

# TRATAMIENTO DE LA SEÑAL DE VOZ USANDO FILTROS ADAPTATIVO EN SUBBANDAS EN AMBIENTES ALTAMENTE RUIDOSOS

Miguel Ángel Castillo

Ing. Paul Bernal

Ing. Daniel Altamirano

DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA  
UNIVERSIDAD DE LAS FUERZAS ARMADAS - ESPE

**Resumen**—El presente artículo propone un sistema para el mejoramiento de la señal de voz, con el fin de disminuir el ruido durante el chequeo e inspección de una aeronave, dado que impide la correcta comunicación por radio entre los operadores y los pilotos. Para ello, se utiliza el algoritmo adaptativo en subbandas (*SAF*) y la herramienta MatLab® Versión 8, a fin de acelerar la velocidad de convergencia, reducir el costo computacional y para superar la desventaja de trabajar en todo el espectro de banda de la señal, especialmente en aplicaciones donde necesitan una mayor longitud del filtro como el habla. Este artículo describe el algoritmo adaptativo, la división de número de bandas, el tamaño del filtro adecuado y el error cuadrático mínimo para la aplicación.

## I. INTRODUCCIÓN

El ruido como un factor de riesgo físico presente en aeropuertos, trae consigo un trabajo molesto, doloroso y estresante para el trabajador, cuyos efectos, se verán reflejados en la incorrecta comunicación entre los miembros expuestos al mismo [1].

En ambiente aeronáutico, el trabajo que realizan los operadores, es el chequeo e inspección de la aeronave antes de cada vuelo. Esta inspección, donde la potencia de los motores va desde un 80% hasta el 100%, generan altas fuentes de ruido ocasionando, una incomodidad en la comunicación verbal durante el chequeo [2] y en ocasiones accidentes lamentables.

Los filtros adaptativos surgen para mejorar el tratamiento de señales no estacionarias como la voz y el ruido en aviones, su procesamiento, ayudará a extraer la

información deseada cuando ésta se encuentre en ambientes muy ruidosos.

Con el fin de minimizar la complejidad de trabajar en todo el espectro de la señal de voz, se propone el filtrado adaptativo con partición de la señal en bandas de bajo orden *SAF*, y así tener una alternativa de solución a la cancelación de ruido durante la comunicación por radio entre los operadores y los pilotos.

## II. CONCEPTOS BÁSICOS

### A. FILTRADO ADAPTATIVO Y EL MEJORAMIENTO DE LA SEÑAL

El objetivo del mejoramiento de la señal es mitigar las señales indeseadas, que impiden que la señal original sea audible o entendible al final de su procesamiento, estas señales indeseadas conocidas como ruido, impiden que la comunicación por radio entre los operadores de vuelo y los pilotos sea confiable y libre de interferencia. En la Figura 1, se puede visualizar el diagrama para el mejoramiento de la señal donde; la señal de voz  $x(k)$  corrompida por el ruido de aeronave  $n(k)$  ingresa por el canal primario, y el mismo ruido  $n(k)$  ingresa por el canal secundario. Los coeficientes adaptativos  $w(k)$  de la estructura interna del bloque del filtrado adaptativo, convergen a la función de transferencia diferencial, resultando una señal de salida  $y(k)$  del cual es una estimación del ruido  $n(k)$  presente en la señal deseada  $d(k)$ , por lo tanto, la señal de error  $e(k)$  se aproxima a la señal de voz  $x(k)$  [3]. Los algoritmos adaptativos LMS (*Least Mean Square*), NLMS (*Normalized Least*

---

Paul Bernal e-mail: cpbernal@espe.edu.ec

Daniel Altamirano e-mail: cdaltamirano@espe.edu.ec

Miguel Ángel Castillo e-mail: macastillo8@espe.edu.ec

Mean Square) y RLS (Recursive Least Square), parte del filtrado adaptativo, permiten variar sus parámetros ajustables o paso de adaptación  $\mu$  a fin de mejorar la velocidad de convergencia y el rendimiento al final del sistema [4].

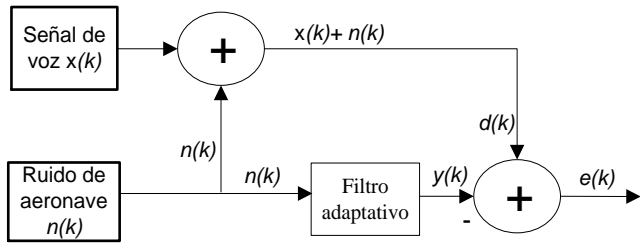


Figura 1: Estructura para el mejoramiento de la señal

### B. FILTRADO ADAPTATIVO EN SUBBANDAS (SAF)

SAF de la Figura 2, divide a la señal de entrada  $x(k)$  y la señal deseada  $d(k)$  en N-bandas mediante banco de filtros llamados de análisis  $F_i(z)$ , para posteriormente, filtrar las señales usando un filtro adaptativo en cada subbanda, donde finalmente reconstruyen la señal usando banco de filtros llamados de síntesis  $G_i(z)$ , todo esto, con el fin de disminuir la complejidad de trabajar en todo el espectro, reducir el costo computacional y para mejorar la velocidad de convergencia, parámetros que el filtrado adaptativo convencional no brinda.

Los bancos de filtros del tipo Pseudo-QMF o coseno modulado (Quadrature Mirror Filter) y DFT o modulación compleja (Discrete Fourier Transform), son usados en SAF para manipular, descomponer y reconstruir la información en cada subbanda dividida a fin de modificar el número de muestras presentes en la señal y reducir errores de procesamiento. A este proceso se integra el factor de diezmado e interpolado  $L$  con la finalidad de reducir el efecto *aliasing* causado por la reducción del muestreo y el efecto *band-edge* producido por la expansión de la tasa de muestreo.

El orden del diezmado  $L$  y el número de bandas  $N$  a dividir en los bancos de filtros, obedece a tres esquemas: Cuando  $L=N$  se habla de un filtro máximamente decimado, es decir, no hay pérdida de información, cuando  $L>N$  la pérdida de información debido al efecto *aliasing* se verá reflejado al momento de reconstruir la señal de origen, y cuando  $L<N$ , se dice que es un banco de filtro no máximamente decimado o sobremuestreado, es decir, evitará el solapamiento del espectro en la salida del filtro de análisis.

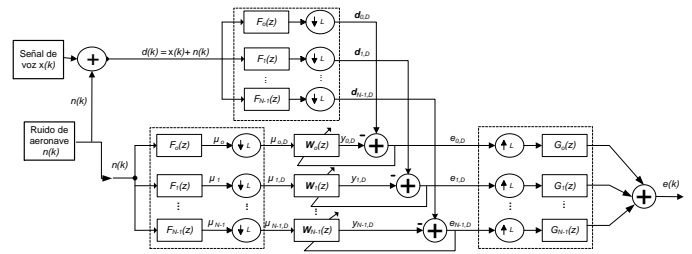


Figura 2: Diagrama de bloques para la cancelación de ruido usando SAF

Actualmente, SAF abre campo al desarrollo de nuevas aplicaciones contribuyendo a la innovación, investigación y desarrollo, todo con el fin de reducir el MSE (Mean Squared Error) y acelerar la velocidad de convergencia de los algoritmos [5].

### III. DESARROLLO SAF

Para el desarrollo SAF, se inició con las mediciones de los niveles de potencia pico de ruido de las aeronaves militares usando el equipo sonómetro SoundPro SE/DL [6], estas mediciones permitieron tener una visión real de los niveles de ruido en que está expuesto un operador.

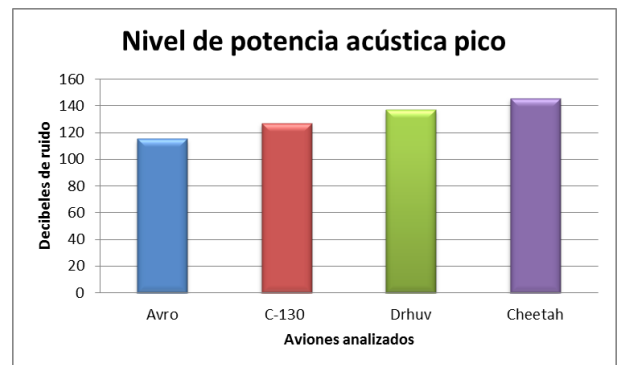


Figura 3: Medición de los niveles de ruido en las aeronaves

En la Figura 3, se visualiza las cuatro aeronaves analizadas (avión Avro, C-130, helicóptero Drhuv y el avión supersónico cheetah) de los cuales, exceden los 110 dB de ruido, ocasionando que la comunicación de voz entre el emisor y el receptor sea errada y desconfiable.

Posterior a las mediciones, se estructuró el filtrado adaptativo de acuerdo a la Figura 1. Sin embargo, para mejorar el desempeño del sistema atacando a las bandas de interés, se estructuró el filtrado adaptativo en subbandas conforme a la Figura 2.

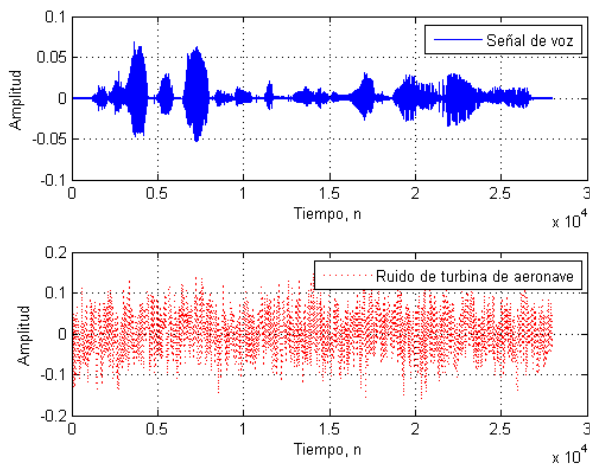


Figura 4: Señal de voz y ruido de aeronave

En la Figura 4, se visualiza la señal de voz de una frase grabada considerada la peor condición y el ruido de una aeronave similar a la de un avión supersónico. Señales que se tomaron como ejemplo para la aplicación de la cancelación de ruido

#### IV. PRUEBAS Y RESULTADOS

Para las pruebas de cancelación de ruido usando *SAF*, se inició con el análisis del comportamiento de los algoritmos LMS, NLMS y RLS, del cual, se resume que el error  $e(k)$  generado en cada simulación del sistema, está estrechamente relacionado con el tamaño de paso  $\mu$ , por tal motivo, de su valor dependerá el tiempo de convergencia del sistema. En conclusión, el algoritmo RLS converge más rápido que LMS y NLMS pero es inestable al aumentar el número de iteraciones, por lo que no fue tomado en cuenta en el desarrollo de la aplicación. Por tal motivo se utilizaron los dos algoritmos restantes.

Posterior al análisis de los algoritmos adaptativos, se estructuró el esquema del filtrado adaptativo de la Figura 1. Para ello, se trabajó con tres archivos (una letra, una instrucción acorde al entorno laboral y una frase) de voz grabados con extensión .wav; con una frecuencia de muestreo de 8 KHz y con una atenuación de SNR (*Signal to noise ratio*) de -40 dB. Cada señal fue corrompida con ruido AWGN (Additive white Gaussian noise) y ruido de una aeronave similar a la de un avión supersónico adquirido para la simulación.

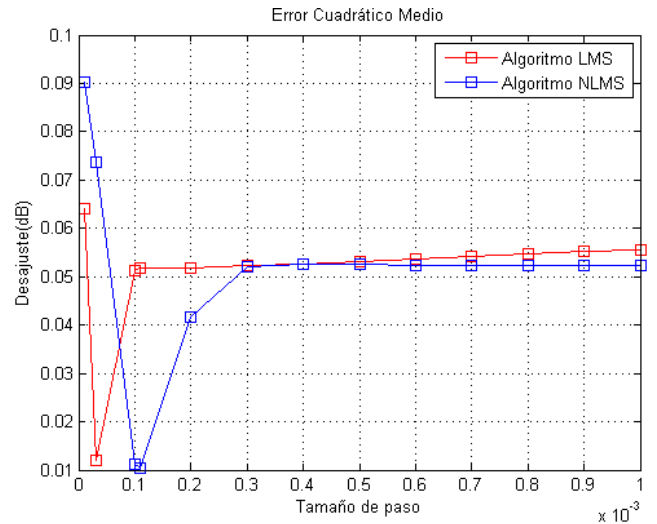


Figura 5: Filtrado de la señal de audio y ruido AWGN

El desajuste o error de estimación es una medida de desempeño del sistema, del cual mide la diferencia entre el valor esperado y el valor obtenido de acuerdo al algoritmo usado. Por tanto, en el filtrado de la señal de voz con ruido AWGN de la Figura 5, se evaluó que con un tamaño de paso pequeño el desajuste disminuye, al aumentar el tamaño de paso tanto para el algoritmo LMS y NLMS no habrá mejoría del MSE. En la Figura 6, se obtuvo datos similares usando el filtrado de la señal de voz con ruido de aeronave.

En resumen, para la comunicación por radio durante el chequeo de una aeronave se recomienda trabajar entre  $0,0001 < \mu < 0,008$  para el LMS y  $0,001 < \mu < 0,008$  para el NLMS. Los valores obtenidos están dentro de los parámetros teóricos recomendados [7], es decir,  $0 < \mu < 0.00059$  para el algoritmo LMS y entre  $0 < \mu < 2$  para el algoritmo NLMS.

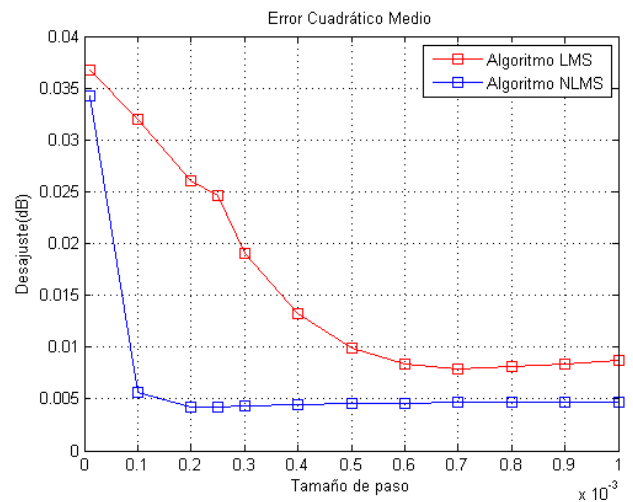


Figura 6: Filtrado de la señal de audio y ruido de aeronave

Una vez fijado el paso de adaptación  $\mu$  óptimo, se estructuró al filtrado adaptativo con partición en bandas de bajo orden denominados *SAF* de la Figura 2. Para aquello se trabajó con tres sistemas: *SAF* con banco de filtros de tipo DFT, y de tipo coseno modulado sobremuestreado [7] y *SAF* con banco de filtros de tipo coseno modulado críticamente muestreado [8]. Las pruebas fueron encaminadas a fin de obtener el número óptimo de bandas ( $N=4$ ,  $N=8$ ,  $N=16$  y  $N=64$ ) a dividir y el tamaño del filtro ( $M=32$ ,  $M=128$  y  $M=256$ ) adecuado para la cancelación de ruido usando el algoritmo LMS y NLMS.

En la Figura 7 y Figura 8, se puede apreciar las tres curvas de aprendizaje con  $N$ -bandas ( $N=4$ ,  $N=8$ ,  $N=16$  y  $N=64$ ) y tamaño de paso  $M=32$  del algoritmo LMS y NLMS. Como se puede observar, la curva de aprendizaje de *SAF* con Pseudo-QMF críticamente muestreado converge rápidamente pero a un nivel más bajo del estado de equilibrio (-20 dB) debido a la presencia del efecto *aliasing* en los banco de filtros, por tanto no está modelando a la aplicación de cancelación de ruido de los aviones.

*SAF* con DFT y Pseudo-QMF coseno modulado sobremuestreado, convergen en el estado de equilibrio aunque su convergencia fue lenta, en especial cuando se aumenta el número de bandas. También se demostró que *SAF* con banco de filtros de tipo DFT garantizaron una mayor estabilidad y menor MSE que *SAF* de tipo coseno modulado, excepto cuando se dividió en 64 bandas, donde el algoritmo no converge.

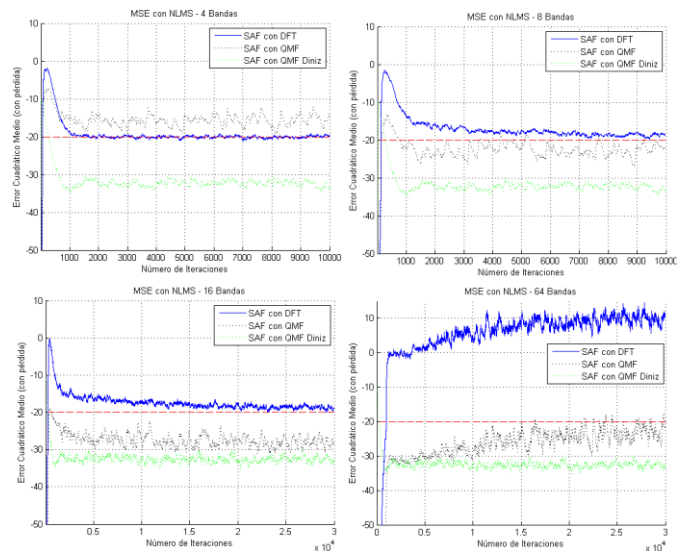


Figura 8: MSE de *SAF* NLMS con 4, 8, 16 y 64 subbandas

En relación al desajuste producido por la variación del número de bandas y el tamaño del filtro de los algoritmos LMS y NLMS, en la Figura 9 y Figura 10 se resume que; se puede garantizar un desajuste mínimo de la señal de voz cuando se tiene una división de  $N=4$  y  $N=8$  bandas con  $M=32$  para *SAF* DFT sobremuestreado. Si se aumenta el número de bandas y el tamaño del filtro no va a existir una mejora del audio al final del sistema *SAF*, excepto cuando se tenga una división de  $N=16$  y  $M=256$  con *SAF* NLMS, pero esto generaría un alto costo computacional.

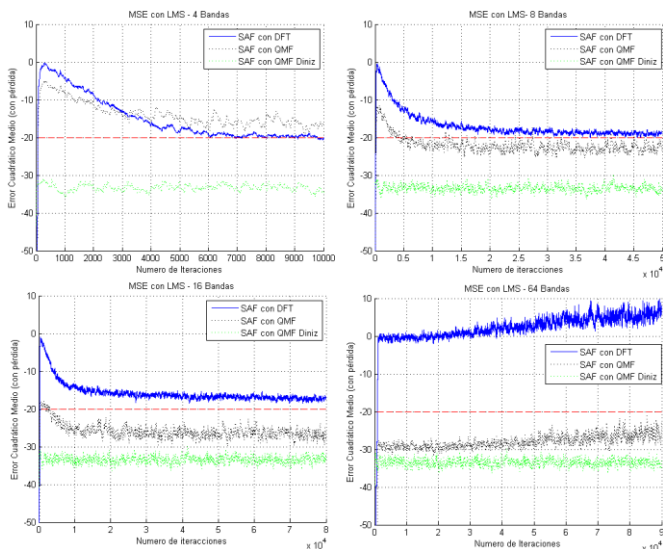


Figura 7: MSE de *SAF* LMS con 4, 8, 16 y 64 subbandas

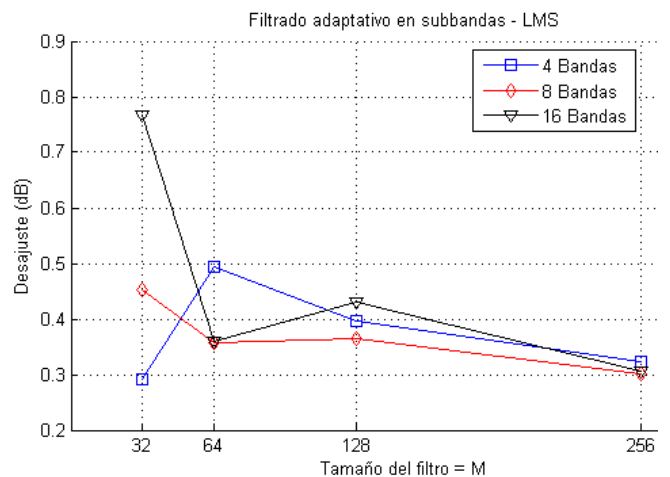


Figura 9: Desajuste (dB) en cada subbandas divididas

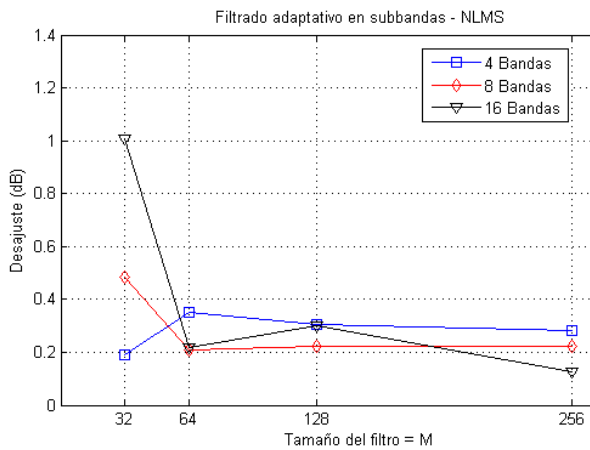


Figura 10: Desajuste (dB) en cada subbandas divididas

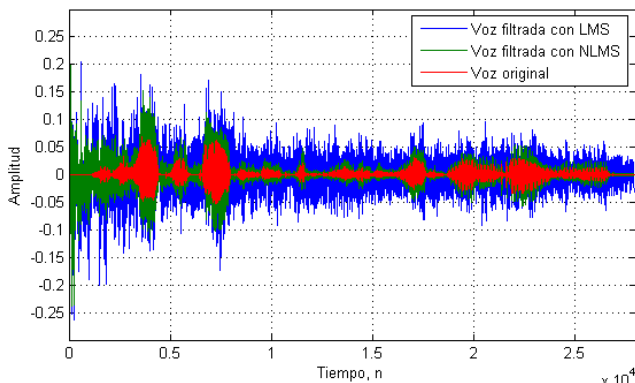


Figura 11: Señal de error del LMS y NLMS usando SAF

En la Figura 11, se verifica el resultado de la voz filtrada con el algoritmo LMS y NLMS después del filtrado SAF DFT sobremuestreado. Como se observa, el algoritmo NLMS a diferencia del algoritmo LMS, trata de parecerse lo más posible a la señal original. No es posible limpiar a la señal en su totalidad, sin embargo, tanto en la Figura como al escuchar, se aprecia la cancelación de ruido en aviones.

A fin de evaluar el desempeño de manera audible al sistema SAF DFT sobremuestreado, se pidió a los operadores de vuelo que califiquen de acuerdo a su subjetividad el grado de entendimiento del sistema SAF DFT sobremuestreado LMS y SAF DFT sobremuestreado NLMS en relación al número de bandas N.

En la Figura 12, el 90 % del personal que labora en el chequeo de la aeronave indicaron que la señal de audio fue comprensible cuando se usa SAF DTF sobremuestreado LMS con N=4 y N=8, aumentando el número de subbandas, el mensaje comenzó a distorsionarse.

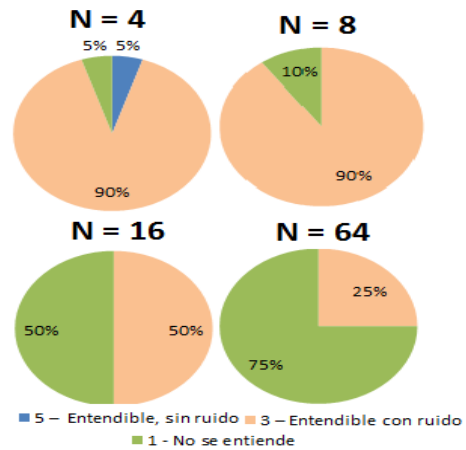


Figura 12: Pruebas de campo variando el número de subbandas con LMS

De la Figura 13 usando SAF DFT sobremuestreado NLMS, se resume que el 95% del personal que trabaja en la inspección de la aeronave indicaron que la señal de voz es comprensible cuando se usó SAF DFT NLMS sobremuestreado con N=4 y N=8 bandas, mientras más se aumente el número de bandas, el mensaje se distorsiona y no es entendible. En resumen, se recomienda usar SAF con el algoritmo NLMS dado que garantiza un mejor entendimiento del mensaje al final del sistema.

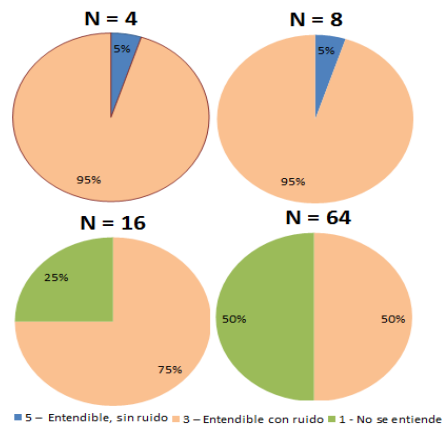


Figura 13: Pruebas de campo variando el número de subbandas con NLMS

La evaluación de desempeño cuantitativa fue realizado con la medición de distorsión espectral mbsd.m [9], donde, se cargó la señal de voz original y la señal de voz resultante después del procesamiento de SAF NLMS a una frecuencia de 4 KHz y con un tamaño de 20 ms. Se ajustó el retardo producido debido al procesamiento del sistema, al final, se obtuvo una medida de SNR en dB,



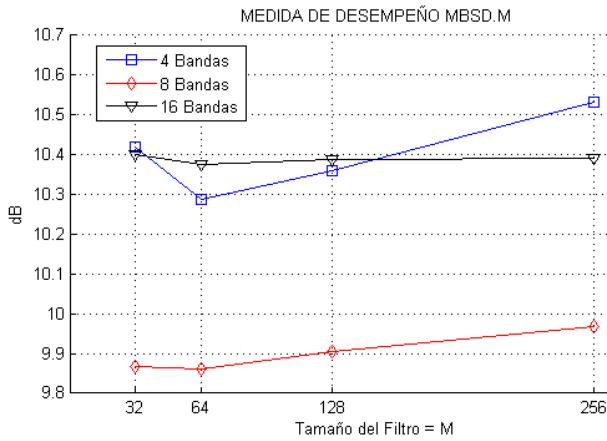


Figura 14: Medición cuantitativa del procesamiento SAF

Como se puede visualizar en la Figura 14, se garantiza una ganancia aproximada de 10,4 dB con N=4 y N=16 bandas, por consiguiente, se valida y se recomienda utilizar el algoritmo SAF DFT sobremuestreado NLMS para la cancelación de ruido en aviones con una división mínima de bandas.

## V. CONCLUSIONES

Se verificó los niveles de ruido en el ambiente real, valores que permitieron tener una perspectiva de la potencia de ruido en que está expuesto un operador y el piloto.

Para la cancelación de ruido usando SAF, se recomienda usar el algoritmo NLMS con un paso de adaptación de  $0,001 < \mu < 0,008$  dado que garantiza una rápida convergencia y una mejor reducción del error cuadrático, estos valores son eficiente para la presente aplicación y pueden ser ineficientes para otras aplicaciones.

Los bancos de filtros con DFT en concordancia con el proceso de diezmado e interpolado, permitieron mejorar la carga computacional y reducir el efecto *aliasing* presente durante el procesamiento de la señal, el factor de decimación e interpolación  $L \leq N$  (no máximamente decimado o sobremuestreado) permitió estabilizarse en el estado de equilibrio de -20 dB garantizando un modelamiento del sistema acorde a la aplicación de cancelación de ruido en aeronaves.

Para la aplicación descrita, la comprensión del audio transmitido se dio entre N=4 bandas divididos uniformemente usando SAF DFT sobremuestreado con LMS y NLMS. Al aumentar el número de bandas divididos a N=64, se distorsionó el mensaje.

Con un tamaño de filtro M=256 y N=16 se tiene una mejora significativa comparado con un tamaño de filtro M=32 y M=64 usando SAF DFT sobremuestreado NLMS, sin embargo, esta ganancia tiene un costo computacional elevado.

Con las mediciones de campo donde el operador de vuelo calificó el grado de comprensión de la voz al final de su procesamiento, se pudo demostrar, que el algoritmo SAF DFT sobremuestreado NLMS presenta una ligera ventaja de comprensión del audio con respecto a SAF DFT sobremuestreado LMS, especialmente cuando se divide a la señal en N=4 y N=8 bandas.

La medición de distorsión espectral mbsd.m permitió demostrar cuantitativamente la ganancia en dB del algoritmo SAF DFT sobremuestreado NLMS con respecto a SAF DFT sobremuestreado LMS, por tal motivo se recomienda utilizar el algoritmo SAF con el algoritmo NLMS con una división N=4 bandas.

Para conocer el buen desempeño del algoritmo, se recomienda hacer la evaluación en el entorno laboral, su implementación en un DSP (*Digital Signal Processor*), FPGA (*Field Programmable gate array*) o VLSI (*Very large scale integration*) validarán los resultados obtenidos en cada simulación.

## VI. REFERENCIAS BIBLIOGRÁFICAS

- [1] Organización Mundial de la Salud. (2013, Noviembre 5). Función de la OMS en la salud pública. Retrieved from <http://www.who.int/about/role/es>
- [2] Muñoz, M. (n.d.). Nociones básicas de vuelo, manualvuelo.com. CQ Counter.
- [3] Darlington, D., & Campbell, D. (1996). Sub-Band Adaptive Filtering Applied to Speech Enhancement. Paisley.
- [4] S. Haykin, *Adaptive Filter Theory*, 2nd Edition, Prentice-Hall, 1991
- [5] GILLOIRE, Andre y VETTRLI, Martin, "Adaptive Filtering in subbands with critical sampling: Analysis, experiments and application to acoustic echo cancellation", august 1992.
- [6] 3M company. (2011). SoundPro User Manual. Copyright
- [7] Kong-Aik Lee, Woon-Seng Gan, Sen M. Kuo, *Subband adaptive Filtering: theory and implementation*, United Kingdom: John Wiley and Sons, Ltd, 2009

- [8] DINIZ S.R. Paulo, *ADAPTIVE FILTERING: Algorithms and Practical Implementation*, Rio de Janeiro, Third Edition, Cap 1, 2008
- [9] WONHO YANG, Modified bark spectral distortion measure, Programa para calcular la medida de desempeño, mbsd.m, marzo 2005.

## VII. BIOGRAFÍAS

### Miguel Ángel Castillo Marcillo



Nació en el Puyo en 1986. Realizó sus estudios en la Unidad educativa experimental militar “Héroes del 41” de Machala obteniendo el título de Físico-Matemático en 2004. Su estudio superior lo realizó en Salinas en la Escuela Superior Militar de Aviación obteniendo el título de Subteniente de la Fuerza Aérea en 2008. Entre 2007 y 2013 estudió en la Universidad de las Fuerzas Armadas-ESPE, en la actualidad está realizando la tesis previo a obtener el título de Ingeniero en Electrónica y Telecomunicaciones con el tema “Tratamiento de la señal de voz usando filtros adaptativos en subbandas para ambientes altamente ruidosos”.

### Paul Bernal



Recibió el título de Ingeniero en electrónica y Telecomunicaciones en el año 2000 en la Universidad de las Fuerzas Armadas ESPE. Obtuvo la maestría en ciencias de Ingeniería Electrónica en el año 2005 en el Instituto Militar de Ingeniería (IME) de Rio de Janeiro, Brazil. Desde el año 2006 está trabajando para la Universidad de las Fuerzas Armadas-ESPE, actualmente, es Director del Departamento de Electrónica y Telecomunicaciones. Sus principales intereses de investigación incluyen el procesamiento digital de señales, filtros adaptativos y sistemas no lineales.

### Daniel altamirano



Nació en Quito el 18 de Julio de 1985. Recibió el título de Ingeniero en Electrónica y Telecomunicaciones de la Escuela Politécnica del Ejército (ESPE), Ecuador, en el año 2008 donde fue el mejor graduado. El grado de MSc. lo obtuvo en la Universidad Estatal de Campinas (UNICAMP), Brasil en el año 2011. Actualmente es Profesor en la Escuela Politécnica del Ejército y colabora en el desarrollo de proyectos en el Centro de Investigaciones Tecnológicas del Ejército. Ha publicado y presentado artículos en revistas de la IEEE. Sus áreas de interés son las comunicaciones inalámbricas, comunicación digital, radio definida por software y los códigos correctores de errores.