

# TRATAMIENTO ROBUSTO DEL SONIDO EN EL INTERIOR DE VEHICULOS

Enrique Masgrau, Alfonso Ortega, Pedro Ramos, Luis Vicente, Eduardo Lleida

Grupo de Tecnologías de la Comunicaciones (GTC). Instituto de Investigación en Ingeniería de Aragón (I3A)

Universidad de Zaragoza  
masgrau@unizar.es

**Abstract- Applications of Information and Communication Technologies in Automotive (Telematics) is a promising market at present. In this paper we show two interesting applications in this field aiming to improve the confort of in-vehicle passengers: a novel Cabin Car Communication System (CCSS) and two Active Control Noise (ANC) schemes. The first provides a high speech reinforcement while keeping an excellent sound quality. So, it outperforms other published solutions. The ANC schemes are based on two novel algorithms: the FxGAL, which provides a fast convergence at the expense of higher complexity; and the  $G_{\mu}$ -FxSLMS, which presents a very reduced computational load while keeping the same convergence performance.**

## I. INTRODUCCIÓN

El uso intensivo de tecnologías TIC en el interior de vehículos representa actualmente un escenario de aplicación muy prometedor y una gran oportunidad de mercado. Particularmente, las tecnologías acústicas basadas en técnicas de tratamiento de señal permiten un mayor confort de los pasajeros tanto mediante la cancelación acústica del ruido ambiente como por la posibilidad de uso de interfaces orales, que facilitan la conversación entre pasajeros y el control oral de comandos (marcación telefónica, radionavegación automática, etc). El diseño de estos sistemas requieren el desarrollo de sofisticados algoritmos de cancelación acústica (CAR) y eléctrica del ruido ambiente y la cancelación de los ecos acústicos producidos en el habitáculo. En este trabajo se presentan dos tipos de aplicaciones bien diferenciadas: la primera aborda una original solución de refuerzo de la señal de voz (denominada CCS: Cabin Car Communications) en el interior del habitáculo del vehículo que permite una natural y cómoda comunicación entre los pasajeros (inter-comunicación) y la de éstos con el exterior (multi-conferencia); la segunda aborda el estudio de dos algoritmos novedosos de Control Activo de Ruido (CAR), directamente dirigidos al aumento de la confortabilidad de los pasajeros mediante la reducción del nivel sonoro en el vehículo.

## II. REFUERZO DE LA SEÑAL DE VOZ

La comunicación oral entre pasajeros en un vehículo se ve dificultada por el ruido existente en el interior del mismo, la distancia entre ellos, los asientos, etc. Esto es especialmente grave cuando las dimensiones del mismo son grandes (mono-volumenes) y/o el nivel de ruido es muy alto

(carreteras en mal estado, alta velocidad, etc). En esta situación resultaría muy cómodo el uso de un sistema electro-acústico de refuerzo o amplificación del nivel sonoro de la señal de voz, que no amplifique el nivel de ruido y, sobretudo, no produzca inestabilidades por acoplo acústico entre los elementos emisores (altavoces) y captadores (micrófonos) de sonido. Estos sistemas, denominados CCS, se componen de los siguientes elementos: un conjunto de micrófonos colocados en el techo del vehículo, sobre las cabezas de los pasajeros; un conjunto de altavoces, que pueden ser los mismos del sistema de audio incorporado de serie; una etapa de amplificación; y un conjunto de dispositivos electrónicos basados en DSPs. Estos últimos son un cancelador de ecos adaptativo (AEC), un filtro supresor de eco residual (ESF) y un filtro reductor de ruido (NRF). Los dos primeros tienen la función de cancelar el eco producido por el acoplo acústico entre los altavoces y los micrófonos y el tercero la de cancelar el ruido ambiente captado por los micrófonos para evitar su amplificación.

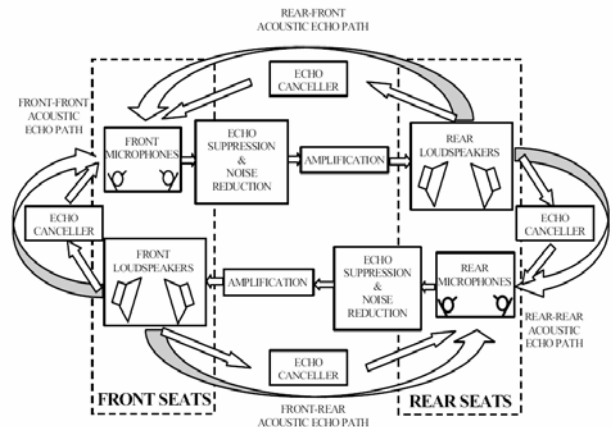


Fig. 1. Esquema del sistema CCS de dos canales.

En la figura 1 se muestra el esquema de un sistema CCS de dos canales. Un canal refuerza la voz de los pasajeros delanteros en la parte trasera del vehículo y el otro realiza la función dual. Este sistema requiere por cada uno de los canales dos micrófonos, dos altavoces, un amplificador, dos AECs, un filtro ESF y un NRF. La duplicidad de AECs por canal es debido a la necesidad de cancelar el acoplo acústico tanto entre los altavoces y micrófonos del propio canal como de los canales cruzados, más peligroso éste último debido a la mayor proximidad entre ambos elementos. Para una mayor sencillez y comprensión analizaremos el caso de un solo canal, cuyo esquema se ilustra en la figura 2.

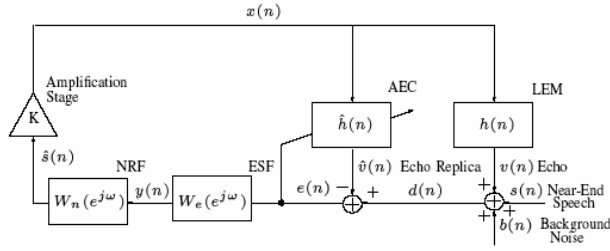


Fig. 2. Diagrama de bloques de un CCS de un canal simple.

Una característica especial de un sistema CCS es que el cancelador de ecos AEC trabaja siempre bajo condiciones de doble conversación, como se denomina la situación que se da en telefonía cuando ambos interlocutores hablan al mismo tiempo. La mezcla de la señal de eco con la conversación del otro interlocutor distorsiona gravemente la estimación del camino del eco (LEM), lo que hace más conveniente la congelación de la adaptación del AEC durante estos pasajes. En un CCS siempre están presentes la señal de voz directa  $s(n)$  y su propio eco  $v(n)$ , además del ruido ambiente  $b(n)$  (ver fig. 2), y la congelación de la adaptación no tiene sentido. Esto hace que la estimación del LEM sea burda y sea necesaria la introducción de un filtro supresor de eco residual ESF que evite, sobretodo, la inestabilidad del sistema. Sin duda, esta situación de doble conversación continua es la característica más singular de un CCS, y la que determina su dificultad y sus prestaciones finales.

### II.1. Estimación del Filtro Supresor de Eco Residual (ESF)

El filtro ESF colocado tras el cancelador de ecos (ACE) debe asegurar la estabilidad, reducir el eco residual y evitar la distorsión de la señal de voz. De acuerdo al esquema de la figura 2, la función de transferencia en lazo cerrado entre la señal de entrada  $s(n)+b(n)$  y la salida  $x(n)$  (señal reforzada) responde a la expresión:

$$P(e^{j\omega}) = \frac{K W_e(e^{j\omega}) W_n(e^{j\omega})}{1 - K W_e(e^{j\omega}) W_n(e^{j\omega}) (H(e^{j\omega}) - \hat{H}(e^{j\omega}))} \quad (1)$$

Como ya se ha dicho antes, el error de estimación del camino del eco puede ser elevado debido a la situación permanente de doble conversación, y el denominador de (1) puede aproximarse a cero dependiendo de la ganancia  $K$  del amplificador. La solución óptima del filtro ESF  $W_e(e^{j\omega})$  es la que evita la inestabilidad, minimiza la distorsión y no amplifica el nivel de ruido, y puede obtenerse forzando a que la función de transferencia  $P(e^{j\omega}) = K W_n(e^{j\omega})$ . Es decir, el lazo cerrado sólo produce amplificación y reducción de ruido mediante el filtro NRF. Puede demostrarse [1] que la expresión óptima del filtro ESF responde a:

$$W_e(e^{j\omega}) = 1 - \sqrt{\frac{S_r(e^{j\omega})}{S_e(e^{j\omega})}} \quad (2)$$

donde  $S_r(e^{j\omega})$  y  $S_e(e^{j\omega})$  son, respectivamente, el espectro del eco residual  $r(n)$  (respuesta a  $x(n)$  del sistema  $h(n)-\hat{h}(n)$ ) y de la señal error  $e(n)$ . La señal  $e(n)$  es accesible directamente, pero no así la señal  $r(n)$ . La estimación del espectro de  $e(n)$  se obtienen mediante métodos de periodograma basados en la cuasi-estacionaridad de la voz en segmentos cortos de 10-20 ms. La estimación del espectro de  $r(n)$  se obtiene a partir del de la señal error  $e(n)$  explotando la correlación que existe entre segmentos adyacentes de ambas señales, causado por

el retardo existente entre ambas e introducido por el camino del eco LEM. Los detalles de esta estimación pueden consultarse en [1].

### II.2. Estimación del Filtro Reductor de Ruido (NRF)

La solución óptima del filtro NRF  $W_n(e^{j\omega})$ , encargado de reducir el ruido  $b(n)$  presente en  $y(n)$ , se obtiene aplicando la solución de Wiener (MMSE). Bajo el supuesto de buen funcionamiento del filtro ESF ( $y(n) \approx s(n)+b(n)$  ó  $r(n) \approx 0$ ), la expresión de  $W_n(e^{j\omega})$  resulta ser:

$$W_n(e^{j\omega}) = \frac{S_s(e^{j\omega})}{S_y(e^{j\omega})} = 1 - \frac{S_b(e^{j\omega})}{S_y(e^{j\omega})} \quad (3)$$

La señal  $y(n)$  está disponible, pero no así la señal  $b(n)$ . Para estimar el espectro de ésta señal debe aprovecharse, de manera similar al caso anterior de estimación del filtro ESF, la correlación existente entre la señal error  $e(n)$  y la señal  $b(n)$  de segmentos adyacentes de ambas señales. Los detalles pueden consultarse también en [1].

### II.3. Evaluación de las prestaciones del sistema CCS

Para evaluar las prestaciones del sistema CCS propuesto se ha simulado un sistema de un canal simple usando una respuesta impulsional  $h(n)$  de un LEM real medido en un vehículo. La evaluación del sistema se ha basado en múltiples medidas acústicas (eco residual-ERLE-, distorsión, refuerzo de la voz, atenuación del ruido, etc) que permiten la selección optimizada de los diversos parámetros del sistema [1]. Además, un sistema real de dos canales fue implementado en un vehículo monovolumen, basado en una placa DSP de coma flotante de altas prestaciones, trabajando a una frecuencia de muestreo de 16 KHz. Las prestaciones de este sistema CCS en tiempo real se han evaluado en base al denominado factor de refuerzo de la voz (SR: speech reinforcement) definido como la relación entre los niveles de la señal de voz con el sistema CCS ON y OFF medido en dBs y una evaluación subjetiva de la distorsión de la voz y del nivel de atenuación del ruido ambiente conseguido.

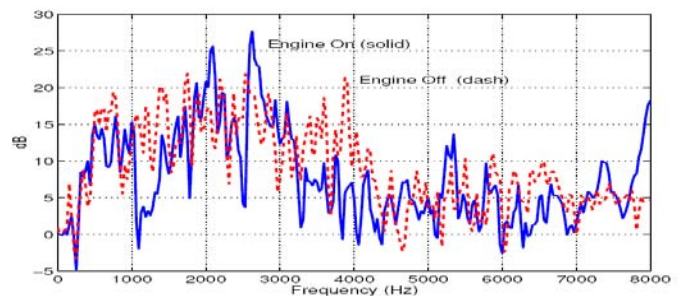


Fig. 3. Nivel de refuerzo del CCS con el motor ON y OFF

En la figura 3 se muestra el resultado del factor SR medio obtenido en uno de los asientos traseros del vehículo con el motor parado y encendido, y cuando diversas sentencias de voz son emitidas por un altavoz situado en la posición de la cabeza del conductor, simulando una conversación. Puede observarse que el refuerzo medio obtenido de la voz es de unos 15 dBs hasta 4 KHz, y de unos 5 dBs en el margen de 4 a 8 KHz en ambas situaciones, no manifestándose diferencias apreciables entre ambas. Los diversos tests subjetivos informales realizados a diversas velocidades y condiciones de conducción y de estado del firme confirmaron el buen comportamiento del sistema.

### III. CONTROL ACTIVO DE RUIDO

El ruido ambiente presente en el interior de un vehículo reduce sobremedida el confort de los pasajeros aún cuando el diseño de aquél haya sido realizado con sumo cuidado para reducirlo. Es bien sabido que, debido a la elevada longitud de onda, es imposible atenuar de manera eficiente con métodos pasivos (apantallamiento) el ruido de baja frecuencia proveniente del exterior de un habitáculo. Para conseguir esto es necesario usar técnicas activas de cancelación de ruido, consistentes en la introducción de un ruido secundario en contrafase generado a partir de una referencia del ruido primario y que cancele a este último. Por otro lado, estas técnicas activas resultan mucho más eficientes a baja frecuencia, lo que las hace complementarias de la técnicas pasivas de aislamiento acústico. Los algoritmos de procesamiento de señal encargados de generar el ruido secundario deben ser capaces de seguir las variaciones naturales de la estadística del ruido manteniendo un nivel alto de cancelación, lo que requiere el uso de algoritmos adaptativos sofisticados, con alta velocidad de convergencia y buena precisión de seguimiento del escenario acústico.

Como es bien sabido, una característica diferencial de una aplicación de control activo respecto a otras aplicaciones de filtrado adaptativo reside en que el ruido secundario es inyectado en el escenario acústico a través de una función de transferencia, denominada secundaria. Esta función de transferencia es la equivalente al camino acústico existente entre los altavoces actuadores y el punto donde se mide el resultado de la cancelación (ruido residual o error), además de la cadena de amplificación, conversión D/A y altavoz. Esto conduce al uso de algoritmos denominados de señal filtrada, entre los que el más popular es el denominado FxLMS, derivado de forma directa del popular LMS. Una de las consecuencias de esta función de transferencia secundaria, cuya estimación es imprescindible, es un cierto retardo que se introduce en la dinámica de adaptación del algoritmo adaptativo y que ralentiza la convergencia, reduce el nivel de cancelación de ruido e incluso puede producir inestabilidades cuando la estimación de esta función de transferencia es demasiado gruesa. Esta penalización sufrida por la capacidad de adaptación del FxLMS puede ser compensada a veces por el uso de algoritmos de convergencia más rápida a cambio de aumentar su complejidad computacional. Otras veces, el grado de complejidad no es asumible y lo que se requiere son algoritmos más simples desde el punto de vista computacional pero que mantengan sus propiedades de convergencia. En este artículo se presentan dos algoritmos totalmente novedosos desarrollados por los autores. El denominado FxGAL pertenece a la categoría de algoritmos rápidos y el segundo, denominado  $G_{\mu}$ -FxSLMS, a la categoría algoritmos de complejidad reducida. Ambos se exponen a continuación.

#### III.1. Algoritmo rápido FxGAL

Es bien conocido que el algoritmo LMS (y del mismo modo ocurre con el FxLMS) presenta una convergencia muy dependiente de la estadística de la señal de entrada, de tal modo que la velocidad de convergencia se reduce cuanto mayor es la dispersión de autovalores de la correlación de esta señal. De este modo, el algoritmo LMS trabaja

adecuadamente con ruido blanco pero se ralentiza enormemente en presencia de ruidos muy coloreados, muy especialmente, en el caso de señales de banda estrecha (p.e., tonos con ruido). Una forma de paliar este inconveniente consiste en blanquear u ortogonalizar previamente la señal de entrada de modo que presente una mínima dispersión de autovalores, seguido de una mera combinación lineal o filtrado FIR adaptativo de los componentes ortogonales de la señal, denominado DRE (Desired Response Estimator). Obviamente, este bloque DRE presenta una rápida convergencia, insensible a la dispersión de autovalores original. Este proceso puede realizarse mediante el uso de un predictor adaptativo en celosía (ALP: Adaptive Lattice Predictor) que presenta la propiedad de ser bastante insensible a la dispersión de autovalores de la señal a ortogonalizar; de otro modo, se estaría únicamente trasladando el problema de la ralentización al blanqueador, pero no resolviéndolo. En la figura 4 se muestra el esquema de un sistema CAR basado en el algoritmo FxGAL, que debe su nombre a que la estructura ALP es actualizada mediante un algoritmo de tipo gradiente (GAL: Gradient Adaptive Lattice). El bloque esclavo SLF (Slave Lattice Filter) presenta en cada instante unos coeficientes copia de los coeficientes de la estructura ALP, realizándose así el mismo filtrado en ambas ramas del algoritmo. Los detalles del algoritmo FxGAL pueden consultarse en [2].

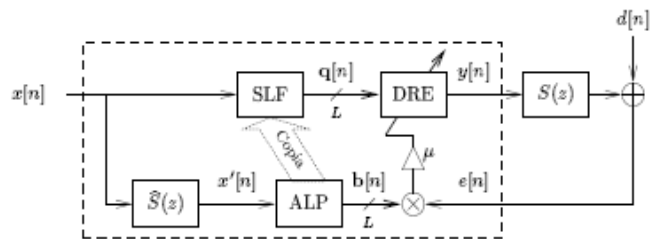


Fig. 4. Diagrama de bloques de un sistema CAR FxGAL

El esquema CAR FxGAL propuesto aquí presenta la particularidad de que el proceso de filtrado de la señal de referencia por la función de transferencia secundaria  $S(z)$  no colorea la señal  $b(n)$  que controla la convergencia del combinador DRE, como pasa en otros esquemas FxGAL [2]. A esta característica se debe las mejores prestaciones obtenidas con nuestro esquema FxGAL, sobretodo en presencia de funciones secundarias  $S(z)$  selectivas. La mayor complejidad de los esquemas FxGAL frente a los FxLMS suponen un incremento apreciable de la carga computacional de los mismos, que queda compensada por sus mejores prestaciones de convergencia. Es en este compromiso donde se debe elegir entre simplicidad o capacidad de cancelación.

A efectos de comparar las prestaciones de un esquema CAR FxLMS y el FxGAL propuesto se implementó un sistema real monocanal en el interior de un vehículo, basado en un DSP TMS320C6701. A fin de un mayor control del grado de coloreado de la señal se simuló el ruido a cancelar, que era inyectado en el habitáculo a través de un altavoz auxiliar. En la figura 5 se muestra la comparación de la velocidad de convergencia de ambos esquemas con tres tipos de señales de referencia: a) ruido blanco, b) dos tonos mas ruido blanco a -10dB y c) idem con ruido a -20 dB. En el primer caso la velocidad de convergencia es similar en ambos casos, como era previsible. En los otros dos, con

mayor dispersión de autovalores, la velocidad de convergencia es muy superior en el caso del FxGAL, especialmente en el caso c), con mucho más alta dispersión de autovalores.

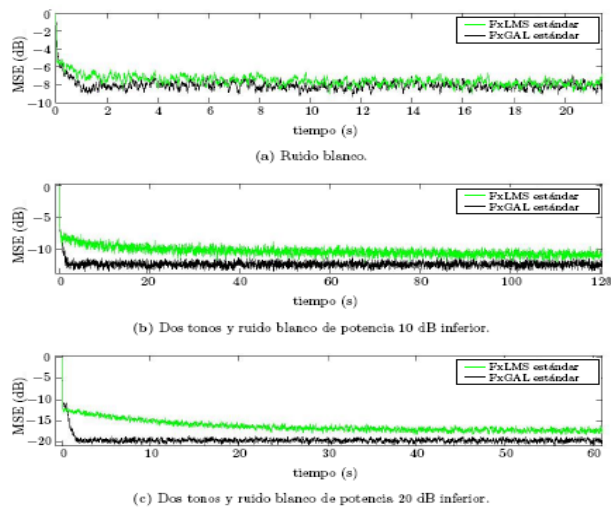


Fig. 5. Convergencia del FxGAL versus FxLMS.

### III.2. Algoritmo $G_{\mu}$ -FxSLMS

El segundo algoritmo de CAR pertenece a la categoría de algoritmo de complejidad reducida, aunque manteniendo las mismas prestaciones de convergencia. Esto último que puede parecer imposible de cumplir a priori, al menos totalmente, es posible de alcanzar en ciertas condiciones bastante usuales en escenarios ruidosos. Este es el caso de situaciones en las que el ruido es fundamentalmente armónico o de banda estrecha con un pedestal reducido de ruido de banda ancha. El algoritmo propuesto pertenece a una conocida clase de algoritmos adaptativos denominados secuenciales o de adaptación parcial, en los que se reduce su complejidad a costa de adaptar de forma simultánea únicamente una fracción ( $1/N$ ) del total de coeficientes, denominada también factor de diezmado  $N$ . Obviamente, esto produciría una ralentización de la convergencia por un factor  $N$  ya que cada coeficiente se adapta sólo cada  $N$  muestras. Esto será así a menos que el parámetro de convergencia del algoritmo pueda incrementarse por un factor similar. En general esto no será posible, pues, o se transgrede la cota máxima del valor de este parámetro con la consiguiente inestabilidad del algoritmo, o si hay margen hasta este valor, se reducirá el nivel de cancelación ya que aumentará el parámetro de desajuste del algoritmo. En el caso del algoritmo propuesto se demuestra que, en presencia de ruido fundamentalmente armónico, en un algoritmo secuencial es posible aumentar el parámetro de convergencia por el factor  $N$  de diezmado usado, manteniendo por consiguiente la velocidad de convergencia y sin reducir el nivel de cancelación o el desajuste. La única condición que debe cumplirse es que la señal a cancelar no debe presentar contenido significativo en ciertas frecuencias, determinadas por la longitud total del filtro cancelador  $L$ , el factor de diezmado  $N$  y la frecuencia de trabajo. Además, estas frecuencias “prohibidas” no se encuentran en el rango de baja frecuencia, donde el ruido presenta su mayor contenido energético, con lo que es relativamente fácil asegurar el cumplimiento de esta condición. Además, siempre es posible ecualizar el ruido

residual, de tal modo que no se pretenda cancelar aquellos armónicos “prohibidos”, asegurando de este modo que no se transgrede nunca la condición de estabilidad. Los detalles analíticos del algoritmo son bastante prolijos, y pueden consultarse en [3]. Para evaluar experimentalmente este algoritmo se ha actuado como con el anterior algoritmo, implementando un sistema CAR monocal canal basado en un DSP TMS320C6701 en el interior de un vehículo NISSAN Vanette. La señal a cancelar es un ruido de motor real pregrabado y que es inyectado en el habitáculo mediante un altavoz auxiliar. Se ha medido el espectro del ruido ambiente con el sistema CAR OFF y ON, y ambas curvas se representan en la figura 6 para el caso de factor de diezmado  $N=8$  y  $L=512$ . Puede observarse el muy buen comportamiento de cancelación de los armónicos del ruido, similar al que se obtendría con un esquema FxLMS pero con una complejidad aproximadamente 8 veces inferior.

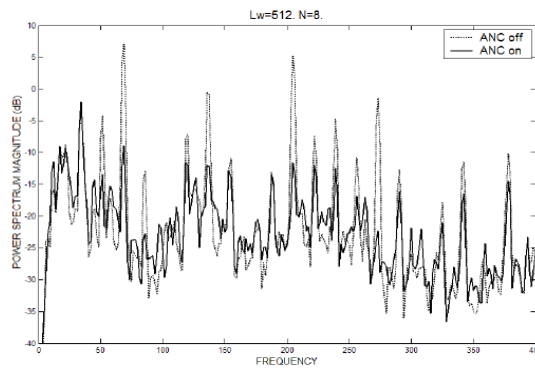


Fig. 6. Cancelación de ruido de motor con el  $G_{\mu}$ -FxSLMS

## II. CONCLUSIONES

En este trabajo se han presentado dos aplicaciones reales de tratamiento robusto de señal de audio en el interior de vehículos, basadas ambas en soluciones y algoritmos totalmente novedosos. La primera es un sistema de refuerzo de la señal de voz (CCS) que facilita la comunicación de los pasajeros de un vehículo entre ellos mismos y con el exterior a través de telefonía móvil, especialmente en escenarios adversos. La segunda aplicación es de control activo (CAR) y se han presentado dos algoritmos de altas prestaciones: el primero, el FXGAL, presenta una alta velocidad de convergencia, lo que permite alta cancelación en situaciones de alta variabilidad del escenario acústico aunque a costa de una alta carga computacional. El segundo, el  $G_{\mu}$ -FxSLMS, es capaz de reducir por un alto factor la carga computacional manteniendo las prestaciones de velocidad de convergencia y nivel de cancelación. Todos los casos han sido evaluados experimentalmente mostrando muy buenas prestaciones.

## AGRADECIMIENTOS

Este trabajo ha sido financiado por el Plan Nacional I+D bajo el proyecto DIHANA ref.: TIC-2002-04103-C03-01.

## REFERENCIAS

- [1] A. Ortega, E. Lleida, E. Masgrau, “Speech reinforcement system for cabin car communications”, IEEE Trans on Speech and Audio Processing, to be appeared in Vol 13, No. 6, November 2005.
- [2] L. Vicente, E. Masgrau, “Performance comparison of two fast algorithms for active control”. Proc. Active 99, pp1089-1100, December 1999.
- [3] P. Ramos, R. Torrubia, A. López, A. Salinas, E. Masgrau, “Computationally efficient implementation of an active noise control system based on partial updates”. Proc. Active 04, September 2004.