

MUESTREO ALEATORIO EN FRECUENCIA APLICADO AL MODELADO DE SISTEMAS NO LINEALES

Alba Pagès-Zamora, Montse Nájjar, Ana Pérez-Neira
Departament TSC, Mòdul D5
ETSE. de Telecomunicació
Apt. 30002, 08080 Barcelona

ABSTRACT

This work deals with the problem of modelling non-linear systems. For this purpose, a new architecture based on the Kolmogorov theorem is presented. It combines such suggestive subjects as Volterra series and random samples in the frequency domain. The main idea is to approximate input/output relation of a non-linear system by means of two truncated Fourier series generated from two different principle harmonics. They are chosen in order to obtain from the Fourier transform of the input/output function nonequidistant samples and, furthermore, in an almost random manner. On the other hand, the mean square error is obtained in order to evaluate the behaviour of the proposed scheme.

I. INTRODUCCION

Desde los años 80 parte del estudio de sistemas no lineales (SNL) se ha centrado en un intento de aglutinar y organizar todas las técnicas que se han ido desarrollando en torno a ellos [1]. Análisis de planos de fase, técnicas basadas en la teoría de ecuaciones diferenciales y simulaciones realizadas en ordenador representan tres de los grupos mencionados. Estos métodos se resienten en uno u otro aspecto bien sea por la falta de generalidad, bien por la poca aplicación en problemas específicos. Por otro lado, el auge de los momentos de orden superior se debe en gran parte a los resultados obtenidos en el área de sistemas no lineales [2].

Una de las técnicas ha sido la representación de Volterra/Wiener de sistemas no lineales. Se basa en la representación de la relación entrada/salida de un sistema no lineal mediante las llamadas "series funcionales de Volterra".

En lo que sigue de trabajo se presenta como los autores utilizan dichas técnicas en el modelado de sistemas no lineales proponiendo una nueva estructura que combina las series funcionales de Volterra, con el muestreo aleatorio en frecuencia y el teorema de Kolmogorov.

II. SERIES DE VOLTERRA EN EL MODELADO DE SISTEMAS NO LINEALES.

Una de las formas más generalizadas de caracterizar un sistema no lineal es mediante la relación entrada/salida.

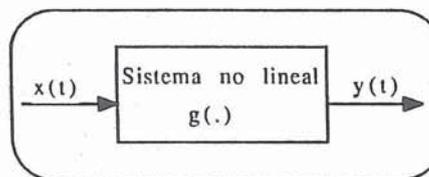


Fig.1. Representación entrada/salida de un SNL.

Así, si denotamos la entrada por $x(t)$ y la salida por $y(t)$ ambas se pueden relacionar tal y como sigue:

$$y(t) = g[x(t)] \quad (\text{Ec.1})$$

Suponiendo que la señal de entrada está limitada en amplitud a un intervalo $[-D/2, D/2]$, la función entrada/salida se puede repetir con periodo D , resultando de ello $\hat{g}(x)$ definida por:

$$\hat{g}(x) = \sum_{n=-\infty}^{+\infty} g(x - nD) \quad (\text{Ec.2})$$

En la notación se omite la dependencia temporal de las señales de entrada y salida para evitar complicar las expresiones. En caso de sustituir la función $g(x)$ por $\hat{g}(x)$ la salida sería la misma y además periódica en el dominio x , y por lo tanto se puede desarrollar en serie de Fourier.

$$y = \hat{g}(x) = \sum_{n=-\infty}^{+\infty} \frac{1}{D} G(n\omega_0) \exp(jn\omega_0 x) \quad \omega_0 = 2\pi/D \quad (\text{Ec.3})$$

Siendo $G(\omega)$ la transformada de Fourier respecto del dominio x de la relación entrada/salida $g(x)$. El resultado ((Ec.3)) puede interpretarse como una serie de Volterra ya que desarrollando la función exponencial en serie de Taylor se obtendría una función polinómica en potencias de la entrada, x .

II. MUESTREO ALEATORIO EN EL DOMINIO FRECUENCIAL.

Supongamos por un momento que se muestrea la transformada de Fourier de la relación entrada/salida denominada $G(\omega)$ no de una forma equiespaciada sino aleatoriamente [3]. Si denotamos \tilde{y} por la función que se obtendría de este proceso,

$$y = \tilde{g}(x) = \sum_{n=-\infty}^{+\infty} G(\omega_n) \exp(j\omega_n x) \quad (\text{Ec.4})$$

El valor esperado de la función \tilde{y} se puede expresar en función de la inversa de la transformada de Fourier de la expresión (Ec.4) tal y como sigue

$$\begin{aligned} E[y] &= E \left[\int_{-\infty}^{+\infty} \left[\sum_{n=-\infty}^{+\infty} G(n\omega_n) \delta(\omega - \omega_n) \right] \cdot \exp(-j\omega x) d\omega \right] = \\ &= \int_{-\infty}^{+\infty} G(\omega) \cdot E \left[\sum_{n=-\infty}^{+\infty} \delta(\omega - \omega_n) \right] \cdot \exp(-j\omega x) d\omega \end{aligned} \quad (\text{Ec.5})$$

Realizando el muestreo en frecuencia de forma que en la expresión anterior (Ec.5) el término en el que se calcula el valor esperado del conjunto de muestras equivaliera a una constante, se deduciría que el valor medio de la función \tilde{y} coincide con la relación entrada/salida del sistema no lineal que se pretende modelar.

$$E \left[\sum_{n=-\infty}^{\infty} \delta(\omega - \omega_n) \right] = \text{cte} \Rightarrow E\{\hat{y}\} = y \quad (\text{Ec.6})$$

III. ESQUEMA DEL SISTEMA PROPUESTO.

En un caso práctico de modelado de sistemas sería necesario truncar el sumatorio de infinitos términos (Ec.5) para poder almacenar los datos en un ordenador.

Alternativamente, los autores proponen obtener ese muestreo no equiespaciado mediante el esquema de la (Fig.2). Éste puede considerarse una aplicación del teorema de Kolmogorov, el cuál prueba la existencia de una función continua que transforma n entradas a m salidas utilizando únicamente funciones de una sola variable. El teorema también describe la arquitectura básica para implementar la citada función [4,5].

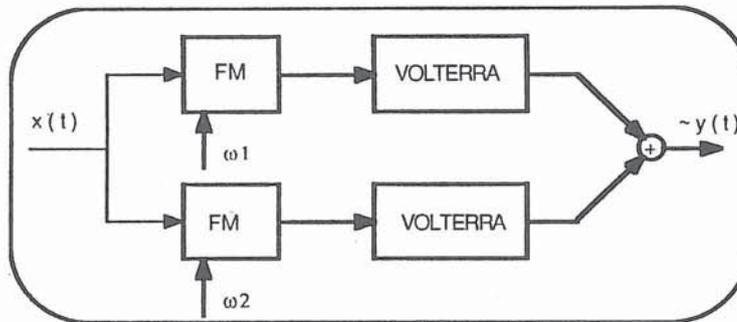


Fig.2. Arquitectura práctica propuesta para el modelado de SNL.

La figura anterior (Fig.2) representa la idea de aproximar la relación entrada/salida del SNL mediante dos series iguales a la descrita en (Ec.4) pero truncadas y utilizando para cada una de ellas dos armónicos fundamentales distintos. Es decir,

$$\tilde{y} = \tilde{g}(x) = \sum_{m=-M}^{+M} \frac{1}{2D_1} G(m\omega_1) \exp(jm\omega_1 x) + \sum_{n=-N}^{+N} \frac{1}{2D_2} G(n\omega_2) \exp(jn\omega_2 x)$$

$$\omega_i = 2\pi/D_i \leq \omega_n \quad i = 1, 2 \quad (\text{Ec.7})$$

La elección de las frecuencias ω_1 y ω_2 es una decisión crítica y cabe destacar que es una cuestión todavía sujeta a estudio. De todas formas avanzamos que estas frecuencias deben ser escogidas en un intento de muestrear de forma "aleatoria" el espectro de $g(x)$.

Antes de entrar en detalles, debe considerarse que para poder realizar las simulaciones mediante ordenador se ha tenido que discretizar el problema. Supongamos de este modo que $g(x)$ se muestrea obteniendo una secuencia de N puntos centrados en el origen que llamaremos $g(n)$.

$$g(n) = g(x) \Big|_{x = (n - \frac{N-1}{2}) \frac{D}{N-1}} \quad n = 0..(N-1); N \text{ impar} \quad (\text{Ec.8})$$

Consecuentemente, en lugar de la expresión (Ec.7) se ha implementado su versión discreta.

$$y_2(n) = \frac{1}{N_1} \sum_{k_1=0}^{NQ_1-1} G(k_1) \cdot \exp(j \frac{2\pi}{N_1} n k_1) + \frac{1}{N_2} \sum_{k_2=0}^{NQ_2-1} G(k_2) \cdot \exp(j \frac{2\pi}{N_2} n k_2)$$

$$G(k) = \sum_{n=0}^{N-1} g(n) \cdot \exp(-j \frac{2\pi}{N_i} n k) \quad N \leq N_i$$

(Ec.9)

Se trata de dos series discretas de Fourier truncadas a NQ_1 y NQ_2 coeficientes respectivamente, siendo los armónicos fundamentales

$$\omega_i = 2\pi/N_i \quad ; N_i \geq N \quad i = 1,2$$

(Ec.10)

Los valores N_1 y N_2 se han escogido primos. Con ello se muestrea $G(\omega)$ de una forma no equiespaciada consiguiendo además que la diferencia entre dos frecuencias consecutivas pueda considerarse una variable aleatoria. Con esto no aseguramos que la condición impuesta por (Ec.6) se cumpla pero, tal y como se comprobará en el siguiente apartado, los resultados obtenidos en términos del error cuadrático medio nos inducen a pensar que una elección adecuada de N_1 y N_2 podría aproximar la citada condición.

RESULTADOS.

En este apartado se presentan algunas de las simulaciones realizadas. Para evaluar la bondad de la arquitectura propuesta (Fig.2, Ec.9) se comparan los resultados obtenidos con los que se obtendrían de aproximar la relación entrada/salida del sistema no lineal con una serie de Fourier truncada convencional.

$$y_0(n) = \frac{1}{N_0} \sum_{k=0}^{NQ_0-1} G(k) \cdot \exp(j \frac{2\pi}{N_0} n k) \quad N_0 \geq N$$

(Ec.11)

Para que la comparación sea válida se ha forzado a que las expresiones (Ec.9) y (Ec.11) tuvieran el mismo nº de coeficientes (por lo tanto, que ocuparan el mismo espacio de memoria al almacenarlas) y que abarcaran el mismo ancho de banda. Es decir,

$$\text{Mismo número de coeficientes} \Rightarrow NQ_0 = NQ_1 + NQ_2$$

$$\text{Mismo ancho de banda} \Rightarrow NQ_0/N_0 = \max(NQ_1/N_1, NQ_2/N_2)$$

(Ec.12)

A continuación se presentan los resultados de las simulaciones realizadas. El sistema no lineal que se ha modelado es la función sigmoide. La primera simulación (Fig.3) consta de tres funciones:

- A: Aproximación obtenida con una serie de Fourier (Ec.11).
- B: Aproximación obtenida con dos series de Fourier (Ec.9)
- C: Función sigmoide a aproximar.

La segunda (Fig.4) representa el error para cada uno de los casos.

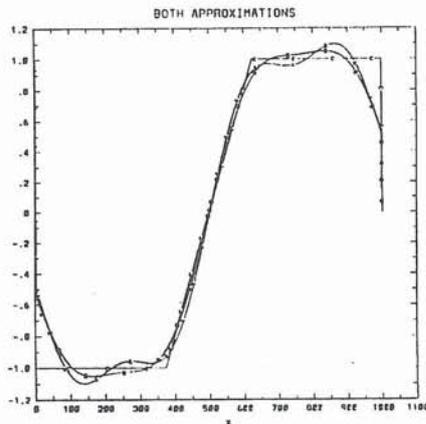


Fig.3

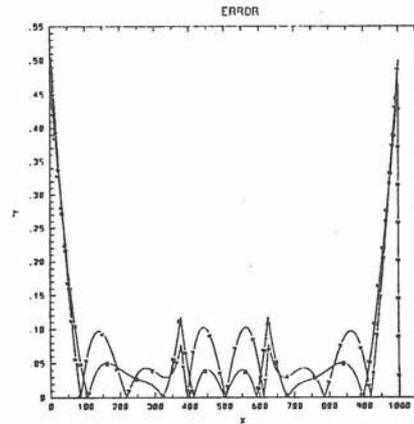


Fig.4

Parámetros	N	N0	NQ0	N1	NQ1	N2	NQ2
Valor	1001	2674	20	1337	10	1577	10

Tal y como se observa la aproximación obtenida en el caso denominado por B es mejor que para el caso A, lo que se puede verificar con la gráfica del error. De la tabla se puede comprobar que se han respetado las condiciones impuestas por (Ec.12).

CONCLUSIONES

En este artículo se propone un esquema que modela sistemas aunque nuestro estudio se ha dedicado a los sistemas no lineales. En determinados casos el esquema presenta un error cuadrático menor que el esquema alternativo propuesto. La línea de actuación actualmente se centra en determinar que valores N1 y N2 son los óptimos para reducir todavía más el error.

REFERENCIAS

- [1] W.J.Rugh: "Nonlinear system theory: The Volterra/Wiener Approach", Johns Hopkins Press, 1981.
- [2] J.S.Bendat: "Nonlinear system analysis & identification from random data", John Wiley & Sons, 1990.
- [3] I.Bilinskis, A.Mikelsons, "Randomised Signal Processing", Prentice-Hall, 1992.
- [4] A.N.Kolmogorov: "On the representation of continuous functions of many variables by superposition of continuous function of one variable and addition". (in Russian), Dokl. Akad. Nauk USSR, 114, pp953-956, 1957.
- [5] Lagunas M.A., Pérez-Neira A., Nájjar M., Pagès-Zamora A.: "The Kolmogorov Signal Processor", New trends in neural computation, Proc. IWANN'93, pp. 495-512, Barcelona, Spain, June 1993.

MECANISMO DE GRADIENTE ADAPTATIVO PARA LA CONMUTACION CA-GO-SO-CFAR EN ENTORNOS DE CLUTTER NO HOMOGENEO

G. de Miguel Vela, J.J. Martínez Madrid, F. Jiménez Rodríguez, J. R. Casar Corredera
Dep. Señales Sistemas y Radiocomunicaciones. ETSI Telecomunicación.
Universidad Politécnica de Madrid
Ciudad Universitaria s/n, 28040-Madrid

ABSTRACT: The contribution proposes a CFAR scheme (with reference in range) which uses an edge detector to look for clutter ridges in the reference window. The algorithm divides the whole reference window into a series of subzones and determines which are occupied by the clutter edge. This information allows us to adapt the reference size thus assuring that CFAR works in the homogeneous clutter zone corresponding to the cell under test.

1. INTRODUCCION

El problema de detección automática en sistemas radar ha recibido mucho interés históricamente. La cantidad de información que llega en un momento dado al operador puede saturar su capacidad de respuesta. Este hecho ha tenido como consecuencia que, a medida que la tecnología lo ha permitido, se hayan implantado extensivamente los dispositivos de detección automática.

Sin embargo, la utilización de detectores automáticos en sistemas radar no es directa. Esto se debe a la multiplicidad de situaciones a que debe hacer frente el dispositivo. El controlar la saturación del operador supone diseñar, en cualquiera de las anteriores situaciones, según el criterio de Neyman-Pearson (probabilidad de falsa alarma (PFA) constante maximizando la probabilidad de detección (PD)). A este tipo de detectores se les denomina CFAR ("Constant False Alarm Rate"). El problema consiste en que un detector posee la propiedad CFAR para unas condiciones de detección (tipo de interferencia, ruido o clutter) y no para otras. En definitiva, hay que dotar al receptor de unos mecanismos adaptativos que permitan la reconfiguración del detector según la situación a que se enfrente.

El dispositivo CFAR más empleado es el de promediado de celdas (CA-CFAR) en distancia. Consiste en estimar la potencia de clutter a partir del promedio de potencia en celdas de resolución adyacentes a la detectada. Esta potencia se multiplica por un factor de seguridad y se emplea como umbral. El sistema solamente funciona correctamente si el ruido dentro de las celdas de la ventana de referencia posee las mismas características estadísticas. Una situación típica en que el CA-CFAR falla en el control de la PFA es cuando aparece un borde de clutter dentro de la ventana de referencia (una parte de las muestras tiene mayor potencia que las otras). Cuando la celda detectada está en el borde del escalón se produce un aumento súbito de la PFA. Para evitar este fenómeno se aplica el GO-CFAR ("Greatest Of CFAR") que suma cada mitad de la ventana de referencia por separado y elige para el cálculo del umbral la mayor. Este dispositivo produce efectos de supresión de blancos al elevar innecesariamente el umbral en determinadas situaciones. Para contrarrestar estos efectos de supresión se puede elegir la mitad menor, dando lugar al SO-CFAR ("Smallest Of CFAR"). Esta última modificación tiene el problema del aumento intolerable de PFA en bordes de clutter.

La contribución presenta un algoritmo de decisión que permite conmutar entre el CA, GO y SO-CFAR dependiendo de un estimador que actúa sobre las celdas de referencia. La principal característica del algoritmo es su sencillez e incremento prácticamente nulo del coste computacional sobre el total de ejecución del CFAR.

En el primer y segundo apartados se describen el funcionamiento y la metodología de diseño del detector. Para tener una idea de la aplicabilidad de este sistema CFAR se ha realizado una comparación de prestaciones en presencia de bordes de clutter con el CA y GO-CFAR. Los resultados de esta comparación se exponen en el tercer apartado. Se concluye la contribución con la sección de conclusiones y bibliografía.

2. DESCRIPCION DEL DETECTOR CFAR

El detector CFAR propuesto está basado en una referencia de celdas espacial de la que se extraerá el nivel local de potencia de clutter. En la figura 1 se representa dicho sistema CFAR con sus respectivos algoritmos y la nomenclatura que se seguirá en el artículo. Con el fin de

facilitar los cálculos, y debido a que no se pierden ninguna de las características cualitativas del detector, se estudiará un sistema con clutter de estadística gaussiana, incorrelado entre celdas, con detector cuadrático y sin el empleo de integración incoherente de pulsos. No obstante, en la figura 1 se expone el esquema general, que incluye integrador incoherente y un tipo genérico de detector.

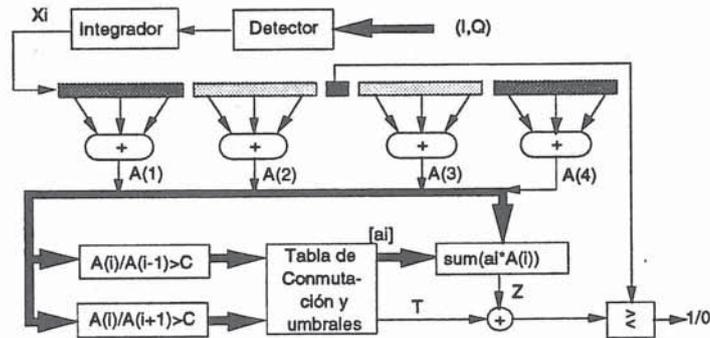


Figura 1: Esquema del detector CFAR.

En primer lugar, la señal en fase y cuadratura pasa por el detector de envolvente. A continuación está colocado el integrador incoherente, que sumará los sucesivos pulsos provenientes de la misma celda de resolución. La señal resultante de estos dos procesos se alimenta a la línea de retardo en distancia sobre la que actúa el CFAR. Hasta aquí se realiza el mismo proceso que en un CFAR clásico (CA, GO o SO-CFAR).

La ventana de referencia se divide en un número par de subventanas iguales, distribuidas en dos mitades a ambos lados de la celda bajo prueba. En la figura 1 se muestra un esquema de cuatro ventanas que será el estudiado en lo que resta de artículo. No obstante, este esquema es generalizable a un número mayor de ventanas. La razón de dividir la referencia en cuatro es que para los tamaños usados constituye un buen compromiso entre complejidad y prestaciones. Las muestras de cada subventana se suman formando las variables $A(i)$. La filosofía del detector está basada en realizar una serie de procesos con la secuencia $A(i)$ que permitan distinguir en que subventanas está situado el borde de clutter y en cuales no.

En primer lugar se concibe la secuencia $A(i)$ como circular, es decir, cuando i sea igual a 4, entonces $i+1$ valdrá 1. Se realiza una búsqueda de bordes ascendentes y otra de descendentes mediante la comparación de los cocientes $A(i)/A(i-1)$ y $A(i)/A(i+1)$ contra un umbral "C". Este umbral constituye el único parámetro de diseño de este detector (si exceptuamos el número de subventanas en la referencia). Su determinación se aborda en el siguiente apartado. Con el resultado de esta detección de bordes (8 valores binarios) se direcciona una tabla en la que se almacenan las subzonas a sumar y el factor de umbral a emplear en la determinación de la potencia de ruido.

La tabla que contiene las zonas a sumar dependiendo de la situación de bordes detectada es el elemento más crítico de este CFAR. Su función es asignar en una determinada situación un tipo de detector (GO, CA o SO principalmente y en algunos casos menos comunes alguna combinación intermedia). Su concepción se ha basado en el análisis de las posibles inhomogeneidades de clutter presentes en un entorno de este tipo. Además de los bordes de clutter, se han tenido en cuenta aumentos bruscos del nivel del ruido en una, dos (contiguas y separadas) y tres subventanas (se consideraron también las posibilidades de disminuciones del nivel de potencia de ruido de las mismas características). En estas condiciones se han evaluado los resultados que se obtendrán en el caso de que los detectores de bordes fuesen ideales. Cuando han existido dudas de en que parte del borde de clutter se encuentra la celda bajo prueba el CFAR empleado ha sido sesgado hacia el GO. Por último, y como se explicará en el

siguiente apartado, la tabla ha sufrido un ajuste de salidas que produzcan exceso de falsas alarmas mediante un estudio estadístico. En los casos en que existan patrones extraños a la salida del detector de bordes estos han sido asignados a un CA-CFAR suponiendo que vienen provocados por la presencia de fuertes inhomogeneidades espaciales equivalentes a que el ruido de fondo no posea estadística gaussiana. Los aumentos de PFA debidos a estos fenómenos deben ser corregidos mediante el incremento del factor de umbral ("T" en la figura 1) y no con un cambio de tipo de CFAR.

3. METODOLOGIA DE DISEÑO

En este apartado proponemos una metodología de diseño para el detector CFAR propuesto. A lo largo de la discusión supondremos clutter de estadística gaussiana e incorrelado entre celdas de resolución. Así mismo, se asumirá que no existe integración incoherente de pulsos. Estas suposiciones no quitan generalidad al método de diseño, mientras que facilitan el manejo matemático. Con otro tipo de estadística, la única diferencia estaría en el cálculo de los factores de umbral del CFAR. La suposición de incorrelación en distancia es realista en entornos radar.

El diseño se basa en la minimización de la probabilidad de falsa alarma (PFA) promedio cuando un borde de clutter desliza por la ventana de referencia. Se promedia la PFA para todas las posibles posiciones del borde del clutter dentro de la referencia. Otros criterios, como la minimización de la PFA de pico o el error cuadrático medio sobre la PFA en ruido, dan un resultado cualitativamente equivalente. Se optó por este criterio debido a que tiene un significado de degeneración global de PFA que puede ser tenido en cuenta en el diseño del sistema (por ejemplo, para el cálculo de sobrecarga del procesador de datos).

Las etapas seguidas en el proceso de diseño, bajo las condiciones asumidas en el primer párrafo, son las siguientes:

1- Cálculo del factor de umbral para los casos en que sean empleadas una, dos, tres o las cuatro subventanas en la composición de la referencia. Su expresión es la del factor de umbral para un CA-CFAR sin integración incoherente [1]:

$$PFA(T) = (1 + T)^{-L} \quad (1)$$

Donde L tomará los valores M/4, M/2, 3M/4 o M según el número de subventanas sumadas en la composición de la referencia (M es el número de celdas totales en la referencia). PFA se refiere a la probabilidad de falsa alarma deseada en condiciones de clutter homogéneo.

2- Creación de la primera versión de la tabla que relaciona las detecciones de bordes con las subventanas sumadas en la referencia. En este primer caso se supondrá un detector de bordes ideal (localiza el borde en la ventana en que se encuentra y en ninguna otra). El resto de casos se supone que son debidos a estadística no gaussiana, y en principio se asignan a un CA-CFAR (se suman las cuatro subventanas).

3- Se realiza una primera optimización del umbral de detección de bordes "C". Para ello se ha empleado un método de simulación semianalítico: se simula solamente la referencia, la detección de bordes y la consulta en la tabla. La PFA para cada posición del borde se obtiene mediante la fórmula [2]:

$$PFA = \prod_{i=1}^L \left[1 + \frac{T \cdot \sigma_i^2}{\sigma^2} \right]^{-1} \quad (2)$$

Donde σ_i^2 indica la potencia de clutter en cada celda de la referencia y σ^2 en la celda bajo prueba. Se halla la PFA promedio para todas las posiciones del borde en función del umbral "C", obteniéndose curvas como las de la figura 2. Este proceso se realiza para varias relaciones

de borde de clutter (en nuestro caso: 5, 10, 15, 20 y 25 dB). Se busca un "C" provisional dentro de los amplios mínimos de dichas curvas.

4- Con el umbral "C" elegido en el apartado anterior se realiza una estadística de los casos de detección de bordes más comunes y que producen aumentos indeseados de la PFA. Estas posiciones de la tabla (que corresponderían en principio a CA-CFAR) se corrigen a detectores de tipo GO-CFAR que controlen bien el aumento de PFA. De esta manera queda establecida la tabla de conmutación definitiva.

5- Se repite la optimización del apartado 3 con la tabla definitiva, obteniéndose las curvas de la figura 2 (para un caso de 24 celdas). Esto se repite para varias relaciones de borde de clutter. Existe un problema debido a que el mínimo de PFA promedio depende de la relación de potencia entre bordes. Para la elección definitiva de un C se ha acudido al comportamiento en probabilidad de detección. Se ha realizado una simulación del mismo tipo que en el punto 3 (en la expresión (2) debe sustituirse T por $T/(1+SN)$, donde SN es la relación señal a ruido en la celda bajo prueba) pero de PD en función de "C" ("C" en el margen en que se encuentran los mínimos de PFA). Se obtiene así la curva de la figura 3. En ella se ve que con un umbral a partir de 3 la mejora en PD es muy lenta. Fijándonos en la figura 2 se puede ver que aumentar más arriba el umbral produce una degeneración considerable en PFA para bordes de clutter bajos.

4. RESULTADOS Y DISCUSION

Para ilustrar el método de diseño y comparar las prestaciones del detector propuesto con las del GO y CA-CFAR respectivamente se han hallado la PFA promedio y PD promedio para los tres sistemas CFAR. La PFA de diseño en condiciones homogéneas ha sido de 10^{-6} y el número de celdas de referencia de 24. Con el fin de ilustrar la influencia del umbral "C" sobre las prestaciones del detector se han representado en la figura 2 y 3 la PFA y PD en función de "C" (para varios valores del escalón de clutter: 5, 10, 15, 20 y 25 dB). Estas figuras corresponden al paso 5 del proceso de diseño. La PFA y PD promedio del CA y GO-CFAR aparecen representadas en dichas figuras con "*" o "o" respectivamente. En el caso de la figura 2 los puntos más altos corresponden a bordes mayores y viceversa en la figura 3.

En la figura 2 puede observarse que existe un mínimo claro de PFA promedio en función de C. Dicho mínimo es estrecho para bordes de clutter pequeños (5 y 10 dB), mientras que va ensanchando para bordes de clutter grandes (alcanzando mínimos inferiores cuanto mayor es el borde). Como se puede observar, la PFA de nuestro detector tienden a la del CA-CFAR cuando C está en torno a 0 o tiende a infinito. Se pretende garantizar con nuestro CFAR unas prestaciones en PFA muy superiores a las del CA y por lo menos equivalentes a las del GO-CFAR. A la vista de la gráfica, esto implica un valor de C entre 2 y 4. Como puede verse, esta elección implica que con bordes pequeños las prestaciones sean peores que las del GO y algo mejores que las del CA. En el caso de bordes grandes las prestaciones en PFA son siempre muy superiores. Esta es la situación donde interesan las mejoras puesto que es cuando se producen aumentos de PFA intolerables.

Con el fin de comparar las prestaciones en PD se han realizado las pruebas cuyos resultados se representan en la figura 3. En dicha figura aparece la PD promediada para todas las posibles posiciones del borde de clutter en función del umbral "C". La relación señal a ruido empleada es de 22.5 dB (corresponde a una PD de aproximadamente 0,9 con un CA-CFAR y condiciones homogéneas). Los resultados para el CA y GO-CFAR se representan de idéntica manera a como se hizo en la figura 2. Debemos puntualizar que la curva superior de la figura corresponde a nuestro CFAR funcionando en condiciones homogéneas (cuantifica las pérdidas respecto al CA-CFAR). Se puede concluir que para obtener unas prestaciones en PD superiores a las del GO se impone un umbral "C" superior a 3. Con respecto al CA obtenemos fuertes mejoras con bordes de clutter grandes, mientras se obtienen leves pérdidas con bordes de clutter pequeños. Las pérdidas en condiciones de homogeneidad con respecto al CA-CFAR están en torno a la décima de decibelio, cantidad despreciable desde el punto de vista práctico.

5. CONCLUSION

Se ha desarrollado un sistema CFAR con referencia en distancia que integra un detector de bordes con el fin de estimar la presencia y posición de bordes de clutter. Dependiendo de la situación del borde, el sistema emplea un CA, GO o SO-CFAR (e incluso alguna modificación de estos).

El detector se diseña mediante un proceso iterativo que permite minimizar la PFA promedio en el caso de que un borde de clutter atraviese la referencia CFAR. Esta minimización permite la determinación del umbral empleado por el detector de bordes y la creación de la tabla de conmutación (asigna un determinado CFAR a una situación de clutter):

El detector propuesto ofrece unas prestaciones en PFA en presencia de bordes de clutter equivalentes, y en algunos casos superiores, al GO-CFAR. Con respecto a la PD promedio (medida en las mismas condiciones que la PFA), puede decirse que es superior al CA y GO con bordes de clutter grandes (superiores a 10 dB). Con bordes de clutter medianos o pequeños las prestaciones están entre el GO y el CA.

6. BIBLIOGRAFIA

- [1] P. P. Gandhi, S. A. Kassam, "Analysis of CFAR Processors in Nonhomogeneous Background", IEEE Trans. on Aerospace and Electronic Systems, Vol 24, N° 4, July 1988.
- [2] A. E. Gibson, "Adaptive Detection Probabilities for Fluctuating target Models in Nonhomogeneous Gaussian Noise", IEEE Trans. on Aerospace and Electronic Systems, Vol 14, N° 1, January 1978.

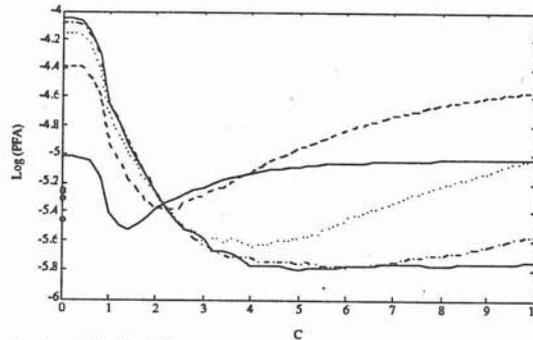


Figura 2: Logaritmo decimal de la PFA promedio en función del umbral de detección de bordes "C" (con la potencia del borde de clutter como parámetro).

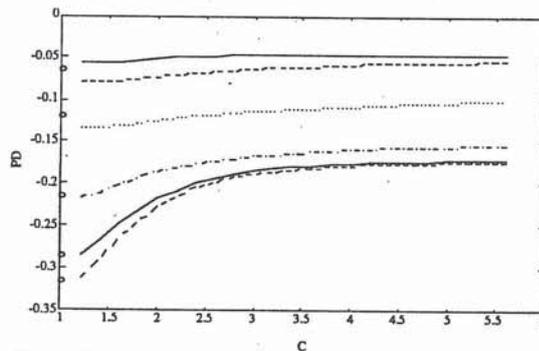


Figura 3: PD promedio en función del umbral de detección de bordes "C" (con la potencia del borde de clutter como parámetro).

SISTEMA DE MEJORA DE VOZ USANDO ESTIMACION AR DE ORDEN SUPERIOR EN AMBIENTES REALES

Josep M.SALAVEDRA*, Xavier JOVE*, Enrique MASGRAU**, Asunción MORENO*

* Departament de Teoria del Senyal i Comunicacions. Universidad Politècnica de Catalunya.
Apdo.. 30002. 08080-BARCELONA. SPAIN. E-mail: mia@tsc.upc.es

** Departamento de Ingeniería Eléctrica y Ordenadores. Universidad de Zaragoza.
María de Luna,3. 50015-ZARAGOZA. SPAIN

ABSTRACT

We study some speech enhancement algorithms based on the iterative Wiener filtering method due to Lim-Oppenheim [2], where the AR spectral estimation of the speech is carried out using a second-order analysis. But in our algorithms we consider an AR estimation by means of cumulant analysis. This work extends some preceding papers due to the authors, providing a behavior comparison between cumulant algorithms and classical autocorrelation one. Some results are presented considering AWGN that allows the best improvement and those noises (diesel engine and reactor noises) that leads to the worst one. An exhaustive empirical listening test shows that cumulant algorithms outperform the original autocorrelation algorithm, specially at low SNR.

1.INTRODUCCION

No es un secreto que en muchas aplicaciones de procesamiento de voz las grandes mejoras obtenidas en laboratorio se degradan drásticamente en ambientes reales por su poca robustez. La solución propuesta aquí alude a un preprocesamiento delante-atrás con el objetivo de mejorar la calidad de voz, es decir, el uso de un modelado paramétrico insensible al ruido. El uso de cumulantes de orden superior en el cálculo del modelado AR produce la deseada descorrelación entre señal y ruido. Este efecto se basa en una propiedad sólo aplicable a procesos Gaussianos, todos los cumulantes de orden superior a dos son idénticamente nulos. Más aun, los procesos no Gaussianos que presentan una p.d.f. simétrica tienen los cumulantes impares nulos. Si consideramos una p.d.f. Gaussiana o simétrica (una buena aproximación de ambientes reales) y las características no-Gaussianas de la voz (principalmente en las tramas sonoras) es posible obtener un modelado AR del espectro de voz más independiente del ruido usando, por ejemplo cumulantes de 3er orden en señal ruidosa. El problema aparece en la mayor distorsión del espectro de orden superior, que presenta la estimación de los cumulantes del modelo AR en comparación con el caso de autocorrelación. Esto se debe a la mayor varianza en la estimación de los cumulantes y que la secuencia de error producida cuando el filtro AR inverso funciona como predictor sobre señal de voz deja de ser plana.

2.FILTRADO DE WIENER ITERATIVO

En el método original de Lim-Oppenheim [2], la voz ruidosa era mejorada usando filtrado iterativo de Wiener, así definido:

$$W(\omega) = \frac{P_s}{P_s + P_r} \quad (1)$$

donde P_r es el espectro de ruido $r(n)$, estimada en las tramas sin señal de voz, y P_s es la estimación del espectro de la voz limpia, señal a obtener. Así, la estimación de los espectros de voz y ruido es necesaria para el diseño del filtro de Wiener en cada trama (ver fig. 1). Un filtrado iterativo de Wiener lleva a una mejor estimación del modelado AR de la voz:



Fig.1: Esquema del algoritmo iterativo de Wiener.

Este trabajo ha sido subvencionado por el TIC 92-0800-C05-04

En nuestros algoritmos la estimación del espectro de ruido se obtiene mediante técnicas de periodograma promediado durante las tramas de silencio. Hemos observado que un incremento en el número consecutivo de tramas de silencio, para estimar el ruido, produce una mejor estimación que repercute en una mejora de unos 0.3dB en las distancias espectrales tras el procesado, especialmente si se consideran las estadísticas de 2º orden o 3er orden.

Después de cada iteración se puede esperar una mejora, ya que la estimación AR de la voz se obtiene a partir de una señal más limpia que en la iteración anterior. Ciertos factores limitadores propios del algoritmo no deben olvidarse, como el número de iteraciones máximo útiles. Evidentemente la voz filtrada presenta un ruido residual menor, pero la distorsión espectral es más acusada. Así pues, el aumento del número de iteraciones no siempre implica una mejora de la estimación. Además se produce un conocido efecto sobre los formantes [3], más estrechos y picudos, dando lugar a un sonido poco natural de la voz.

En [4] se muestra un análisis detallado de la convergencia de este algoritmo. Se demuestra que la estimación del filtro de Wiener tiende a cancelar todas las frecuencias con SNR inferior a 4.77dB, y una atenuación adicional, proporcional al nivel de ruido, afecta a las frecuencias con SNR superior, en comparación con el filtro óptimo de Wiener. Sólo en las frecuencias de señal no contaminadas la atenuación permanece nula.

3.EVALUACION EMPIRICA

En este apartado se consideran varias versiones modificadas del algoritmo original de Lim-Oppenheim. Todas ellas usan estimaciones de orden superior para el cálculo del modelado AR de la voz. Cada algoritmo se compara con el clásico de segundo orden. Los resultados de usar diferentes frases de locutores masculinos y femeninos (incluida la frase propuesta por ESCA), y diferentes ruidos: AWGN y varios ruidos reales (motor diesel y ruidos de reactor) se muestran en esta sección, usando diferentes SNR globales de 0dB a 18dB. Más detalles de nuestro algoritmo se pueden encontrar en [5].

La mejora se evalúa en términos de las medidas espectrales estándar, tales como las distancias de Itakura, Cosh y Cepstrum. También se comentan test de audición realizados, donde se evalúa principalmente la calidad e inteligibilidad resultantes. Las medidas de calidad se basan en la sensación de placer en el oyente y las de inteligibilidad en la cantidad de información que puede ser extraída de la voz mejorada, sin tener en cuenta la calidad. Un incremento en la calidad representa como ventajas: mejor comprensión, efecto de bienestar y minimización de la fatiga del locutor. Desafortunadamente estos logros conllevan un precio: la pérdida de inteligibilidad. Existe pues, un compromiso entre la mejora en calidad y la pérdida de inteligibilidad.

3.1 Algoritmo de tercer-orden

En modelado AR (fig.1) de la estimación del espectro de voz se calcula con cumulantes de tercer-orden mediante el uso de covarianzas:

$$C_k(i,j) = \sum_{n=p+1}^N x(n-k)x(n-i)x(n-j) \quad ,, \quad 0 \leq k,i,j \leq p \quad (2)$$

donde $p=10$ es el orden del filtro. Los coeficiente del filtro de Wiener a_k se calculan resolviendo las siguientes ecuaciones [1]:

$$\sum_{k=0}^p a_k \cdot C_k(i,j) = 0 \quad ,, \quad 1 \leq i \leq p ; 0 \leq j \leq i \quad (3)$$

Mediante este modelado AR de tercer orden se espera conseguir dos beneficios principales: Primero, una aceleración en la convergencia del algoritmo iterativo y por tanto una reducción en la complejidad de cálculo y en la pérdida de inteligibilidad; Segundo, parametrizar la señal de voz de una forma menos contaminada. En comparación a la estimación mediante estadísticas en 2º orden, obtenemos una sustancial mejora, aunque los formantes acusan más el efecto de estrechez y picado.

Cuando el ruido aditivo es AWGN a SNR=0dB, la mejora respecto al algoritmo de 2º orden es muy apreciable para cualquier número de iteraciones (tabla 1). Por supuesto, el éxito más rotundo se obtiene cuando el ruido es Gaussiano debido a las propiedades de la estimación de los cumulantes. Mientras que la mejora de 2º orden es

contante en cada iteración, aunque lenta, 3er orden da una gran mejora, unos 3dB, en tan solo dos iteraciones, obteniendo una convergencia mucho más rápida.

En las pruebas de audición se aprecia una mayor calidad en tercer orden con referencia al modelo clásico de segundo orden. Las dos primeras iteraciones usando el algoritmo de 2º orden contiene los efectos desagradables del ruido residual, convirtiéndose en las iteraciones posteriores en la aparición de un *ruido musical*. Este *ruido musical* residual se reduce un poco al pasar de la tercera a la cuarta iteración, pero se mantiene en un cierto nivel mientras la inteligibilidad disminuye en cada iteración. En comparación, el algoritmo de tercer orden contiene los efectos desagradables del ruido residual tan solo en la primera iteración. En la segunda iteración aparece *ruido musical* (a un nivel similar a la tercera iteración del caso de 2º orden). La voz mejorada tras la tercera iteración contiene sólo un poco de *ruido musical* y la pérdida de inteligibilidad es parecida a la de las cuatro primeras iteraciones con 2º orden, el *ruido musical* desaparece tras cuatro iteraciones. Resuminedo, tercer orden cancela los efectos desagradables del ruido residual y el *ruido musical* muy pronto, mientras 2º orden mantiene *ruido musical* después de varias iteraciones. Cuando el ruido aditivo proviene de un motor diesel, las pruebas de audición dan una mejora similar con una pérdida leve de calidad: perdura algo de *ruido musical* después de cuatro iteraciones \neq 1 tercer orden (SNR=0dB).

Estos efectos pueden verse en la tabla 1. La reducción del *ruido musical* se refleja especialmente en la reducción de la distancia espectral de Itakura, ya que esta pondera más las resonancias espectrales y estas frecuencias son poco distorsionadas en la aproximación por cumulantes. La pérdida de inteligibilidad se evalúa en la distancia Cepstral, donde una reducción del *ruido musical* se corresponde con una mejora en distorsión, ya que esta distancia pondera todo el espectro de manera más uniforme y además es más sensible a la distorsión en los valles y las zonas llanas del espectro de voz, y el mencionado efecto de picado del método de filtrado iterativo de Wiener [3] causa una mayor distorsión en esas zonas (fig.3).

3.2 Algoritmo híbrido

En [4], analizamos la convergencia del algoritmo de tercer orden de modo similar a segundo orden. La independencia del ruido introducida en el análisis por cumulantes de tercer orden acelera la convergencia, pero el filtrado de Wiener causa una distorsión espectral mayor, ya que la estimación del filtro de Wiener para tercer orden tiende a cancelar todas las frecuencias con SNR inferior a 5 ó 6 dB, en función de la independencia entre voz y ruido dada por el análisis de cumulantes. Además, el resto de las frecuencias se atenúan más que en el algoritmo de 2º orden. Esto agudiza el efecto de estrechez o picado de los formantes de la voz, limitando el número de iteraciones.

Afortunadamente los algoritmos de tercer orden dan una importante mejora con sólo 1 ó 2 iteraciones. El algoritmo híbrido propuesto en [4],[5] consiste en utilizar de 1 a 3 iteraciones usando modelado de autocorrelación después de una primera iteración basada en modelado AR de cumulantes de 3er orden (fig.3). Este método híbrido intenta combinar las ventajas presentadas por ambos métodos: buena convergencia y poca distorsión. Los test de audición, combinando dos iteraciones de tercer orden seguidas por dos iteraciones de segundo orden dan un nivel de *ruido musical* parecido al obtenido con tres iteraciones con el algoritmo de tercer orden y mejor inteligibilidad.

3.3 Algoritmo de tercer orden parametrizado

Otro camino que permite combinar las ventajas de ambos algoritmos, la introducción de dos parámetros α y β en el cálculo del filtro de Wiener. Así se considera la siguiente ecuación:

$$W_f(w) = \left(\frac{P_y}{P_y + \beta \cdot P_r} \right)^\alpha \quad (4)$$

donde los coeficientes a_k se calculan resolviendo la ecuación (3).

El parámetro α varía en el margen 0.5 a 1.4 (pasos de 0.1) y el parámetro β en el rango de 0.2 a 2.0 (paso=0.2). Fig.2 muestra la distancia Cepstral mínima para cada combinación de α y β dentro del intervalo (SNR=0dB, AGWN). Obtenemos varios mínimos locales en distancia [6]. Cuando α, β son pequeños, la mejora es muy pobre (la convergencia es muy lenta). Algunas veces llegamos a mínimos locales para valores pequeños de α y grandes de β , pero el número de iteraciones necesario para alcanzar el mínimo es alto (>10), ya que un parámetro α alto corresponde a una convergencia más rápida, aunque la distorsión crece también. Los mejores resultados utilizando un número razonable de iteraciones corresponden a valores de α en el margen .8 a 1.0 y de β en 1.0 a 1.4 [6].

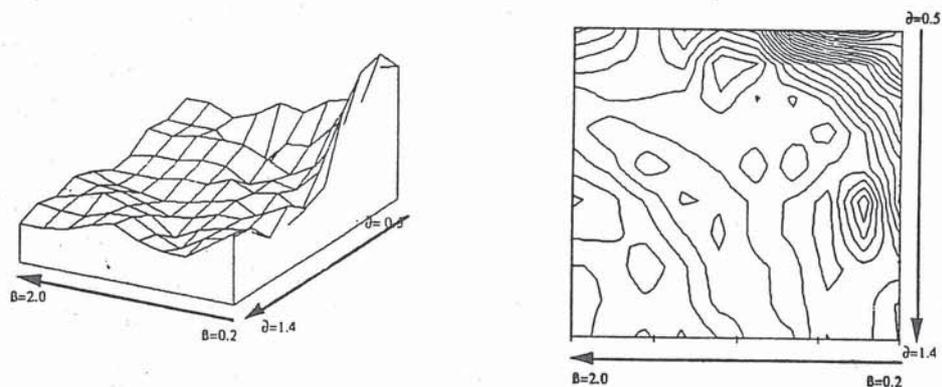


Fig.2: Distancias Cepstrales mínimas a SNR=0dB (AWGN). Valor máximo 9.11dB, valor mínimo 7.59dB.

3.4 Algoritmo de cuarto orden

Los coeficientes se calculan de manera similar a la presentada en la expresión (3) pero con el uso de cumulantes de cuarto orden en lugar de los de tercer orden. El algoritmo de cuarto orden mejora la señal de voz ruidosa con una convergencia tan rápida como el caso híbrido y un poco más lento que tercer orden, pero la distorsión resulta menos importante cuando se hace el test de audición, como consecuencia de la mejor conservación de las componentes simétricas de la voz (ver fig.3 y fig.4).

En la fig.4, se representa la distancia Cepstral frente a la SNR en el rango de 0dB a 18dB, considerando la primera iteración de los algoritmos anteriores. Segundo orden consigue una mejora uniforme independientemente del nivel y del tipo de ruido añadido a la frase de voz pronunciada libremente. El algoritmo de cuarto orden da una mayor mejora que 2º orden cuando la SNR es baja o media, para cualquier clase de ruido. La mejora del algoritmo de tercer orden disminuye cuando la SNR aumenta debido al dominio del efecto de distorsión espectral frente a la reducción de ruido y en ciertos casos da un resultado peor incluso al caso de 2º orden o no filtrado, siempre en un conjunto reducido de locutores considerando ruido de motor diesel.

4. CONCLUSIONES

Se han propuesto varios algoritmos basados en el filtrado iterativo de Wiener. La estimación del espectro de voz se obtiene mediante el modelado AR de voz basado en el análisis por cumulantes de tercer y cuarto orden para alcanzar el deseado desacople voz-ruido. Varias aproximaciones del algoritmo de Lim-Oppenheim usando estimación AR de cumulantes han sido comparadas con el método clásico de autocorrelación. Los algoritmos basados en cumulantes dan mejores resultados cuando el ruido es AGWN, aunque para ruido de motor la pérdida de calidad es poco notoria. Así el método híbrido y la parametrización del tercer orden son buenos en velocidad de convergencia, efecto distorsionante y complejidad computacional. El de cuarto orden necesita sólo de una primera iteración para obtener el mismo resultado que el método clásico de autocorrelación después de tres iteraciones, y a veces la distorsión implícita en el filtrado iterativo lleva a una menor mejora para cualquier número de iteraciones usando este método. Finalmente, la convergencia de los algoritmos iterativos basados en la estimación de los cumulantes AR es más rápido.

REFERENCIAS

- [1] C.L. Nikias, M.R. Raghuveer. "Bispectrum Estimation: A Digital Signal Processing Framework". Proc. of IEEE, pp 869-891. Julio 1987.
- [2] J.S. Lim and A.V. Oppenheim. "All-Pole Modeling of Degraded Speech". IEEE Trans. on ASSP, pp 197-210. Junio 1978.
- [3] J.H.L. Hansen and M.A. Clements. "Constrained Iterative Speech Enhancement with Applications to Speech Recognition". IEEE Trans. on Signal Processing, pp 795-805. Abril 1991.
- [4] E. Masgrau, J.M. Salavedra, A. Moreno, A. Ardanuy. "speech Enhancement by Adaptive Wiener Filtering based on Cumulant AR Modeling". Proc. ESCA Workshop on Speech Processing in Adverse Conditions, pp 143-146. Canes, France. Nov'92.

- [5] J.M. Salavedra, E. Masgrau, A. Moreno, X. Jové. "Comparison of different order cumulants in a Speech Enhancement System by adaptive Wiener Filtering". Proc. IEEE Signal Processing Workshop on HOS, pp 61-65. Junio'93. South Lake Tahoe, CA, USA.
- [6] J.M. Salavedra, E. Masgrau, A. Moreno, X. Jové. "Speech Enhancement based on parameterized third-order cumulant AR modeling". Proc. Inter. Conference on Signal Processing'93. Octubre 26-30, 1993. Beijing, China. To be published.

a)	SNR	SEGSN	ITAKU	COSH	CEPST
0 iter.	0.00	0.79	9.57	11.67	12.02
1 iter.	7.36	4.38	9.21	10.71	11.01
2 iter.	8.83	5.92	8.86	10.17	9.90
3 iter.	9.04	6.16	7.30	9.04	9.34
4 iter.	9.11	6.25	6.42	8.45	9.20

b)	SNR	SEGSN	ITAKU	COSH	CEPST
0 iter.	0.00	0.79	9.57	11.67	12.02
1 iter.	7.96	4.87	8.73	10.23	10.15
2 iter.	7.81	5.41	6.25	8.44	8.67
3 iter.	7.85	5.73	5.63	7.91	8.27
4 iter.	7.62	5.75	5.46	7.83	8.35

c)	SNR	SEGSN	ITAKU	COSH	CEPST
0 iter.	0.00	0.79	9.57	11.67	12.02
1 iter.	7.96	4.87	8.73	10.23	10.15
2 iter.	8.79	5.97	7.31	9.15	9.33
3 iter.	9.00	6.29	6.01	8.15	8.82
4 iter.	8.88	6.33	5.62	7.87	8.65

d)	SNR	SEGSN	ITAKU	COSH	CEPST
0 iter.	0.00	0.79	9.57	11.67	12.02
1 iter.	7.47	4.53	8.97	10.49	10.53
2 iter.	7.39	4.95	7.88	9.65	9.30
3 iter.	7.37	5.11	6.55	8.65	8.80
4 iter.	7.77	5.49	5.52	7.91	8.47

Tabla 1. Medidas usando algoritmos basados en: a) estadística 2º orden; b) cumulantes 3er orden; c) híbrido; d) cumulantes cuarto orden con SNR=0dB (AWGN).

Fig.3: Distancia Cepstral frente al número de iteraciones (SNR=0dB) usando algoritmos basados en: a) 2º orden (AR2), b) tercer orden (AR3), c) híbrido (AR3H), d) cuarto orden (AR4).

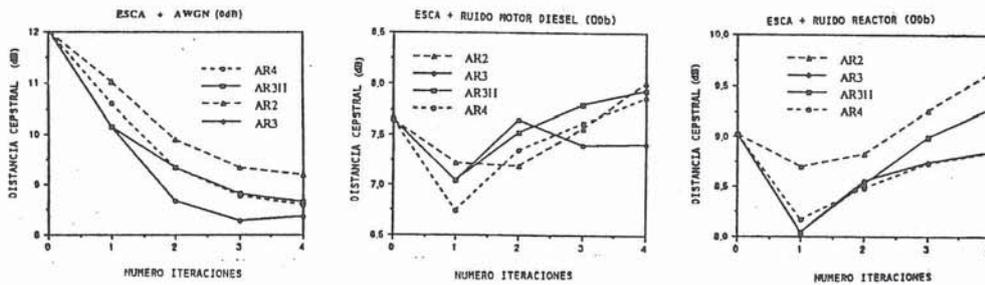


Fig.4: Distancia Cepstral (1ª iteración) frente a SNR usando algoritmos basados en: a) 2º orden (AR2), b) 3er orden (AR3), d) cuarto orden (AR4).

