o nada", o se recupera la señal prácticamente limpia, o será incapaz de mejoras pequeñas.

En los Cuadros 3 y 4 se aprecia que los resultados son similares al ruido blanco aunque la no

estacionariedad del ruido afectó a los filtros adaptables y a la sustracción espectral. Asimismo se hace evidente que la técnica de sustracción espectral tiene un pobre desempeño con ruidos artificiales.

En el caso de los ruidos reales, como se puede observar en los Cuadros 5 a 10, las mejoras son mayores al ponderar en la escala A, probablemente debido a que en estos casos el ruido no ocupa toda la banda de frecuencias. También se aprecia que mejora el desempeño de la sustracción espectral, especialmente cuando la SNR no es tan adversa.

En los datos de las tablas se observa que la mejora que proporciona ICA para el ruido de oficina es la menor para esta técnica, sin embargo sigue estando por encima de las demás técnicas y de diversos umbrales de inteligibilidad.

En el caso de ruido de conversaciones, la sustracción espectral tiene un desempeño equipara-

**Cuadro 4.** Estadísticas de las mejoras a la SNR ponderadas en la escala A obtenidas para los registros de voz contaminados con ruido no estacionario de banda limitada (dBA).

	SNR -9dB		SNR -6dB		SNR -3dB		SNR 0dB		SNR +3dB		SNR +6dB	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
RACE	4.718	0.606	3.307	0.390	2.622	0.310	2.144	0.400	1.070	0.696	-0.927	1.096
S. Esp.	-2.407	1.188	-2.453	1.167	-2.521	1.140	-2.614	1.108	-2.742	1.071	-2.909	1.033
GSVD	4.052	0.879	1.778	0.881	-0.201	0.940	-2.006	1.006	-3.709	0.964	-5.553	1.024
ICA	59.060	8.962	57.161	11.951	51.585	11.290	51.129	13.332	47.841	12.890	45.841	12.852

**Cuadro 5.** Estadísticas de las mejoras a la SNR obtenidas para los registros de voz contaminados con ruido de oficina (dB).

	SNR -9dB		SNR -6dB		SNR -3dB		SNR 0dB		SNR +3dB		SNR +6dB	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
RACE	7.072	0.173	4.570	0.247	2.293	0.348	0.304	0.477	-1.391	0.625	-2.893	0.786
S. Esp.	5.957	0.524	5.662	0.516	5.288	0.525	4.857	0.550	4.361	0.601	3.787	0.622
GSVD	8.216	0.378	6.192	0.490	4.352	0.594	2.522	0.702	0.500	0.779	-1.833	0.857
ICA	39.205	2.097	38.158	3.009	38.353	3.109	37.506	3.657	36.351	4.996	35.165	5.910

**Cuadro 6.** Estadísticas de las mejoras a la SNR ponderadas en la escala A obtenidas para los registros de voz contaminados con ruido de oficina (dBA).

	SNR -9dB		SNR -6dB		SNR -3dB		SNR 0dB		SNR +3dB		SNR +6dB	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
RACE	9.821	0.231	7.256	0.347	4.825	0.510	2.493	0.719	0.157	0.946	-2.304	1.141
S. Esp.	9.862	1.038	9.465	1.022	9.006	1.006	8.500	0.984	7.929	0.951	7.256	0.895
GSVD	10.555	0.551	8.440	0.670	6.325	0.778	3.989	0.873	1.380	0.936	-1.460	0.994
ICA	38.375	1.436	38.039	1.792	38.080	1.960	37.764	2.092	37.146	3.308	36.474	3.724

**Cuadro 7.** Estadísticas de las mejoras a la SNR obtenidas para los registros de voz contaminados con ruido de conversaciones indistintas (dB).

	SNR -9dB		SNR -6dB		SNR -3dB		SNR 0dB		SNR +3dB		SNR +6dB	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
RACE	7.338	0.145	4.938	0.212	2.799	0.308	0.986	0.431	-0.497	0.571	-1.758	0.718
S. Esp.	5.454	1.091	5.235	1.038	4.975	0.973	4.665	0.897	4.289	0.797	3.838	0.705
GSVD	8.087	0.318	5.969	0.447	4.028	0.570	2.135	0.672	0.081	0.787	-2.237	0.919
ICA	45.861	3.049	45.234	3.599	44.222	4.339	42.918	5.136	40.444	7.052	39.478	6.823

**Cuadro 8.** Estadísticas de las mejoras a la SNR ponderadas en la escala A obtenidas para los registros de voz contaminados con ruido de conversaciones indistintas (dBA).

	SNR -9dB		SNR -6dB		SNR -3dB		SNR 0dB		SNR +3dB		SNR +6dB	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
RACE	10.156	0.300	7.785	0.405	5.599	0.544	3.549	0.709	1.495	0.873	-0.731	1.000
S. Esp.	9.010	1.297	8.693	1.280	8.338	1.270	7.939	1.270	7.481	1.265	6.947	1.267
GSVD	9.834	0.860	7.668	0.956	5.548	1.016	3.298	1.028	0.770	1.025	-1.976	1.051
ICA	45.732	2.292	45.495	2.745	45.068	3.318	44.452	3.955	42.911	5.527	42.459	5.424

**Cuadro 9.** Estadísticas de las mejoras a la SNR obtenidas para los registros de voz contaminados con ruido de un medio de transporte (dB).

	SNR -9dB		SNR -6dB		SNR -3dB		SNR 0dB		SNR +3dB		SNR +6dB	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
RACE	7.103	0.174	4.611	0.249	2.347	0.351	0.369	0.481	-1.317	0.630	-2.818	0.791
S. Esp.	6.680	0.751	6.270	0.700	5.789	0.658	5.253	0.614	4.666	0.574	4.009	0.522
GSVD	7.765	0.355	5.517	0.469	3.408	0.606	1.412	0.755	-0.645	0.848	-3.065	1.098
ICA	60.029	14.156	56.820	14.112	52.356	12.972	51.198	13.338	47.984	13.594	44.717	14.536

**Cuadro 10.** Estadísticas de las mejoras a la SNR ponderadas en la escala A obtenidas para los registros de voz contaminados con ruido de un medio de transporte (dBA).

	SNR -9dB		SNR -6dB		SNR -3dB		SNR 0dB		SNR +3dB		SNR +6dB	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
RACE	9.691	0.220	7.127	0.332	4.697	0.490	2.372	0.694	0.054	0.919	-2.379	1.118
S. Esp.	12.062	1.081	11.341	1.085	10.574	1.080	9.774	1.058	8.943	1.024	8.042	0.983
GSVD	9.870	0.481	7.511	0.618	5.238	0.719	2.850	0.835	0.203	0.970	-2.711	1.186
ICA	67.095	14.642	63.883	14.275	59.423	13.297	58.264	13.656	55.051	13.931	51.781	14.958

ble con los filtros adaptables, y es más robusta a la variación de SNR, pues cuando ésta es positiva, los filtros empeoran la señal en vez de mejorarla.

En los Cuadros 9 y 10 se aprecian los resultados para el caso del ruido de medio de transporte, éste es prácticamente estacionario y de banda limitada, tal vez esa sea la razón por la que todas las técnicas obtuvieron mejores resultados, en especial la sustracción espectral que es en esta situación en la que mejor se desempeñó.

De los resultados obtenidos sobresale la técnica de ICA por su buen desempeño y porque, haciendo ciertas consideraciones, su complejidad permitiría su implementación en un DSP. En las simulaciones incluso al incrementar el nivel de energía del ruido de tal forma que la SNR fuera de -15 dB, mientras que las otras tres técnicas no lograban mejorar la inteligibilidad, con ICA se obtenían resultados similares a los mostrados en los cuadros. Sin embargo, hay que hacer hincapié en los problemas de

acondicionamiento de la señal que hay que solventar antes de una implementación real. Las condiciones asumidas en (12) pueden ser inválidas debido a efectos no lineales de los transductores.

Tomando en cuenta esta última consideración, se realizó un experimento real donde se notó en el osciloscopio que los micrófonos introducen una distorsión en la fase de la señal que está en función de la frecuencia y a pesar de ser micrófonos de un mismo modelo, la función de transferencia es distinta.

Al observar esto, se simuló filtrando cada canal de la mezcla de voz y ruido con un filtro IIR pasabandas con la misma respuesta en amplitud pero respuesta de fase no lineal distinta. Al simular con las señales filtradas de la manera descrita anteriormente, ICA no logró mejoras evidentes.

Asimismo, tal vez sea conveniente señalar que el desempeño de los filtros adaptables no está tan alejado de uno en condiciones ideales. En las Figuras 1 y 2 se muestra la media de las mejoras obtenidas con un filtro LMS adaptable al cual se le proporcionó la señal limpia de voz como señal deseada (condiciones ideales).

Si bien la implementación de las técnicas de ICA o filtrado por GSVD en un auxiliar auditivo portátil sería una tarea difícil, su implementación en tiempo real en una computadora personal se simplifica bastante con bibliotecas de funciones como LAPACK, MKL o IT++, lo que permitiría un aparato auxiliar auditivo de infraestructura, esto es, que en lugares como una escuela, una oficina se puede utilizar una computadora personal para el procesamiento de señales necesario para adecuar la señal a la persona con pérdida auditiva conectando un audífono y micrófonos colocados en lugares especiales.

## CONCLUSIONES

En el presente trabajo se han evaluado y comparado cuatro técnicas de procesamiento digital de señales para reducir el ruido ambiental, el cual es el principal problema identificado por los usuarios de auxiliares auditivos. Dos de ellas son de baja complejidad y de hecho han sido implementadas en prototipos de auxiliares auditivos. Las otras dos utilizan un procesamiento más complejo y aunque se pueden implementar en tiempo real en una computadora personal, lo cual permitiría la construcción de los auxiliares auditivos de "infraestructura" que se han mencionado, su implementación en auxiliares portátiles se antoja más complicada.

Una de las ventajas que ofrecerían los auxiliares auditivos digitales, es que no se limitarían a implementar un solo algoritmo de reducción de ruido, sino que se podría incluir varios "programas" y seleccionar el que resulte más adecuado para circunstancias específicas. Probablemente con los datos obtenidos en este comparativo, se pueda hacer un estimado acerca de cuál tipo de técnica es más conveniente en un determinado escenario.

El algoritmo que obtuvo mejoras sustanciales a lo largo de las pruebas fue el análisis de componentes independientes, sin embargo, tal vez el principal de sus problemas es la alta sensibilidad al desfasamiento entre los dos canales de entrada. El requerimiento de linealidad en la mezcla de los componentes puede suponer problemas en ambientes específicos y tampoco hay garantía de que alguno de los dos componentes independientes que encuentra sea la señal de voz deseada.

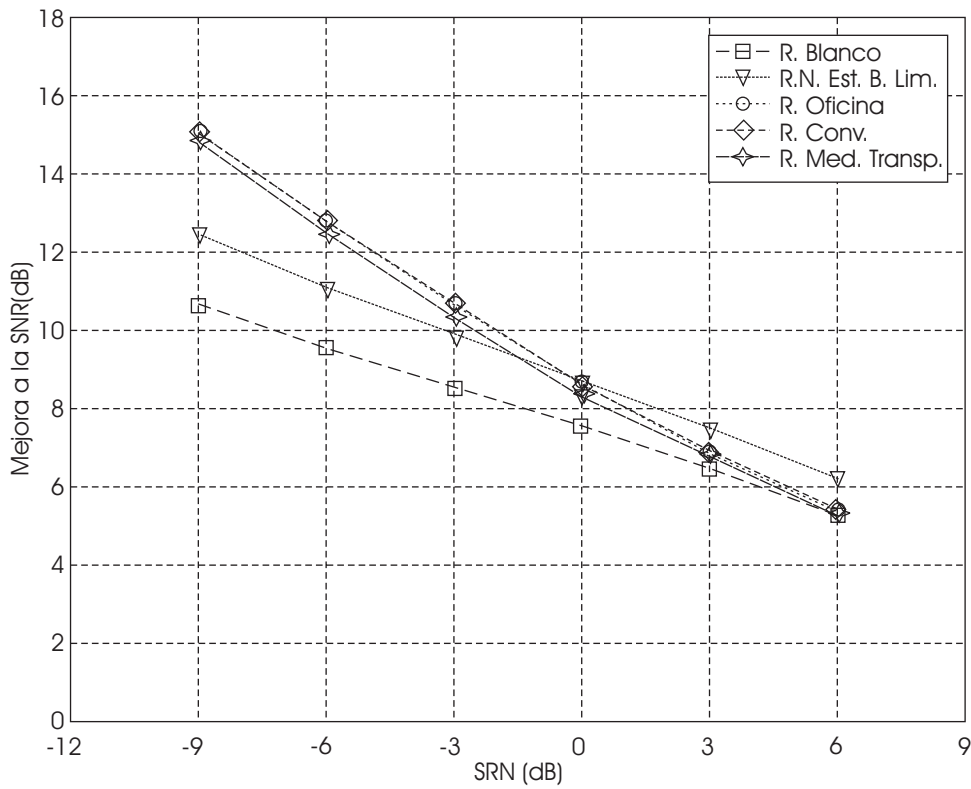
La técnica de filtrado por GSVD tiene un costo computacional demasiado alto para el beneficio obtenido. Cabe mencionar que la mejora de esta técnica es muy evidente en una situación donde el nivel de ruido es igual en ambos canales y el nivel de voz es diferente, sin embargo, en dichas circunstancias la reducción de ruido se podría lograr con la diferencia de los canales y un posterior ajuste de ganancia.

La sustracción espectral tiene un desempeño aceptable únicamente cuando la SNR es positiva o cercana a cero, si además de esto se considera la necesidad de un detector de actividad de voz robusto, pone en duda el beneficio real que pudiera proporcionar en una implementación real.

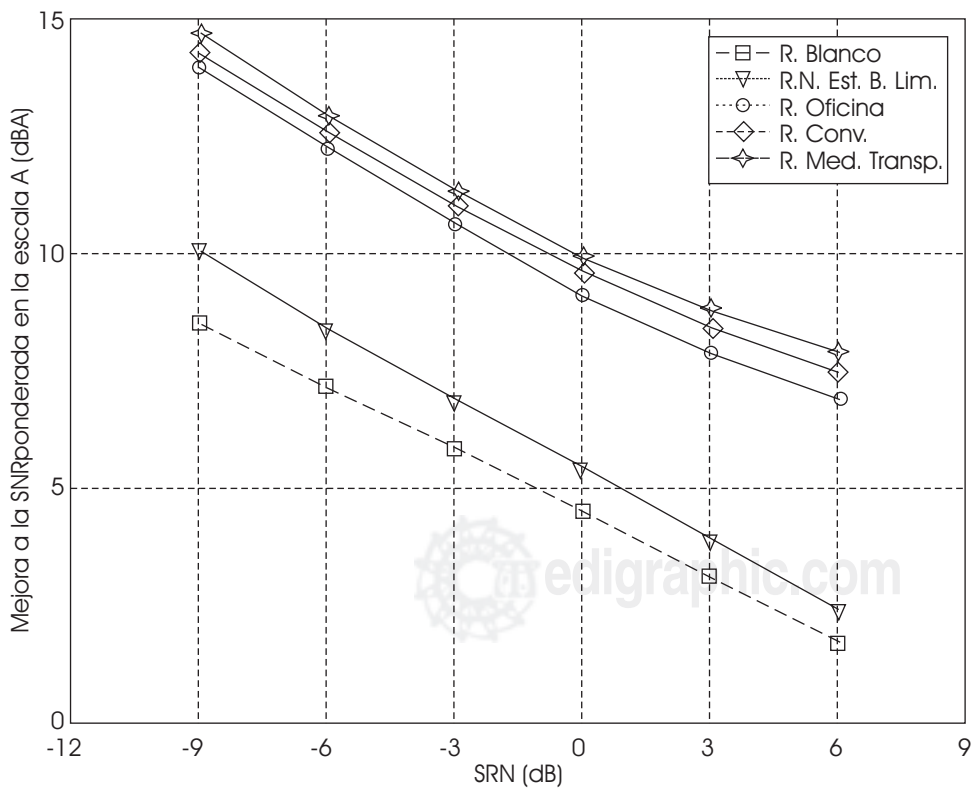
Es importante destacar que el filtro RACE aunque no logró resultados notables, tuvo un desempeño bastante consistente a lo largo de las pruebas y si se considera su sencillez esquemática y su baja complejidad computacional, lo hace un buen candidato como un algoritmo de "respaldo" para técnicas más complicadas, ya sea que otra técnica se tope con condiciones adversas o bien en ambientes donde el ruido no es tan desfavorable y una técnica sencilla podría alargar la duración de la batería.

De las observaciones anteriores se pueden sugerir líneas de investigación del trabajo a futuro. La primera y tal vez la más obvia, es la implementación en tiempo real del algoritmo de análisis de componentes independientes en un DSP u otro tipo de procesador embebido.

Otra línea puede ser tipificar las características y posicionamiento de los transductores fonocaptadores



**Figura 1.** Media de la mejora a la SNR obtenida con un filtro adaptable LMS.



**Figura 2.** Media de la mejora a la SNR ponderada en la escala A obtenida con un filtro adaptable LMS.

para satisfacer las condiciones necesarias para la convergencia del algoritmo de ICA, asimismo, para complementar esta última propuesta, una línea paralela de investigación es encontrar técnicas de preprocesamiento de la señal que compensen condiciones adversas al algoritmo.

## REFERENCIAS

1. Fitzgerald RG, Parkes CM. Blindness and Loss of Other Sensory and Cognitive Functions. *BMJ* 1998; 316(4): 1160–1163.
2. Choi YB, Krause JS, Capitan KE. The Applications of Human Factors Associated with Hearing Impairments: Issues and Recent Technological Developments in Telecommunications IEEE Proceedings of the 38th Hawaii International Conference on System Sciences 2005; 1–10.
3. Instituto Nacional de Estadística Geografía e Informática. XII Censo General de Población y Vivienda. Base de Datos del INEGI, Tabulador de Discapacidades 2000.
4. Raina P, Wong M, Massfeller H. The Relationship between Sensory Impairment and Functional Independence among Elderly. *BMC Geriatrics* 2004; 4(3).
5. Engebretson AM. Benefits of Digital Hearing Aids. *IEEE Engineering in Medicine and Biology Magazine* 1994; 13(2): 238–248.
6. Wouters J, Geurts L, Laneau J, Luts H, Maj JB, Moonen M et al. Digital Hearing Aids and Future Directions for Hearing Aids. *Acta Oto-Rhino-Laryngologica Belg* 2002; 56: 357–361.
7. Levitt H. Processing of Speech Signals for Physical and Sensory Disabilities. *Proc Natl Acad Sci USA* 1995; 92(10): 9999–10006.
8. Schneider T, Brennan R. A Multichannel Compression Strategy for a Digital Hearing Aid. *IEEE Proceedings of the 1997 International Conference on Acoustics, Speech, and Signal Processing ICASSP '97, 1997*; 411–414.
9. Gao R, Liu Y, Basseas S, Tsoukalas LH. Neurofuzzy Approaches for Advanced Hearing Devices. *11th IEEE International Conference on Tools with Artificial Intelligence* 1999; 327.
10. Radhakrishnan A, Viswanathan V, Gao R, Tsoukalas LH, Basseas S. New Generation Intelligent Hearing Prosthetics. *IEEE Proceedings of the International Symposium on Bio-Informatics and Biomedical Engineering BIBE '00, 2000*; 270.
11. Levitt H. Noise Reduction in Hearing Aids: A Review. *Journal of Rehabilitation Research and Development* 2001; 38(1): 111–121.
12. Magotra N, Hamill T, Swartz B. Digital Signal Processing of Speech for the Hearing-Impaired. *IEEE Proceedings of the Conference Record of the Twenty-Ninth Conference on Signals, Systems and Computers ASILOMAR '95, 1995*; 1209–1213.
13. Itoh K, Mizushima M. Environmental Noise Reduction Based on Speech/Non-Speech Identification for Hearing Aids. *IEEE Proceedings of the 1997 International Conference on Acoustics, Speech, and Signal Processing ICASSP '97, 1997*; 419–422.
14. Maj JB, Royackers L, Moonen M, Wouters J. SVD-Based Optimal Filtering for Noise Reduction in dual Microphone Hearing Aids: A Real Time Implementation and Perceptual Evaluation. *IEEE Transactions on Biomedical Engineering* 2005; 52(9): 1563–1573.
15. Sheller L. (Ed.) *Bernafo AG Audiology Insight*, 2005; 1(3).
16. Carnegie Securities Research, *Entering the Unit Growth Era. Carnegie's 6th Annual Hearing Aid Industry Report*, 2004.
17. Hyvärinen A, Oja E. Independent Component Analysis: Algorithms and Applications. *Neural Networks* 2000; 13(4): 411–430.
18. Stetzler GT, Magotra N, Gelabert P, Kasthuri P, Bangalore S. Low-Power Real-Time Programmable DSP Development Platform for Digital Hearing Aids. *Texas Instruments Application Report SPRA657, 2000*.
19. Sovka P, Pollak P, Kybic J. Extended Spectral Subtraction. *Proceedings of the European Signal Processing Conference EUSIPCO 96, Trieste, Italy 1996*; 963–966.
20. Malca Y, Wulich D. Improved Spectra Subtraction for Speech Enhancement. *Proceedings of the European Signal Processing Conference EUSIPCO 96, Trieste, Italy 1996*; 975–978.
21. Tanyer SG, Özer H. Voice Activity Detection in Nonstationary Noise. *IEEE Transactions on Speech Audio Processing* 2000; 8(4): 478–482.
22. Doclo S, Moonen M. GSVD-Based Optimal Filtering for Single and Multi-Microphone Speech Enhancement. *IEEE Transactions on Signal Processing* 2002; 50(9): 2230–2244.
23. Hyvärinen A. Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. *IEEE Transactions on Neural Networks* 1999; 10(3): 626–634.
24. Hyvärinen A, Oja E. A Fast Fixed-Point Algorithm for Independent Component Analysis. *Neural Computation* 1997; 9(7): 1483–1492.
25. Conover WJ. *Practical Nonparametric Statistics*. 3rd Ed. Wiley, New York, USA, 1998.
26. Judge GG, Hill RC, Griffiths WE, Lütkepohl H, Lee TC. *Introduction to the Theory and Practice of Econometrics*. 2nd Ed. Wiley, New York, USA, 1988.
27. Andraka R. A Survey of CORDIC Algorithms for FPGAs. *Proceedings of the ACM/SIGDA 6th Intl. Symposium on Field Programmable Gate Arrays FPGA '98, Monterey, CA, 1998*; 191–200.
28. Golub GH, Van Loan CF. *Matrix Computations* 3rd Ed. Johns Hopkins University Press, Baltimore, Maryland, USA, 1996.