

Umbral automático y segmentación en el "denoising"

S. Cerdá y J. Romero.

Laboratorio de Acústica. Departamento Física Aplicada. Universidad de Valencia.

Resumen

Para eliminar el ruido de una señal tenemos que tener en cuenta dos aspectos: el nivel de ruido y el tamaño de la señal. El nivel de ruido va a determinar que establezcamos un umbral que nos permita seleccionar la parte coherente de la señal y la parte "ruidosa". El tamaño de la señal es importante pues señales de larga duración son necesariamente tratadas con métodos de truncación de la señal. Al trabajar con trozos de la señal, suelen aparecer discontinuidades en el dominio temporal que producen molestos "clicks". En este trabajo establecemos un mecanismo automático de determinación del umbral de separación así como una forma de eliminar los indeseados efectos que produce la truncación de la señal.

Introducción

La eliminación de ruido en una señal digital, se plantea como la extracción a partir de la señal inicial $y_i = f_i + \lambda z_i$, de lo que denominamos la propia señal f_i . Para conseguir esto en el dominio temporal, necesitamos aplicar mecanismos de suavizado que si bien son efectivos para las componentes de altas frecuencias, no lo son tanto para las frecuencias bajas. Es el dominio de frecuencias el ruido blanco tiene una peculiaridad importante: su espectro está homogéneamente distribuido. El mecanismo general de eliminación de ruido (denoising)[1], consta de tres pasos:

1. Transformar la señal y en el dominio de frecuencias.
2. Actuar en ese dominio (denoising).

3. Transformar los coeficientes

tratados en el dominio temporal.

El procedimiento en general consiste en establecer un umbral de selección, el cual nos va a servir para extraer la parte de señal coherente [2] (es decir la señal que no es en sí misma ruido). El problema que analizamos se encuentra al aplicar el método general con señales de larga duración, la transformación es indispensable: es muy costoso operacionalmente y las peculiaridades locales de la señal se enmascararán entre sí. La utilización de ventanas trae consigo el problema de la generación de discontinuidades (extremos de las ventanas), que una vez realizado el proceso de denoising, producen "clicks" y extraños ruidos que bajan la calidad en el proceso [4]. En este trabajo ofrecemos un método de determinación local del umbral así como un mecanismo efectivo de eliminación de los efectos molestos que produce el trabajar con ventanas en señales de larga duración.

Caracterización del ruido blanco.

El ruido blanco está caracterizado por seguir una distribución normal $N(0,1)$. Denotamos entonces una muestra de ruido blanco como:

$$z_i \sim N(0,1) \quad (1)$$

Al trabajar con ruido blanco, si aplicamos una transformación ortogonal al vector de N valores, el resultado es otro ruido blanco [3]. La diferencia es una reescala de la distribución. Es decir ahora tendremos en general:

$$z_i \sim \lambda N(0,1) \quad (2)$$

Otra característica especial es que al realizar una transformada en frecuencias, obtenemos que la amplitud media en todo el rango de frecuencias es constante. Este último hecho, que realmente es consecuencia directa del primero, es el que vamos a utilizar para caracterizar al ruido blanco.

Dada una muestra de ruido blanco de tamaño N , podemos estimar los parámetros estadísticos fundamentales: la media y su dispersión.

$$\text{med}(x) = 1/N \sum x_i \quad (3)$$

$$\text{dis}(x) = (1/N \sum (x_i - \text{med}(x))^2)^{1/2} \quad (4)$$

Nuestro interés es establecer un nivel mínimo que nos permita decir que cualquier valor que esté distribuido con una distribución normal del tipo $\lambda N(0,1)$, esté por debajo de este límite.

Vamos a considerar exclusivamente aquellos valores que resulten ser picos en la muestra. Es decir dados tres elementos contiguos cualesquiera de la muestra seleccionaremos sólo aquellos que satisfagan:

$$x_{i-1} < x_i > x_{i+1} \quad (5)$$

La distribución de los picos así seleccionados la modelizamos con el siguiente razonamiento. Podemos estimar la distribución de los picos máximos como

$$Y \sim 1/3 F^2(y) f(y) \quad (6)$$

En la figura 3 mostramos la distribución $N(0,1)$ y la distribución de los picos (6).

Las características estadísticas de la distribución de los picos son:

$$\text{med}(y) = 0.8622435$$

$$\text{dis}(y) = 0.7457826 \quad (7)$$

En general trabajamos con muestras de una distribución del tipo $\lambda N(0,1)$, donde el parámetro λ es desconocido. Esto indica que este factor de amplitud nos va a modificar la media y la desviación estándar. Para esquivar este problema introducimos un parámetro ξ independiente de la amplitud:

$$\xi = \text{dis}(y) / \text{med}(y) \quad (8)$$

En tal caso el valor teórico de este parámetro para la distribución de los picos es:

$$\xi_{\text{Theor.}} = 0.87 \quad (9)$$

Los algoritmos de "denoising".

En 1984, Donoho y Johnstone establecieron el algoritmo de "shrinkage" como una nueva herramienta de eliminación de ruido [3]. El algoritmo general se puede describir en tres pasos:

1. Se descompone la señal en su serie de wavelet, utilizando el algoritmo de multiresolución.
2. Se aplica un umbral no lineal $\eta_t(y)$ a la señal y .
3. Se invierte el algoritmo de descomposición.

En la aplicación del umbral se toma [3]:

$$\eta_t(y) = \text{sign}|y| (|y| - t)_+ \quad (10)$$

Donde $\text{sign}(y)$ es el signo del valor de y ; $(|y| - t)_+$ está definido por

$$\begin{cases} 0 & \text{si } |y| \leq t \\ |y| - t & \text{si } |y| > t \end{cases} \quad (11)$$

Como vemos esto difiere del procedimiento de umbral simple donde se toma:

$$\begin{cases} 0 & \text{si } |y| \leq t \\ y & \text{si } |y| > t \end{cases} \quad (12)$$

Cuando para $\eta_t(y)$ toma la forma (14) tenemos el denominado soft-thresholding. Para $\eta_t(y)$ definida por (15) tenemos el Hard-thresholding. Existen otras elecciones de $\eta_t(y)$ que se desarrollaron después: Posteriormente se

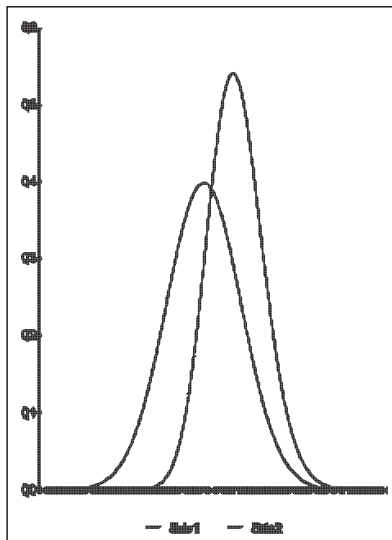


Figura 1.
Serie 1: distribución $N(0,1)$
Serie 2: distribución picos (8).

introdujeron otras elecciones para: el denominado non-negative garrote thresholding [6]; y otro mecanismo que tiene dos umbrales y recibió el nombre de firm-thresholding [5].

El valor teórico que se suele tomar para el umbral depende del tamaño de la muestra n y del nivel de ruido λ [7]:

$$t_n = (2\log(n))^{1/2} \lambda \quad (13)$$

Si consideramos este umbral y tenemos una muestra aleatoria de una distribución $N(0,1)$ de tamaño n , entonces

$$P\{\max_{i=1,n} |x_i| > (2\log(n))^{1/2}\} \rightarrow 0$$

cuando $n \rightarrow \infty$

Por tanto el algoritmo de "denoising" queda bien definido con los tres pasos especificados más arriba y tomando como umbral el valor (18).

Los problemas de los extremos.

Si queremos eliminar el ruido de una señal de unos 4 ó 5 minutos, utilizando una frecuencia de muestreo de 11025 Hz, nos vamos a encontrar con ficheros de unas $3 \cdot 10^6$ muestras. Así es necesario segmentar la señal en subseñales más pequeñas (ventanas). El algo-

ritmo de "denoising" se aplicará a estas ventanas. Al trabajar en el dominio de frecuencias, es conveniente para aplicar algoritmos rápidos trabajar con señales de tamaño potencia de 2. Como la señal va evolucionando en el tiempo, si tomamos ventanas muy grandes, las partes coherentes de la señal normalmente se superpondrán generando un enmascaramiento que se puede confundir con la presencia de ruido. En la práctica, trabajamos con ventanas de tamaño 1024.

El procedimiento más sencillo consiste por tanto en segmentar la señal en ventanas disjuntas (y si es necesario completar con ceros la última ventana). Aplicar el algoritmo de "denoising" a cada ventana. Unir las ventanas para obtener la señal completa tratada. El problema es la multiplicación de una señal continua por una función discontinua [4]. Esto provoca que tras el proceso de "denoising", uno oiga periódicos "clicks" [4].

Umbral automático y segmentación.

Hemos visto cómo podemos caracterizar el ruido blanco. En el desarrollo teórico que realizamos en el punto 2, vimos que la distribución de los máximos correspondía a una distribución normal con una media y una desviación (9). Introducimos un parámetro ξ independiente del nivel de ruido. Esto nos sirve para saber si la señal es ruido puro o no. Además nos va a permitir establecer el umbral automático para el proceso de "denoising". Para calcular el valor de este parámetro necesitamos seguir los pasos:

paso 1. Seleccionamos de la señal transformada solo aquellos valores que satisfacen la relación (5).

paso 2. Calculamos la media aritmética y la desviación media de los valores seleccionados en paso1.

Con estos dos pasos podemos establecer el valor del parámetro ξ . Pero además obtenemos un valor que nos servirá de gran utilidad: la media de los valores que satisfacen (5). Seguimos con:

paso 3. Si el parámetro ξ es menor o igual que el valor teórico establecido

en la sección 2, tratamos con señal ruidosa y la eliminamos completamente. En caso contrario aplicamos el soft-thresholding (u otro método uni-umbral), utilizando como valor del umbral: $t = 2 * \text{media} (14)$.

Por último aplicamos el algoritmo inverso de descomposición para obtener la señal limpia.

Una vez establecido el proceso de denoising que vamos a seguir, nos queda por tratar el problema de la segmentación. Nosotros proponemos el siguiente procedimiento:

paso 1. Seleccionamos un conjunto de 1024 muestras de la señal y_i .

paso 2. A partir de la muestra seleccionadas construimos un conjunto de muestras de doble tamaño (2048) de la siguiente forma:

$$\begin{aligned} x_{512+j} &= -y_{1+j} \\ j &= 0, 511; x_{512+j} = y_j \\ j &= 1, 1024; x_{1536+j} = -y_{1024-j} \\ j &= 0, 511 \end{aligned}$$

paso 3. Aplicamos una ventana de análisis de tipo Hanning a la señal obtenida en el paso 2 modificada. La mo-

dificación de dicha ventana consiste en que en vez de anularse en los extremos se introduzca un valor ϵ de forma que la ventana no tome un valor nulo en ningún punto.

paso 4. Aplicamos el algoritmo completo de denoising.

paso 5. Como hemos modificado la ventana de análisis y no se anula en ningún valor, multiplicamos la señal obtenida en el paso 4 por la inversa de la ventana de análisis.

paso 6. Finalmente nos quedamos con los 1024 valores centrados.

paso 7. Volvemos al paso 1 con las 1024 muestras consecutivas.

En el procedimiento descrito, hacemos un pretratamiento a la señal. En vez de analizar un trozo real de señal lo hacemos con la extensión impar de doble tamaño. Supongamos que estamos analizando una señal que es ruido blanco. Cuando construimos dicha extensión y pasamos a aplicar el algoritmo de "denoising", ¿obtendremos un valor del parámetro ξ que coincida con el teórico? La respuesta es no, pues la extensión impar no es realmente un ruido blanco. Tendremos por tanto

que corregir el valor teórico del parámetro ξ . En nuestro caso obtenemos $\xi = 0.67 \pm 0.04$.

Conclusiones.

En este trabajo presentamos dos nuevas opciones a la hora de realizar "denoising".

1. Una selección automática del umbral que permite determinar si una señal es totalmente ruidosa.
2. Un método de segmentación que elimina los efectos indeseados que dicho procedimiento conlleva.

En nuestros experimentos con señales de larga duración (más de una decena de segundos), los resultados son satisfactorios. Se reduce el nivel de ruido manteniendo las características que el ruido no enmascara totalmente. Se evitan los problemas de la segmentación haciendo casi inaudibles dichos efectos. Sin embargo, cuando hablamos de "denoising", no hay nada más eficaz que la escucha del resultado.

BIBLIOGRAPHY.

- [1] Donoho D. (1992) "De-noising by soft-thresholding". Technical Report Standford Univ.
- [2] Coifman R and Majid F (1993) "Adapted waveform analysis and denoising". in Meyer Y Roques S. (Eds) Progress in Wavelet Analysis and Applications. Ed Frontieres pp. 63-76.
- [3] Donoho D and Johnstone I. (1994) "Ideal Denoising in an orthogonal basis chosen from a library of bases". Comptes R Acad Sci Paris Serie I. vol 319 pp. 1317-1327.
- [4] Berger et al. (1994) "Removing noise from music using local trigonometric bases and wavelet packets". J. audio Eng Soc. vol 42 pp. 808-818.
- [5] Gao H and Bruce A. (1996) "Waveshrink with Firm shrinkage". Statistica Simica Research Report No 39. <ftp://ftp.statsci.com/pub/gao>
- [6] Gao H. (1997). "Wavelet shrinkage Denoising using the non-negative garrote". Preprint. <ftp://ftp.statsci.com/pub/gao>
- [7] Gao H. (1997) "Threshold Selection in WaveShrink". preprint <ftp://ftp.statsci.com/pub/gao>.