

Denoising de señales mediante modelos probabilísticos

Raúl Benítez Iglesias*

*Departament d'Enginyeria de Sistemes, Automàtica i Informàtica Industrial
Escola Universitària d'Enginyeria Tècnica Industrial de Barcelona
Universitat Politècnica de Catalunya
c/ Compte d'Urgell 187
08028 Barcelona*

La transformada wavelet discreta representa las señales como una expansión en términos de translaciones y dilaciones de una onda básica de soporte finito¹. Este tipo de representación no solo permite capturar las características espectrales de la señal, sino también eventos puntuales localizados en el tiempo en forma de transitorios o perturbaciones.

Como consecuencia, el análisis wavelet tiene numerosas aplicaciones en compresión de señal y eliminación de ruido (denoising), puesto que permite realizar una estimación de la componente ruidosa de la señal y separarla de la señal sin ruido. Una vez identificados los coeficientes de la expansión que corresponden al ruido, pueden eliminarse de la representación para recuperar la señal limpia de ruido mediante la transformación wavelet inversa (wavelet shrinkage)^{3,4}. Este procedimiento puede aplicarse a datos con ruido gaussiano aditivo, y la forma habitual para separar señal de ruido consiste en definir un umbral sobre los coeficientes wavelet a partir de una cierta función de riesgo estadística.

El objetivo de este trabajo es presentar un método robusto y adaptativo que permite estimar la componente ruidosa de la señal mediante la utilización de modelos probabilísticos basados en mezclas gaussianas². En caso que la señal solamente contenga ruido, los coeficientes wavelet se distribuirán de forma gaussiana. Si por contra, los datos contienen señal y ruido, la distribución de los coeficientes será la mezcla de una gaussiana (ruido) y una distribución uniforme (señal). Considerando estos dos modelos, es posible definir el proceso de estimación señal/ruido como un problema de selección de modelos probabilísticos en un contexto de inferencia estadística.

En particular, utilizamos el algoritmo de Expectación-Maximización para determinar los coeficientes de la mezcla gaussiana que maximizan la verosimilitud (likelihood) respecto a los datos, y el índice BIC (Bayes Information Criterion) para seleccionar el modelo que mejor describe los datos. Una vez definido el modelo óptimo, los coeficientes wavelet son clasificados como señal o ruido siguiendo un determinado criterio de decisión basado en las probabilidades de pertenencia a cada uno de los componentes de la mezcla. Diferentes criterios de selección proporcionan estimaciones con diversos grados de sensibilidad y especificidad en la eliminación del ruido.

El método ha sido testeado de forma cuantitativa mediante simulaciones Monte-Carlo con un conjunto de señales de prueba que representan una gran variedad de situaciones (transitorios, spikes, alta densidad espectral, etc), presentando un rendimiento superior al de los métodos tradicionales basados en estimación de umbrales (Umbrales universales de Donoho³, Stein's Unbiased Risk Estimate, False Discovery Rate, Minimax, etc).

* raul.benitez@upc.edu

¹ S. Mallat, "A Wavelet Tour of Signal Processing" (2nd Edition), Academic Press, 1999.

² R. Benítez, Z. Nenadic, "Robust Unsupervised Detection of Action Potentials with Probabilistic Models", *IEEE Transactions in Biomedical Engineering* (2007), in press.

³ D.L. Donoho, I.M. Johnstone (1994), "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, pp. 425-455.

⁴ D.L. Donoho, I.M. Johnstone (1995), "Adapting to unknown smoothness via wavelet shrinkage via wavelet shrinkage," *JASA*, vol. 90, 432, pp. 1200-1224.