



UNIVERSIDAD DE JAÉN
Escuela Politécnica Superior de Linares

Trabajo Fin de Grado

**SEPARACIÓN DE FUENTES SONORAS
EMITIDAS POR EL CORAZÓN Y PULMONES
EN UN SER HUMANO APLICADO AL
DIAGNÓSTICO DE ENFERMEDADES
CARDIOVASCULARES Y RESPIRATORIAS.**

Alumno: Albadalejo Lijarcio, Juan Luis

Tutor: Cañadas Quesada, Francisco Jesús
Depto.: Ingeniería Telecomunicación

Septiembre, 2016



UNIVERSIDAD DE JAÉN

Escuela Politécnica Superior de Linares

Trabajo Fin de Grado

SEPARACIÓN DE FUENTES SONORAS EMITIDAS POR EL CORAZÓN Y PULMONES EN UN SER HUMANO APLICADO AL DIAGNÓSTICO DE ENFERMEDADES CARDIOVASCULARES Y RESPIRATORIAS.

RESUMEN

En los seres humanos, tanto el corazón como los pulmones juegan un papel fundamental en el sistema respiratorio y cardiovascular. Uno de los métodos para analizar el estado de salud de los pulmones y corazón de posibles enfermedades que existen sobre ellos de manera no invasiva es el análisis de las señales sonoras emitidas por los mismos. Así, una señal emitida por los pulmones o corazón sin interferencia de ninguna otra fuente aumenta la exactitud de un buen diagnóstico por parte del especialista en medicina. Sin embargo, tal situación es difícil de conseguir en la vida real ya que la mayor parte del tiempo el sonido del corazón se puede considerar interferencia en la escucha del sonido del pulmón y viceversa. Este Trabajo Fin de Grado está centrado en el procesado de audio, concretamente en la separación de fuentes sonoras. Se implementará un sistema que sea capaz de separar la señal sonora del corazón de la señal sonora de los pulmones emitida por un ser humano. Para tal fin, se utilizarán herramientas de procesado de señal aplicadas a la señal mezcla de sonidos de corazón y pulmón. El proyecto se implementará utilizando el entorno MATLAB.

Índice de contenido

ÍNDICE DE FIGURAS.....	3
ÍNDICE DE TABLAS.....	8
SIGLAS	9
1. INTRODUCCIÓN	11
1.1. MOTIVACIÓN	11
1.2. CONCEPTOS PREVIOS.....	13
1.2.1. <i>Relación señal a ruido</i>	13
1.2.2. <i>Información en el dominio temporal</i>	14
1.2.2.1. Correlación cruzada y autocorrelación	14
1.2.3. <i>Información en el dominio de la frecuencia</i>	16
1.2.3.1. Transformada Discreta de Fourier	16
1.2.3.2. Densidad Espectral de Potencia	17
1.2.4. <i>Información en el dominio tiempo-frecuencia</i>	18
1.2.4.1. Transformada de Fourier de Tiempo Corto	18
1.2.4.2. Máscara	20
1.3. SONIDO.....	23
1.3.1. <i>Frecuencia fundamental, altura, intensidad, duración y timbre</i>	23
1.3.2. <i>Sonido monoaural, estereofónico y multicanal</i>	24
1.3.3. <i>Sonidos del corazón</i>	24
1.3.4. <i>Sonidos del pulmón</i>	30
2. OBJETIVO.....	36
3. ESTADO DEL ARTE	37
3.1. FILTRADO PASO-ALTO.....	37
3.2. FILTRADO ADAPTATIVO.....	37
3.2.1. <i>Gradiente estocástico</i>	40
3.2.2. <i>Estimación de mínimos cuadrados</i>	41
3.3. NLM.....	41
3.4. FILTRADO EN EL DOMINIO TRANSFORMADO	42
3.4.1. <i>Basado en STFT</i>	42
3.4.2. <i>Basado en DWT</i>	43
4. IMPLEMENTACIÓN.....	44
4.1. ALE.....	44
4.1.1. <i>Desarrollo matemático</i>	44
4.1.1.1. Filtro de Wiener	44
4.1.1.2. Método de máxima pendiente.....	48
4.1.1.3. LMS.....	49
4.1.1.4. VSS-LMS.....	51
4.1.1.5. Señal de referencia	52
4.2.2. <i>Señal sinusoidal + AWGN</i>	53
4.2.3. <i>Sonidos cardiacos y pulmonares</i>	56

4.2.	NLM.....	65
4.2.1.	<i>Desarrollo matemático</i>	65
4.2.2.	<i>Coste computacional</i>	66
4.2.2.1.	Reducción radio de búsqueda	67
4.2.2.2.	Ritmo cardiaco.....	68
4.2.3.	<i>Sonidos cardiacos y pulmonares</i>	69
5.	EVALUACIÓN	84
5.1.	BASE DE DATOS.....	84
5.2.	MÉTRICAS.....	85
5.3.	OPTIMIZACIÓN	86
5.3.1.	<i>Optimización ALE VSS-LMS y ALE LMS</i>	86
5.3.2.	<i>Optimización NLM</i>	92
5.4.	RESULTADOS	95
6.	CONCLUSIONES.....	99
7.	LÍNEAS FUTURAS.....	101
8.	INTERFAZ GRÁFICA	102
8.1.	SEPARACION_HSLS.....	102
8.2.	CALIDAD_SEPARACION_HSLS.....	106
9.	BIBLIOGRAFÍA.....	108

Índice de figuras

Figura 1.1 - Estetoscopio acústico modelo 3M™ Littmann® Classic III™ con un coste aproximado de 100 € (izquierda) y estetoscopio electrónico modelo 3M™ Littmann® Electronic Stethoscope Model 3100 con un coste aproximado de 300 € (derecha) [1]	11
Figura 1.2 - Señal temporal (izquierda) y su autocorrelación (derecha). Se aprecia como aparecen los máximos en aquellos desplazamientos para los cuales coinciden los distintos pulsos, es decir, en los periodos de la señal.	15
Figura 1.3 - Señal temporal de 1 y 10 Hz (izquierda) y el módulo de su FFT (derecha). Se aprecia como aparecen dos deltas, es decir, la FFT de una senoide, en cada una de las frecuencias de las sinusoides que componen la señal de entrada, en este caso 1 y 10 Hz.	17
Figura 1.4 - Señal temporal de 10 Hz (izquierda) y su PSD (derecha). Como se puede observar en la PSD, toda la potencia se concentra en la frecuencia de 10 Hz, que es la frecuencia de la señal temporal.	18
Figura 1.5 - Espectrograma de la señal temporal vista en la Figura 1.3. Se puede observar a partir del espectrograma, como la mayor parte de la energía (blanco) aparece en las frecuencias de 1 y 10 Hz, información que se podía obtener a partir de la FFT, pero con la ventaja de poder ver además como varía esta energía entre las distintas frecuencias a lo largo del tiempo.....	19
Figura 1.6 - Separación de fuentes usando una máscara tiempo-frecuencia. A) Espectrograma de la señal mezcla. B) Máscara Hard (o binaria). C) Espectrograma de la Fuente 1 obtenido al multiplicar la máscara por el espectrograma de la señal mezcla. D) Espectrograma de la Fuente 2 obtenido a partir de la resta del espectrograma de la señal mezcla y el espectrograma de la Fuente 1 obtenida anteriormente.....	20
Figura 1.7 - A) Máscara Wiener de la Fuente 1. B) Máscara Wiener de la Fuente 2. C) Máscara Hard de la Fuente 1. D) Máscara Hard de la Fuente 2.	22
Figura 1.8 - Anatomía del corazón. En la figura de la izquierda se muestran las válvulas aurículo-ventriculares (tricúspide y mitral) que regulan el paso de sangre desde las aurículas hacia los ventrículos mientras que en la figura de la derecha se muestran las válvulas semilunares (pulmonar y aórtica) que controlan el paso de sangre desde los ventrículos hacia el pulmón y la arteria aorta respectivamente [7].	25
Figura 1.9 - Sonidos S1 y S2. A) Señal en el dominio del tiempo en la que aparecen remarcados los sonidos S1 y S2, así como el tiempo que los separa, denominados sístole y diástole. B) Espectrograma (el blanco indica la presencia de energía). C) Zoom del espectrograma entre 0 y 1000 Hz.....	27
Figura 1.10 - Sonido S3. A) Señal en el dominio del tiempo en la que aparece remarcado el sonido S3. B) Espectrograma. C) Zoom del espectrograma entre 0 y 1000 Hz.	27
Figura 1.11 - Sonido S4. A) Señal temporal en la que aparece remarcado el sonido S4. B) Espectrograma. C) Zoom del espectrograma entre 0 y 1000 Hz.	28
Figura 1.12 - Clic de eyección. A) Señal temporal en la que aparece remarcado el clic de eyección. B) Espectrograma. C) Zoom del espectrograma entre 0 y 1000 Hz.	28

Figura 1.13 - Chasquido de apertura. A) Señal temporal en la que aparece remarcado el chasquido de apertura. B) Espectrograma. C) Zoom del espectrograma entre 0 y 1000 Hz.	29
Figura 1.14 – Soplo. A) Señal temporal en la que aparece remarcado el soplo. B) Espectrograma. C) Zoom del espectrograma entre 0 y 1000 Hz.	29
Figura 1.15 - Anatomía del pulmón. Se pueden observar las vías aéreas (tráquea, bronquios, ...) por donde circula el aire y cuyo flujo genera una serie de turbulencias que originan vibraciones percibidas como sonidos [13].	30
Figura 1.16 - Sonido vesicular. A) Señal temporal. B) Espectrograma. C) Zoom del espectrograma entre 0 y 2000 Hz.	32
Figura 1.17 - Sonido bronquial. A) Señal temporal. B) Espectrograma. C) Zoom del espectrograma entre 0 y 2000 Hz.	32
Figura 1.18 - Sonido broncovesicular. A) Señal temporal. B) Espectrograma. C) Zoom del espectrograma entre 0 y 2000 Hz.	33
Figura 1.19 – Crepitaciones. A) Señal temporal. B) Espectrograma. C) Zoom del espectrograma entre 0 y 2000 Hz.	33
Figura 1.20 – Sibilancias. A) Señal temporal. B) Espectrograma. C) Zoom del espectrograma entre 0 y 2000 Hz.	34
Figura 1.21 – Roncus. A) Señal temporal. B) Espectrograma. C) Zoom del espectrograma entre 0 y 2000 Hz.	34
Figura 1.22 - Roce pleural. A) Señal temporal. B) Espectrograma. C) Zoom del espectrograma entre 0 y 2000 Hz.	35
Figura 3.1 - Diagrama de bloques ANC [16]. En este esquema, la señal deseada “Primary input” está formada por la suma del sonido cardíaco “Signal source” y el sonido pulmonar “Noise source”. Se toma como señal de referencia “Reference input” una señal parecida al sonido cardíaco y que por tanto este ampliamente correlacionada con este. A) Salida del filtro donde se obtiene el sonido cardíaco recuperado. B) Salida del sistema (señal de error) donde se obtiene el sonido pulmonar recuperado.	38
Figura 3.2 - Diagrama de bloques ALE [16]. Se observa como a diferencia del esquema ANC, en este caso únicamente se dispone de una señal de entrada “Primary input”, formada por la mezcla del sonido cardíaco “Signal source” y el sonido pulmonar “Noise source”, a partir de la cual se obtiene la señal de referencia introduciendo un retardo Δ . A) Salida del filtro donde se obtiene el sonido cardíaco recuperado. B) Salida del sistema donde se obtiene el sonido pulmonar recuperado.	39
Figura 3.3 - Señal ECG (izquierda) y señal ECG modificada (derecha) [19]. Puesto que el sonido cardíaco, como se vio en el capítulo 1, está formado principalmente por los sonidos S1 y S2, es necesario realizar una modificación a la señal ECG original de manera que la señal de referencia esté formada por dos picos que coincidan con los dos principales sonidos del sonido cardíaco.	40
Figura 4.1 - Diagrama de bloques filtro de Wiener [36]. La señal deseada $x[n]$ se corresponde con la señal mezcla y la señal de referencia $u[n]$ debe ser una señal con un gran parecido al sonido cardíaco de la señal mezcla, de este modo, la señal de salida $y[n]$ se corresponderá con el sonido cardíaco recuperado y la señal de error $e[n]$ se corresponderá con el sonido pulmonar recuperado.	45

Figura 4.2 - Señal sinusoidal de 10 y 30 Hz (arriba) y señal ruidosa SNR=5dB (abajo)	53
Figura 4.3 - Autocorrelación de la señal sinusoidal de 10 y 30 Hz (arriba) y autocorrelación del ruido AWGN (abajo).....	54
Figura 4.4 – Señal sinusoidal recuperada (arriba) y ruido recuperado (abajo) utilizando un filtro ALE VSS-LMS.....	54
Figura 4.5 - Respuesta en frecuencia del filtro ALE VSS-LMS con L=100 con dos lóbulos principales en las frecuencias de 10 y 30 Hz.....	55
Figura 4.6 - Respuesta en frecuencia del filtro ALE VSS-LMS con L=50 con dos lobulos principales en las frecuencias de 10 y 30 Hz.....	55
Figura 4.7 - Sonido cardiaco (arriba), sonido pulmonar (medio) y señal mezcla SNR=5dB (abajo). Estos ficheros de audio se adjuntan en la carpeta de nombre “Software y Ficheros de Audio” anexa a la memoria.	56
Figura 4.8 – Autocorrelación del sonido cardiaco visto en la Figura 4.7	57
Figura 4.9 - Autocorrelación del sonido pulmonar visto en la Figura 4.7	57
Figura 4.10 - Sonido cardiaco recuperado (arriba) y sonido pulmonar recuperado (abajo) utilizando un filtro ALE VSS-LMS.....	58
Figura 4.11 - Variación del parámetro μ utilizando un filtro ALE VSS-LMS	58
Figura 4.12 - Máscara de corazón (arriba) y máscara de pulmón (abajo) obtenidas a partir de las salidas tras aplicar un filtro ALE VSS-LMS	60
Figura 4.13 - Sonido cardiaco recuperado (arriba) y sonido pulmonar recuperado (abajo) tras aplicar una máscara Wiener a la salida del filtro ALE VSS-LMS. Estos ficheros de audio se adjuntan en la carpeta de nombre “Software y Ficheros de Audio” anexa a la memoria	61
Figura 4.14 - PSD del sonido de corazón original, sonido de corazón recuperado y sonido de corazón recuperado tras aplicar una máscara Wiener a la salida del filtro ALE VSS-LMS	62
Figura 4.15 - PSD del sonido de pulmón original, sonido de pulmón recuperado y sonido de pulmón recuperado tras aplicar la máscara Wiener a la salida del filtro ALE VSS-LMS	62
Figura 4.16 - Espectrogramas del sonido de corazón original (arriba) y sonido de pulmón original (abajo) vistos en la Figura 4.7	63
Figura 4.17 - Espectrogramas del sonido de corazón recuperado (arriba) y sonido de pulmón recuperado (abajo) utilizando un filtro ALE VSS-LMS.....	63
Figura 4.18 - Espectrogramas del sonido de corazón recuperado (arriba) y sonido de pulmón recuperado (abajo) tras aplicar una máscara Wiener a la salida del filtro ALE VSS-LMS	64
Figura 4.19 - Ejemplo de dos segmentos a comparar para P=2 muestras.....	66
Figura 4.20 - Ejemplo de la reducción del radio de búsqueda para M=4 muestras.....	67
Figura 4.21 - Sonido cardiaco (arriba), sonido pulmonar (medio) y mezcla SNR=5dB (abajo). Estos ficheros de audio se adjuntan en la carpeta de nombre “Software y Ficheros de Audio” anexa a la memoria.	69
Figura 4.22 - Señal mezcla mostrada en la Figura 4.21 filtrada entre 20 y 150 Hz	70
Figura 4.23 - Autocorrelación de la señal mezcla filtrada vista en la Figura 4.22	70
Figura 4.24 - Periodos de la señal mezcla vista en la Figura 4.21 utilizando el ritmo cardiaco	71

Figura 4.25 - Sonido de corazón recuperado (arriba) y sonido de pulmón recuperado (abajo) utilizando un filtro NLM	72
Figura 4.26 - Máscara del sonido de corazón (arriba) y del sonido de pulmón (abajo) obtenidas a partir de las salidas tras aplicar un filtro NLM	73
Figura 4.27 - Sonido de corazón recuperado (arriba) y sonido de pulmón recuperado (abajo) tras aplicar una máscara Wiener a la salida del filtro NLM. Estos ficheros de audio se adjuntan en la carpeta de nombre “Software y Ficheros de Audio” anexa a la memoria.	73
Figura 4.28 - PSD del sonido de corazón original, sonido de corazón recuperado utilizando un filtro NLM y del sonido de corazón recuperado tras aplicar una máscara Wiener a la salida del filtro NLM.....	74
Figura 4.29 - PSD del sonido de corazón original, sonido de corazón recuperado utilizando un filtro NLM y del sonido de corazón recuperado tras aplicar una máscara Wiener a la salida del filtro NLM.....	74
Figura 4.30 - Espectrograma del sonido de corazón original (arriba) y espectrograma del sonido de pulmón original (abajo).....	75
Figura 4.31 - Espectrograma corazón recuperado (arriba) y espectrograma pulmón recuperado (abajo) utilizando NLM.....	75
Figura 4.32 - Espectrograma corazón recuperado (arriba) y espectrograma pulmón recuperado (abajo) utilizando NLM tras aplicar la máscara	76
Figura 4.33 - Sonido cardiaco (arriba), sonido pulmonar (medio) y mezcla SNR=5dB (abajo). Estos ficheros de audio se adjuntan en la carpeta de nombre “Software y Ficheros de Audio” anexa a la memoria.	77
Figura 4.34 - Autocorrelación del sonido cardiaco visto en la Figura 4.33	77
Figura 4.35 - Autocorrelación de la señal temporal vista en la Figura 4.33 tras eliminar el primer segmento	78
Figura 4.36 – División en periodos de la señal mezcla vista en la Figura 4.33 utilizando el primer método (arriba) y utilizando el segundo método (abajo).	79
Figura 4.37 - Solapamiento de dos de los “periodos” obtenidos tras realizar la separación del sonido cardiaco visto en la Figura 4.33	79
Figura 4.38 - Sonido cardiaco recuperado (arriba) y sonido pulmonar recuperado (abajo) utilizando un filtro NLM.....	80
Figura 4.39 - Sonido cardiaco recuperado (arriba) y sonido pulmonar recuperado (abajo) utilizando un filtro NLM tras aplicar la máscara. Estos ficheros de audio se adjuntan en la carpeta de nombre “Software y Ficheros de Audio” anexa a la memoria	80
Figura 4.40 - PSD del sonido de corazón original, sonido de corazón recuperado y sonido de corazón recuperado tras aplicar una máscara Wiener a la salida del filtro NLM	81
Figura 4.41 - PSD del sonido de pulmón original, sonido de pulmón recuperado y sonido de corazón recuperado tras aplicar una máscara Wiener a la salida del filtro NLM	81
Figura 4.42 - Espectrogramas del sonido de corazón original (arriba) y sonido de pulmón original (abajo)	82
Figura 4.43 - Espectrogramas del sonido de corazón recuperado (arriba) y sonido de pulmón recuperado (abajo) utilizando un filtro NLM	82

Figura 4.44 - Espectrogramas del sonido de corazón recuperado (arriba) y sonido de pulmón recuperado (abajo) tras aplicar una máscara de Wiener a la salida del filtro NLM	83
Figura 5.1 - Optimización del parámetro Δ en el método ALE VSS-LMS	87
Figura 5.2 - Optimización del parámetro L en el método ALE VSS-LMS (en la derecha, el eje de ordenadas se corresponde con el tiempo que tarda en procesar un segundo de la señal de entrada)	88
Figura 5.3 - Optimización del parámetro α en el método ALE VSS-LMS	89
Figura 5.4 - Optimización del parámetro γ en el método ALE VSS-LMS	89
Figura 5.5 - Optimización del parámetro μ_{min} en el método ALE VSS-LMS	90
Figura 5.6 - Optimización del parámetro μ_{max} en el método ALE VSS-LMS	91
Figura 5.7 - Optimización del parámetro h en el método NLM	93
Figura 5.8 - Optimización del parámetro P en el método NLM (en la derecha, el eje de ordenadas se corresponde con el tiempo que tarda en procesar un segundo de la señal de entrada)	93
Figura 5.9 - Optimización del parámetro T en el método NLM (en la derecha, el eje de ordenadas se corresponde con el tiempo que tarda en procesar un segundo de la señal de entrada)	94
Figura 5.10 - Resultados finales de los métodos ALE LMS, ALE VSS-LMS, NLM y comparación con NMF No Supervisado	96
Figura 8.1 - Pantalla principal de la interfaz Separacion_HSLS	102
Figura 8.2 - Cargar archivo de sonido en la interfaz Separacion_HSLS	103
Figura 8.3 - Selección método, máscara y parámetros en la interfaz Separacion_HSLS ..	104
Figura 8.4 - Realizar filtrado en la interfaz Separacion_HSLS	104
Figura 8.5 - Sonidos de corazón y de pulmón recuperados en la interfaz Separacion_HSLS	105
Figura 8.6 - Representación en el dominio de la frecuencia en la interfaz Separacion_HSLS	105
Figura 8.7 - Pantalla principal de la interfaz Calidad_Separacion_HSLS	106
Figura 8.8 - Cargar archivos en la interfaz Calidad_Separacion_HSLS	107
Figura 8.9 - Resultados en la interfaz Calidad_Separacion_HSLS	107

Índice de tablas

Tabla 4.1 - Resumen de los parámetros utilizados en el filtro ALE LMS y ALE VSS-LMS.....	64
Tabla 4.2 - Resumen de los parámetros utilizados en el filtro NLM.....	83
Tabla 5.1 - Valor óptimo del parámetro Δ en el método ALE VSS-LMS.....	87
Tabla 5.2 - Valor óptimo del parámetro L en el método ALE VSS-LMS.....	88
Tabla 5.3 - Valor óptimo del parámetro α en el método ALE VSS-LMS.....	89
Tabla 5.4 - Valor óptimo del parámetro γ en el método ALE VSS-LMS.....	90
Tabla 5.5 - Valor óptimo del parámetro $\mu_{\min}$ en el método ALE VSS-LMS.....	90
Tabla 5.6 - Valor óptimo del parámetro $\mu_{\max}$ en el método ALE VSS-LMS.....	91
Tabla 5.7 - Valores óptimos de los parámetros en el método ALE LMS.....	92
Tabla 5.8 - Valor óptimo del parámetro h en el método NLM.....	93
Tabla 5.9 - Valor óptimo del parámetro P en el método NLM.....	94
Tabla 5.10 - Resultados finales de los métodos ALE LMS, ALE VSS-LMS y NLM (SDR, SIR y SAR).....	96
Tabla 5.11 - Resultados finales de los métodos ALE VSS-LMS y NLM (tiempo que se tarda en procesar un segundo de la señal de entrada).....	98

Siglas

ALE	Adaptive Line Enhancer
ANC	Adaptive Noise Canceller
AWGN	Additive White Gaussian Noise
DFT	Discrete Fourier Transform
DWT	Discrete Wavelet Transform
ECG	ElectroCardioGram
EMSE	Excess Mean Squared Error
FIR	Finite Impulse Response
FOS	Fourth Order Statistics
ICA	Independent Component Analysis
IDFT	Inverse Discrete Fourier Transform
IDWT	Inverse Discrete Wavelet Transform
ISTFT	Inverse Short Time Fourier Transform
LECG	Laplacian Electrocardiogram
LMS	Least Mean Squares
MSE	Mean Squared Error
NLM	Non Local Mean
NLMS	Normalized Least Mean Squares
NMF	Non negative Matrix Factorization
OMS	Organización Mundial de la Salud
PSD	Power Spectrum Density
RLS	Recursive Least Squares

ROKF	Reduced Order Kalman Filter
SAR	Source to Artifacts Ratio
SDR	Sources to Distortion Ratio
SIANC	Single Input Adaptive Noise Canceller
SIR	Source to Interference Ratio
SNR	Signal to Noise Ratio
STFT	Short Time Fourier Transform
VSS-LMS	Variable Step Size Least Mean Squares

CAPÍTULO 1

1. Introducción

1.1. Motivación

Las enfermedades cardiovasculares son, según la OMS (*Organización Mundial de la Salud*), la principal causa de muerte en todo el mundo. Se estima que en 2012 murieron alrededor de 17,5 millones de personas por esta causa. Los motivos más importantes del incremento de este tipo de enfermedades son, entre otros, una mala dieta, inactividad física y un elevado consumo de tabaco y alcohol.

Entre las principales herramientas médicas para el diagnóstico de este tipo de enfermedades se encuentra el estetoscopio (ver **Figura 1.1**). Gracias a este utensilio, el médico puede, mediante una técnica conocida como auscultación, escuchar los sonidos emitidos por los órganos, especialmente corazón y pulmones, y detectar posibles anomalías.



Figura 1.1 - Estetoscopio acústico modelo 3M™ Littmann® Classic III™ con un coste aproximado de 100 € (izquierda) y estetoscopio electrónico modelo 3M™ Littmann® Electronic Stethoscope Model 3100 con un coste aproximado de 300 € (derecha) [1]

La auscultación mediante estetoscopio ha sido y sigue siendo un método muy usado en medicina para realizar una primera observación del paciente. Entre sus principales ventajas, destacan el hecho de que es un método no invasivo y bajo coste. No obstante, existen algunos inconvenientes que pueden dificultar en gran medida esta tarea.

La principal dificultad es que ambas señales se encuentran solapadas temporal y espectralmente, por lo que la técnica utilizada debe ser capaz de solventar este problema y realizar la separación sin perder información de la señal deseada que puede ser importante para el diagnóstico del paciente. Este solapamiento tiempo-frecuencia puede provocar el enmascaramiento de partes útiles de la señal de interés, desviando de este modo la atención del especialista. Evidentemente, no es una solución adecuada que el paciente aguante la respiración puesto que puede llevar a un cambio de la actividad cardíaca, además, esta solución sería inviable en algunas personas, principalmente en niños.

El análisis computarizado permitiría, en un futuro, realizar un diagnóstico de los problemas cardíacos y pulmonares de un paciente a partir de los sonidos obtenidos en la auscultación. Los sonidos cardíacos y pulmonares, que han de ser previamente separados, podrían ser analizados, determinando de este modo si el paciente presenta una anomalía y, de ser así, de qué tipo de anomalía podría tratarse. Esto ayudaría en gran medida al médico a la hora de llevar a cabo el diagnóstico, convirtiendo la auscultación en una tarea algo menos subjetiva.

Todo esto ha llevado a realizar numerosas investigaciones con el objetivo de desarrollar métodos que permitan llevar a cabo la separación de ambos sonidos y a la creación de estetoscopios electrónicos (ver **Figura 1.1**), los cuales digitalizan y procesan la señal de entrada optimizando de este modo la auscultación. Entre las principales ventajas de los estetoscopios electrónicos se pueden citar las siguientes [2]:

- **Sonido amplificado:** Pueden amplificar la señal sonora permitiendo escuchar aquellos sonidos más leves y que serían realmente difíciles de detectar con un estetoscopio acústico.
- **Reducción del ruido ambiente:** Permiten reducir el ruido ambiente pudiendo escuchar más claramente los sonidos del paciente.
- **Grabación y reproducción:** Ofrecen la posibilidad de realizar una grabación de la auscultación, así como la posibilidad de reproducirla posteriormente.

- **Separación de sonidos cardiacos y pulmonares:** La aparición de este tipo de estetoscopios posibilita la separación de los sonidos cardiacos y pulmonares, lo que podría suponer una enorme mejora en la auscultación y por tanto en la precisión del diagnóstico. Este trabajo se centrará en el desarrollo de algunos de los métodos para la separación de ambos sonidos a partir de la señal que se obtendría con un estetoscopio electrónico.
- **Información visual:** Algunos modelos cuentan con una pantalla a través de la cual se puede mostrar información relevante de la auscultación como, por ejemplo, el ritmo cardiaco del paciente.

Anteriormente, se han descrito las principales ventajas del estetoscopio electrónico, sin embargo, estos estetoscopios también presentan una serie de desventajas, entre las cuales se pueden citar las siguientes [2]:

- **Peso y movilidad:** En la actualidad resulta difícil entrar a un centro sanitario y no ver a un médico o enfermero con un estetoscopio siempre a mano, normalmente, colgado del cuello. Un estetoscopio electrónico necesita de un mayor número de componentes que un estetoscopio acústico lo que lo convierte en un elemento más pesado, reduciendo notablemente su movilidad.
- **Coste:** A pesar de haber reducido notablemente el precio en los últimos años, los estetoscopios electrónicos siguen presentando un precio muy alto respecto a los estetoscopios acústicos.

1.2. Conceptos previos

En este apartado se describirán varios conceptos necesarios para entender el desarrollo de este trabajo, así como algunos de los procedimientos que serán llevados a cabo más adelante.

1.2.1. Relación señal a ruido

La SNR (*Signal-to-Noise Ratio*) no es más que una medida para cuantificar la diferencia entre el nivel de potencia de la señal y el nivel de potencia del ruido. Se mide en *dB* y se calcula utilizando la siguiente ecuación:

$$SNR(dB) = 10 \log_{10} \frac{\int_{-\infty}^{\infty} \text{señal}^2(t) dt}{\int_{-\infty}^{\infty} \text{ruido}^2(t) dt} \quad (1.1)$$

En el caso de trabajar con señales discretas de tamaño N , la SNR se calcularía de la siguiente forma:

$$SNR(dB) = 10 \log_{10} \frac{\sum_{n=1}^N \text{señal}^2[n]}{\sum_{n=1}^N \text{ruido}^2[n]} \quad (1.2)$$

Se considera ruido a toda aquella señal no deseada que aparece en cualquier otra señal que si resulta de interés. En este trabajo se considerará la SNR como la relación entre la potencia del sonido cardiaco y el sonido pulmonar. Es importante conocer este concepto pues será utilizado más adelante de manera que se pueda comprobar el funcionamiento de cada uno de los métodos desarrollados ante un determinado valor de SNR.

1.2.2. Información en el dominio temporal

1.2.2.1. Correlación cruzada y autocorrelación

La correlación cruzada es una operación que ofrece una medida de la similitud existente entre dos señales. Sean dos señales $x(t)$ e $y(t)$, se calcula la correlación cruzada de la siguiente forma:

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} x(t) \cdot y(t - \tau) dt \quad (1.3)$$

En el caso de trabajar con señales discretas $x[n]$ e $y[n]$, la correlación cruzada pasaría a calcularse del siguiente modo:

$$R_{xy}[\tau] = \sum_{n=-\infty}^{\infty} x[n] \cdot y[n - \tau] \quad (1.4)$$

Cuando en lugar de hacer la comparación entre dos señales distintas se hace con una única señal se le denomina autocorrelación y permite conocer el parecido de una señal consigo misma a medida que se va desplazando τ muestras. Dada una señal discreta $x[n]$ se calcula su autocorrelación de la siguiente forma:

$$R_x[\tau] = \sum_{n=-\infty}^{\infty} x[n] \cdot x[n - \tau] \quad (1.5)$$

La correlación cruzada y la autocorrelación son realmente útiles para encontrar patrones repetitivos en señales ruidosas. Más concretamente, en este trabajo se hará uso de la autocorrelación para encontrar fragmentos similares de sonidos de corazón en una señal formada por la mezcla de sonidos cardiacos y pulmonares y poder, de esta forma, determinar la distancia entre los distintos pulsos cardiacos.

En la figura siguiente se muestra una señal temporal correspondiente a un tren de pulsos junto con su autocorrelación. Se observa como la función de autocorrelación es máxima para aquellos desplazamientos en los cuales coinciden los pulsos, es decir, los valores máximos de la función de autocorrelación se dan en los distintos periodos de la señal.

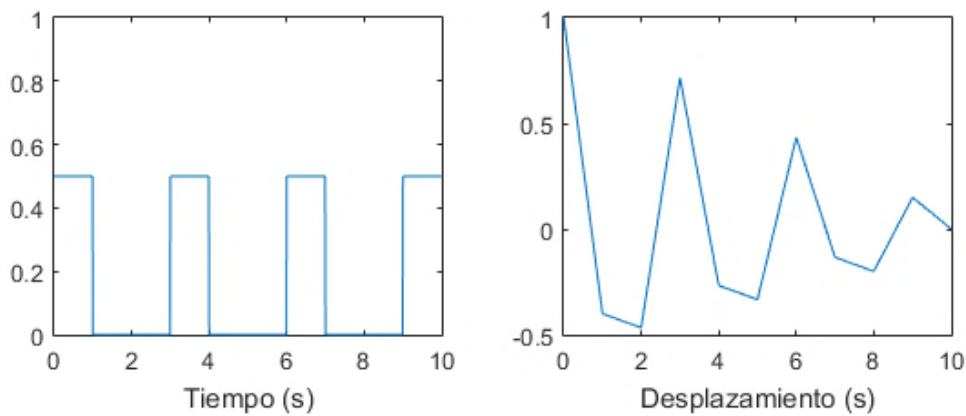


Figura 1.2 - Señal temporal (izquierda) y su autocorrelación (derecha). Se aprecia como aparecen los máximos en aquellos desplazamientos para los cuales coinciden los distintos pulsos, es decir, en los periodos de la señal.

Como se puede ver en la figura anterior, los valores de los picos de la autocorrelación no tienen el mismo valor. Esto es debido a que la señal de entrada tiene una duración finita, por lo tanto, a medida que se desplaza la señal, habrá una menor parte de la señal desplazada que se encuentre solapada con la señal sin desplazar y, por tanto, estas muestras serán multiplicadas por cero. Por esta razón, a medida que aumenta el desplazamiento, los picos de la función de autocorrelación se van aproximando a cero.

1.2.3. Información en el dominio de la frecuencia

Hasta ahora, se han visto señales en el dominio del tiempo, pero, sin embargo, en procesamiento de la señal, existen otros dominios que pueden aportar información mucho más útil que el dominio temporal, principalmente el dominio de la frecuencia.

1.2.3.1. Transformada Discreta de Fourier

El objetivo de la DFT (*Discrete Fourier Transform*) es visualizar la información de una determinada señal discreta y de duración finita en el dominio de la frecuencia. Se trata de una transformada reversible, es decir, se puede volver a obtener la señal en el dominio temporal mediante la IDFT (*Inverse Discrete Fourier Transform*).

$$x[n] \xleftrightarrow{DFT} X[k] \quad (1.6)$$

Dada una señal $x[n]$ discreta de tamaño N , se obtiene su DFT como:

$$DFT\{x[n]\} = X[k] = \sum_{n=0}^{N-1} x[n] \cdot e^{\frac{-j \cdot 2\pi \cdot k \cdot n}{N}} \quad k = 0, \dots, N-1 \quad (1.7)$$

De igual forma, se puede volver a obtener $x[n]$ aplicando la IDFT:

$$IDFT\{X[k]\} = x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] \cdot e^{\frac{j \cdot 2\pi \cdot k \cdot n}{N}} \quad n = 0, \dots, N-1 \quad (1.8)$$

Debido al enorme coste computacional para obtener la DFT se recurre normalmente al uso de la FFT (*Fast Fourier Transform*) que no es más que un método optimizado para el cálculo de la DFT. A continuación, en la **Figura 1.3** se representa una señal en el dominio del tiempo junto con el módulo de su FFT:

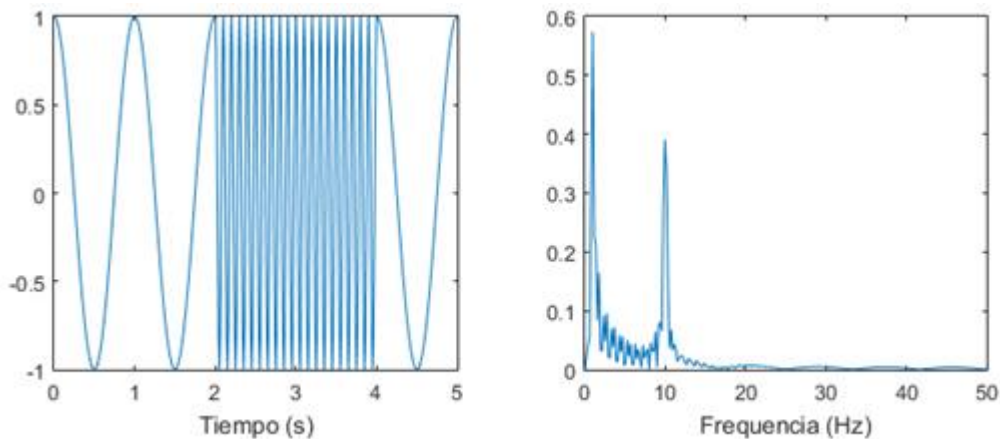


Figura 1.3 - Señal temporal de 1 y 10 Hz (izquierda) y el módulo de su FFT (derecha). Se aprecia como aparecen dos deltas, es decir, la FFT de una senoide, en cada una de las frecuencias de las sinusoides que componen la señal de entrada, en este caso 1 y 10 Hz.

Como se puede ver en la figura anterior, a partir de la información que aporta el módulo de la FFT, se sabe que la señal temporal está formada por la composición de dos señales sinusoidales de 1 y 10 Hz. Además, viendo la amplitud de la FFT, se puede ver como hay más información contenida en 1 Hz que en 10 Hz.

1.2.3.2. Densidad Espectral de Potencia

La PSD (*Power Spectrum Density*) ofrece información sobre cómo está repartida la potencia de una señal para las distintas bandas de frecuencia. Se calcula como la DFT de la función de autocorrelación.

Su representación es muy utilizada para determinar visualmente la calidad de la separación de sonidos cardiacos y pulmonares. Cuanto mayor sea el parecido entre la PSD del sonido recuperado y del sonido original, mayor será la calidad de la separación.

En la **Figura 1.4** se muestra una señal sinusoidal de 10 Hz junto con su PSD:

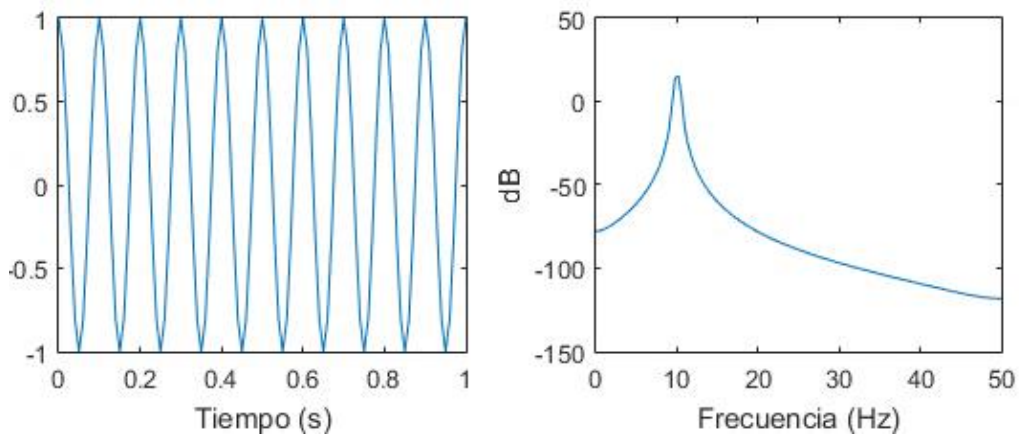


Figura 1.4 - Señal temporal de 10 Hz (izquierda) y su PSD (derecha). Como se puede observar en la PSD, toda la potencia se concentra en la frecuencia de 10 Hz, que es la frecuencia de la señal temporal.

1.2.4. Información en el dominio tiempo-frecuencia

Hasta ahora se ha trabajado con señales en el dominio del tiempo y en el dominio de la frecuencia, sin embargo, a pesar de que este último es realmente útil, no se dispone de suficiente resolución temporal, algo que resulta de utilidad cuando se trabaja con señales no estacionarias en las cuales la frecuencia no permanece constante con el tiempo.

1.2.4.1. Transformada de Fourier de Tiempo Corto

Como se puede ver en la **Figura 1.3**, se sabe que la señal está compuesta a su vez por señales sinusoidales de 1 y 10 Hz, sin embargo, no se tiene suficiente resolución en el dominio temporal.

El objetivo de la STFT (*Short Time Fourier Transform*) es solucionar el problema anterior, para ello, se calcula la FFT de aquellas muestras que se encuentren dentro de una ventana $w[n]$ (Rectangular, Hanning, Hamming, etc...) que se va desplazando temporalmente. Una ventana de tamaño grande da lugar a una mayor resolución en frecuencia mientras que una ventana de tamaño menor ofrece una mayor resolución en el tiempo.

Como resultado, con el módulo al cuadrado de la STFT se obtiene lo que se conoce como espectrograma, que indica, no solo en que frecuencias está contenida la energía sino como varía esa energía a lo largo del tiempo. Se puede mostrar en escala de grises de negro a blanco, de menor a mayor energía. La STFT es también una transformada reversible:

$$x[n] \xLeftrightarrow{STFT} X[n, k] \quad (1.9)$$

Dada una señal $\mathbf{x}[n]$ de tamaño N discreta se obtiene su STFT como:

$$X[n, k] = \sum_{m=0}^{N-1} x[m] \cdot w[n-m] \cdot e^{\frac{-j \cdot 2\pi \cdot k \cdot m}{N}} \quad (1.10)$$

De igual forma, se puede volver a obtener $\mathbf{x}[n]$ aplicando la ISTFT (*Inverse Short Time Fourier Transform*):

$$x[n] = \frac{1}{Nw[0]} \sum_{k=0}^{N-1} X[n, k] \cdot e^{\frac{j \cdot 2\pi \cdot k \cdot n}{N}} \quad (1.11)$$

Como se puede ver en la **Figura 1.5**, no solo se sabe en qué frecuencias está concentrada la mayor parte de la energía sino también en que instantes temporales se está produciendo.

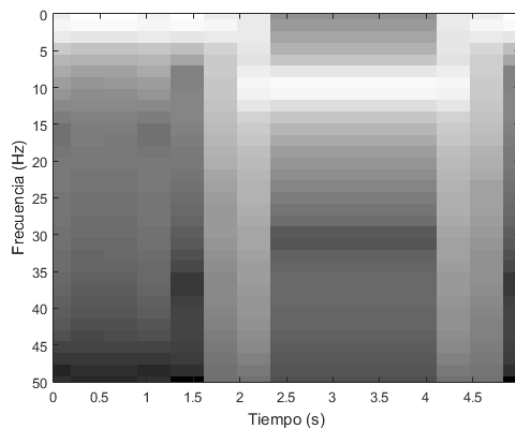


Figura 1.5 - Espectrograma de la señal temporal vista en la Figura 1.3. Se puede observar a partir del espectrograma, como la mayor parte de la energía (blanco) aparece en las frecuencias de 1 y 10 Hz, información que se podía obtener a partir de la FFT, pero con la ventaja de poder ver además como varia esta energía entre las distintas frecuencias a lo largo del tiempo.

1.2.4.2. Máscara

Uno de los métodos utilizados para mejorar la separación de fuentes acústicas es el uso de máscaras en el dominio tiempo-frecuencia. Una máscara no es más que una matriz de números de manera que al multiplicarla por la transformada STFT de una determinada señal mezcla se obtienen las transformadas STFT de cada una de las fuentes por separado.

En la **Figura 1.6** se puede ver una representación del funcionamiento básico de una máscara en la separación de dos fuentes sonoras.

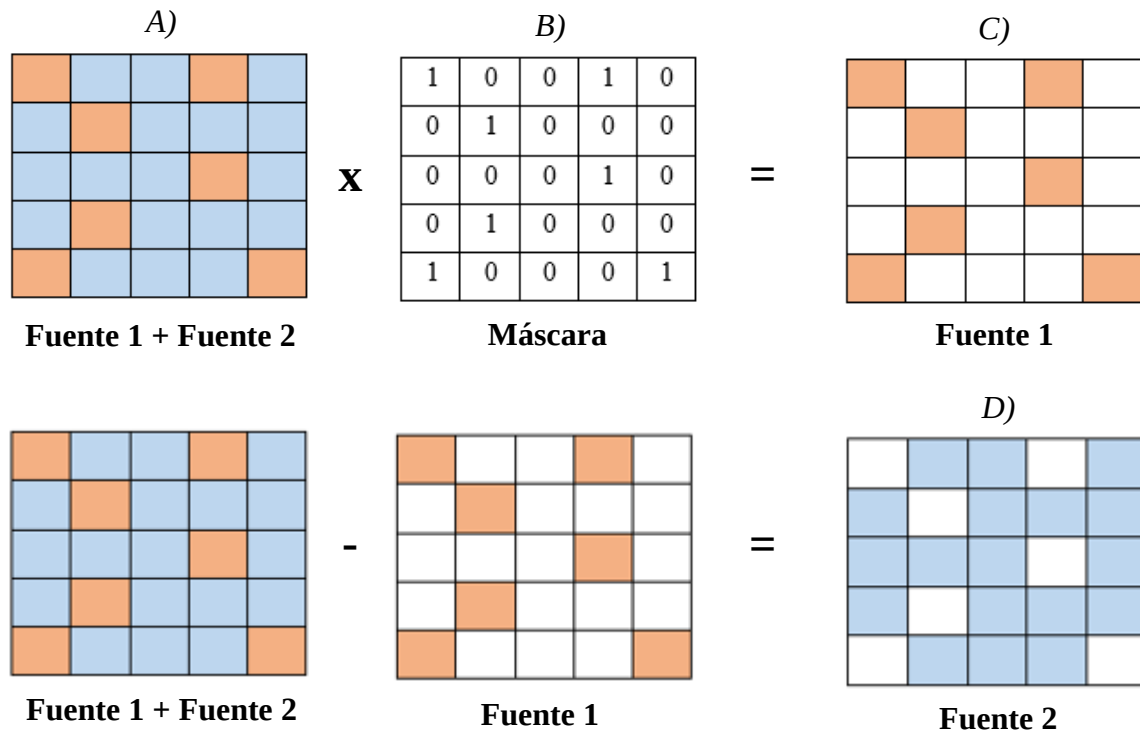


Figura 1.6 - Separación de fuentes usando una máscara tiempo-frecuencia. A) Espectrograma de la señal mezcla. B) Máscara Hard (o binaria). C) Espectrograma de la Fuente 1 obtenido al multiplicar la máscara por el espectrograma de la señal mezcla. D) Espectrograma de la Fuente 2 obtenido a partir de la resta del espectrograma de la señal mezcla y el espectrograma de la Fuente 1 obtenida anteriormente.

Como se muestra en la figura anterior, la máscara se multiplica por el espectrograma de la señal mezcla y como resultado se obtienen los espectrogramas de cada una de las fuentes separadas. Además, se debe utilizar la información de la fase de la señal mezcla, necesaria para realizar la ISTFT de los espectrogramas obtenidos y volver al dominio temporal. Otra solución consiste en multiplicar la máscara directamente por la transformada

STFT de la señal mezcla, que se tratará de una matriz compuesta por números complejos, cada uno con su correspondiente módulo y fase. Al realizar la multiplicación de la máscara por la transformada STFT únicamente se variará el módulo por lo que la información de fase (necesaria para volver al dominio temporal) permanecerá intacta. Posteriormente, se realiza la transformada ISTFT de las dos fuentes y se obtienen ambas señales en el dominio del tiempo.

No obstante, a pesar de que el procedimiento anterior resulta simple, en realidad, la dificultad recae en el hecho de encontrar la máscara adecuada que logre realizar la separación de esas fuentes sonoras. En este trabajo se utilizarán previamente dos métodos que serán desarrollados más adelante, de manera que, para cada uno de ellos, se obtenga una señal de salida correspondiente al sonido cardiaco y otra señal que se corresponde con el sonido pulmonar.

Una vez realizado el paso anterior se puede llegar a pensar que ya se ha realizado la separación pues se tienen ambos sonidos por separado, sin embargo, los métodos implementados pueden no separar completamente ambas señales y, por tanto, dejar restos de la señal de corazón en la señal de pulmón y viceversa, por lo que la aplicación de una máscara podría solventar este problema.

De esta forma, una vez se tienen ambas señales bastaría con aplicar la transformada STFT a cada una de ellas y crear la máscara a partir de los espectrogramas obtenidos. Existen distintos tipos de máscara en función de la forma de obtener la matriz. En [3] se describen los siguientes tipos:

- Máscara Wiener: También se le conoce como máscara *soft* (suave). La matriz está formada por valores comprendidos entre cero y uno. Al elevar el espectrograma al cuadrado, se eliminan de la Fuente 1 recuperada aquellos componentes que aparecen con mayor intensidad en la Fuente 2 recuperada mientras que se mantienen el resto. Por otro lado, en la Fuente 2 recuperada ocurrirá justo lo contrario.

$$Máscara_{Fuente1} = \frac{\|F_1\|^2}{\|F_1\|^2 + \|F_2\|^2} \quad (1.12)$$

- Máscara Hard (o binaria): Se la conoce de este modo puesto que la matriz está formada por unos y ceros. Surge como resultado de aplicar un redondeo a la máscara Wiener.

$$Máscara_{Fuente1} = \text{redondeo} \left(\frac{\|F_1\|^2}{\|F_1\|^2 + \|F_2\|^2} \right) \quad (1.13)$$

donde $\|F_1\|$ y $\|F_2\|$ se corresponden con la magnitud de las transformadas STFT correspondientes a la Fuente 1 y a la Fuente 2 recuperadas respectivamente.

De las dos máscaras vistas con anterioridad, en este trabajo se implementará la máscara Wiener puesto que, como ya se ha comentado anteriormente, los sonidos cardiacos y pulmonares se encuentran solapados temporal y espectralmente por lo que habrá zonas en el espectrograma con información de ambas fuentes. En el caso de trabajar con una máscara Hard, es decir, con una matriz de unos y ceros, esto provocará que toda la energía de una determinada zona del espectrograma vaya a una de las fuentes recuperadas mientras que la otra fuente no presentará nada de energía en esa zona del espectrograma, algo que no ocurriría utilizando una máscara Wiener, la cual realiza un mejor equilibrio de la energía entre ambas fuentes y por tanto la que mejores resultados ofrece. En la **Figura 1.7** se muestra un ejemplo sencillo de estos dos tipos de máscara.

A)	0.37	0.73	0.62	0.24	0.51
	0.09	0.16	0.65	0.37	0.67
	0.73	0.42	0.79	0.68	0.55
	0.11	0.52	0.18	0.87	0.74
	0.68	0.22	0.66	0.43	0.13
B)	0.63	0.27	0.38	0.76	0.49
	0.91	0.84	0.35	0.63	0.33
	0.27	0.58	0.21	0.32	0.45
	0.89	0.48	0.82	0.13	0.26
	0.32	0.78	0.34	0.57	0.87
C)	0	1	1	0	1
	0	0	1	0	1
	1	0	1	1	1
	0	1	0	1	1
	1	0	1	0	0
D)	1	0	0	1	0
	1	1	0	1	0
	0	1	0	0	0
	1	0	1	0	0
	0	1	0	1	1

Figura 1.7 - A) Máscara Wiener de la Fuente 1. B) Máscara Wiener de la Fuente 2. C) Máscara Hard de la Fuente 1. D) Máscara Hard de la Fuente 2.

1.3. Sonido

Se denomina sonido a la propagación de ondas por un medio elástico producidas como consecuencia de la vibración de un cuerpo sonoro. Estas ondas elásticas son convertidas en ondas mecánicas por el oído y enviadas al cerebro para su interpretación [4]. Cada sonido presenta una serie de características que lo hacen diferenciarse del resto y que se describen a continuación.

1.3.1. Frecuencia fundamental, altura, intensidad, duración y timbre

- Frecuencia fundamental: Es la frecuencia (número de vibraciones de una onda por segundo) [4] más baja del espectro de manera que el resto de frecuencias pueden considerarse múltiplos enteros de esta. A estas frecuencias se les denominan armónicos [5].
- Altura: Es la propiedad, determinada por la frecuencia fundamental, que diferencia entre sonidos graves o agudos. Los sonidos de mayor frecuencia son considerados sonidos agudos mientras que los sonidos de frecuencia menor son considerados sonidos graves. El oído humano es capaz de detectar sonidos comprendidos entre 20 y 20.000 Hz aproximadamente. A los sonidos de frecuencia superior se les denominan ultrasonidos [4].
- Intensidad: Esta propiedad depende de la cantidad de energía que se desplaza por el medio debido a la propagación de la onda acústica. De ello dependerá que un sonido sea más fuerte o suave. Se mide en *dB* [4].
- Duración: Es el tiempo durante el cual persiste la onda sonora producida por la vibración de un cuerpo sonoro y que determina que un sonido sea largo o corto [4].
- Timbre: Es la propiedad que permite distinguir dos sonidos con igual altura, intensidad y duración. Dependerá de la cantidad de armónicos de un sonido y de la envolvente o relación entre dichos armónicos [4].

1.3.2. Sonido monoaural, estereofónico y multicanal

Los sonidos también se pueden clasificar según el número de canales utilizados para transmitir la información:

- Sonido monoaural: Se abrevia comúnmente como “mono”. Se caracteriza por la utilización de un único canal. Surge como resultado de la suma de todos los canales de audio en uno solo. Se carece de información espacial por lo que es similar al sonido escuchado con un solo oído.
- Sonido estereofónico: Abreviado como “estéreo”, es un término utilizado para referirse a la transmisión de sonido utilizando dos canales. A diferencia del sonido monoaural, la presencia de dos canales permite al cerebro ubicar los sonidos espacialmente gracias a la variación de intensidad en cada canal.
- Sonido multicanal: Son aquellos sonidos que utilizan tres o más canales para ser transmitidos.

1.3.3. Sonidos del corazón

Se podría decir que el corazón es el órgano más importante del ser humano pues de su correcto funcionamiento depende el de todo el organismo. Está situado en el tórax y su función principal es bombear sangre al resto de órganos, llegando a latir entre 60 y 100 veces por minuto en reposo en un paciente sano [6].

La estructura del corazón (ver **Figura 1.8**) está constituida por dos cavidades superiores, denominadas aurículas, y dos cavidades inferiores, denominadas ventrículos. El control del flujo de sangre se realiza mediante la apertura y cierre de una serie de válvulas. Las válvulas aurículo-ventriculares (tricúspide y mitral) controlan el paso de sangre desde las aurículas hacia los ventrículos mientras que las válvulas semilunares (aórtica y pulmonar) regulan el paso de sangre desde el corazón hacia el exterior [7].

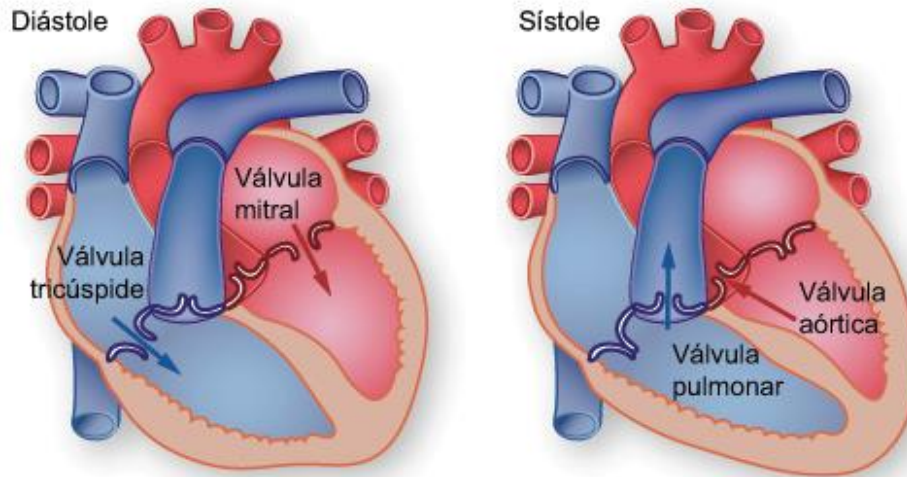


Figura 1.8 - Anatomía del corazón. En la figura de la izquierda se muestran las válvulas aurículo-ventriculares (tricúspide y mitral) que regulan el paso de sangre desde las aurículas hacia los ventrículos mientras que en la figura de la derecha se muestran las válvulas semilunares (pulmonar y aórtica) que controlan el paso de sangre desde los ventrículos hacia el pulmón y la arteria aorta respectivamente [7].

El ciclo cardíaco está formado por una fase de contracción, denominada sístole, en la cual se produce el bombeo de sangre hacia el exterior, y una fase de relajación, denominada diástole, donde la sangre retorna al corazón. Como consecuencia del flujo de sangre entrando y saliendo del corazón y de las estructuras involucradas en el control de este flujo, el corazón emite una serie de sonidos que permiten a un especialista determinar el estado de salud de un individuo [8].

Los principales sonidos que pueden ser escuchados en la auscultación cardíaca son los siguientes [9] [10]:

- **S1**: Es un sonido agudo que marca el final de la diástole y el comienzo de la sístole. Se produce como consecuencia de las vibraciones ocasionadas por el cierre de las válvulas tricúspide y mitral. Ver **Figura 1.9**.
- **S2**: Es un sonido agudo que indica el final de la sístole y el comienzo de la diástole, se produce como consecuencia del cierre de las válvulas aórtica y pulmonar. Ver **Figura 1.9**.

En algunas personas, es posible escuchar otros sonidos además de los ya mencionados anteriormente [9] [10]:

- S3: Es un sonido grave que ocurre justo después del S2. Es producido por la vibración del músculo ventricular cuando recibe sangre desde las aurículas alcanzando su límite elástico. Se encuentra en condiciones normales en niños y jóvenes, sin embargo, su presencia en adultos suele ser patológica. Ver **Figura 1.10**.
- S4: Es un sonido grave y de baja intensidad que ocurre justo antes del S1. Se produce debido a la vibración ocasionada por la contracción de las aurículas al final de la diástole. Normalmente no se escucha en la auscultación y su presencia suele significar, casi siempre, la existencia de una anomalía. Ver **Figura 1.11**.
- Clic de eyección: Es un sonido agudo que se produce justo después del sonido S1. Está relacionado con una estenosis (estrechamiento) de las válvulas semilunares, esta estrechez dificulta la apertura de la válvula generando este sonido. Ver **Figura 1.12**.
- Chasquido de apertura: Es un sonido agudo que se produce tras el sonido S2. Al igual que el clic de eyección, está provocado como consecuencia de una estenosis de las válvulas aurículo-ventriculares. Ver **Figura 1.13**.
- Soplo: Se origina como consecuencia de un paso turbulento de sangre a través de una válvula cardíaca que puede estar dañada. Existen soplos sistólicos y diastólicos, y soplos agudos y graves. Su presencia suele ser algo normal en niños. Ver **Figura 1.14**.

A continuación, se muestra la representación tanto en el dominio del tiempo como en el dominio tiempo-frecuencia de cada uno de los sonidos cardiacos que han sido explicados anteriormente:

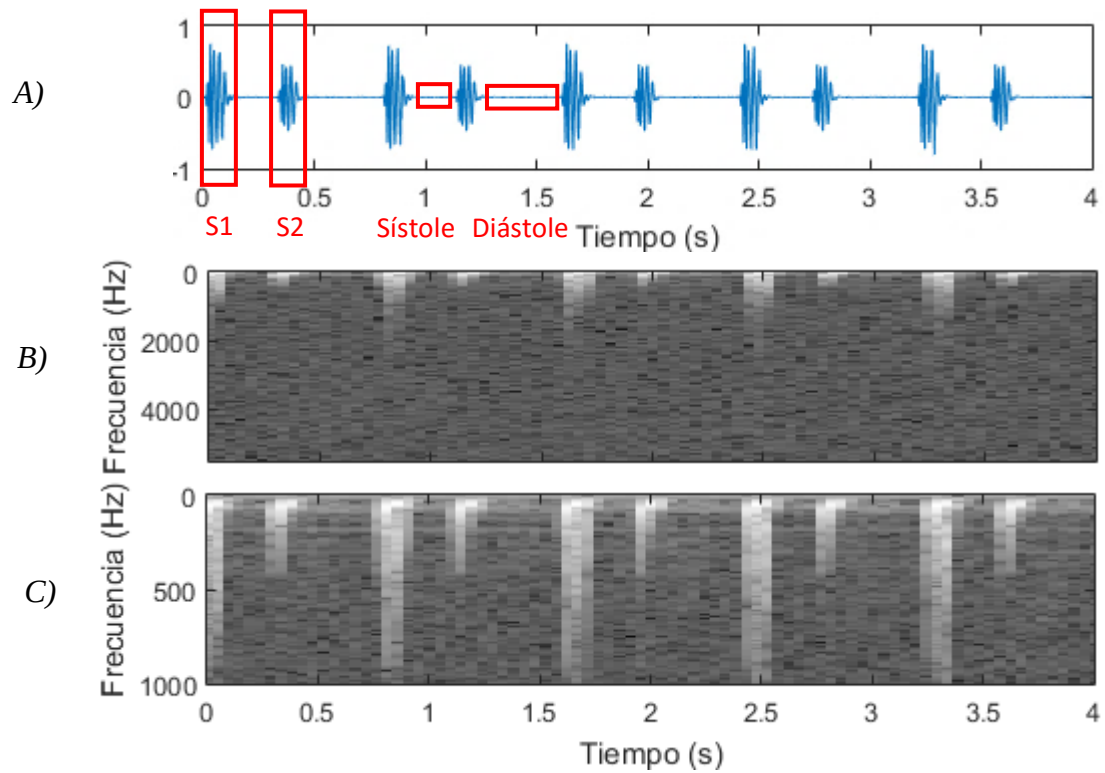


Figura 1.9 - Sonidos S1 y S2. A) Señal en el dominio del tiempo en la que aparecen remarcados los sonidos S1 y S2, así como el tiempo que los separa, denominados sístole y diástole. B) Espectrograma (el blanco indica la presencia de energía). C) Zoom del espectrograma entre 0 y 1000 Hz.

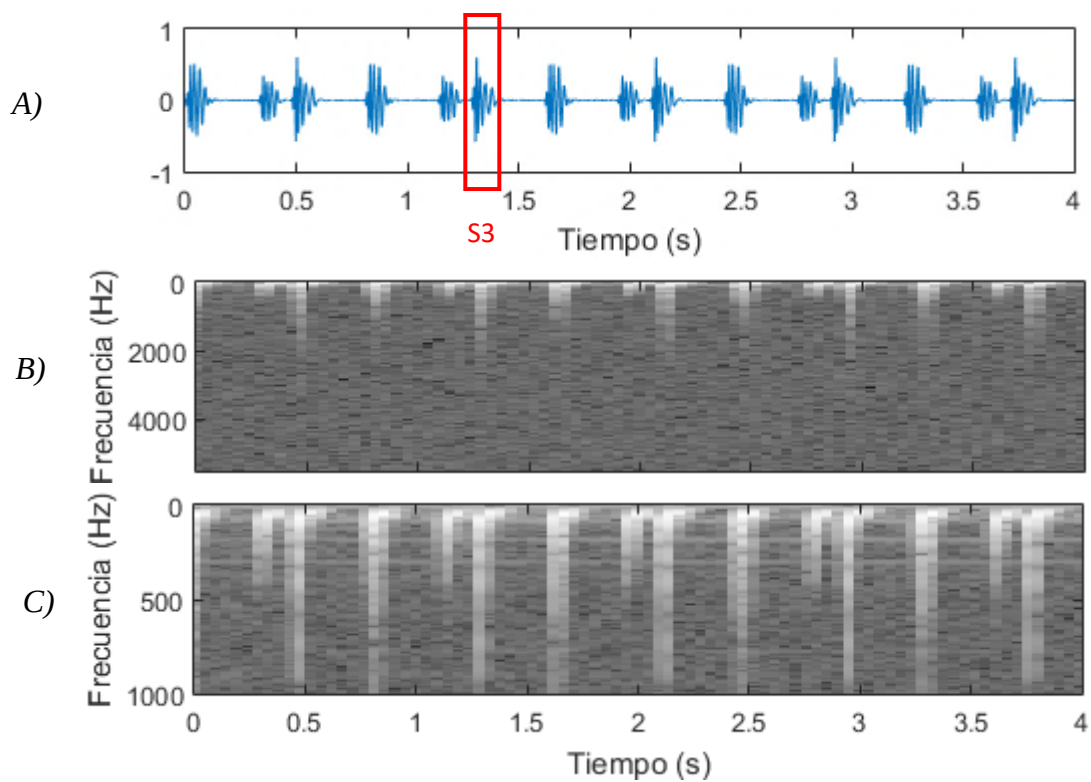


Figura 1.10 - Sonido S3. A) Señal en el dominio del tiempo en la que aparece remarcado el sonido S3. B) Espectrograma. C) Zoom del espectrograma entre 0 y 1000 Hz.

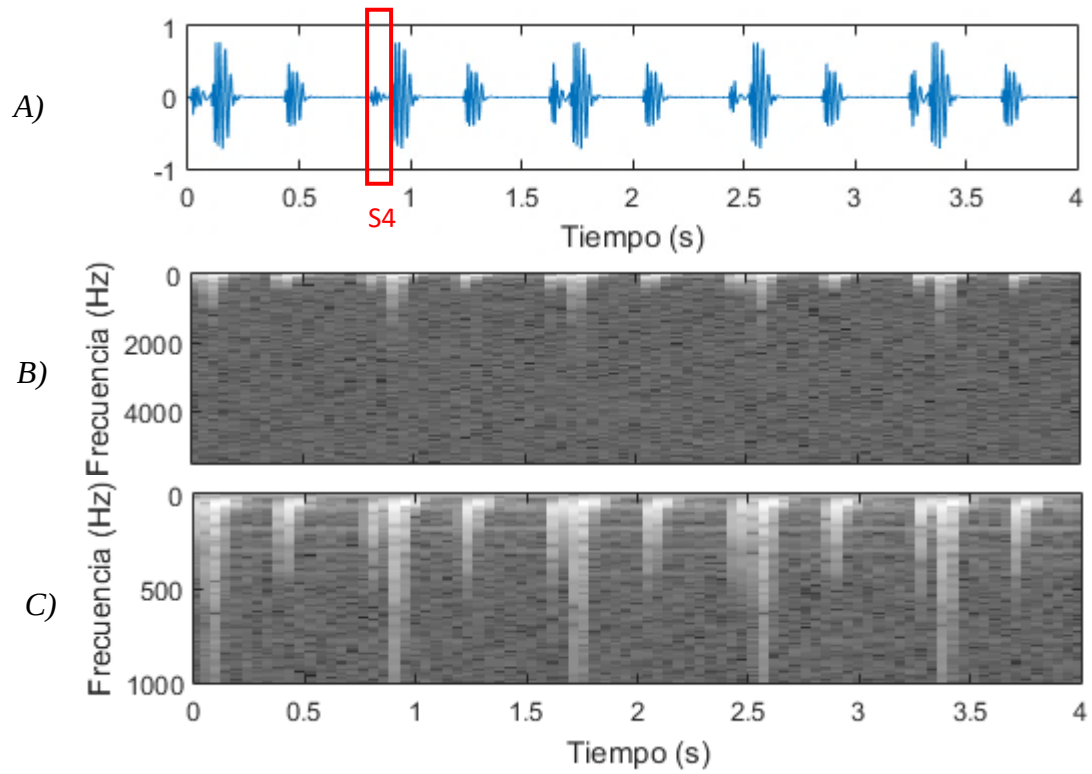


Figura 1.11 - Sonido S4. A) Señal temporal en la que aparece remarcado el sonido S4. B) Espectrograma. C) Zoom del espectrograma entre 0 y 1000 Hz.

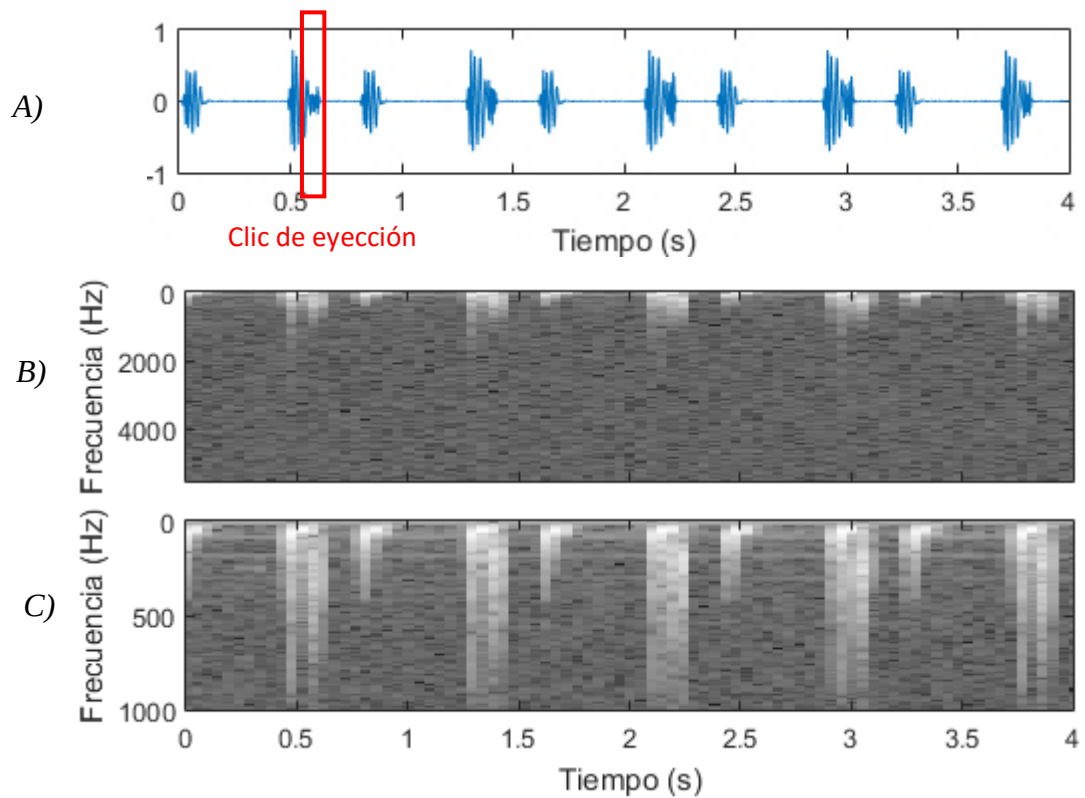


Figura 1.12 - Clic de eyección. A) Señal temporal en la que aparece remarcado el clic de eyección. B) Espectrograma. C) Zoom del espectrograma entre 0 y 1000 Hz.

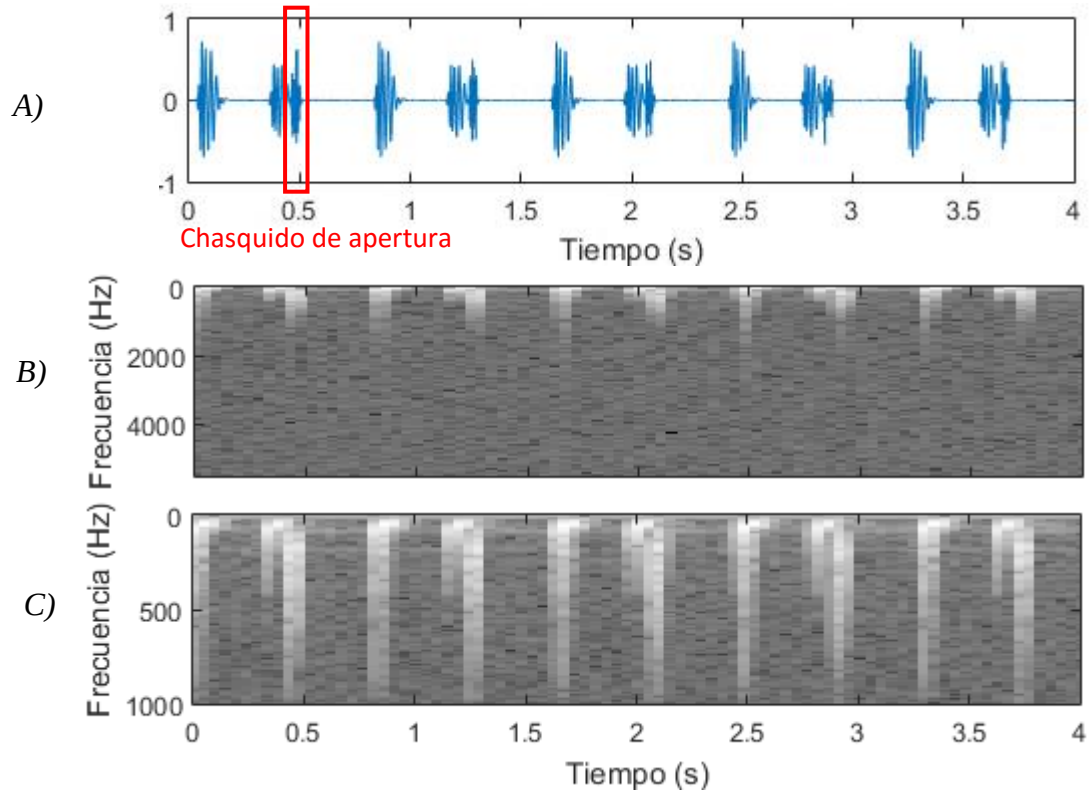


Figura 1.13 - Chasquido de apertura. A) Señal temporal en la que aparece remarcado el chasquido de apertura. B) Espectrograma. C) Zoom del espectrograma entre 0 y 1000 Hz.

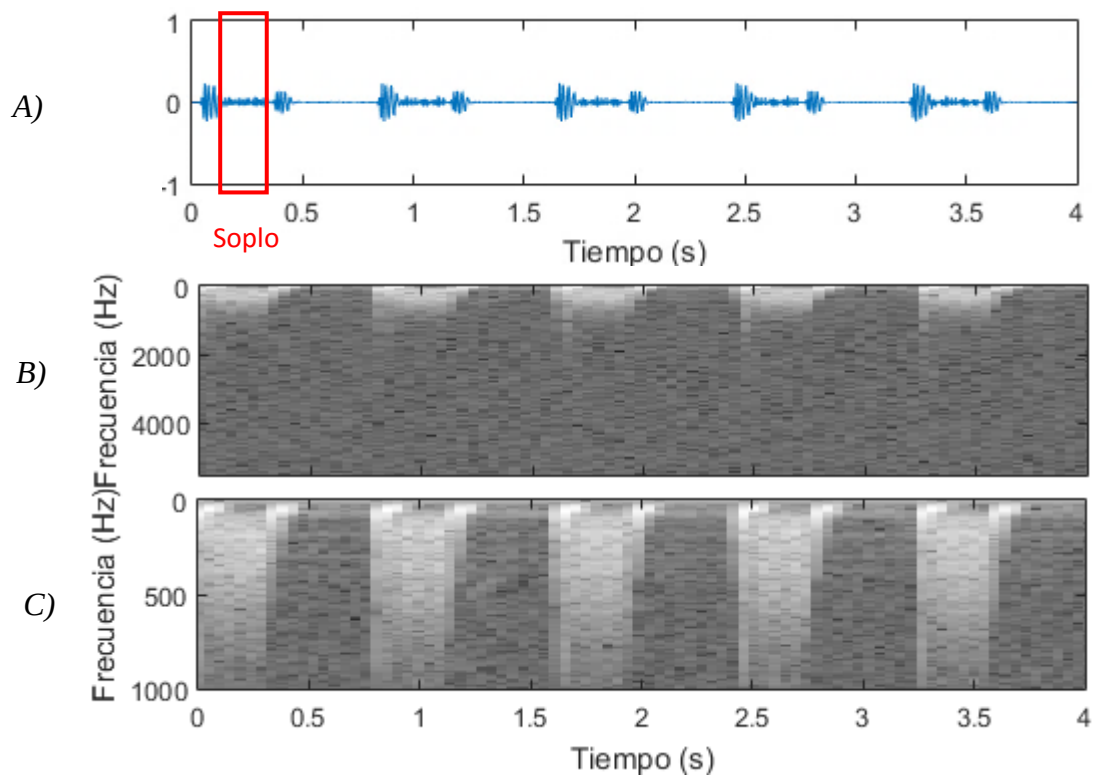


Figura 1.14 – Soplo. A) Señal temporal en la que aparece remarcado el soplo. B) Espectrograma. C) Zoom del espectrograma entre 0 y 1000 Hz.

Tal y como se ha visto en las figuras anteriores, se considera que la mayor parte de la información de los sonidos del corazón se encuentra en el rango de 30 a 150 Hz (S1: 30-120 Hz; S2: 70-150 Hz; S3: 10-70 Hz; S4: 20-70 Hz; ruido sistólico: 50-600 Hz; ruido diastólico: 120-800 Hz) [11].

1.3.4. Sonidos del pulmón

Los pulmones (ver **Figura 1.15**) forman parte del aparato respiratorio. Al igual que el corazón, están situados en el tórax. Su función es introducir el oxígeno obtenido durante la fase de inspiración en la sangre, así como extraer el dióxido de carbono presente en la sangre durante la espiración [12].

Del mismo modo que el corazón, los pulmones también generan sonidos que pueden determinar el estado de salud de una persona. El flujo de aire produce turbulencias a su paso por las vías aéreas, estas turbulencias dan lugar a las vibraciones que son percibidas como sonidos [8].

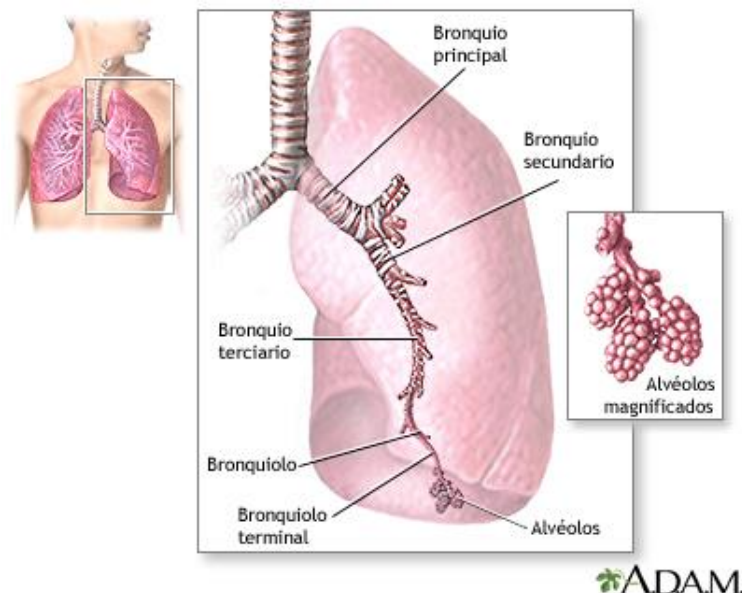


Figura 1.15 - Anatomía del pulmón. Se pueden observar las vías aéreas (tráquea, bronquios, ...) por donde circula el aire y cuyo flujo genera una serie de turbulencias que originan vibraciones percibidas como sonidos [13]

Entre los sonidos fisiológicos del pulmón, es decir, aquellos cuya presencia en la auscultación es considerada normal, se encuentran los distintos tipos descritos a continuación [14] [15]:

- Vesicular: Suaves y de tono bajo, son los sonidos de pulmón más comúnmente auscultados. Son escuchados en la mayor parte de la superficie del pulmón. Ver **Figura 1.16**.
- Bronquial: Son sonidos graves y huecos. Pueden ser auscultados sobre la tráquea debido al paso de aire por ella. Ver **Figura 1.17**.
- Broncovesicular: Sonidos normales en la parte media y posterior del tórax. Son una combinación de los sonidos bronquiales y vesiculares. Ver **Figura 1.18**.

Los sonidos patológicos del pulmón, también conocidos como “ruidos adventicios”, se encuentran superpuestos a los sonidos anteriormente mencionados. Su presencia significa, en todos los casos, la existencia de algún tipo de anomalía en el paciente. Existen 4 grupos principales [14] [15]:

- Crepitaciones: Sonidos burbujeantes que son producidos debido a la presencia de fluidos en cualquier parte de la vía respiratoria. Ver **Figura 1.19**.
- Sibilancias: Son sonidos similares a un silbido. Se producen como consecuencia del estrechamiento de las vías respiratorias pequeñas. Tienen lugar tanto en la inspiración como en la espiración. Ver **Figura 1.20**.
- Roncus: Es un sonido similar a un ronquido y se produce debido a la obstrucción de aire a través de las vías respiratorias grandes. Ver **Figura 1.21**.
- Roce pleural: Se produce debido a la presencia de líquido entre las superficies pleurales. Ver **Figura 1.22**.

Se muestra a continuación, al igual que se ha hecho con los sonidos de corazón, una representación en el dominio temporal junto con su espectrograma para cada uno de los sonidos de pulmón vistos anteriormente:

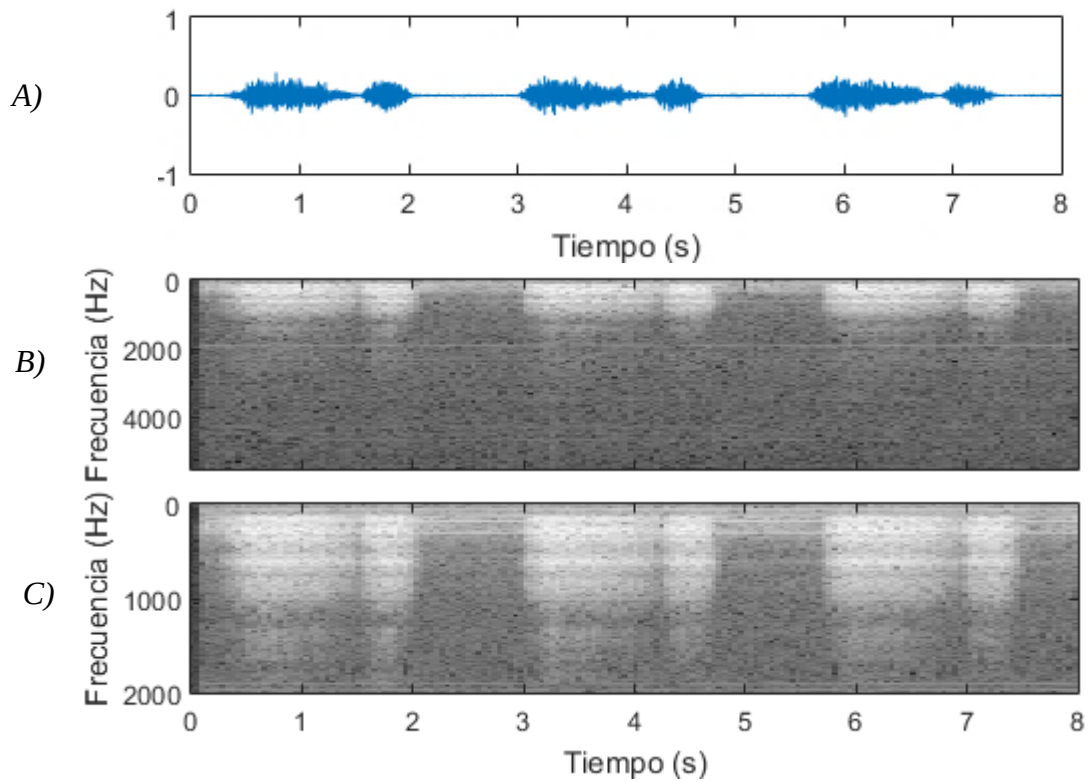


Figura 1.16 - Sonido vesicular. A) Señal temporal. B) Espectrograma. B) Zoom del espectrograma entre 0 y 2000 Hz.

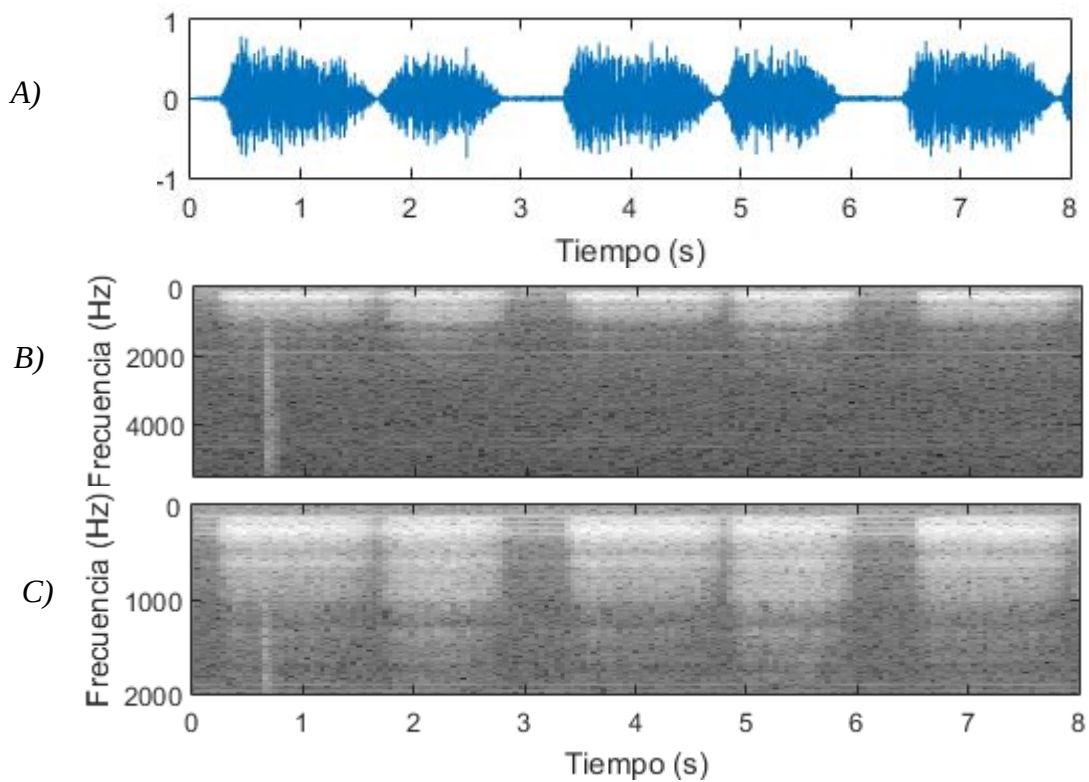


Figura 1.17 - Sonido bronquial. A) Señal temporal. B) Espectrograma. C) Zoom del espectrograma entre 0 y 2000 Hz.

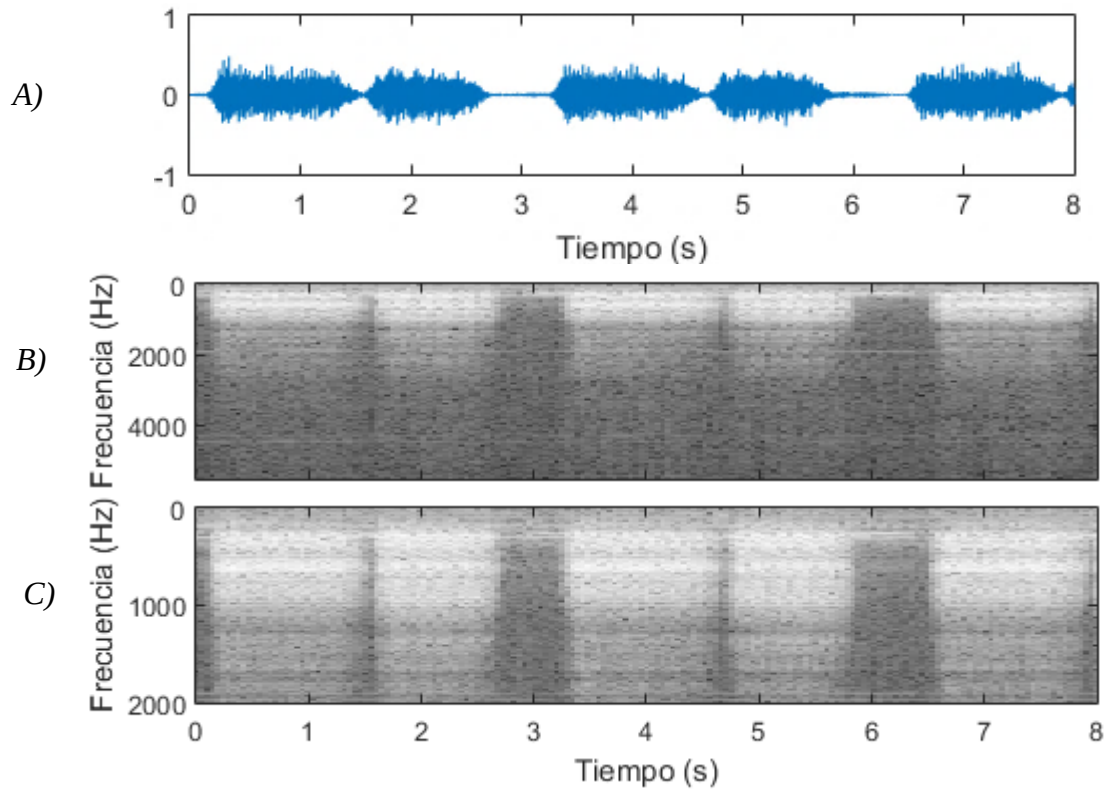


Figura 1.18 - Sonido broncovesicular. A) Señal temporal. B) Espectrograma. C) Zoom del espectrograma entre 0 y 2000 Hz.

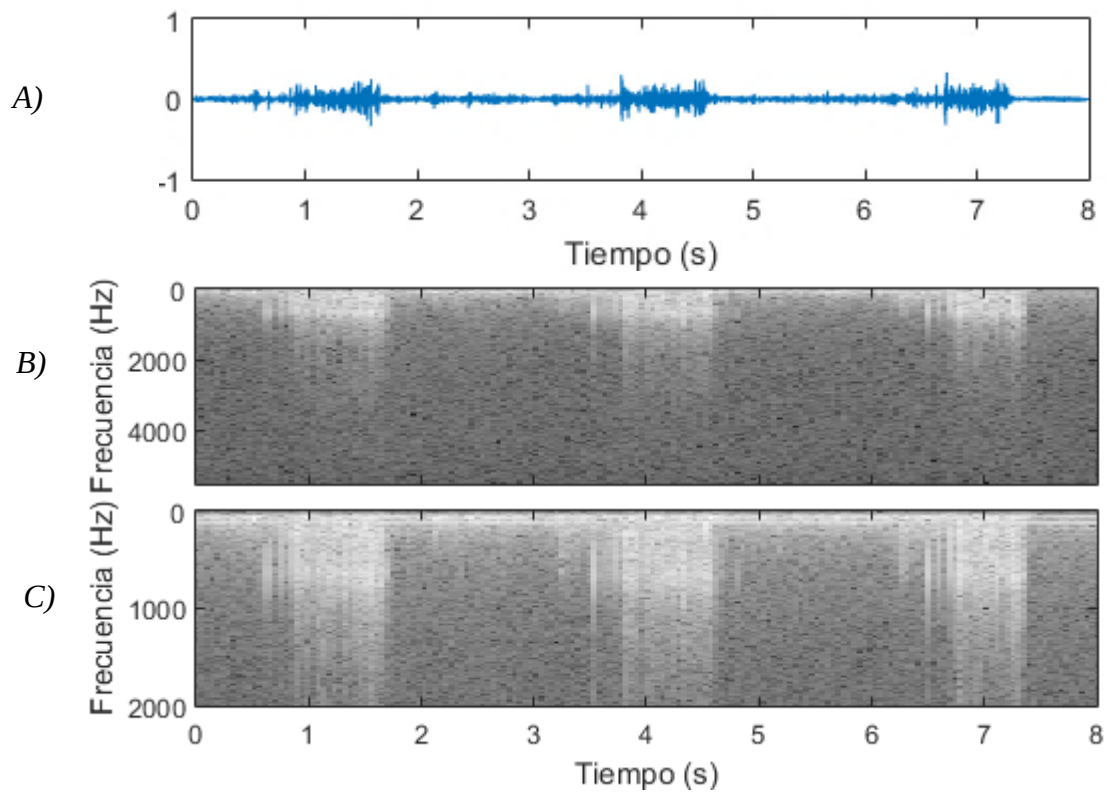


Figura 1.19 – Crepitaciones. A) Señal temporal. B) Espectrograma. C) Zoom del espectrograma entre 0 y 2000 Hz.

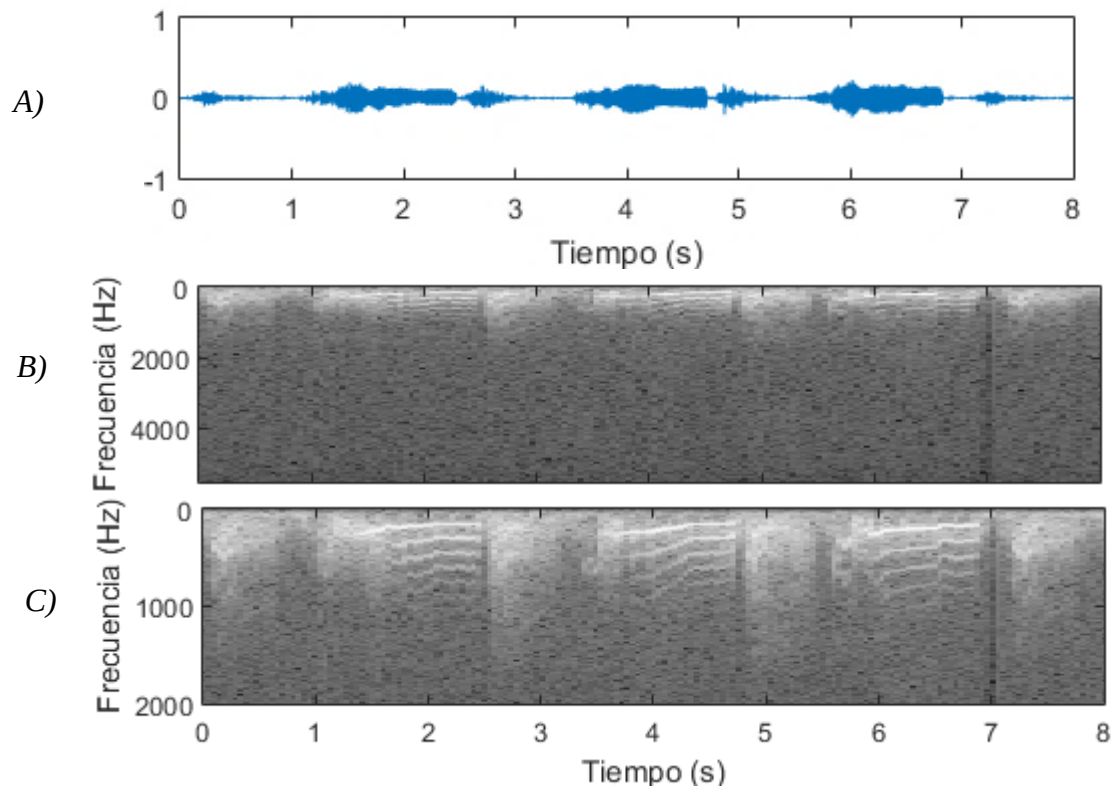


Figura 1.20 – Sibilancias. A) Señal temporal. B) Espectrograma. C) Zoom del espectrograma entre 0 y 2000 Hz.

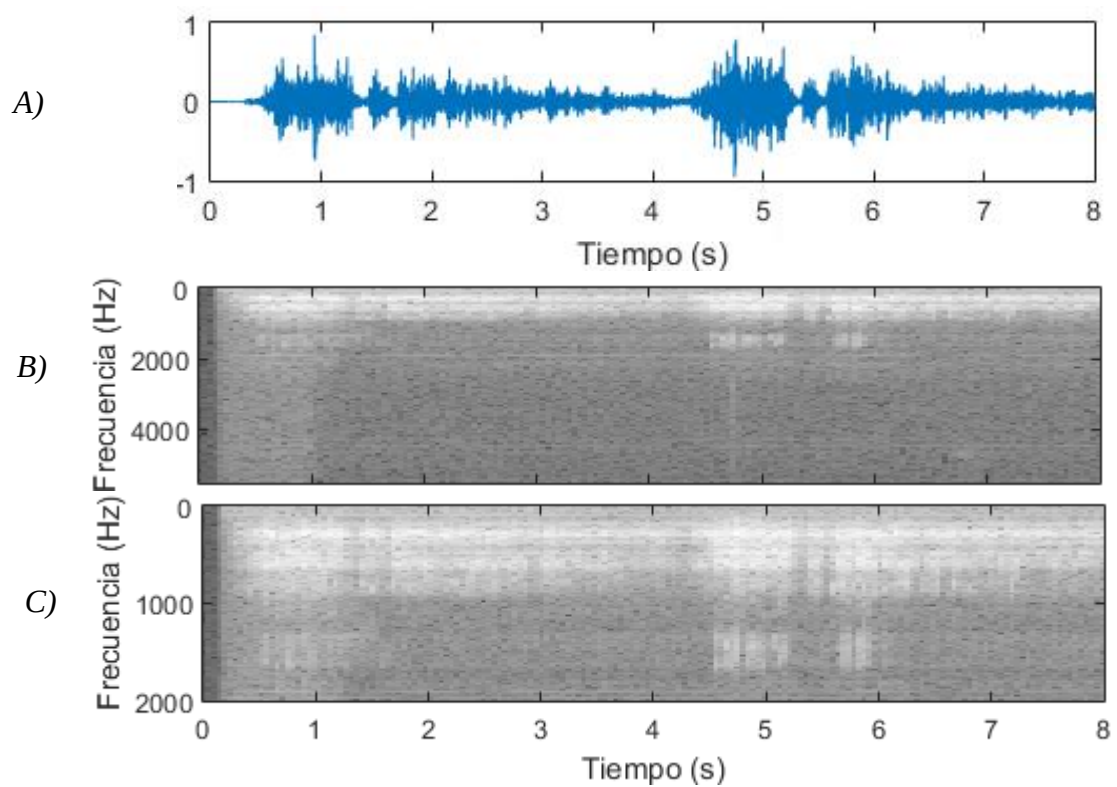


Figura 1.21 – Roncus. A) Señal temporal. B) Espectrograma. C) Zoom del espectrograma entre 0 y 2000 Hz.

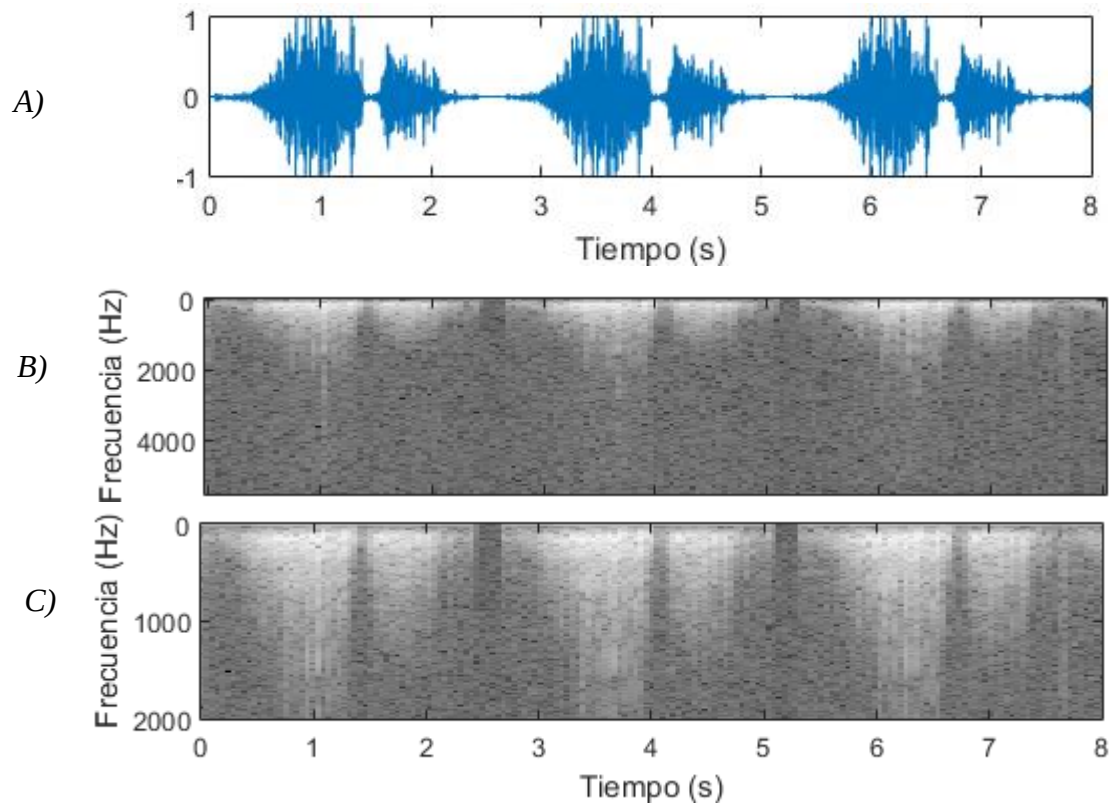


Figura 1.22 - Roco pleural. A) Señal temporal. B) Espectrograma. C) Zoom del espectrograma entre 0 y 2000 Hz.

Es necesario mencionar, al igual que se ha hecho en los sonidos del corazón, las características de los sonidos pulmonares en el dominio de la frecuencia. Se considera que gran parte de los sonidos del pulmón se encuentran comprendidos en el rango de 20 a 1200 Hz [11].

CAPÍTULO 2

2. Objetivo

Como ya se ha explicado en el capítulo anterior, el estetoscopio tradicional siempre ha sido una herramienta realmente útil para realizar una primera observación del enfermo. Tanto los sonidos del corazón como los sonidos del pulmón ofrecen información determinante sobre el funcionamiento de ambos órganos y el estado de salud del paciente.

Sin embargo, tanto la cercanía de ambos órganos como las características de los sonidos que originan, implica que sea realmente difícil auscultar uno de ellos sin la presencia del otro. Todo esto dificulta en gran medida la tarea del médico para realizar un diagnóstico preciso.

El problema anterior ha llevado a la aparición de varios estetoscopios digitales que incluyen una serie de características adicionales como la posibilidad de grabar, visualizar o editar, amplificar el sonido y eliminar el ruido presente en la auscultación.

El objetivo de este trabajo es implementar dos de entre todos los métodos usados para la separación de sonidos de pulmón y de corazón. Ambos métodos deben ser capaces de separar las señales a partir de una señal de entrada monoaural, pues muchos estetoscopios electrónicos no son capaces de ofrecer una salida de más canales.

Para tal propósito, una vez implementados ambos algoritmos, cada uno se someterá a una fase de evaluación, permitiendo de esta forma, obtener los parámetros que permitirán realizar una separación de la manera más óptima posible atendiendo a unas métricas de calidad que se detallarán más adelante.

Por último, se realizará una interfaz gráfica que permita al usuario introducir una señal de entrada y realizar tareas como: separar los sonidos de corazón y de pulmón, reproducir, guardar y visualizar ambas señales en el tiempo, así como su densidad espectral de potencia y espectrograma, variar los parámetros de ambos algoritmos fijados por defecto en su valor óptimo y evaluar la calidad de la separación en caso de disponer de las fuentes sonoras originales.

CAPÍTULO 3

3. Estado del arte

En la actualidad existen una multitud de métodos que permiten llevar a cabo la separación de sonidos cardiacos y pulmonares, intentando aprovechar, cada uno de ellos, alguna de las características de ambas señales. En este capítulo se realizará una revisión bibliográfica de los métodos recogidos en [8] [16], así como una breve descripción de algunos de ellos.

3.1. Filtrado paso-alto

Esta fue una de las primeras soluciones que se propusieron para la separación de sonidos cardiacos y pulmonares. Un filtro paso alto es un tipo de filtro cuya respuesta en frecuencia permite atenuar en una determinada señal aquellas frecuencias que estén situadas por debajo de una determinada frecuencia de corte. Se denomina frecuencia de corte a aquella frecuencia en la cual existe una atenuación de 3 dB, es decir, una reducción a la mitad de la potencia original. El problema de este tipo de filtros es, como ya se ha comentado anteriormente, el solapamiento espectral de los sonidos cardiacos y pulmonares entre los 20 y 150 Hz aproximadamente, por lo que se plantea una difícil elección entre mantener las bajas frecuencias del sonido del pulmón o eliminar por completo los sonidos del corazón. Este procedimiento podría eliminar componentes importantes de la señal de corazón y pulmón pudiendo llevar a un diagnóstico erróneo del paciente. En [17] [18] se escoge un valor arbitrario de la frecuencia de corte situado en ese rango, pero si una fuerte base científica para dicha elección.

3.2. Filtrado adaptativo

Los filtros adaptativos lineales surgen como una solución al problema planteado en el punto anterior. Este tipo de filtros está formado por cuatro componentes principales: la señal de entrada o señal de referencia, la señal primaria o señal deseada (formada por la mezcla

del sonido del corazón y del pulmón), la señal de salida y la estimación del error, que es la diferencia entre la señal deseada y la señal obtenida a la salida del filtro. A diferencia de los filtros fijos, los coeficientes de un filtro adaptativo lineal no permanecen invariantes en el tiempo, sino que se van modificando gracias a la realimentación producida por la señal de error obtenida.

Una de las primeras divisiones que se pueden realizar en este tipo de métodos es según el esquema utilizado. Uno de los más utilizados en primer lugar para la separación de sonidos de corazón y pulmón era el esquema ANC (*Adaptive Noise Canceller*) (ver **Figura 3.1**), para el cual se necesitan dos señales, la señal deseada formada por el sonido cardiaco y sonido pulmonar que se quiere separar y una señal de referencia que esté ampliamente correlacionada con el sonido cardiaco, de este modo el filtro únicamente podrá seguir las variaciones del sonido cardiaco, el cual, será obtenido a la salida junto con una señal de error que se corresponderá con el sonido pulmonar.

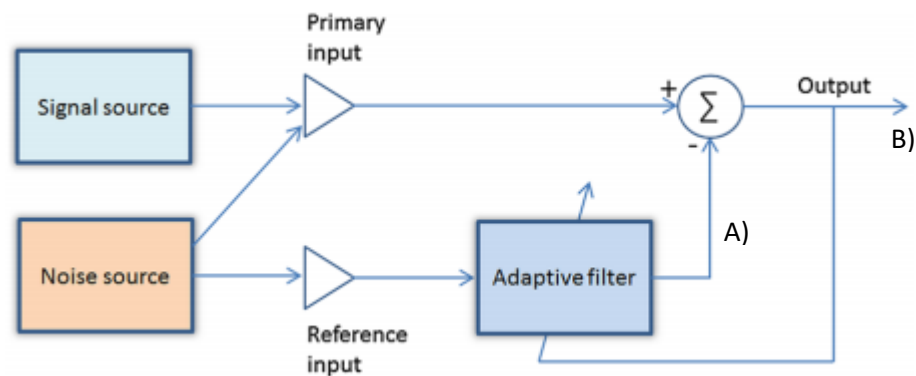


Figura 3.1 - Diagrama de bloques ANC [16]. En este esquema, la señal deseada “*Primary input*” está formada por la suma del sonido cardiaco “*Signal source*” y el sonido pulmonar “*Noise source*”. Se toma como señal de referencia “*Reference input*” una señal parecida al sonido cardiaco y que por tanto este ampliamente correlacionada con este. A) Salida del filtro donde se obtiene el sonido cardiaco recuperado. B) Salida del sistema (señal de error) donde se obtiene el sonido pulmonar recuperado

Uno de los principales inconvenientes de la estructura ANC es, como se ha podido comprobar en la figura anterior, la necesidad de dos sensores que recojan las señales de entrada, la señal deseada y la señal de referencia, para llevar a cabo la separación de ambos sonidos, mientras que un estetoscopio electrónico únicamente es capaz de captar una única señal.

Como solución, se planteó el uso de una estructura SIANC (*Single Input Adaptive Noise Canceller*). Este método, a diferencia del ANC, únicamente requiere la señal de entrada, formada por la mezcla del sonido cardiaco y el sonido pulmonar, ya que, como se detallará en los ejemplos mostrados más adelante, el mismo se encarga de generar una señal de referencia útil aplicando diversos procedimientos a la señal mezcla.

Otro de los métodos muy usados en la actualidad es el esquema ALE (*Adaptive Line Enhancer*) (ver **Figura 3.2**), a diferencia del anterior, este método no necesita aplicar ningún procedimiento a la señal de entrada para generar la señal de referencia, puesto que usa como señal de referencia una versión retardada de esta misma señal mezcla. La idea de introducir este retardo en la señal de referencia se basa en el hecho de que el sonido pulmonar presenta una menor correlación con las muestras retardadas que en el caso del sonido cardiaco, de este modo, eligiendo un valor adecuado de retardo se puede tener una señal de referencia que esté muy correlacionada con el sonido cardiaco y ligeramente correlacionada con el sonido pulmonar. Esto será explicado en mayor profundidad en el capítulo 4 correspondiente a la implementación.

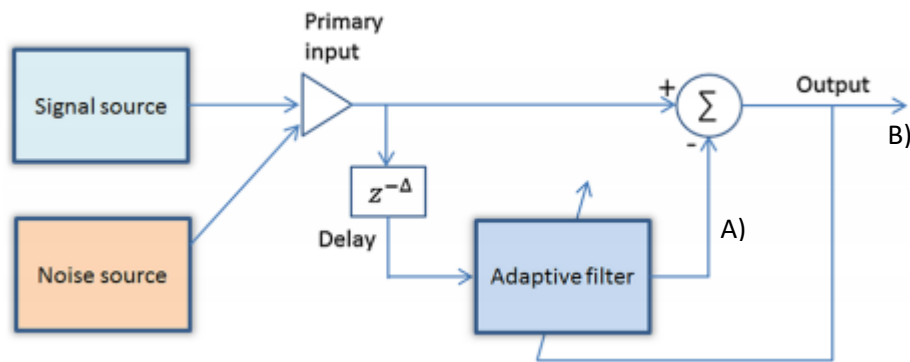


Figura 3.2 - Diagrama de bloques ALE [16]. Se observa como a diferencia del esquema ANC, en este caso únicamente se dispone de una señal de entrada “*Primary input*”, formada por la mezcla del sonido cardiaco “*Signal source*” y el sonido pulmonar “*Noise source*”, a partir de la cual se obtiene la señal de referencia introduciendo un retardo Δ . A) Salida del filtro donde se obtiene el sonido cardiaco recuperado. B) Salida del sistema donde se obtiene el sonido pulmonar recuperado.

Como se puede ver en las figuras anteriores, el filtro se basa en una realimentación para, de este modo, actualizar los coeficientes en cada iteración con el objetivo de reducir la señal de error, obteniéndose a la salida el sonido cardiaco y pulmonar por separado. La forma de actualizar estos coeficientes dependerá del tipo de algoritmo empleado.

3.2.1. Gradiente estocástico

Uno de los algoritmos más empleados en la implementación de filtros adaptativos para la separación de sonidos cardiacos y pulmonares es LMS (*Least Mean Squares*). Existen varios estudios en los cuales se han utilizado distintas señales de referencia.

En [19] se utiliza un esquema ANC-LMS utilizando como señal de referencia una señal ECG (*ElectroCardioGram*) modificada (ver **Figura 3.3**). Una señal ECG solo tiene un pico que se corresponde con el primer sonido del corazón S1, sin embargo, la señal sonora presenta los sonidos S1 y S2, de ahí que a la señal ECG se le añada la misma señal con un retardo similar a la distancia entre los sonidos S1 y S2.

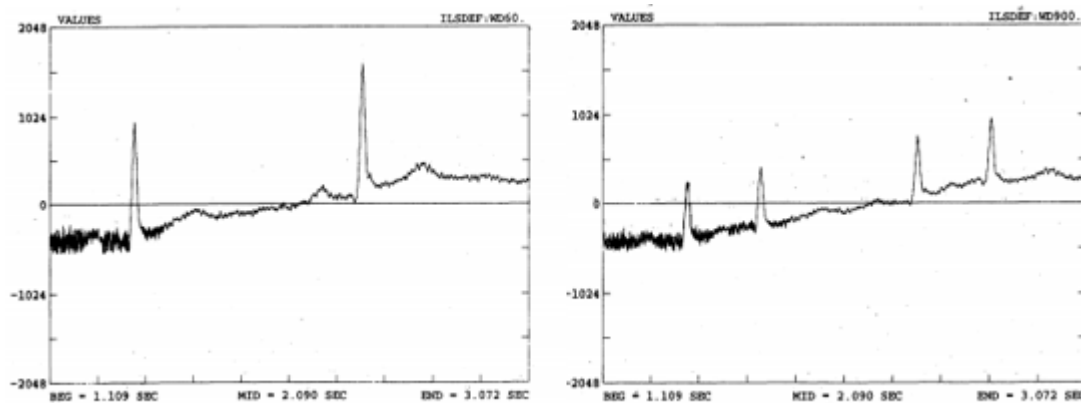


Figura 3.3 - Señal ECG (izquierda) y señal ECG modificada (derecha) [19]. Puesto que el sonido cardíaco, como se vio en el capítulo 1, está formado principalmente por los sonidos S1 y S2, es necesario realizar una modificación a la señal ECG original de manera que la señal de referencia esté formada por dos picos que coincidan con los dos principales sonidos del sonido cardíaco.

En [20] se hace uso de un esquema SIANC-LMS, donde se usa como señal de referencia la misma señal obtenida del estetoscopio tras ser previamente procesada (se le aplica un filtrado paso bajo, se eleva al cuadrado y se le vuelve a aplicar un filtrado paso bajo) con el objetivo de obtener una señal que contiene impulsos en aquellos intervalos de tiempo en los cuales aparecen los sonidos del corazón. En [21], se hace uso de un detector de sonidos de corazón que permite eliminar gran parte de los sonidos cardiacos de la señal mezcla y usar la señal obtenida como señal de referencia. En [22] se emplea una señal LECG (*Laplacian ElectroCardioGram*) con la ventaja de que solo se necesita el uso de un solo sensor para capturar ambas señales a diferencia del método visto en [19].

En [23] se usa un esquema ALE-LMS, utilizando como señal de referencia la misma señal obtenida del estetoscopio retarda y en [24] se hace lo mismo, pero utilizando una versión ligeramente modificada de LMS conocida como NLMS (*Normalized Least Mean Squares*).

Por otro lado, en [25] se utiliza una estructura ANC-FOS (*Fourth Order Statistics*), con una rápida convergencia y muy poco error de exceso, pero con un mayor coste computacional que un esquema LMS básico.

3.2.2. Estimación de mínimos cuadrados

Otro de los algoritmos usados es RLS (*Recursive Least Squares*). En [26] se usa un esquema ANC-RLS utilizando como señal de referencia una versión modificada de la señal deseada filtrada paso banda y en [27] se utiliza el mismo algoritmo, pero en este caso empleando un esquema ALE-RLS. Por otro lado, en [28] se hace uso de ROKF (*Reduced Order Kalman Filtering*), que a diferencia de RLS, puede operar tanto en ambientes estacionarios como no estacionarios sin la necesidad de realizar modificaciones al algoritmo.

3.3. NLM

La utilidad del algoritmo NLM (*Non Local Mean*) fue descubierta y usada inicialmente en imágenes [29] aunque se ha demostrado también su buen funcionamiento en la separación de sonidos de corazón y de pulmón [11] [30]. Los resultados de la aplicación de este método en señales de auscultación reales son prometedores.

El objetivo de este método es aprovechar el hecho de que existen pequeños segmentos de la señal libre de ruido, es decir, el sonido de corazón, similares en diferentes instantes temporales. El funcionamiento se basa en comparar cada una de las muestras con el resto de muestras de la señal, de manera que aquellos segmentos que presenten un cierto parecido con el resto permanecerán en la señal de salida.

Para encontrar los fragmentos similares es necesario recorrer muestra a muestra toda la señal de entrada, de ahí que el principal inconveniente de este método sea su elevado coste computacional, para el cual se han propuesto diversas soluciones. Una de ellas consiste en reducir la distancia de búsqueda de segmentos similares, de esta forma, cuando se esté procesando un determinado segmento de la señal de entrada, no será necesario recorrer toda

la señal para encontrar segmentos similares, sino que se limitará la búsqueda a aquellas muestras cercanas donde existe una gran probabilidad de encontrar patrones similares, reduciendo la dependencia del coste computacional con el tamaño de la señal de entrada como se muestra en [31].

3.4. Filtrado en el dominio transformado

Los métodos de filtrado utilizando dominios distintos al tiempo, se basan fundamentalmente en el empleo de STFT y DWT (*Discrete Wavelet Transform*). Ambas transformadas han sido ampliamente usadas en los esquemas de cancelación de ruido donde están implicados sonidos cardiacos y pulmonares.

3.4.1. Basado en STFT

El estudio mostrado en [32] se basa en considerar el espectrograma de la señal obtenida del estetoscopio como una imagen y eliminar el sonido de corazón utilizando procesamiento digital de imágenes. Para localizar los sonidos de corazón, se establece un umbral en el espectrograma entre los 10 y los 110 Hz para cada fase de la respiración incorporando la media y la desviación típica de la potencia media calculada para sonidos pulmonares durante la inspiración y la espiración (de forma separada). Una vez detectados los segmentos que corresponden a sonidos de corazón, se eliminan directamente del espectrograma o se les aplica un filtro elimina banda situado en la banda de 20 a 300 Hz, como consecuencia, se crean una serie de huecos en el espectrograma. Con el objetivo de rellenar los huecos generados anteriormente, se aplica en cada uno de ellos una interpolación bidimensional basándose en los sonidos de pulmón que se producen antes y después del hueco que se está tratando. Tras aplicar la ISTFT se obtienen las señales separadas en el dominio del tiempo.

Otro procedimiento realizado en [33] consiste en la aplicación de un algoritmo basado en ICA (*Independent Component Analysis*) aprovechando el hecho de que tanto los sonidos de corazón como los sonidos de pulmón son señales de fuentes independientes. Una vez se han obtenido las dos grabaciones realizadas con un estetoscopio en dos localizaciones diferentes del pecho se realiza el espectrograma de ambas señales aplicando la STFT. Tras obtener ambos espectrogramas se aplica una técnica basada en ICA y se obtienen los espectrogramas de ambas señales separadas. Por último, se realiza la ISTFT y se obtienen las correspondientes señales en el dominio temporal.

En [34] se hace uso de una técnica conocida como NMF (*Non-negative Matrix Factorization*). El funcionamiento se basa en descomponer el espectrograma de la señal obtenida en componentes independientes y no redundantes utilizando NMF. Para agrupar los componentes descompuestos en las fuentes originales se hace uso de una técnica no supervisada, posteriormente se utiliza una máscara tiempo-frecuencia para recuperar las señales separadas.

3.4.2. Basado en DWT

El procedimiento realizado en [35] obtiene algunas ideas introducidas en [32], pero, a diferencia de este, se basa en una DWT en lugar de la STFT y hace uso de una predicción lineal en lugar de una interpolación. La idea de este algoritmo es aprovechar el comportamiento distinto tanto de los sonidos de corazón como de los sonidos de pulmón en el dominio wavelet. Mientras que en con los sonidos de corazón la magnitud de la transformada wavelet aumenta a medida que lo hace la escala, en el caso de los sonidos de pulmón esta magnitud decrece. Una vez se ha obtenido la transformada wavelet y se ha descompuesto la señal en 3 niveles se realiza un producto multiescala. Las muestras de ese producto multiescala que supere un determinado umbral permanecerán en las tres descomposiciones realizadas anteriormente o de lo contrario se eliminarán. Los huecos creados son reconstruidos usando una predicción lineal basada en un modelo autorregresivo o de media móvil. Tras aplicar la IDWT (*Inverse Discrete Wavelet Transform*) se obtiene la señal de pulmón separada

CAPÍTULO 4

4. Implementación

En este apartado se describirá todo el proceso que se ha seguido en la implementación de los dos métodos para la separación de sonidos cardiacos y pulmonares. El software utilizado ha sido MATLAB. Se analizará con mayor profundidad su desarrollo matemático y se detallaran los distintos problemas que surgieron, así como las soluciones que se iban proponiendo.

4.1. ALE

Uno del método que serán desarrollados en este trabajo consiste en un esquema especial de filtro adaptativo denominado ALE junto con el algoritmo LMS. Es necesario desarrollar previamente algunos conceptos detallados en [36], en los cuales, se basa el funcionamiento de este tipo de filtros.

4.1.1. Desarrollo matemático

4.1.1.1. Filtro de Wiener

El filtro de Wiener, que recibe este nombre gracias a su creador, Nobert Wiener, es uno de los filtros lineales óptimos más importantes. Comprender su funcionamiento es básico para entender los filtros adaptativos. No debe confundirse con la máscara Wiener que fue detallada en la introducción.

El objetivo de este filtro es obtener una señal de salida $y[n]$ lo más parecida posible a la señal deseada $x[n]$ a partir de una señal de referencia $u[n]$, con el fin de que la señal de error $e[n]$ sea cero. Esta señal de referencia debe presentar una amplia correlación con aquella parte de la señal deseada que se quiere recuperar a la salida. Por ejemplo, si se tiene una señal deseada formada por la mezcla de sonidos cardiacos y sonidos pulmonares, suponiendo que se quiere recuperar el primer sonido, la señal de referencia debe ser una señal que este ampliamente correlacionada, es decir, que sea muy parecida, a la señal

correspondiente al sonido cardiaco. Para conseguir esto, el filtro de Wiener plantea como solución reducir el MSE (*Mean Square Error*) entre la señal estimada a la salida y la señal deseada como se verá más adelante.

En la **Figura 4.1** se muestra una representación del diagrama de bloques del filtro de Wiener.

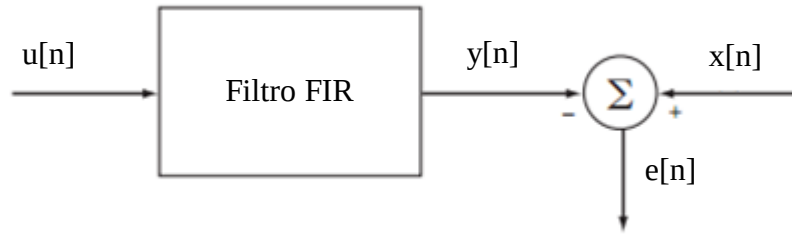


Figura 4.1 - Diagrama de bloques filtro de Wiener [36]. La señal deseada $x[n]$ se corresponde con la señal mezcla y la señal de referencia $u[n]$ debe ser una señal con un gran parecido al sonido cardiaco de la señal mezcla, de este modo, la señal de salida $y[n]$ se corresponderá con el sonido cardiaco recuperado y la señal de error $e[n]$ se corresponderá con el sonido pulmonar recuperado.

Tras la introducción, se realizará a continuación el desarrollo matemático de este método hasta alcanzar la solución final, de manera que sea más fácil entender su funcionamiento.

Se considera un filtro FIR de L coeficientes del siguiente modo [36]:

$$\mathbf{w}_n = [w(0) \ w(1) \ \dots \ w(L-1)]^T \quad (4.1)$$

$$W(z) = w(0) + w(1)z^{-1} + \dots + w(L-1)z^{-L+1}$$

donde $\mathbf{w}_n$ son los L coeficientes del filtro FIR en la muestra n . Se obtendrá, por tanto, una salida en el filtro que vendrá dada por la siguiente ecuación [36]:

$$y[n] = \sum_{k=0}^{L-1} w_n(k) \cdot u[n-k] \quad (4.2)$$

$$y[n] = [w(0) \ w(1) \ \dots \ w(L-1)] \begin{bmatrix} u[n] \\ u[n-1] \\ \vdots \\ u[n-L+1] \end{bmatrix}$$

$$y[n] = w_n^T u_n$$

en la cual, $\mathbf{u}_n$ se corresponde con un vector de tamaño $\mathbf{L} \times \mathbf{1}$ con las últimas $\mathbf{L}$ muestras de la señal de referencia $\mathbf{u}[n]$ que entran al filtro en el instante temporal $\mathbf{n}$.

Como ya se ha explicado anteriormente, el filtro de Wiener tenderá por tanto a minimizar una función de coste definida por el MSE de la siguiente forma [36]:

$$J(w) = E[(e[n])^2] \quad (4.3)$$

$$J(w) = E[(x[n] - w_n^T u_n)^2]$$

$$J(w) = E[(x[n])^2 + (w_n^T u_n)^2 - 2(w_n^T u_n)x[n]]$$

$$J(w) = \sigma_x^2 + w_n^T E[u_n u_n^T] w_n - 2w_n^T E[u_n x[n]]$$

donde σ_x^2 es la varianza de la señal deseada, $E[\mathbf{u}_n \mathbf{u}_n^T]$ es la matriz de autocorrelación de la señal de referencia de tamaño $\mathbf{L} \times \mathbf{L}$ obtenida de la siguiente forma [36]:

$$E[u_n u_n^T] = E \left[\begin{bmatrix} u[n] \\ u[n-1] \\ \vdots \\ u[n-L+1] \end{bmatrix} \begin{bmatrix} u[n] & u[n-1] & \dots & u[n-L+1] \end{bmatrix} \right] \quad (4.4)$$

$$E[u_n u_n^T] = R_u$$

y $E[\mathbf{u}_n \mathbf{x}[n]]$ es un vector de tamaño $L \times 1$ correspondiente a la correlación cruzada de la señal de referencia y la señal deseada [36]:

$$E[u_n d[n]] = E \left[\begin{bmatrix} u[n] \\ u[n-1] \\ \vdots \\ u[n-L+1] \end{bmatrix} x[n] \right] \quad (4.5)$$

$$E[u_n x[n]] = p$$

Si se introducen los resultados obtenidos en (4.4) y (4.5) en (4.3), se obtiene una función de coste final dada por la siguiente ecuación [36]:

$$J(w) = \sigma_x^2 + w_n^T R_u w_n - 2w_n^T p \quad (4.6)$$

Se quiere, por tanto, obtener los coeficientes óptimos del filtro que minimicen esta función de coste, para lo cual, se calcula el gradiente de la función de coste obtenida en (4.6) y se iguala a cero de la siguiente forma [36]:

$$\nabla J(w) = 2R_u w_{opt} - 2p = 0 \quad (4.7)$$

$$R_u w_{opt} = p$$

$$w_{opt} = R_u^{-1} p$$

Alcanzado el valor óptimo obtenido en (4.7), se daría un valor mínimo de la función de coste (potencia del error) dado por la siguiente ecuación [36]:

$$J_{min} = \sigma_x^2 - w_{opt}^T \cdot p \quad (4.8)$$

Hasta ahora, se ha visto cómo se pueden obtener los coeficientes óptimos del filtro que permiten obtener la potencia de error mínima aplicando el gradiente de la función de coste, sin embargo, en la práctica, este método carece de utilidad puesto que las características estadísticas (R_u y p) de la señal de entrada pueden cambiar con el tiempo y, por tanto, habría que volver a resolver todas las ecuaciones cada vez que se produzca este cambio.

4.1.1.2. Método de máxima pendiente

El objetivo de este trabajo es realizar la separación de sonidos cardiacos y pulmonares, ambas son señales no estacionarias, es decir, sus estadísticos variaran a lo largo del tiempo, por lo tanto, como ya se comentó en el punto anterior, el filtro de Wiener no es una solución adecuada para este tipo de señales.

El algoritmo de máxima pendiente es el método propuesto que permite alcanzar los coeficientes óptimos del filtro de forma iterativa y adaptarse de este modo a cualquier cambio en la señal de entrada. Este algoritmo es básico para entender las distintas formas en que la adaptación basada en gradiente puede ser implementada en la práctica.

El gradiente de la función de coste evaluado en un punto indica la dirección de máximo incremento de la función en ese punto, por lo tanto, si se quiere alcanzar el valor mínimo de la función de coste y obtener de este modo los coeficientes óptimos del filtro, hay que moverse en la dirección contraria al gradiente $-\nabla J(\mathbf{w})$. De manera que partiendo de $\mathbf{w}_0$ cuyo valor se suele inicializar a cero, los coeficientes del filtro se actualizarán con el objetivo de que la función de coste sea menor en la siguiente iteración del siguiente modo [36]:

$$w_{n+1} = w_n + \frac{1}{2} \mu (-\nabla J(w)) \quad (4.9)$$

$$w_{n+1} = w_n + \mu(p - R_u w_n)$$

donde μ controla el tamaño del paso para alcanzar la solución óptima $\mathbf{w}_{opt}$. Debido a la realimentación, existe la posibilidad de que el filtro se vuelva inestable, para evitar esto y conseguir que el algoritmo converja se debe escoger un valor de μ acotado de la siguiente forma [36]:

$$0 < \mu < \frac{2}{\lambda_{max}} \quad (4.10)$$

donde λ_{max} se corresponde con el máximo autovalor de la matriz $\mathbf{R}_u$.

4.1.1.3.LMS

Hasta ahora, se ha visto cómo se realiza el cálculo de la función de coste y como se puede reducir iterativamente utilizando el algoritmo del máximo descenso, sin embargo, para actualizar los coeficientes del filtro tal y como se vio en (4.9) es necesario conocer tanto la matriz de autocorrelación $\mathbf{R}_u$ de la señal de referencia como la correlación cruzada $\mathbf{p}$ de la señal de referencia con la señal deseada.

El algoritmo LMS, inventado en 1960 por Bernard Widrow [37], no intenta reducir el MSE, donde aparecen correlaciones como se ha visto anteriormente, sino únicamente el cuadrado del error instantáneo. el gradiente de la función de coste pasaría a ser lo que se conoce como gradiente estocástico [36]:

$$\widehat{\nabla}J(w) = 2u_n u_n^T w_{opt} - 2u_n x[n] = 0 \quad (4.11)$$

$$\widehat{\nabla}J(w) = 2u_n (y[n] - x[n])$$

$$\widehat{\nabla}J(w) = -2u_n e[n]$$

y la actualización de los coeficientes del filtro vista en (4.9) pasaría a ser [36]:

$$\widehat{w}_{n+1} = \widehat{w}_n + \mu u_n e[n] \quad (4.12)$$

El algoritmo convergerá y no se volverá inestable si se cumple que el valor de μ está comprendido en un rango dado en [36] como:

$$0 < \mu < \frac{2}{L \cdot \sigma_u^2} \quad (4.13)$$

En la elección del parámetro μ , un valor grande supondrá que los coeficientes del filtro varíen más bruscamente, es decir, una mayor velocidad de convergencia y por tanto un menor tiempo en alcanzar los coeficientes óptimos del filtro. El algoritmo convergirá geométricamente con una constante de tiempo dada en [36] por:

$$\tau_{LMS} \cong \frac{1}{2\mu\lambda_{min}} \quad (4.14)$$

donde λ_{min} se corresponde con el mínimo autovalor de la matriz de autocorrelación del vector de entrada. Sin embargo, como se ha mencionado anteriormente, el filtro LMS, a diferencia del filtro de Wiener, se basa en estimaciones, por lo tanto, no llega a alcanzar el valor mínimo de la función de coste J_{min} obtenido en (4.8) sino que alcanzará un valor superior que se denominará $J(\infty)$. Una medida útil para cuantificar esta diferencia es el desajuste D que viene determinado por la siguiente ecuación [36]:

$$D = \frac{J(\infty) - J_{min}}{J_{min}} \quad (4.15)$$

$$D = \frac{\mu L \sigma_u^2}{2}$$

Como se puede observar en la ecuación anterior, el desajuste aumentará a medida que aumenta el valor de μ , por lo que se debe intentar escoger un valor lo más pequeño posible, sin embargo, este valor tampoco debe ser excesivamente pequeño puesto que como se vio en (4.14) a medida que se reduce este valor se reducirá la velocidad de convergencia. Por tanto, surge una difícil decisión en la elección del parámetro μ : incrementar la velocidad

de convergencia a cambio de aumentar el desajuste o reducir la velocidad de convergencia consiguiendo de este modo una reducción del desajuste.

4.1.1.4. VSS-LMS

Como se ha descrito anteriormente, la elección del parámetro μ conlleva una seria dificultad puesto que se ha de mantener un equilibrio entre la velocidad de convergencia y el desajuste. Como solución, existen algoritmos de tamaño de paso variable, es decir, su valor no permanece constante en el tiempo, sino que cambia en función de las circunstancias. Cuando la señal de error sea grande, quiere decir que los coeficientes actuales del filtro están muy lejos de ser los coeficientes óptimos, por tanto, el tamaño de paso aumentará para incrementar la velocidad de convergencia y, una vez que el algoritmo haya convergido y se reduzca la señal de error, se disminuirá el valor del tamaño de paso con el objetivo de reducir el desajuste. En definitiva, la principal ventaja de VSS-LMS respecto a LMS es que se puede conseguir un desajuste reducido sin necesidad de renunciar a una elevada velocidad de convergencia, esto se traduce en una mejor adaptación a los posibles cambios en la señal formada por la mezcla de los sonidos cardiacos y sonidos pulmonares que se quiere filtrar.

En este trabajo se implementará el algoritmo VSS-LMS descrito en [38]. De esta forma, el tamaño paso pasaría a actualizarse en cada iteración de la siguiente forma:

$$\begin{aligned} \mu[n+1] &= \alpha\mu[n] + \gamma e^2[n] & 0 < \alpha < 1 \\ & & \gamma > 0 \end{aligned} \quad (4.16)$$

donde α controla la memoria respecto al valor de μ anterior, por lo que un valor de α pequeño provocará una mayor fluctuación en los valores de μ , y γ regula el incremento del tamaño de paso en función de la señal de error obtenida. Para evitar que el valor de μ este fuera de los límites establecidos en (4.13) se realiza el siguiente control [38]:

$$\mu[n+1] = \begin{cases} \mu_{max} & \mu[n+1] > \mu_{max} \\ \mu_{min} & \mu[n+1] < \mu_{min} \\ \mu[n+1] & \mu_{min} < \mu[n+1] < \mu_{max} \end{cases} \quad (4.17)$$

La actualización de los coeficientes del filtro vista en **(4.12)** pasaría a ser:

$$\hat{w}_{n+1} = \hat{w}_n + \mu[n]u_n e[n] \quad (4.18)$$

4.1.1.5. Señal de referencia

Hasta ahora, se ha visto todo el procedimiento matemático que sigue un filtro adaptativo para obtener los coeficientes del filtro en función de la señal de error que se recibe en cada instante de tiempo. Para que un filtro adaptativo funcione se debe cumplir que la señal de referencia presente una cierta correlación con la parte de la señal deseada que se quiere recuperar, pero, en ocasiones, como ya se comentó en el capítulo 3 correspondiente al estado del arte, resulta realmente complicado obtener esta señal de referencia.

El ALE es un tipo especial de filtro adaptativo en el cual se utiliza como señal de referencia la misma señal deseada con un determinado retardo Δ . Se considera que la señal deseada es la siguiente:

$$x[n] = c[n] + p[n] \quad (4.19)$$

es decir, está formada por la mezcla del sonido cardiaco $c[n]$, que queremos recuperar, y el sonido pulmonar $p[n]$. Este tipo de filtro funciona cuando la señal deseada está formada por la suma de una señal de banda estrecha y una señal de banda ancha debido a que una señal de banda estrecha presentará una mayor correlación con las muestras anteriores a diferencia de una señal de banda ancha. La salida del ALE pasaría a ser, a diferencia de la vista en **(4.2)**, la siguiente:

$$y[n] = \sum_{k=0}^{L-1} w_n(k) \cdot x[n - k - \Delta] \quad (4.20)$$

Por tanto, se debe seleccionar un valor de Δ lo suficientemente grande para que, al comparar la señal mezcla $x[n]$ consigo misma retardada $x[n - \Delta]$, el sonido pulmonar

presente en la señal mezcla $p[n]$ no presente ningún parecido con el sonido pulmonar presente en la señal mezcla retardada $p[n - \Delta]$, de este modo, será imposible determinar los valores futuros del sonido pulmonar en base a muestras pasadas. Mientras que el sonido cardiaco presente en la señal mezcla $c[n]$ presentará un gran parecido con el sonido cardiaco presente en la señal mezcla retardada $c[n - \Delta]$ por lo que se podrán determinar las muestras futuras del sonido de corazón en base a muestras pasadas. De este modo, el filtro será capaz de recuperar el sonido cardiaco a la salida introduciendo únicamente un retardo a la señal mezcla.

4.2.2. Señal sinusoidal + AWGN

Una vez se ha realizado el desarrollo matemático del método, así como una explicación básica de su funcionamiento se realizarán algunas pruebas para ver el funcionamiento del algoritmo en función de los distintos parámetros que se han comentado en los apartados anteriores. Para ello, se trabajará en primer lugar con una señal sinusoidal a la que se le añadirá ruido AWGN de manera que se puedan ver más claramente los resultados obtenidos.

Se parte una señal de entrada formada por la combinación de dos señales sinusoidales de 10 y 30 Hz a la que se le añade ruido AWGN con una SNR de 5dB como se puede ver en la figura mostrada a continuación:

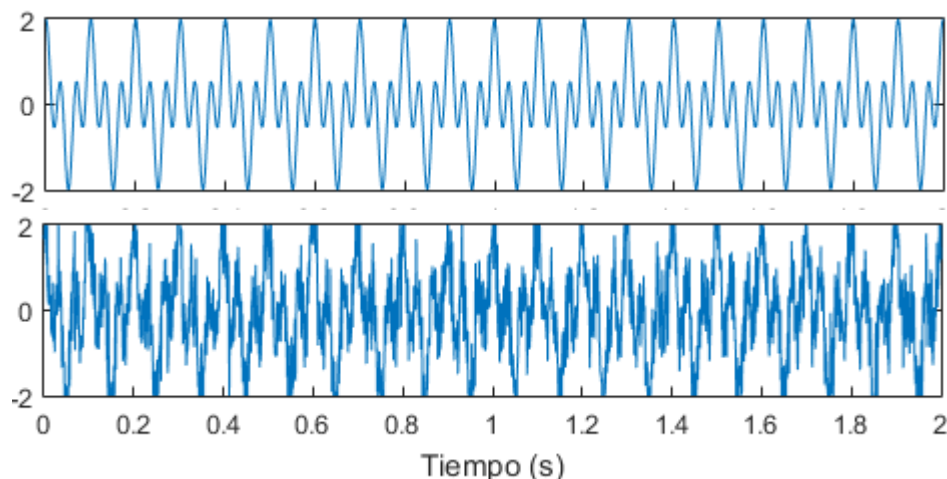


Figura 4.2 - Señal sinusoidal de 10 y 30 Hz (arriba) y señal ruidosa SNR=5dB (abajo)

Se muestran a continuación las autocorrelaciones tanto de la señal sinusoidal como de la señal ruidosa de manera que se pueda demostrar como la señal de banda estrecha, es

decir la señal sinusoidal presenta una mayor correlación con las muestras anteriores que una señal de banda ancha, como es el caso del ruido AWGN.

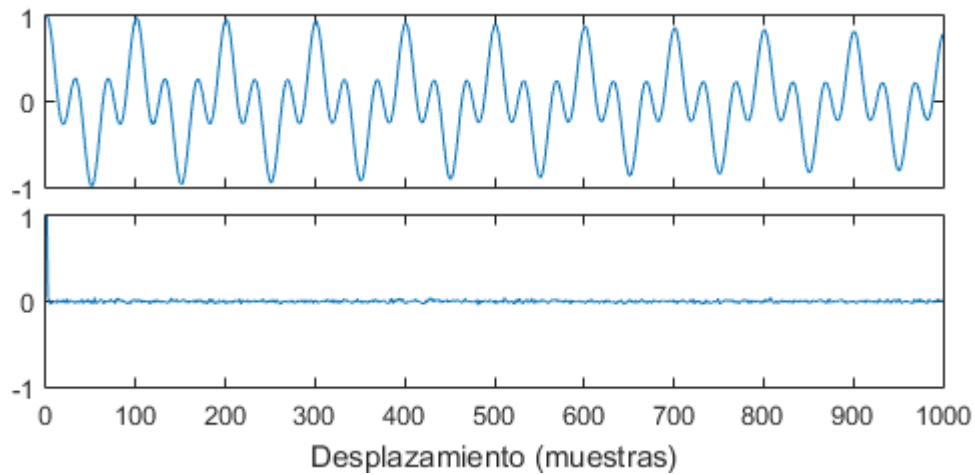


Figura 4.3 - Autocorrelación de la señal sinusoidal de 10 y 30 Hz (arriba) y autocorrelación del ruido AWGN (abajo)

Una vez vistas las gráficas anteriores, se deduce, por tanto, que se debe escoger un valor de Δ pequeño, pues, para ese valor, el ruido presente en la señal deseada no se encontrará correlacionado por lo que el filtro será incapaz de seguir sus cambios y, por tanto, la salida se corresponderá con la señal sinusoidal sin ruido AWGN como se muestra a continuación:

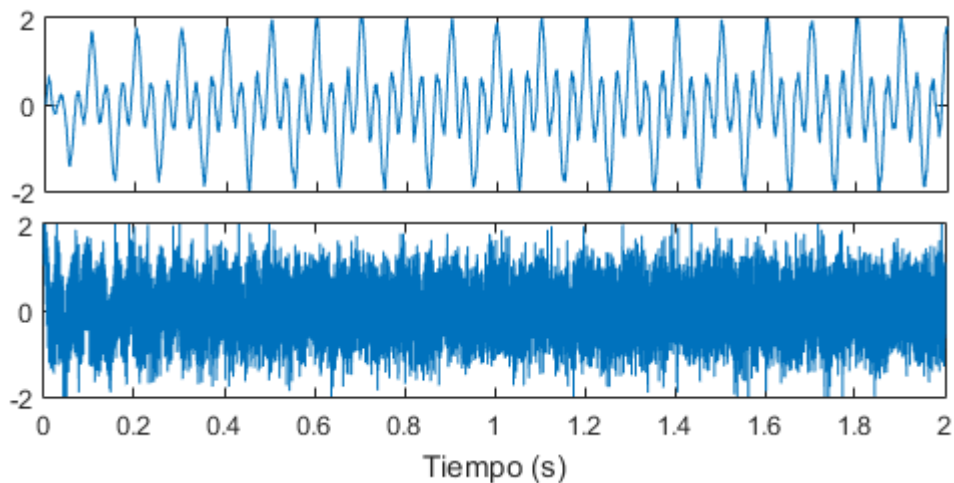


Figura 4.4 – Señal sinusoidal recuperada (arriba) y ruido recuperado (abajo) utilizando un filtro ALE VSS-LMS

Se representa a continuación la respuesta en frecuencia del filtro una vez ha convergido. Se puede ver como aparecen dos lóbulos localizados en las frecuencias de la señal sinusoidal, es decir, en 10 y 30 Hz.

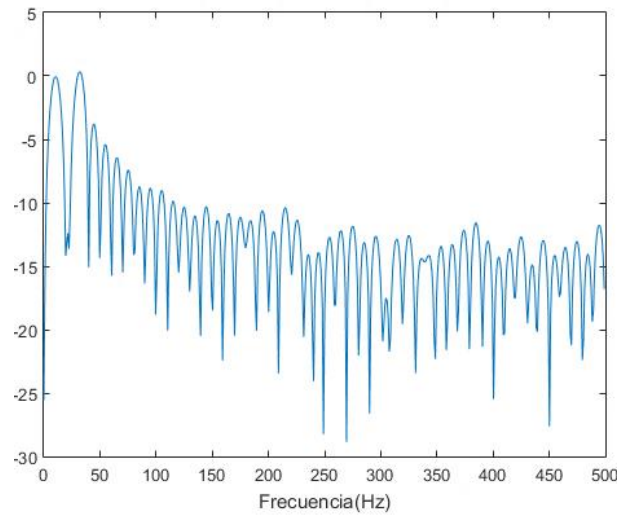


Figura 4.5 - Respuesta en frecuencia del filtro ALE VSS-LMS con $L=100$ con dos lóbulos principales en las frecuencias de 10 y 30 Hz.

Si en lugar de utilizar un filtro de $L = 100$ se utiliza uno de $L = 50$ se puede ver cómo se reduce la resolución en frecuencia del filtro:

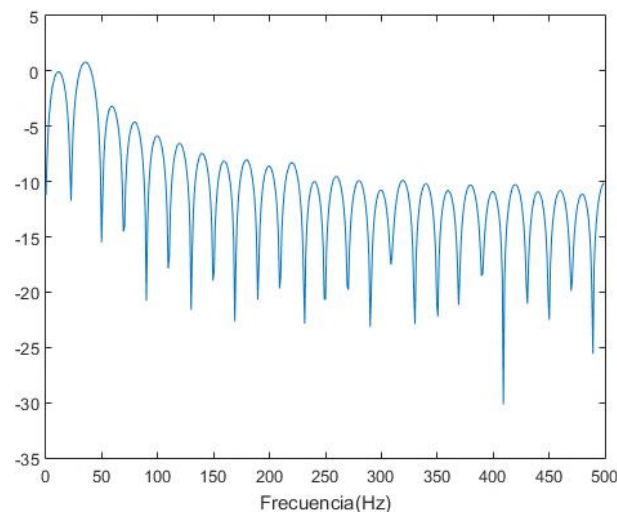


Figura 4.6 - Respuesta en frecuencia del filtro ALE VSS-LMS con $L=50$ con dos lobulos principales en las frecuencias de 10 y 30 Hz.

4.2.3. Sonidos cardiacos y pulmonares

En este apartado se comprobará el funcionamiento del filtro ALE en la separación de sonidos cardiacos y pulmonares. Como ya se explicó en el capítulo introductorio, la mayor parte de los sonidos cardiacos están situados en el rango de 30 a 150 Hz mientras que los sonidos de pulmón están situados entre 20 y 1200 Hz.

De lo explicado anteriormente, se deduce como las características de ambos sonidos los hace muy similares al ejemplo explicado para una señal sinusoidal y ruido AWGN, por lo que, a priori, se trata de un caso apto para la utilización del filtro ALE.

Las señales utilizadas a continuación han sido obtenidas de la misma base de datos que será utilizada para la optimización de los parámetros en el capítulo siguiente. Se parte de una señal de entrada formada por la combinación de sonidos cardiacos y pulmonares con una SNR de 5dB como se puede ver a continuación:

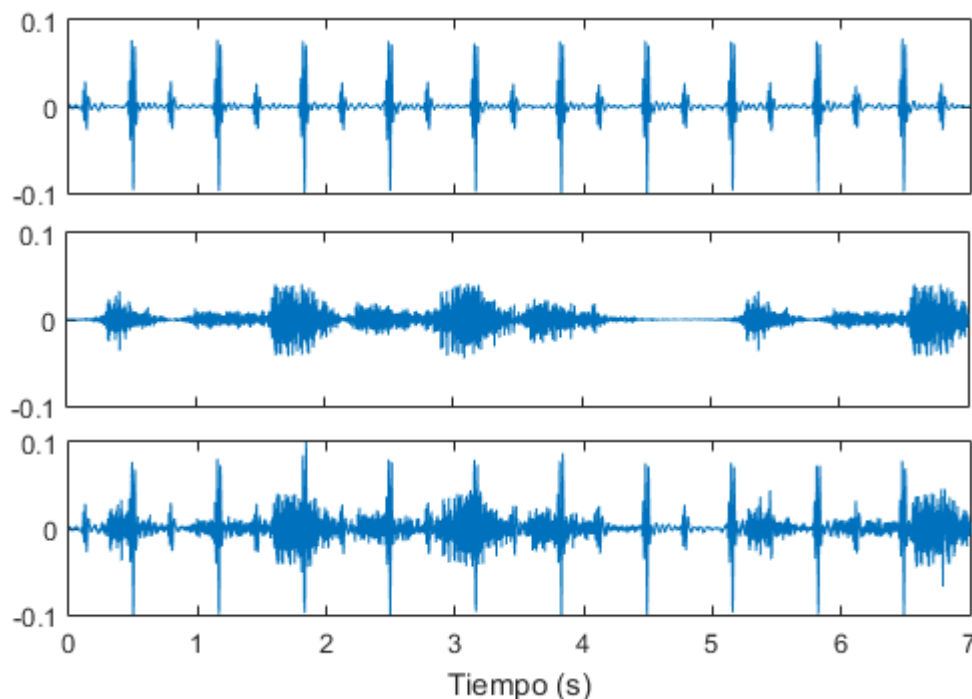


Figura 4.7 - Sonido cardiaco (arriba), sonido pulmonar (medio) y señal mezcla SNR=5dB (abajo). Estos ficheros de audio se adjuntan en la carpeta de nombre “Software y Ficheros de Audio” anexa a la memoria.

Se muestra a continuación, al igual que se hizo para la señal sinusoidal y ruido AWGN, las autocorrelaciones del sonido cardiaco y del sonido pulmonar de manera que se pueda comprobar que efectivamente cumplen con los requisitos establecidos con anterioridad para el correcto funcionamiento del filtro ALE.

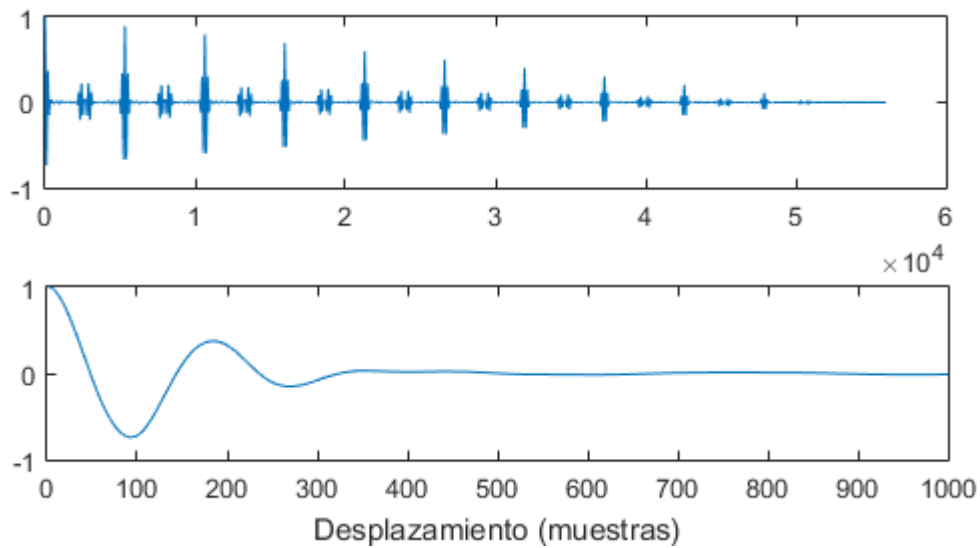


Figura 4.8 – Autocorrelación del sonido cardiaco visto en la Figura 4.7

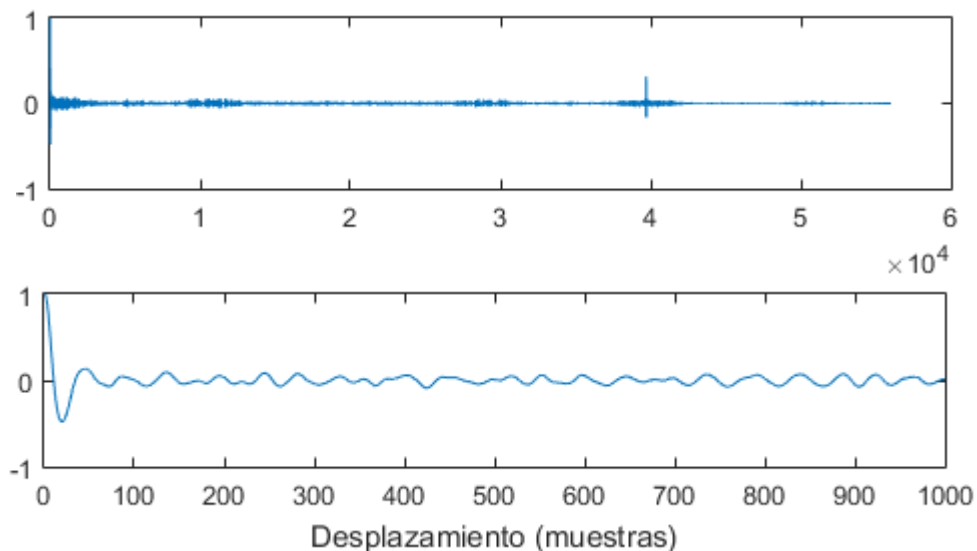


Figura 4.9 - Autocorrelación del sonido pulmonar visto en la Figura 4.7

Como se puede ver en la **Figura 4.9**, la autocorrelación del sonido pulmonar cae de forma más pronunciada que en el caso de la autocorrelación del sonido cardiaco que se

muestra en la **Figura 4.8** por lo que, por tanto, se cumplen con las condiciones para realizar el filtrado utilizando un esquema ALE.

Se muestra a continuación el sonido cardiaco y sonido pulmonar recuperados tras aplicar el filtrado a la señal mezcla:

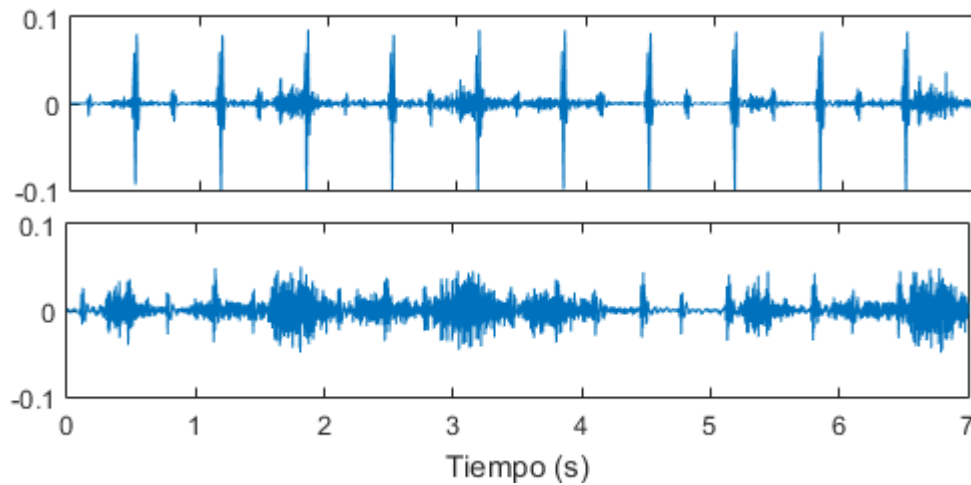


Figura 4.10 - Sonido cardiaco recuperado (arriba) y sonido pulmonar recuperado (abajo) utilizando un filtro ALE VSS-LMS

A continuación, se puede ver la variación en el valor del tamaño de paso del algoritmo VSS-LMS. Se observa como su valor se incrementa cada vez que aparece un sonido cardiaco lo que le permite de esta forma seguir más fácilmente los cambios en la señal y conseguir de esta forma que el filtro se adapte más rápidamente y su funcionamiento sea independiente de la elección del parámetro μ .

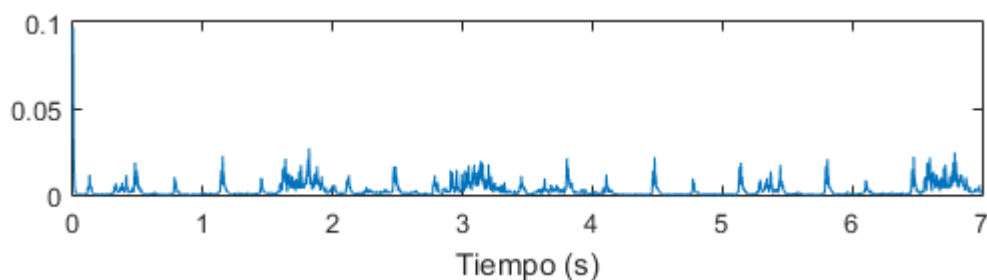


Figura 4.11 - Variación del parámetro μ utilizando un filtro ALE VSS-LMS

Como se ha visto en las figuras anteriores, el algoritmo funciona adecuadamente puesto que logra separar el sonido cardiaco y pulmonar, sin embargo, aún sigue habiendo restos del sonido pulmonar en el sonido cardiaco y viceversa, por lo que se plantea como solución el uso de una máscara tiempo-frecuencia.

Una vez se han obtenido las señales $\mathbf{c}_r$ y $\mathbf{p}_r$, que se corresponden con el sonido cardiaco y el sonido pulmonar recuperados respectivamente, se calcula la transformada STFT de cada uno de ellos y de la señal mezcla. Se obtienen, por tanto, 3 matrices complejas: la matriz $\mathbf{X}$, que se corresponde con la transformada STFT de la señal mezcla, la matriz $\mathbf{C}_r$, que se corresponde con la transformada STFT del sonido de corazón recuperado, y la matriz $\mathbf{P}_r$ que se corresponde con la transformada STFT del sonido de pulmón recuperado. Posteriormente, se calcula la máscara de Wiener utilizando la ecuación vista en (1.12) del siguiente modo:

$$M_c = \frac{\|\mathbf{C}_r\|^2}{\|\mathbf{C}_r\|^2 + \|\mathbf{P}_r\|^2} \quad (4.21)$$

$$M_p = 1 - M_c$$

Tras aplicar la operación anterior se obtiene dos matrices reales: $\mathbf{M}_c$ que se corresponde con la máscara de corazón y una matriz $\mathbf{M}_p$ que se corresponde con la máscara de pulmón. Se muestra a continuación una representación de las máscaras obtenidas:

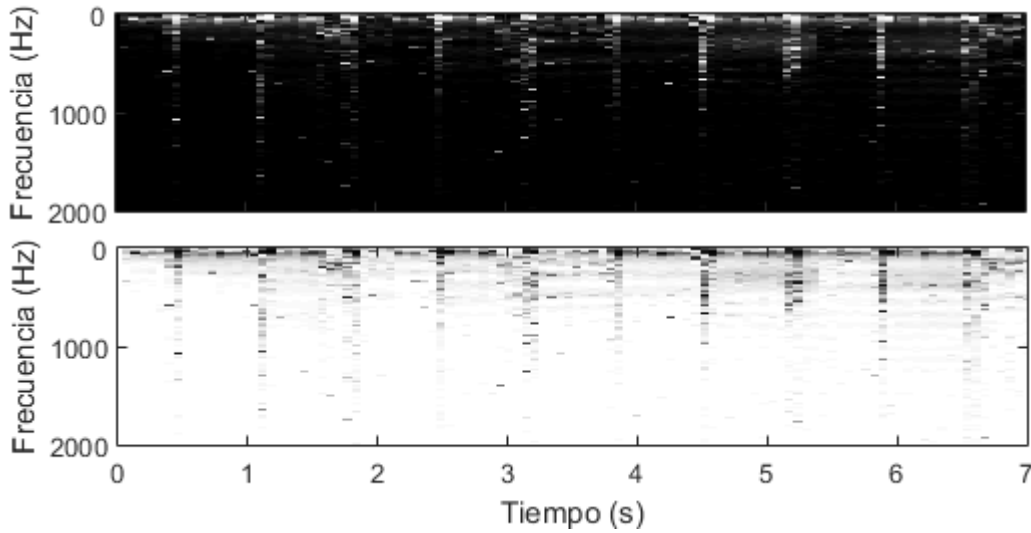


Figura 4.12 - Máscara de corazón (arriba) y máscara de pulmón (abajo) obtenidas a partir de las salidas tras aplicar un filtro ALE VSS-LMS

Como se puede ver en la **Figura 4.12**, la máscara de corazón presenta valores más cercanos a uno (blanco) en aquellas zonas donde predomina el sonido de corazón frente al sonido pulmonar y valores más próximos a cero (negro) en aquellas zonas en las que, por el contrario, predomina el sonido pulmonar frente al sonido cardiaco. En el caso de la máscara de pulmón se produce la situación contraria a la máscara de corazón.

Posteriormente, se multiplican las máscaras obtenidas por la transformada STFT de la señal mezcla de la siguiente forma:

$$C_r^M = X \cdot M_c \quad (4.22)$$

$$P_r^M = X \cdot M_p$$

Al multiplicar las máscaras obtenidas por la transformada STFT de la señal mezcla se obtiene la matriz compleja C_r^M correspondiente a la transformada STFT del sonido cardiaco tras aplicar la máscara y la matriz compleja P_r^M que se corresponde con la transformada STFT del sonido pulmonar tras aplicar la máscara. Posteriormente, se aplica la transformada ISTFT y se obtienen las señales c_r^M y p_r^M en el dominio temporal que se muestran a continuación:

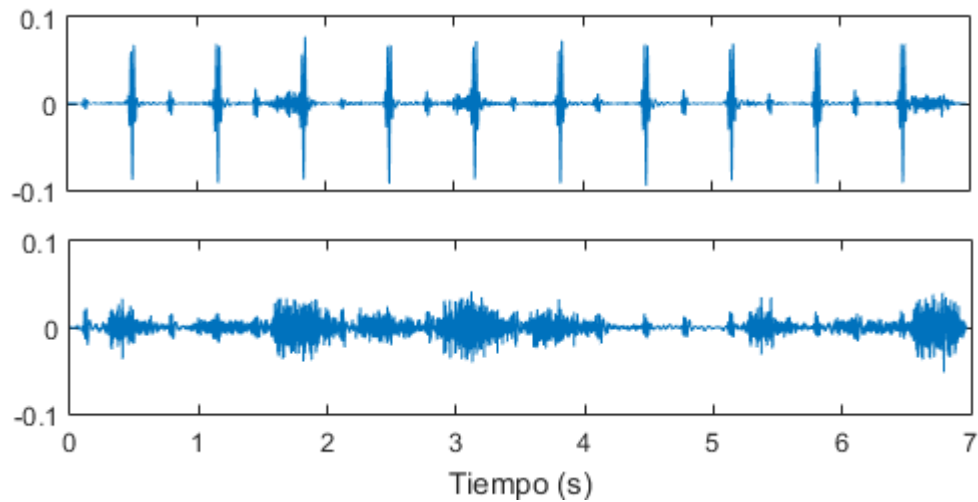


Figura 4.13 - Sonido cardiaco recuperado (arriba) y sonido pulmonar recuperado (abajo) tras aplicar una máscara Wiener a la salida del filtro ALE VSS-LMS. Estos ficheros de audio se adjuntan en la carpeta de nombre “Software y Ficheros de Audio” anexa a la memoria

Tras ver la figura anterior, se puede apreciar cómo se reducen notablemente el sonido de corazón presente en el sonido pulmonar y viceversa tras aplicar la máscara, aunque aún siguen quedando algunos pequeños restos.

A continuación, se muestra la densidad espectral de potencia de los sonidos de corazón y de pulmón originales junto con los sonidos recuperado antes y después de aplicar la máscara.

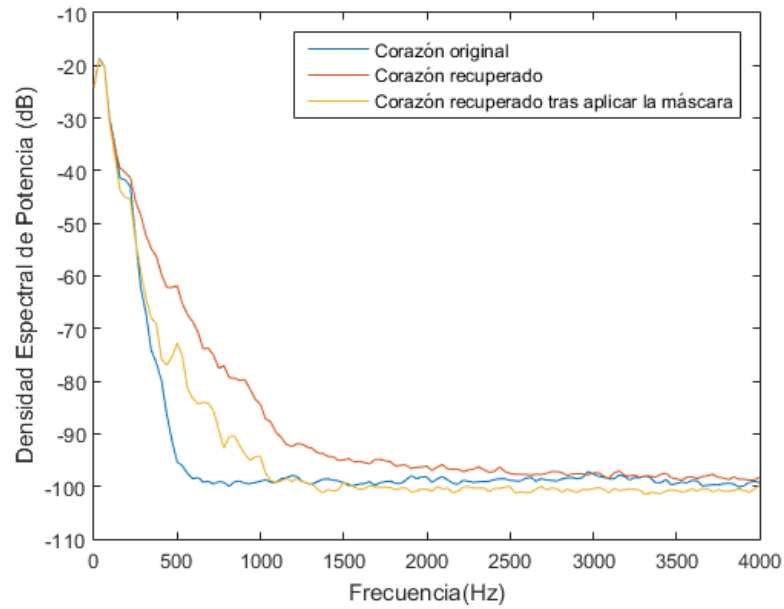


Figura 4.14 - PSD del sonido de corazón original, sonido de corazón recuperado y sonido de corazón recuperado tras aplicar una máscara Wiener a la salida del filtro ALE VSS-LMS

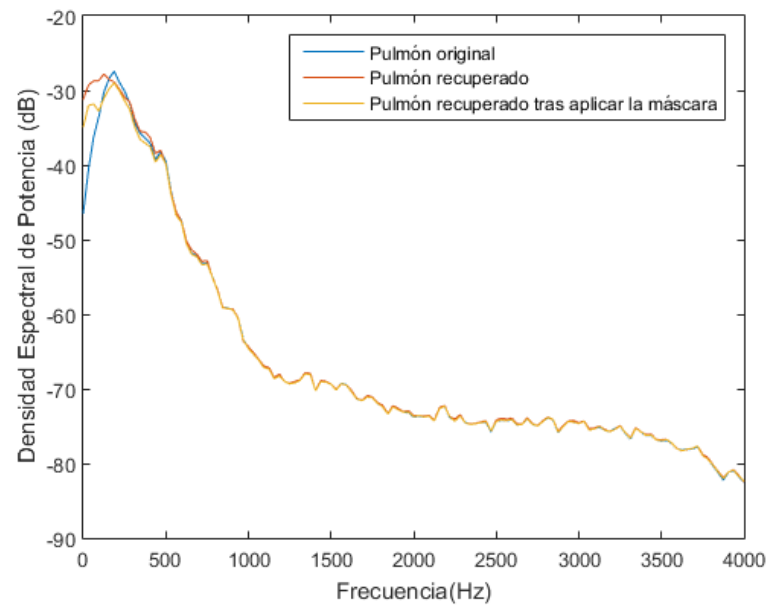


Figura 4.15 - PSD del sonido de pulmón original, sonido de pulmón recuperado y sonido de pulmón recuperado tras aplicar la máscara Wiener a la salida del filtro ALE VSS-LMS

A partir de las figuras anteriores se puede comprobar como las señales obtenidas a la salida del filtro se parecen en gran medida a las señales originales antes de ser mezcladas. Se muestra a continuación los espectrogramas de las señales originales junto con los de las señales recuperadas:

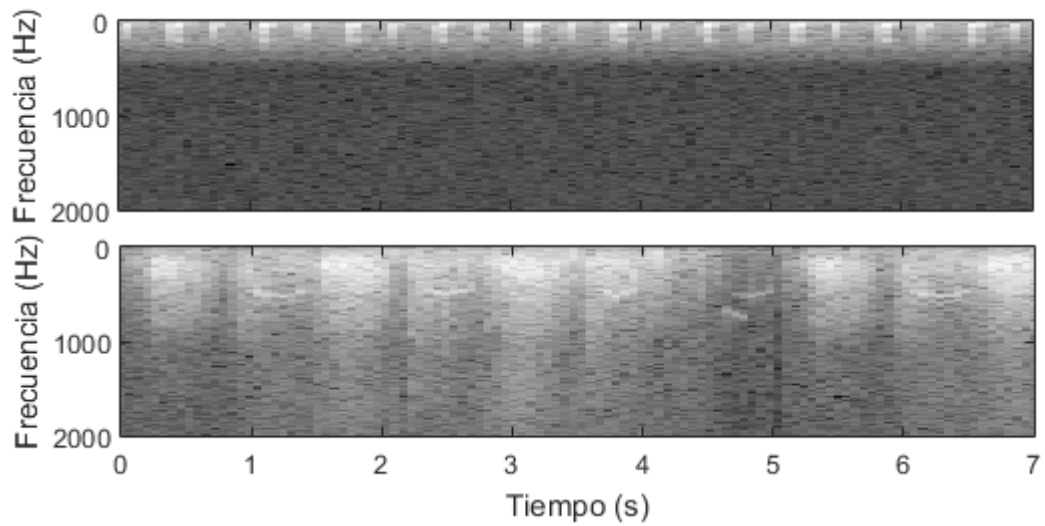


Figura 4.16 - Espectrogramas del sonido de corazón original (arriba) y sonido de pulmón original (abajo) vistos en la Figura 4.7

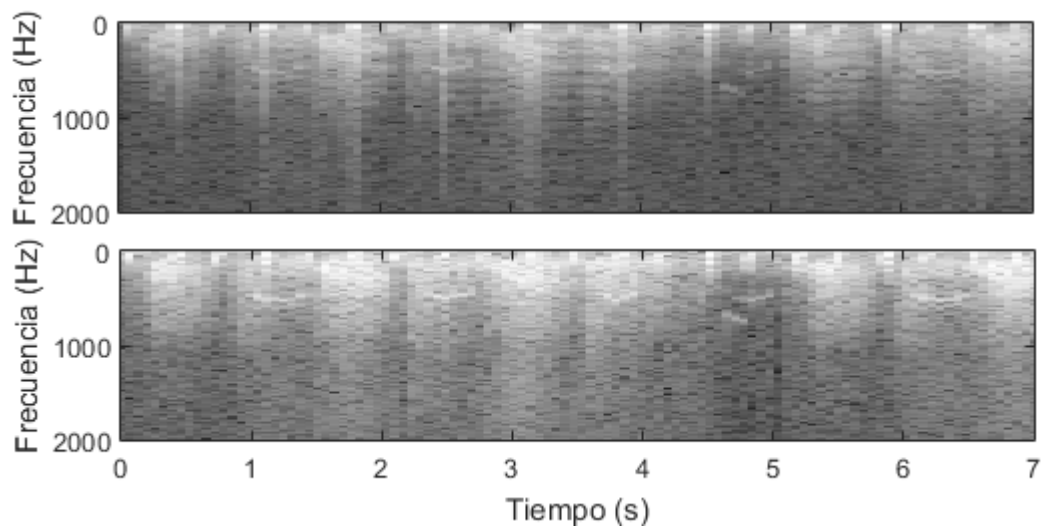


Figura 4.17 - Espectrogramas del sonido de corazón recuperado (arriba) y sonido de pulmón recuperado (abajo) utilizando un filtro ALE VSS-LMS

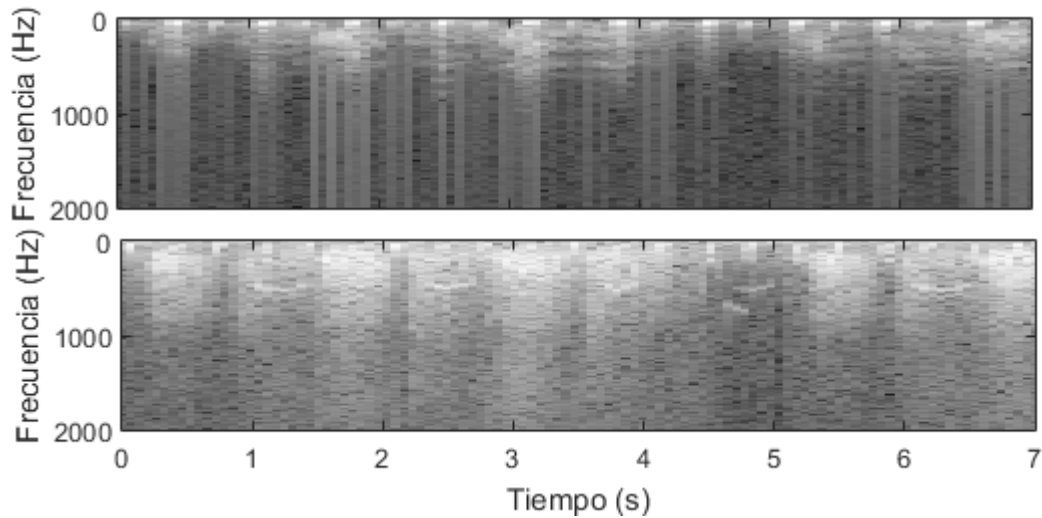


Figura 4.18 - Espectrogramas del sonido de corazón recuperado (arriba) y sonido de pulmón recuperado (abajo) tras aplicar una máscara Wiener a la salida del filtro ALE VSS-LMS

Una vez se ha comprobado el correcto funcionamiento del filtro ALE, únicamente faltaría realizar la optimización de los parámetros de manera que se obtengan aquellos valores que permitan realizar la separación de los sonidos cardíacos y pulmonares de la mejor manera posible. Se muestra a continuación, a modo de resumen, una tabla con los parámetros de los métodos ALE LMS y ALE VSS-LMS:

Parámetros ALE LMS		Parámetros ALE VSS-LMS	
L	Número de coeficientes del filtro FIR	L	Número de coeficientes del filtro FIR
Δ	Retardo de la señal de referencia respecto a la señal deseada	Δ	Retardo de la señal de referencia respecto a la señal deseada
μ	Valor del tamaño de paso	α	Factor de olvido respecto al valor de tamaño de paso anterior
		γ	Variación del tamaño de paso en función del error obtenido
		μ_{min}	Valor mínimo del tamaño de paso
		μ_{max}	Valor máximo del tamaño de paso

Tabla 4.1 - Resumen de los parámetros utilizados en el filtro ALE LMS y ALE VSS-LMS

4.2. NLM

4.2.1. Desarrollo matemático

El objetivo del algoritmo NLM es aprovechar la información redundante de los sonidos cardiacos, es decir, la existencia de segmentos de estos sonidos libres de ruido que se repetirán en distintos instantes temporales. Por tanto, si se hace uso del parecido entre estos segmentos se podrá eliminar toda aquella señal ruidosa mediante el siguiente desarrollo.

Se parte de una señal de entrada $\mathbf{x}[\mathbf{n}]$:

$$x[n_i] = c[n_i] + p[n_i] \quad i = 1, \dots, N \quad (4.23)$$

compuesta por la suma de una señal de interés $\mathbf{c}[\mathbf{n}]$, que se corresponderá con el sonido del corazón y por una señal ruidosa $\mathbf{p}[\mathbf{n}]$ que será el sonido del pulmón, siendo $\mathbf{N}$ el número de muestras de la señal de entrada $\mathbf{x}[\mathbf{n}]$ formada por la mezcla de ambas.

El objetivo es obtener una estimación de la señal de interés $\mathbf{c}[\mathbf{n}]$ a partir de la señal de entrada $\mathbf{x}[\mathbf{n}]$, para ello, este método plantea la siguiente solución [6]:

$$c[n_i] = \sum_j w(i, j) * x[n_i] / \sum_j w(i, j) \quad (4.24)$$

donde $\mathbf{w}(\mathbf{i}, \mathbf{j})$ no es más que la similitud entre dos segmentos de la señal $\mathbf{x}[\mathbf{n}]$ como se verá más adelante. Una vez que cada muestra haya sido multiplicada por su correspondiente peso, se divide por un factor de normalización $\sum_j \mathbf{w}(\mathbf{i}, \mathbf{j})$ de manera que la muestra de salida se ajuste a los valores de la señal original. El cálculo de la similitud entre dos segmentos de la señal se realiza utilizando la distancia Euclidea del siguiente modo [6]:

$$\|u_i - u_j\| = \sqrt{(u_i - u_j)^T (u_i - u_j)} \quad (4.25)$$

donde u_i y u_j se corresponden con dos segmentos de la señal $x[n]$ de tamaño $2P + 1$ centrados en las muestras temporales n_i y n_j respectivamente. En la figura siguiente se muestra un ejemplo utilizando un valor de $P = 2$ *muestras*.

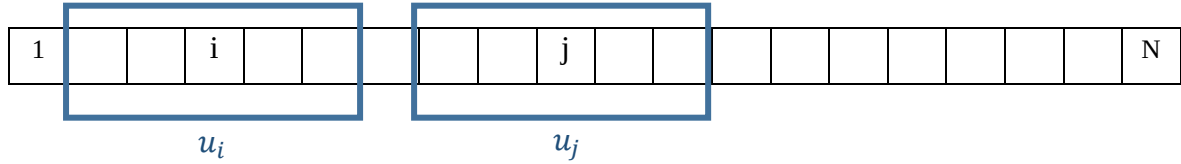


Figura 4.19 - Ejemplo de dos segmentos a comparar para $P=2$ muestras.

Una vez hecho esto se calculará el valor del peso correspondiente de la siguiente manera:

$$w(i, j) = \exp(-\|u_i - u_j\|^2 / 2(2P + 1)h^2) \quad (4.26)$$

en la que el valor de h controla el decaimiento de la señal exponencial y por tanto el grado de filtración de la señal, un valor grande provocará que haya un mayor parecido entre dos fragmentos y, por tanto, que se elimine una mayor cantidad de ruido mientras que también se perderán los detalles de la señal original. El termino $(2P + 1)$ elimina la dependencia del valor del peso $w(i, j)$ con el tamaño P elegido puesto que al aumentar este valor también aumentará el de $\|u_i - u_j\|^2$.

4.2.2. Coste computacional

Tras implementar el método siguiendo el desarrollo matemático visto en el punto anterior y realizar algunas pruebas se observó que, tal y como ya se había explicado, uno los principales inconvenientes de este método es el elevado coste computacional. El hecho de que haya que comparar cada una de las muestras con el resto de muestras de la señal eleva considerablemente el tiempo de procesado a medida que aumenta el tamaño de la señal de entrada. Dada una señal de entrada de tamaño N , habría que realizar N^2 comparaciones para obtener los pesos de todas las muestras.

Se han pensado pues, diversas soluciones que permitan reducir el elevado coste computacional. Una de las soluciones se basa en reducir el radio de búsqueda de segmentos similares como se muestra en [26]. Esta última solución ha sido implementada en este trabajo.

4.2.2.1.Reducción radio de búsqueda

Como consecuencia de la implementación de la solución propuesta anteriormente surge un nuevo parámetro que se denominará M y cuyo objetivo fundamental es reducir el radio de búsqueda de patrones similares. Por ejemplo, si se tiene una señal de entrada de tamaño $N = 10000$ *muestras* y se sabe que aparecerán segmentos similares cada aproximadamente 1000 *muestras*, se puede realizar la comparación de cada segmento de la señal con aquellos segmentos que se encuentren en un radio de 1000 *muestras* en torno al segmento actual, con la certeza de que se encontrarán segmentos similares. De esta forma, se reducirá en gran medida la dependencia del coste computacional con la longitud de la señal de entrada. Se muestra a continuación un ejemplo de reducción de radio de búsqueda para un valor de $M = 4$ *muestras*.

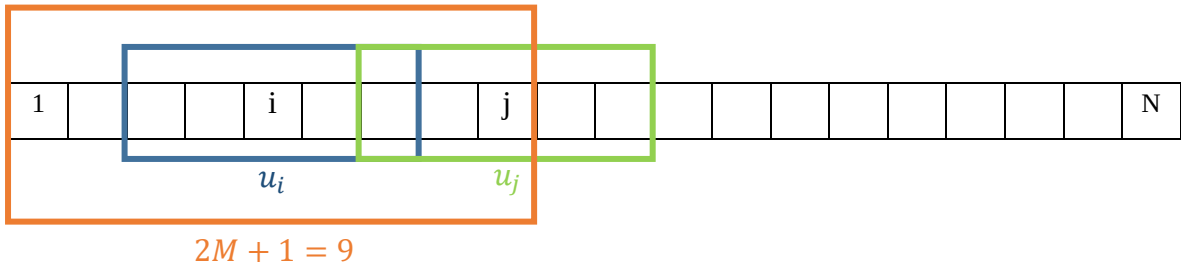


Figura 4.20 - Ejemplo de la reducción del radio de búsqueda para $M=4$ muestras.

Como se puede ver en la figura anterior, para obtener el valor de la muestra i a la salida, únicamente habría que realizar la comparación de esa muestra con las nueve muestras que la rodean (incluyéndose a sí misma), mientras que, en el caso de no haber reducido el radio de búsqueda, se tendrían que realizar veinte comparaciones, que es el número de muestras de la señal de entrada.

Sin embargo, la solución propuesta con anterioridad no es definitiva, puesto que, además de reducir el coste computacional, también se está reduciendo la búsqueda de

segmentos similares a un número cercano de muestras en torno a la muestra que se está procesando, por lo que, el término “no local” puede dejar de tener sentido.

La periodicidad de los sonidos de corazón depende del ritmo cardiaco, cuyos niveles normales están aproximadamente entre uno y dos latidos por segundo, por lo que, si se quiere comparar un segmento con otros segmentos similares se deberá introducir un valor de M , como mínimo, similar a la frecuencia de muestreo.

En este trabajo, por ejemplo, se utilizarán señales muestreadas a 8000 Hz , por tanto, si se escoge un valor de $M = 8000\text{ muestras}$ habría que realiza 16001 comparaciones para obtener el valor de una única muestra, por lo que, se puede deducir, que el algoritmo sigue presentando un enorme coste computacional. En algunas pruebas realizadas, con un valor de $M = 8000\text{ muestras}$ el algoritmo necesitaba, al menos, una hora para procesar una señal de entrada de siete segundos.

4.2.2.2. Ritmo cardiaco

En los pasos anteriores se ha dicho que el algoritmo intenta aprovechar los distintos segmentos del sonido cardiaco que aparecen repetidos en distintos instantes temporales, sin embargo, como se ha visto anteriormente, es necesario acotar el radio de búsqueda para reducir el enorme coste computacional.

La introducción del parámetro M mejora enormemente el coste computacional, pero, sin embargo, surge ahora una difícil decisión entre aumentar el valor de M para encontrar de esta forma una mayor cantidad de fragmentos similares o reducirlo para mejorar el coste computacional. En este trabajo se presenta una nueva solución basada en aprovechar la periodicidad de los sonidos cardiacos para conseguir reducir el coste computacional.

Se parte de un sonido de corazón con un ritmo cardiaco de 60 ppm (pulsaciones por minuto) muestreado a 8000 Hz , por lo que, los distintos sonidos de corazón aparecerán repetidos cada segundo (ocho mil muestras) en la señal de entrada. Si se quisiera comparar cada muestra con, al menos, dos repeticiones del sonido cardiaco utilizando la reducción mediante el parámetro M , el algoritmo seguiría siendo computacionalmente costoso.

Sin embargo, si se sabe que la señal aparece repetida cada ocho mil muestras es evidente que no tiene sentido realizar la comparación con todas las muestras dentro de ese trozo puesto que no habrá ningún segmento similar presente en ellas. Se compara entonces la muestra i con las $2M + 1$ muestras situadas en torno a la muestra $j = i \pm 8000$ o $j =$

$i \pm 16000$ en función del periodo que se esté teniendo en cuenta, asegurando de esta forma, que se encontrarán patrones repetitivos y reduciendo el número de muestras a comparar.

Se introduce, por tanto, un nuevo parámetro que se denominará T . Este parámetro definirá con qué porcentaje de los periodos de la señal se comparará cada una de las muestras de la señal de entrada. Se puede deducir como un valor mayor de T aumentará la no localidad y también elevará el tiempo de computación, necesitando realizar $T \cdot n^{\circ} \text{ pulsos cardiacos de la señal} \cdot (2M + 1)$ comparaciones para obtener el valor de una de las muestras de la señal de salida. Por ejemplo, suponiendo que se tiene una señal de entrada con 10 pulsos cardiacos y se utiliza un $T = 20\%$, cada muestra de la señal de entrada se comparará con la muestra equivalente localizada en dos periodos distintos de la señal.

4.2.3. Sonidos cardiacos y pulmonares

Una vez implementado el algoritmo y realizados los cambios descritos anteriormente, se comprobará su correcto funcionamiento en la separación de sonidos de corazón y de pulmón aplicando los procedimientos explicados con anterioridad.

Se parte de un sonido cardiaco y un sonido pulmonar mezclados con una SNR de 5dB del siguiente modo:

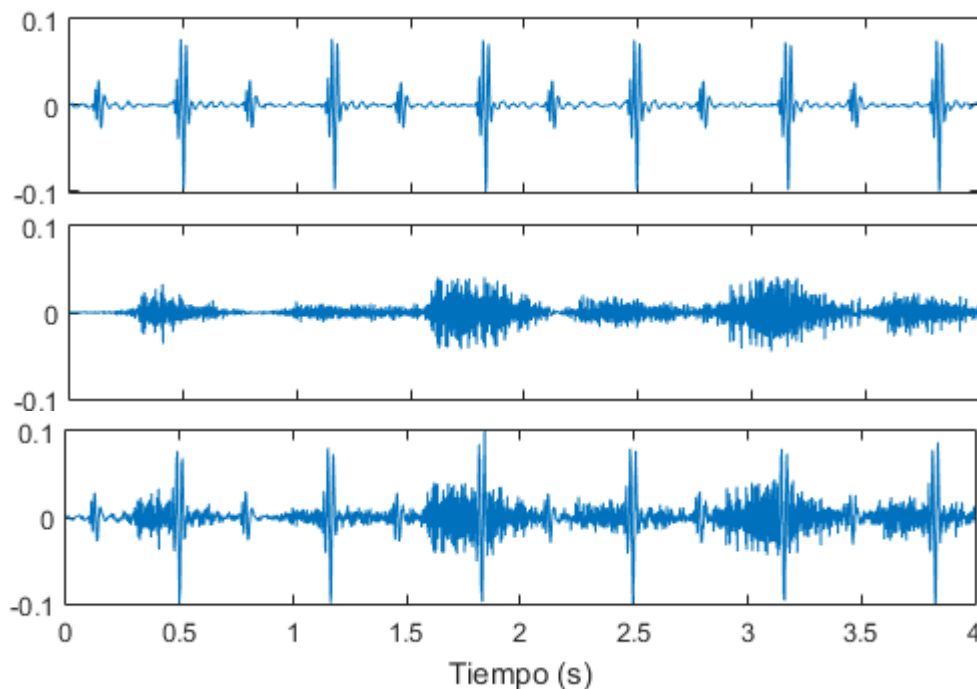


Figura 4.21 - Sonido cardiaco (arriba), sonido pulmonar (medio) y mezcla SNR=5dB (abajo). Estos ficheros de audio se adjuntan en la carpeta de nombre "Software y Ficheros de Audio" anexa a la memoria.

El primer paso se basa en encontrar el periodo de la señal mezcla. Para ello se filtra la señal mezcla entre 20 y 150 Hz con el objetivo de eliminar la mayoría de componentes que no se correspondan con sonidos de corazón y poder, de este modo, localizar el máximo más claramente.

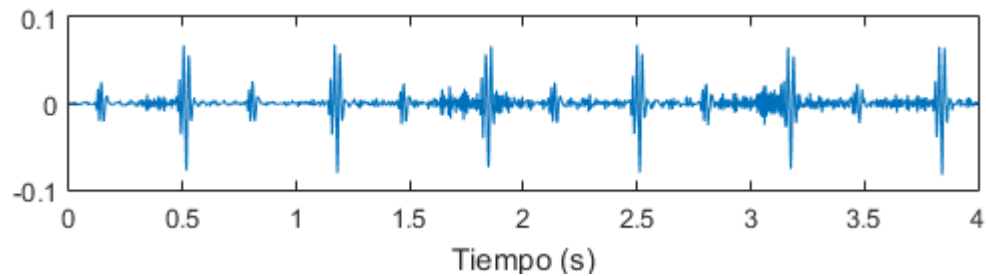


Figura 4.22 - Señal mezcla mostrada en la Figura 4.21 filtrada entre 20 y 150 Hz

Posteriormente se recorta un segmento correspondiente a los dos primeros segundos de la señal de entrada, se elige este valor puesto que se sabe con certeza que aparecerán como mínimo dos periodos del sonido cardiaco. Para localizar la posición del valor máximo se establece un umbral mínimo de $fs/2$ muestras para rechazar los máximos en 0 muestras y alrededores. Este valor es escogido puesto que se sabe que no aparecerán dos periodos de corazón en un intervalo inferior a medio segundo.

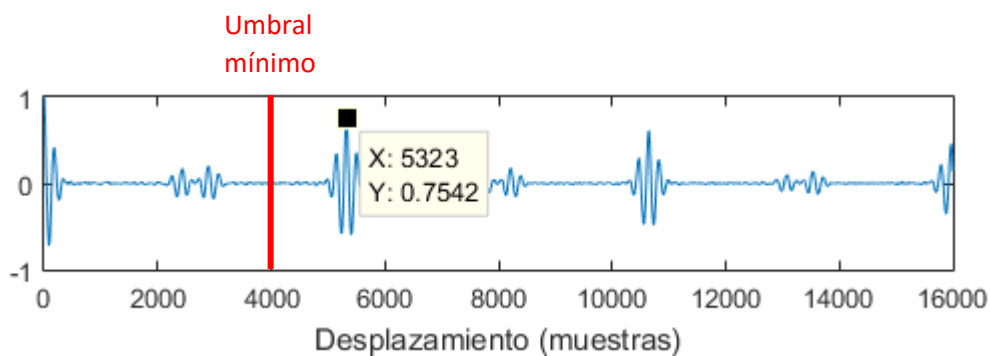


Figura 4.23 - Autocorrelación de la señal mezcla filtrada vista en la Figura 4.22

Se observa a partir de la figura anterior como la función de autocorrelación es máxima para un desplazamiento de 5323 muestras, por lo que, a partir de esta información se sabe que aparecerá un sonido cardiaco tras un número de muestras que estará situado en

torno al valor anterior. Por tanto, si se está procesando un segmento centrado en la muestra i , este se deberá comparar con las $2M + 1$ muestras centradas en torno a $j = i \pm 5323$, $j = i \pm 2 \cdot 5323$, $j = i \pm 3 \cdot 5323 \dots$

A continuación, se muestra la división en trozos de la señal mezcla vista en la **Figura 4.21** para que se pueda ver de una forma más clara los distintos periodos del sonido cardiaco.

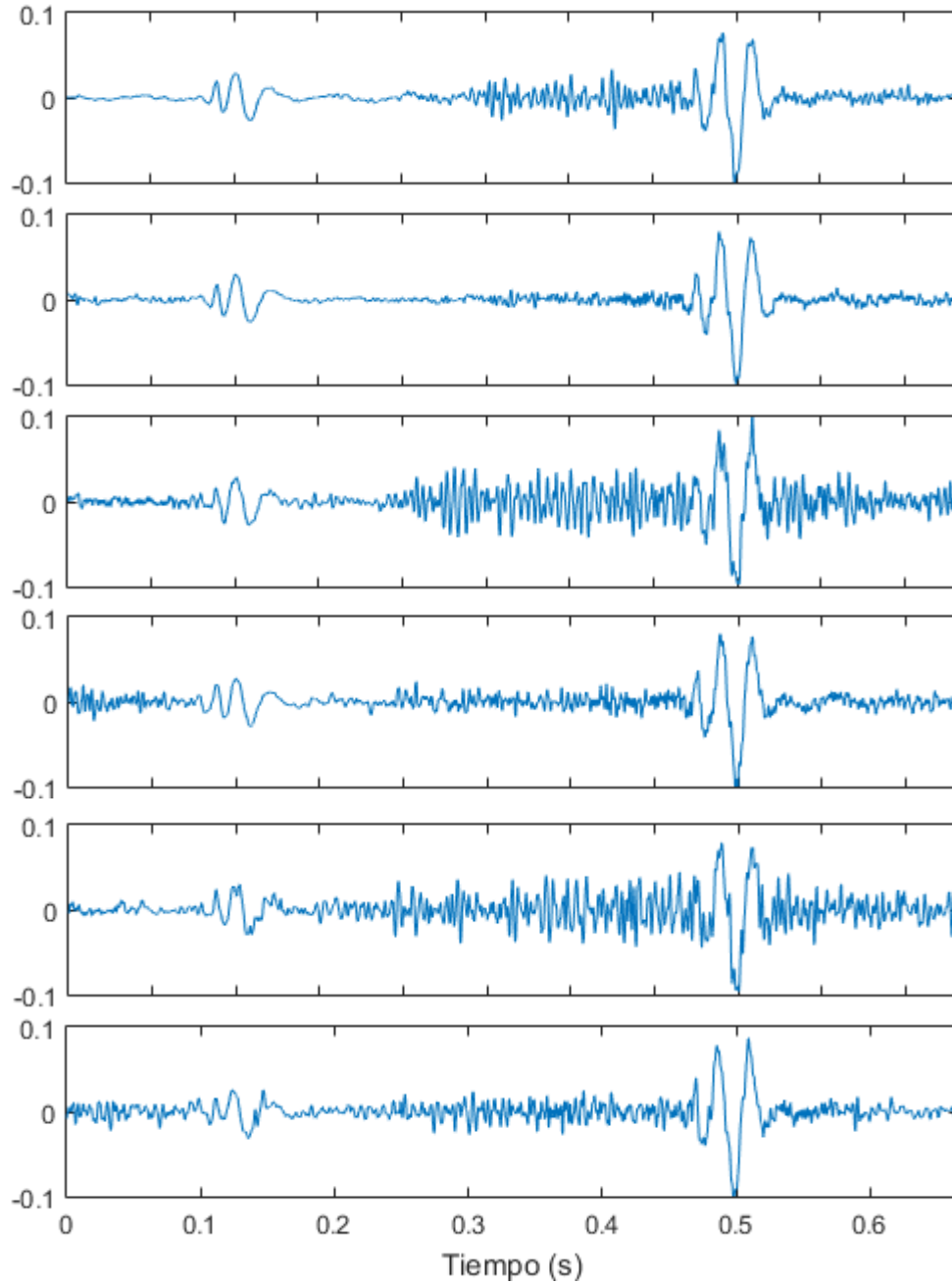


Figura 4.24 - Periodos de la señal mezcla vista en la Figura 4.21 utilizando el ritmo cardiaco

Se puede observar a partir de la figura anterior como los sonidos de corazón son prácticamente iguales en los distintos periodos mientras que no ocurre lo mismo con la señal ruidosa. Bastaría por tanto con hacer una media entre esos segmentos parecidos en los distintos periodos y eliminar de esta forma la señal ruidosa.

Se muestra a continuación la señal obtenida a la salida del filtro junto con la señal ruidosa:

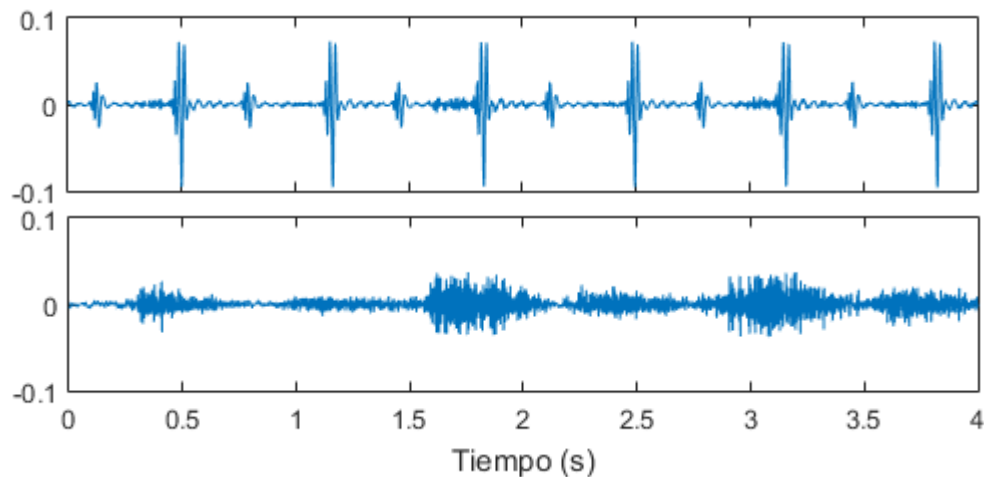


Figura 4.25 - Sonido de corazón recuperado (arriba) y sonido de pulmón recuperado (abajo) utilizando un filtro NLM

Una vez obtenidas las señales anteriores, se aplica, al igual que se hizo en el filtro ALE, una máscara de manera que se logren eliminar aquellos restos de la señal de pulmón presente en la señal de corazón recuperada y viceversa. Una vez aplicada la ecuación **(1.12)** se obtienen las siguientes máscaras:

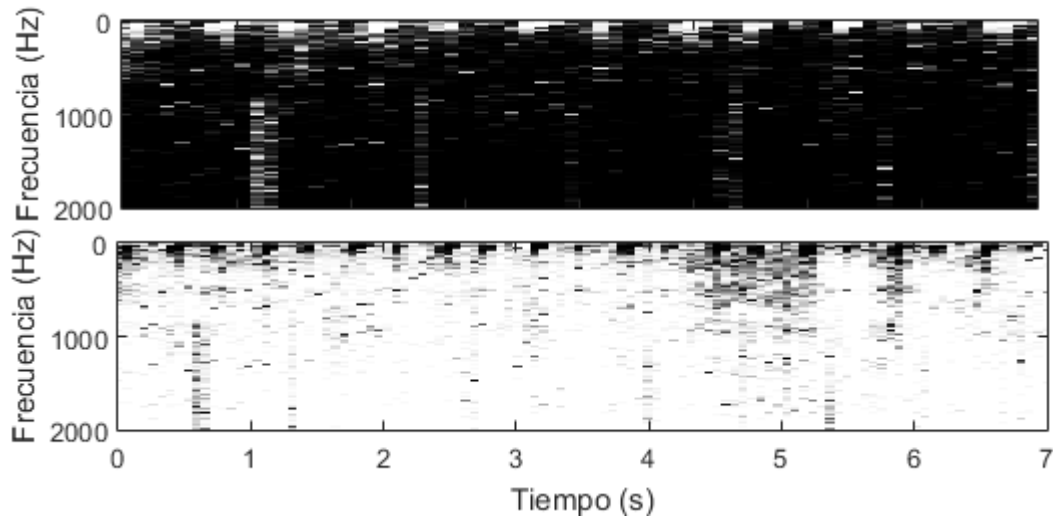


Figura 4.26 - Máscara del sonido de corazón (arriba) y del sonido de pulmón (abajo) obtenidas a partir de las salidas tras aplicar un filtro NLM

A continuación, se muestran las señales obtenidas tras aplicar las máscaras:

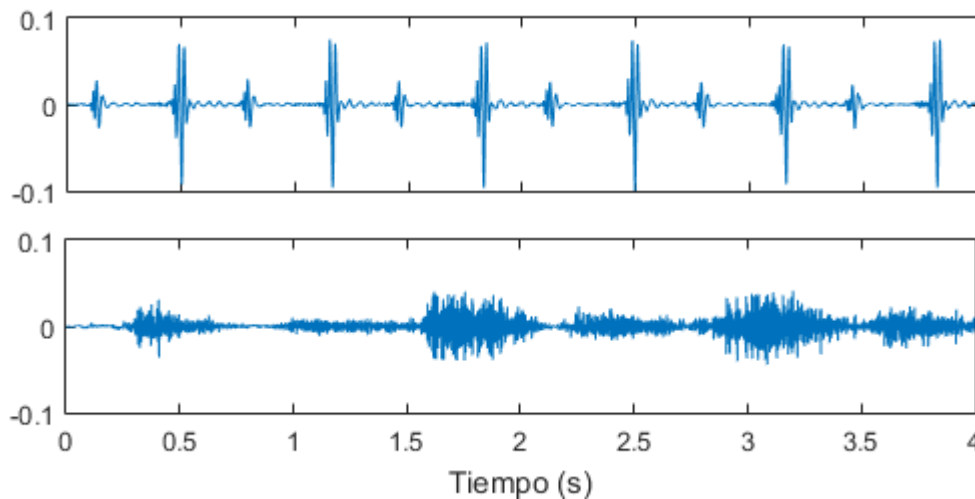


Figura 4.27 - Sonido de corazón recuperado (arriba) y sonido de pulmón recuperado (abajo) tras aplicar una máscara Wiener a la salida del filtro NLM. Estos ficheros de audio se adjuntan en la carpeta de nombre “Software y Ficheros de Audio” anexa a la memoria.

Se puede observar a partir de la figura anterior como se han logrado eliminar los restos de sonido de pulmón presentes en el sonido de corazón y viceversa tras aplicar la máscara. A continuación, se muestra la representación en frecuencia de cada una de las señales obtenidas de manera que se pueda apreciar mejor los resultados obtenidos:

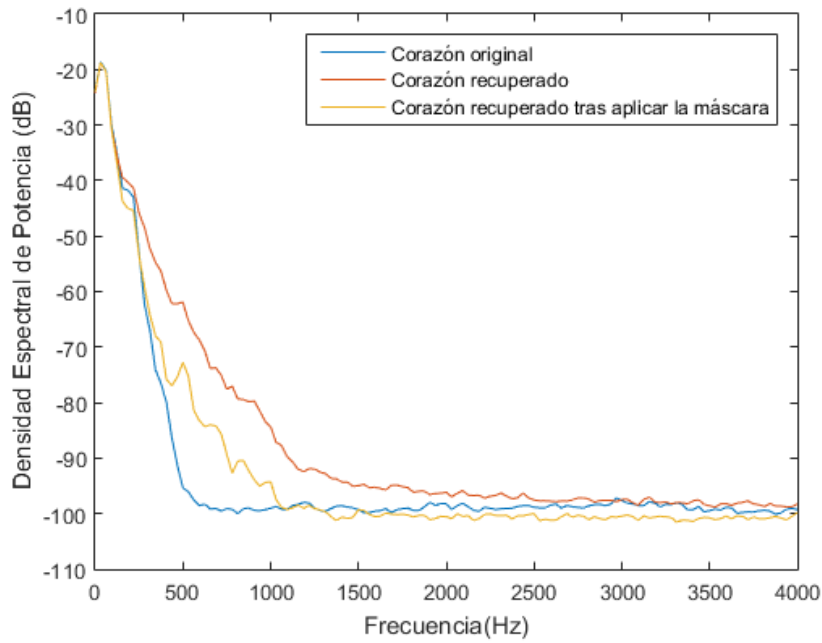


Figura 4.28 - PSD del sonido de corazón original, sonido de corazón recuperado utilizando un filtro NLM y del sonido de corazón recuperado tras aplicar una máscara Wiener a la salida del filtro NLM

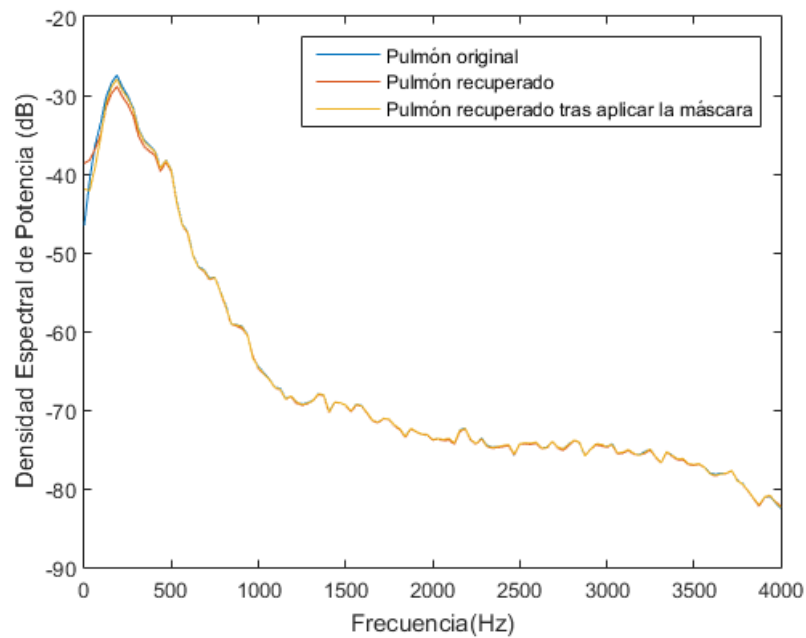


Figura 4.29 - PSD del sonido de corazón original, sonido de corazón recuperado utilizando un filtro NLM y del sonido de corazón recuperado tras aplicar una máscara Wiener a la salida del filtro NLM

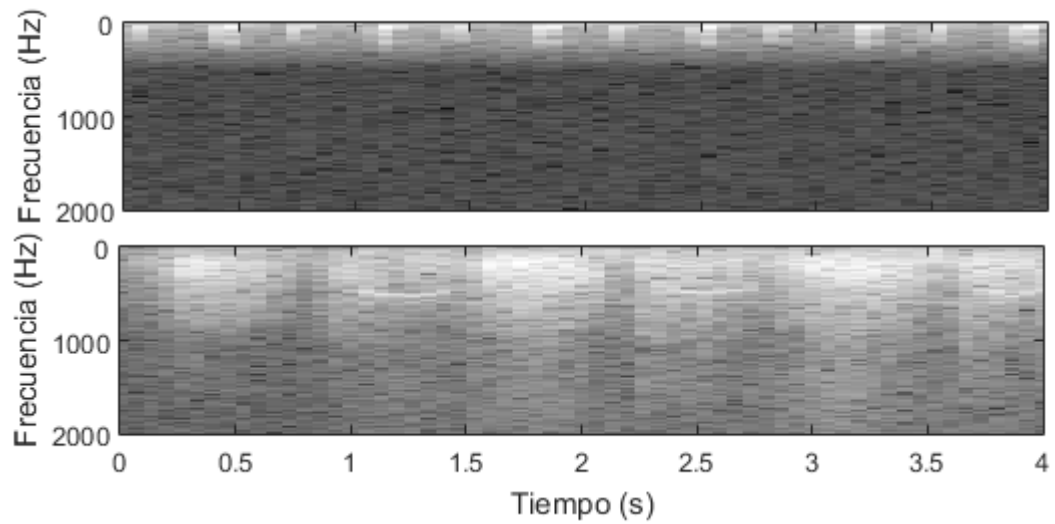


Figura 4.30 - Espectrograma del sonido de corazón original (arriba) y espectrograma del sonido de pulmón original (abajo)

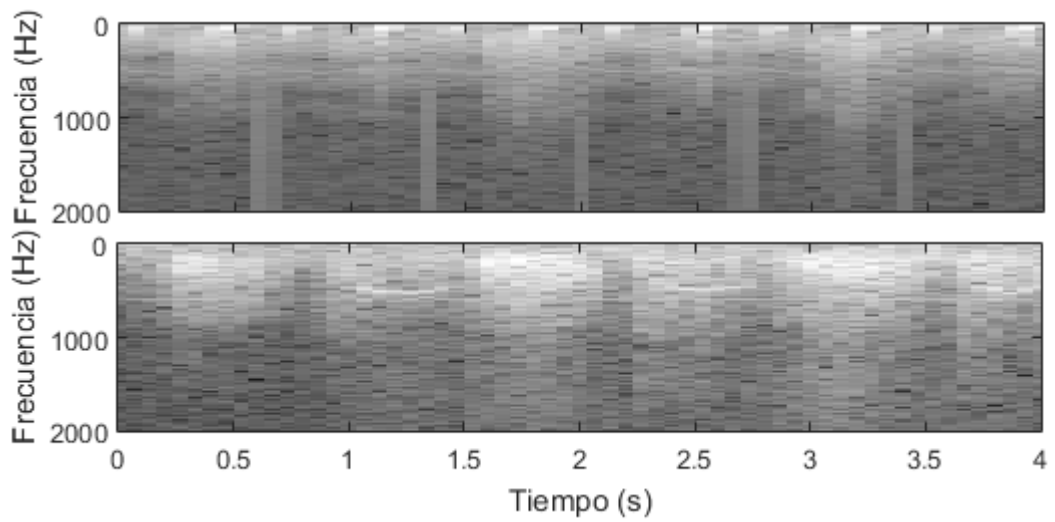


Figura 4.31 - Espectrograma corazón recuperado (arriba) y espectrograma pulmón recuperado (abajo) utilizando NLM

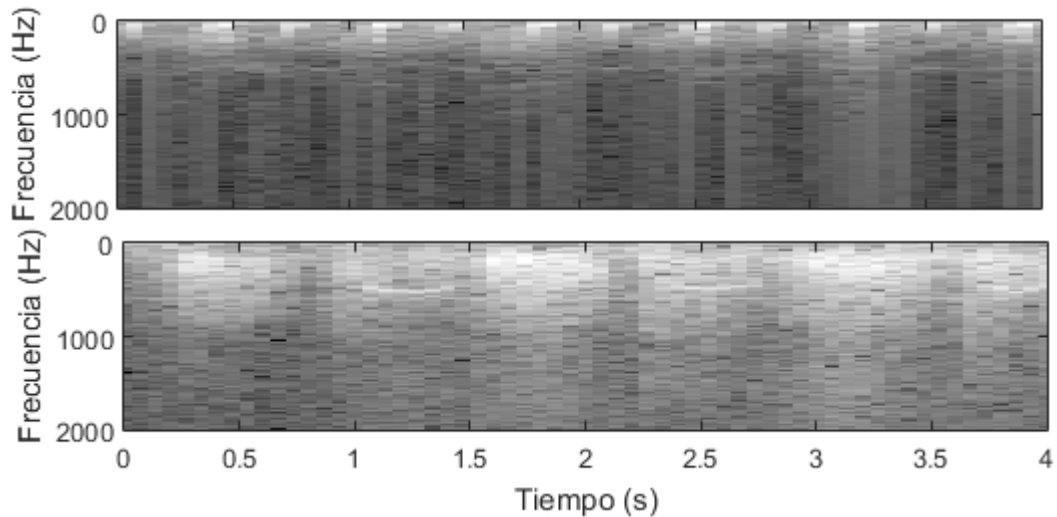


Figura 4.32 - Espectrograma corazón recuperado (arriba) y espectrograma pulmón recuperado (abajo) utilizando NLM tras aplicar la máscara

Como se puede ver a partir de las figuras anteriores, se obtienen unos muy buenos resultados en la utilización del filtro NLM para la separación de sonidos cardiacos y pulmonares.

Hasta ahora, se ha supuesto que los sonidos de corazón son periódicos, es decir, que los distintos pulsos cardiacos se van a repetir con exactitud cada x *muestras* calculadas al inicio como la posición correspondiente al valor máximo de la autocorrelación, sin embargo, la mayor parte de las veces esto no ocurre así. Los sonidos de corazón son cuasiperiodicos, es decir, se repetirán en torno a un valor de muestras, por lo tanto, hay que introducir algunas correcciones al método anterior cómo se desarrolla a continuación.

Se parte de un sonido cardiaco y un sonido pulmonar mezclados con una SNR de 5dB del siguiente modo:

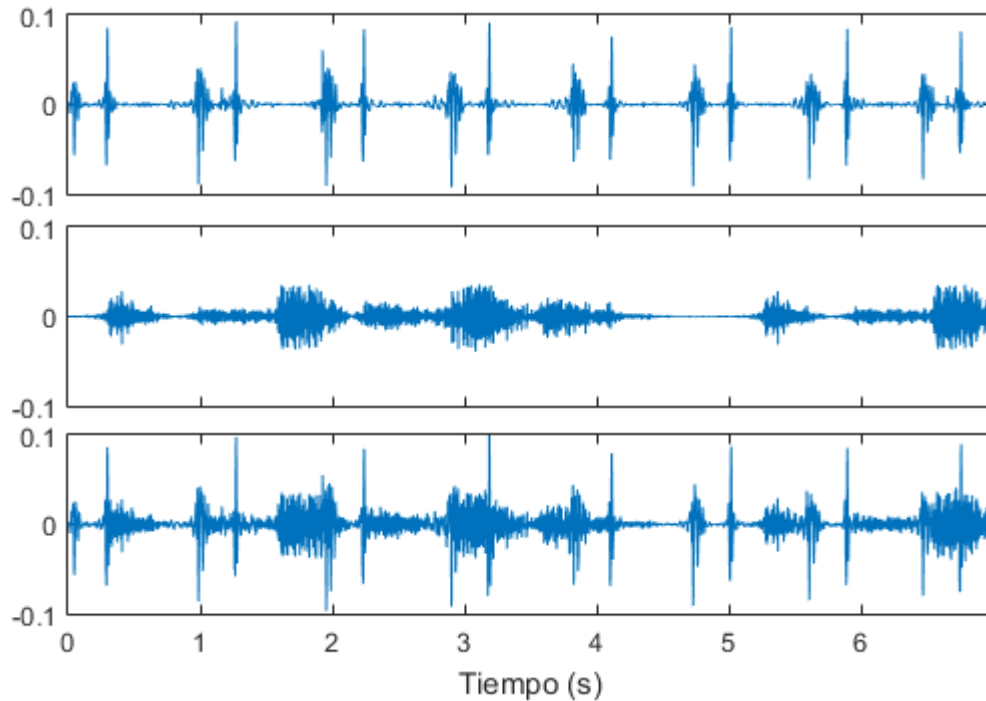


Figura 4.33 - Sonido cardíaco (arriba), sonido pulmonar (medio) y mezcla SNR=5dB (abajo). Estos ficheros de audio se adjuntan en la carpeta de nombre “Software y Ficheros de Audio” anexa a la memoria.

El primer paso es similar al caso anterior, es decir, se filtra la señal de entrada entre 20 y 150 Hz, se recorta el segmento correspondiente a los dos primeros segundos y se localiza la posición máxima de la autocorrelación con un umbral mínimo de $f_s/2$ muestras.

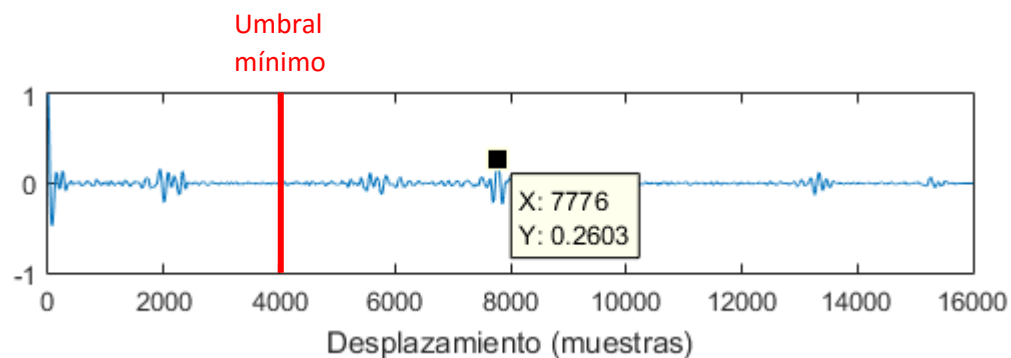


Figura 4.34 - Autocorrelación del sonido cardíaco visto en la Figura 4.33

Una vez detectada la posición máxima de la autocorrelación, se eliminan las primeras 7775 muestras de la señal de entrada, que se corresponderán con el primer “periodo”, quedando el resto de la señal mezcla filtrada entre 20 y 150 Hz. Posteriormente se realiza nuevamente la autocorrelación de la señal mezcla restante y se localiza la posición del

máximo dentro de un rango cuyos valores mínimo y máximo se corresponderán con el 90% y el 110% del valor obtenido en la autocorrelación anterior (6998 y 8554 en este caso). Se han elegido estos valores debido a que tras realizar numerosas pruebas se ha observado que la variación entre pulsos cardiacos consecutivos suele estar dentro de estos márgenes, de esta forma, se evita que se seleccione un máximo que se de en cualquier otra posición errónea.

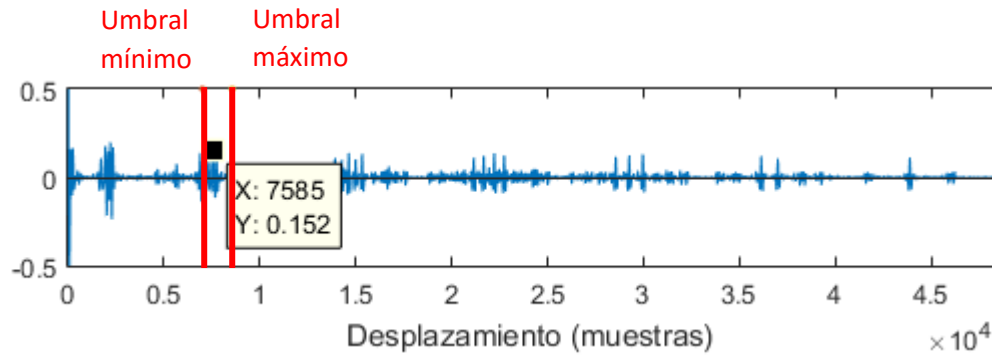


Figura 4.35 - Autocorrelación de la señal temporal vista en la Figura 4.33 tras eliminar el primer segmento

Tras realizar este paso se observa cómo se obtiene un valor de *7585 muestras*, por lo que, se eliminan las primeras *7484 muestras* y se vuelve a calcular la autocorrelación de la señal restante. Este paso se repite hasta que la longitud de la señal mezcla restante sea menor que el valor superior del umbral, lo que querrá decir que se corresponde con el último “periodo”. Tras hacer esto, se habrá dividido la señal de entrada en distintos “periodos” que se corresponderán con los pulsos del corazón, independientemente de las variaciones del ritmo cardiaco. Se deben almacenar las posiciones de inicio de cada “periodo” en un vector que se denominara *pos*. En el ejemplo anterior, *pos(1)=1*, *pos(2)=7776*, *pos(3)=7585+pos(2)=15361...*

En la imagen siguiente se muestra la división en “periodos” de la señal mezcla utilizando el primer método y utilizando el segundo método. Se observa como el primer método desarrollado no es útil para aquellas señales que presentan una variación en el ritmo cardiaco puesto que no se realiza correctamente la división en “periodos” de la señal de entrada. Sin embargo, el segundo método es capaz de adaptarse correctamente a las variaciones, evitando de esta forma, que pueda afectar a la calidad del filtrado.

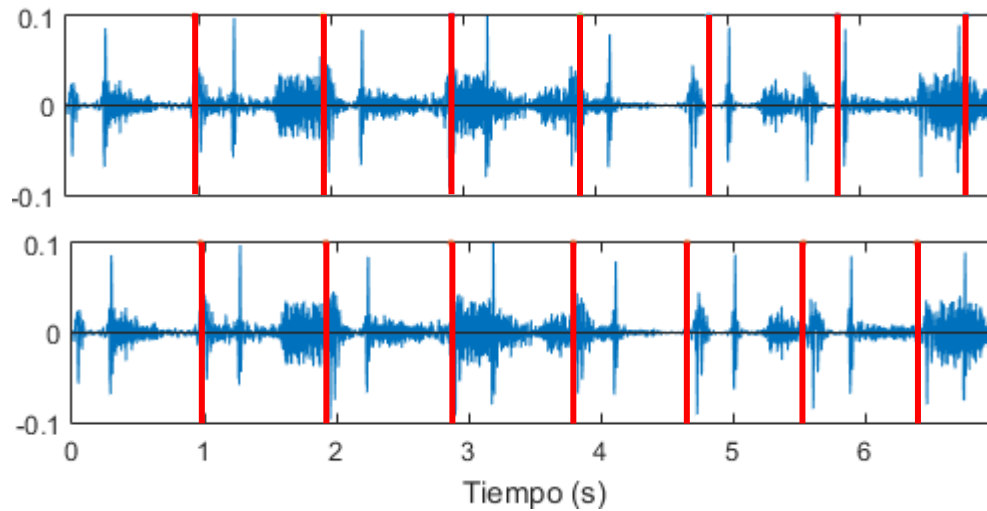


Figura 4.36 – División en periodos de la señal mezcla vista en la Figura 4.33 utilizando el primer método (arriba) y utilizando el segundo método (abajo).

Una vez hecho esto, es necesario asegurar que cada “periodo” estará alineado con el resto de “periodos”. Se deberán calcular los retardos entre cada uno de ellos, los cuales, serán almacenados en una matriz $\mathbf{D}$ cuyo tamaño, en este caso, será de 8×8 (*8 periodos del sonido cardiaco*). Por ejemplo, $\mathbf{D}(2, 5)$ se corresponderá con el retardo existente entre el “periodo” dos y el “periodo” cinco, su valor será positivo si el “periodo” cinco se encuentra retardado respecto al “periodo” dos y negativo en el caso contrario.

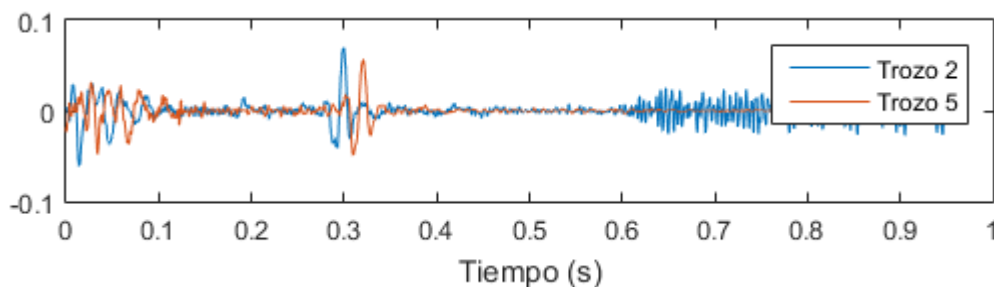


Figura 4.37 - Solapamiento de dos de los “periodos” obtenidos tras realizar la separación del sonido cardiaco visto en la Figura 4.33

Se debe calcular entonces el retardo $\mathbf{D}$ existente entre ambos “periodos”. En este caso, el “periodo” cinco se encontrará retardado respecto al “periodo” dos, por lo que el valor de $\mathbf{D}(2, 5)$ será positivo. Por tanto, suponiendo que se está procesando un segmento centrado en la muestra $i = 7800$ (superior a 7776 e inferior a 15361, por lo que se corresponde con

una muestra del segundo “periodo”) se deberá comparar con las $2M + 1$ muestras centradas en torno a la muestra $j = (7800 - 7776) + pos(5) + D(2, 5)$ de la señal de entrada que se corresponderá con el quinto “periodo”. Por otro lado, un segmento centrado en torno a una muestra i de la señal de entrada correspondiente al quinto “periodo”, se deberá comparar con las $2M + 1$ muestras centradas en torno a la muestra $j = (i - pos(5)) + pos(2) + D(5, 2)$ correspondiente al segundo “periodo”.

A continuación, se muestran los sonidos cardiacos y pulmonares recuperados aplicando la configuración anterior:

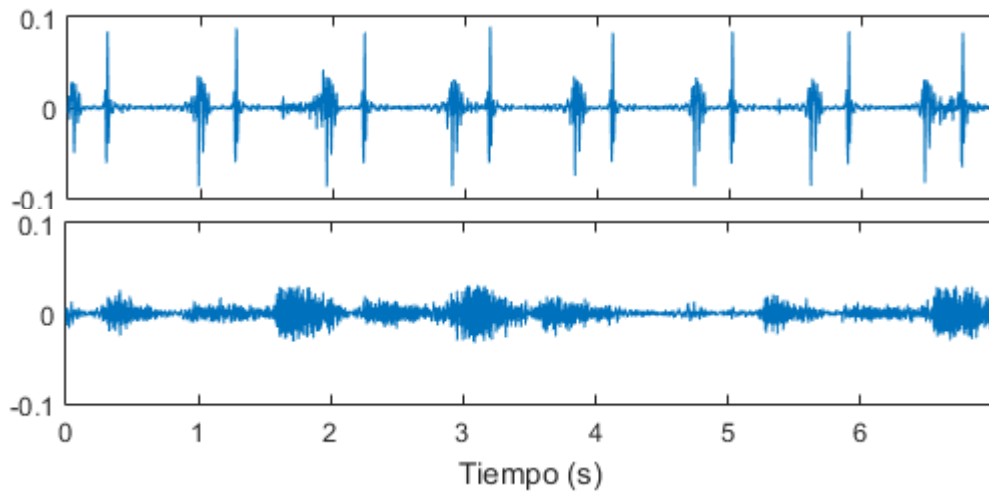


Figura 4.38 - Sonido cardiaco recuperado (arriba) y sonido pulmonar recuperado (abajo) utilizando un filtro NLM

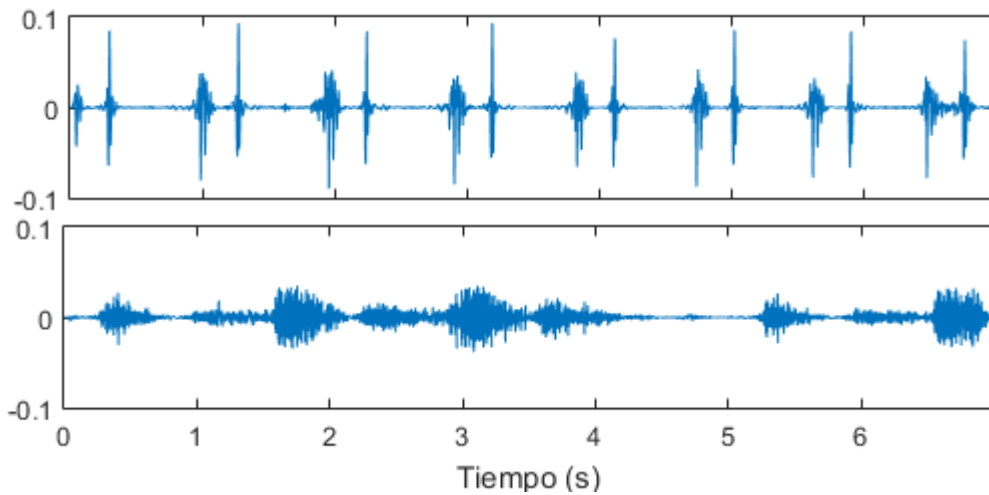


Figura 4.39 - Sonido cardiaco recuperado (arriba) y sonido pulmonar recuperado (abajo) utilizando un filtro NLM tras aplicar la máscara. Estos ficheros de audio se adjuntan en la carpeta de nombre “Software y Ficheros de Audio” anexa a la memoria

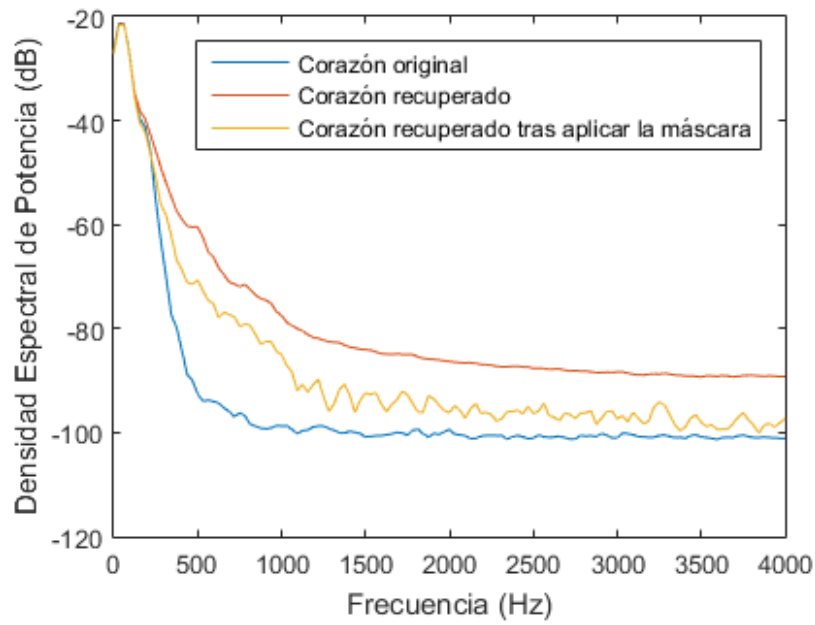


Figura 4.40 - PSD del sonido de corazón original, sonido de corazón recuperado y sonido de corazón recuperado tras aplicar una máscara Wiener a la salida del filtro NLM

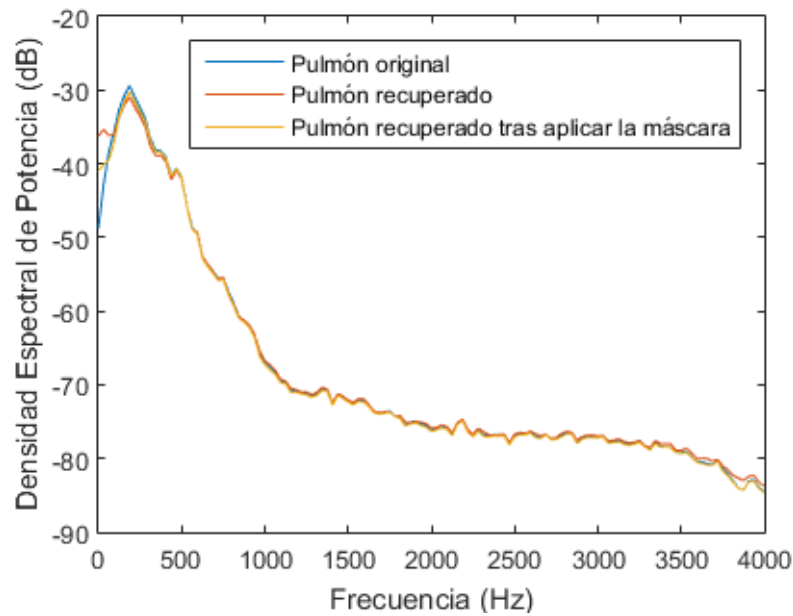


Figura 4.41 - PSD del sonido de pulmón original, sonido de pulmón recuperado y sonido de corazón recuperado tras aplicar una máscara Wiener a la salida del filtro NLM

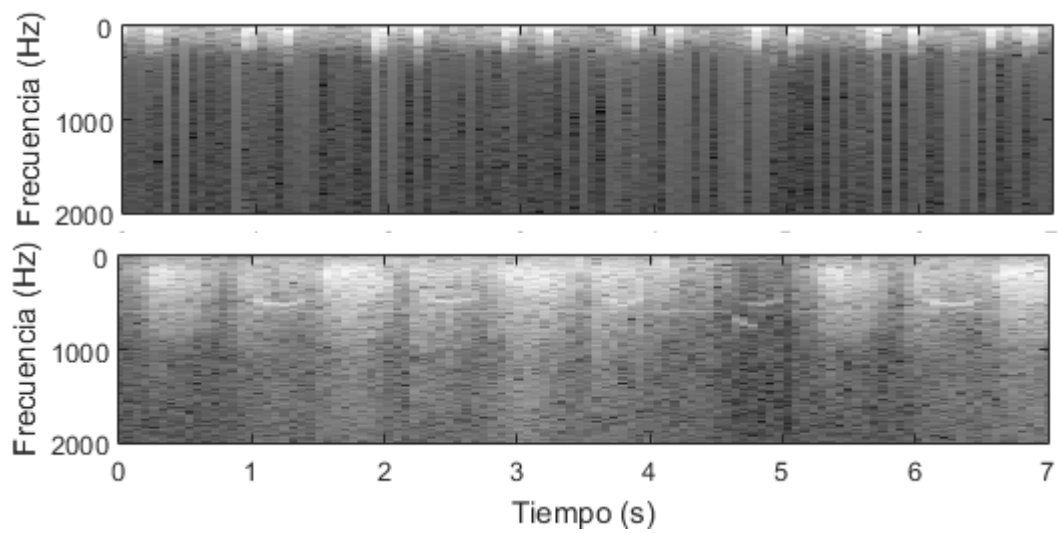


Figura 4.42 - Espectrogramas del sonido de corazón original (arriba) y sonido de pulmón original (abajo)

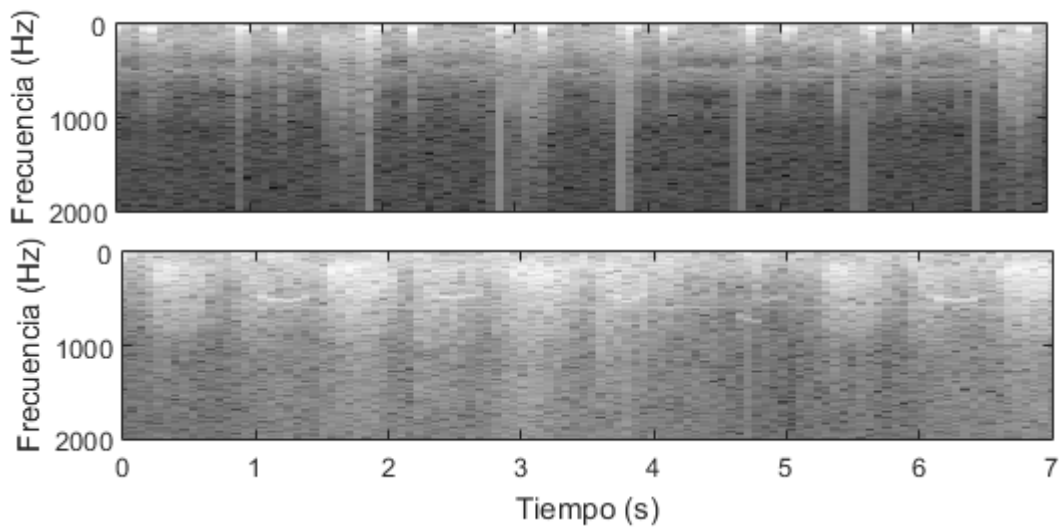


Figura 4.43 - Espectrogramas del sonido de corazón recuperado (arriba) y sonido de pulmón recuperado (abajo) utilizando un filtro NLM

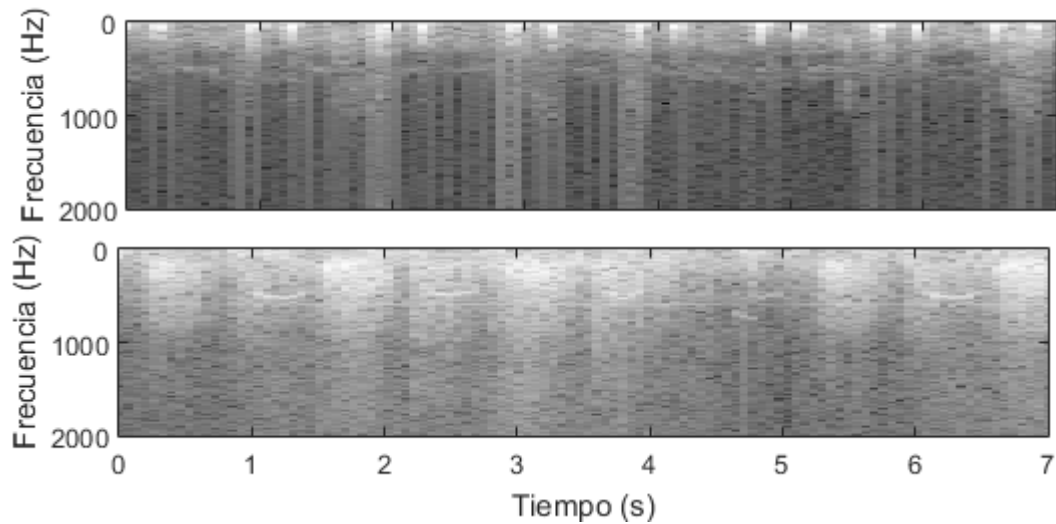


Figura 4.44 - Espectrogramas del sonido de corazón recuperado (arriba) y sonido de pulmón recuperado (abajo) tras aplicar una máscara de Wiener a la salida del filtro NLM

Se puede ver a partir de las figuras anteriores como a pesar de que la señal de entrada no es exactamente periódica esto no influye notablemente en los resultados obtenidos. Una vez se ha comprobado el correcto funcionamiento del filtro NLM, únicamente faltaría realizar la optimización de los parámetros de manera que se obtengan aquellos valores que permitan realizar la separación de los sonidos cardiacos y pulmonares de la mejor manera posible.

Se muestra a continuación, a modo de resumen, una tabla con los parámetros de este método:

Parámetros NLM	
h	Controla el grado de filtrado
P	Tamaño de los segmentos a comparar
M	Radio de búsqueda de segmentos similares
T (%)	Porcentaje de periodos para realizar la búsqueda de segmentos similares

Tabla 4.2 - Resumen de los parámetros utilizados en el filtro NLM

CAPÍTULO 5

5. Evaluación

Una vez desarrollados ambos métodos y demostrado su correcto funcionamiento es necesario realizar una optimización de sus parámetros de manera que se obtengan los valores óptimos que permitan realizar la separación de sonidos de corazón y de pulmón de la mejor manera posible.

5.1. Base de datos

Para realizar la evaluación de ambos métodos es necesario contar con una base de datos de sonidos de corazón y de pulmón, formada tanto por los sonidos cardiacos y pulmonares separados como por la mezcla de ambos con una determinada SNR. La base de datos utilizada es la misma que en [39].

La base de datos cuanta con un total de sesenta y tres directorios en cuyo interior se encuentran los archivos *x_original1.wav*, que se corresponde con el sonido de corazón, y *x_original2.wav*, que se corresponde con el sonido de pulmón, ambos, muestreados a 8000 Hz.

Puesto que se va a utilizar una base de datos con una $SNR=5dB$, es necesario calcular la relación existente entre la potencia de los dos archivos anteriores utilizando la ecuación vista en (1.2) y ajustar la potencia de la señal del pulmón de manera que se cumplan la relación SNR deseada.

Por último, de las sesenta y tres señales que forman parte de la base de datos se seleccionarán dos tercios para la fase de optimización mientras que el tercio restante será utilizado para obtener los resultados de los parámetros obtenidos en la etapa anterior. Debido a que las señales mezcla de la base de datos presentaban distintos niveles de potencia, se comprobó que esto afectaba a la elección de los valores óptimos de cada parámetro en ambos métodos. Como solución, cada método normalizaba la potencia de la

señal de entrada a una potencia de $10^{-4} W$ como paso previo al proceso de filtrado. Se eligió esta potencia de referencia puesto que la potencia media de las señales de la base de datos estaba en torno a este valor. Tras realizar el filtrado, se restauraban las señales obtenidas a su nivel original y se evaluaba el filtrado.

5.2. Métricas

Con el objetivo de seleccionar aquellos parámetros que realicen una mejor separación de los sonidos cardiacos y pulmonares es necesario utilizar unas métricas que den un valor objetivo de la calidad de las señales separadas obtenidas tras realizar el filtrado respecto a las señales originales. En este caso, las métricas utilizadas serán las siguientes [40]:

- **SDR (Source-to-Distortion Ratio):** Es una comparación de los sonidos recuperados con los sonidos originales, es decir, da una medida de la cantidad de sonido original recuperado frente al error total.

$$SDR(dB) = 10\log_{10} \frac{\|s_{target}\|^2}{\|e_{interf} + e_{noise} + e_{artif}\|^2} \quad (5.1)$$

- **SIR (Source-to-Interference Ratio):** Proporciona una medida sobre la presencia de otros sonidos en el sonido recuperado, es decir, compara la cantidad de sonido original recuperado frente al sonido restante de otras fuentes.

$$SIR(dB) = 10\log_{10} \frac{\|s_{target}\|^2}{\|e_{interf}\|^2} \quad (5.2)$$

- **SAR (Source-to-Artifacts Ratio):** Da una medida sobre los artefactos presentes en la fuente separada como consecuencia del proceso de separación.

$$SAR(dB) = 10 \log_{10} \frac{\|s_{target} + e_{interf} + e_{noise}\|^2}{\|e_{artif}\|^2} \quad (5.3)$$

siendo $\hat{s}_j$ la estimación de la fuente original s_j y que se puede descomponer del siguiente modo:

$$\hat{s}_j = s_{target} + e_{interf} + e_{noise} + e_{artif} \quad (5.4)$$

donde s_{target} se corresponde con una versión de s_j modificada por una distorsión permitida y e_{interf} , e_{noise} y e_{artif} se corresponden con las componentes de interferencia, ruido y artefactos respectivamente.

5.3. Optimización

Para llevar a cabo la optimización de ambos métodos se hará uso de un análisis hiperparamétrico, que consiste, básicamente, en variar uno de los parámetros del algoritmo mientras el resto se mantienen fijos. El intervalo de valores iniciales es escogido tras haber realizado numerosas pruebas previamente, asegurando de este modo, que el valor óptimo del parámetro se encontrará en ese rango. Una vez hecho esto, se seleccionarán aquellos parámetros que den un mayor valor medio de SDR correspondiente al sonido de corazón. Se utilizará la base de datos descrita anteriormente.

5.3.1. Optimización ALE VSS-LMS y ALE LMS

En este apartado se llevará a cabo la optimización de los parámetros vistos en la **Tabla 4.1** correspondientes al método de separación de sonidos cardiacos y pulmonares basado en un filtro ALE VSS-LMS. Se mostrarán, señalados en rojo, los valores óptimos obtenidos para cada uno de los parámetros.

Se comenzará optimizando, en primer lugar, el parámetro Δ . Su valor variará en un intervalo establecido entre 1 y 100 mientras que el resto de parámetros permanecerán fijos.

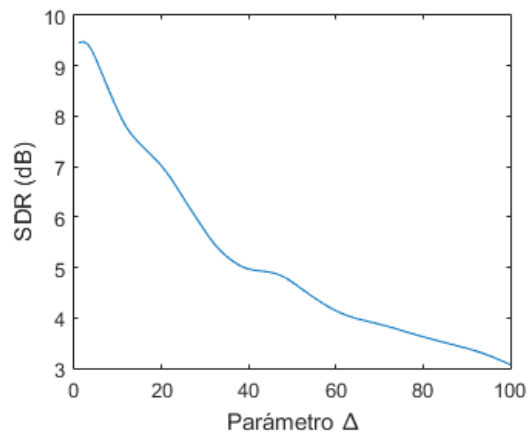


Figura 5.1 - Optimización del parámetro Δ en el método ALE VSS-LMS

ALE VSS-LMS					
Optimización parámetro Δ					
Δ	L	α	γ	μ_{min}	μ_{max}
2	100	0.99	10^1	10^{-4}	10^1

Tabla 5.1 - Valor óptimo del parámetro Δ en el método ALE VSS-LMS

Tras realizar la optimización, se obtiene un valor del parámetro óptimo de $\Delta = 2$. Posteriormente, se fija el valor obtenido y se procede a realizar la optimización del parámetro L cuyos valores variaran en el rango de 1 a 10000. Se muestran los resultados a continuación:

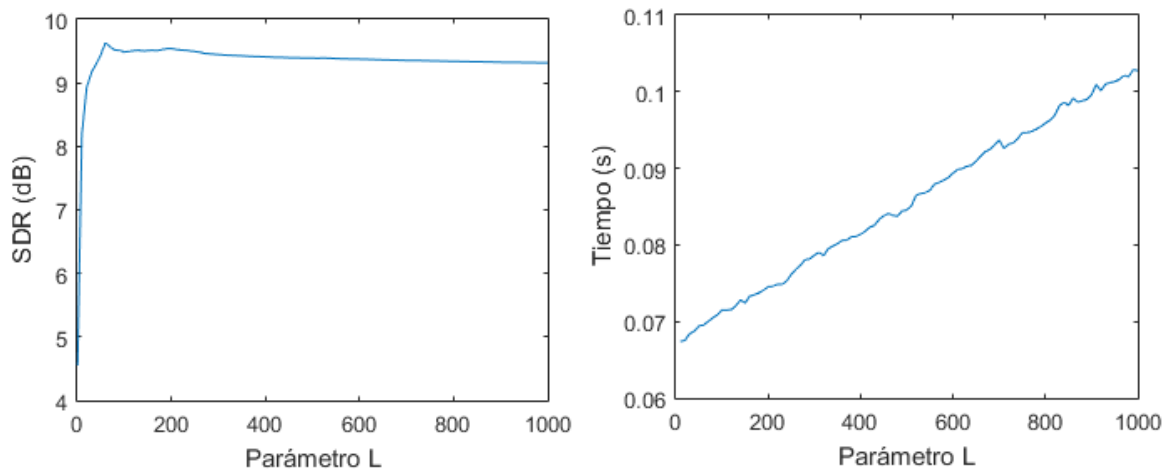


Figura 5.2 - Optimización del parámetro L en el método ALE VSS-LMS (en la derecha, el eje de ordenadas se corresponde con el tiempo que tarda en procesar un segundo de la señal de entrada)

ALE VSS-LMS					
Optimización parámetro L					
Δ	L	α	γ	μ_{min}	μ_{max}
2	60	0.99	10^1	10^{-4}	10^1

Tabla 5.2 - Valor óptimo del parámetro L en el método ALE VSS-LMS

Una vez realizada la optimización, se obtiene un valor óptimo de $L = 60$. Como se ha visto en la **Figura 5.2**, a medida que se incrementa el valor de L se incrementa ligeramente el coste computacional. Se realiza entonces la optimización del parámetro α que variará dentro del rango de 0.01 a 0.99 . A continuación, se muestran los resultados:

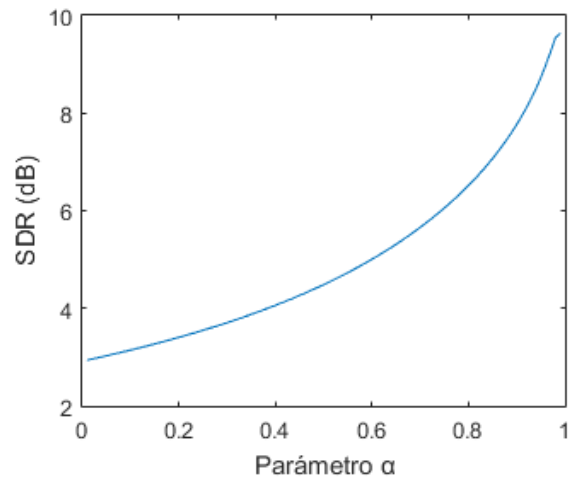


Figura 5.3 - Optimización del parámetro α en el método ALE VSS-LMS

ALE VSS-LMS					
Optimización parámetro α					
Δ	L	α	γ	μ_{min}	μ_{max}
2	60	0.99	10^1	10^{-4}	10^1

Tabla 5.3 - Valor óptimo del parámetro α en el método ALE VSS-LMS

Tras realizar la optimización, se obtiene un valor óptimo de $\alpha = 0.99$, el cual se fija para llevar a cabo, a continuación, la optimización del parámetro γ , cuyo valor variará dentro del intervalo de 10^{-1} a 10^2 . Se muestran los resultados a continuación:

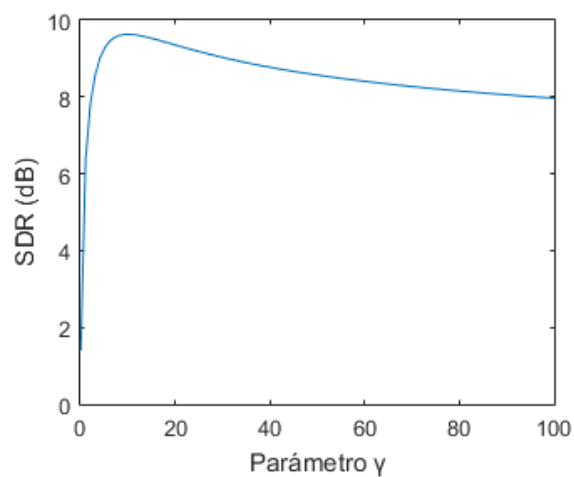


Figura 5.4 - Optimización del parámetro γ en el método ALE VSS-LMS

ALE VSS-LMS					
Optimización parámetro γ					
Δ	L	α	γ	μ_{min}	μ_{max}
2	60	0.99	10^1	10^{-4}	10^1

Tabla 5.4 - Valor óptimo del parámetro γ en el método ALE VSS-LMS

Se puede ver a partir de las figuras anteriores, como se obtiene un valor óptimo de $\gamma = 10^1$. Se fija su valor y se realiza la optimización del parámetro μ_{min} , cuyos valores variaran dentro del rango de 10^{-5} a 1 . Se muestran a continuación los resultados:

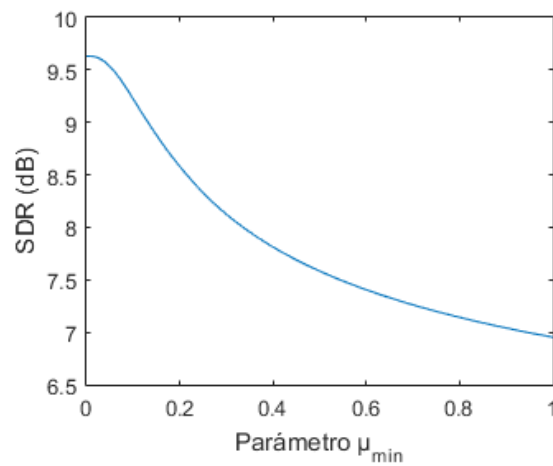


Figura 5.5 - Optimización del parámetro μ_{min} en el método ALE VSS-LMS

ALE VSS-LMS					
Optimización parámetro μ_{min}					
Δ	L	α	γ	μ_{min}	μ_{max}
2	60	0.99	10^1	10^{-5}	10^1

Tabla 5.5 - Valor óptimo del parámetro μ_{min} en el método ALE VSS-LMS

Una vez realizada la optimización, se obtiene un valor óptimo de $\mu_{min} = 10^{-5}$. A continuación, se fija el valor obtenido y se realiza la optimización del parámetro μ_{max} dentro del rango de 10^{-4} y 10^1 :

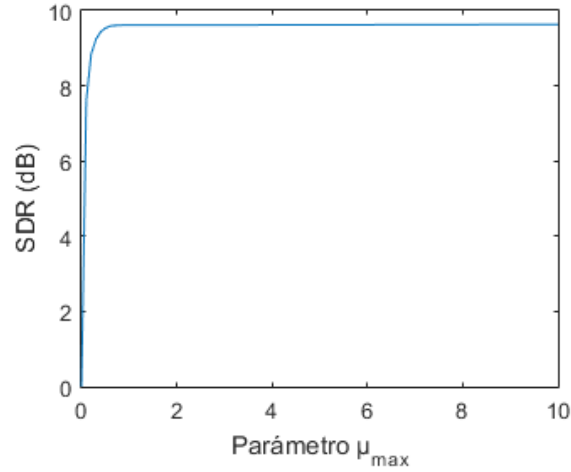


Figura 5.6 - Optimización del parámetro μ_{max} en el método ALE VSS-LMS

ALE VSS-LMS					
Optimización parámetro μ_{max}					
Δ	L	α	γ	μ_{min}	μ_{max}
2	60	0.99	10^1	10^{-5}	1

Tabla 5.6 - Valor óptimo del parámetro μ_{max} en el método ALE VSS-LMS

Como se puede ver en la **Figura 5.6**, hay un valor de μ_{max} a partir del cual el valor medio de SDR se mantiene constante. Esto es debido a que el algoritmo no llega a alcanzar nunca ese valor máximo, sin embargo, se ha escogido como valor óptimo $\mu_{max} = 1$ para evitar que el filtro se vuelva inestable.

Siguiendo todo el procedimiento anterior, se ha llevado a cabo también la optimización del esquema ALE LMS con el objetivo de obtener los valores óptimos de cada uno de los parámetros que se vieron en **Tabla 4.1**. En apartado de resultados, se compararán

los resultados de los dos métodos para demostrar la mejora del esquema ALE VSS-LMS respecto al ALE-LMS.

ALE LMS		
Valores óptimos parámetros		
Δ	L	μ
1	200	10^{-1}

Tabla 5.7 - Valores óptimos de los parámetros en el método ALE LMS

Una vez realizada la optimización de todos los parámetros de los métodos ALE LMS y ALE VSS-LMS, únicamente queda comprobar los resultados que se obtienen utilizando los parámetros obtenidos para cada uno de los métodos.

5.3.2. Optimización NLM

En este apartado, al igual que se ha hecho en el punto anterior, se llevará a cabo la optimización de los parámetros vistos en la **Tabla 4.2** correspondientes al método NLM. En este método, como ya se explicó en el capítulo de desarrollo, el coste computacional juega un papel muy importante. Se debe ser cuidadoso al variar el valor de un determinado parámetro pues, como se puede ver en las figuras que serán mostradas a continuación, una ligera variación de alguno de estos parámetros puede aumentar en gran medida el tiempo de cálculo necesario. Como consecuencia, se ha debido limitar el rango máximo de estos parámetros a un valor en el que aún resultase computacionalmente eficiente.

Se comenzará optimizando, en primer lugar, el parámetro h . Su valor variará en un intervalo establecido entre 0.001 y 0.01 mientras que el resto de parámetros permanecerán fijos. Se muestran a continuación los resultados obtenidos:

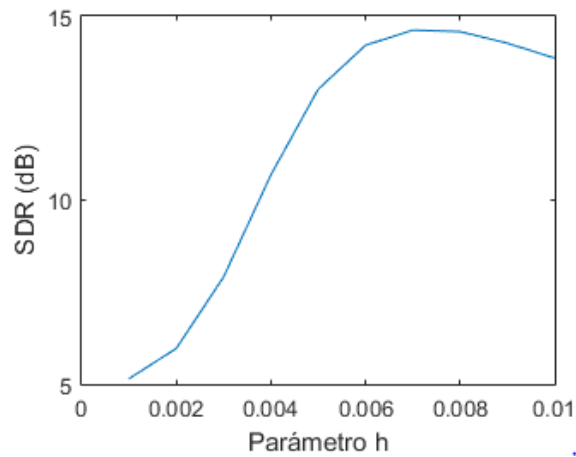


Figura 5.7 - Optimización del parámetro h en el método NLM

NLM			
Optimización parámetro h			
h	P	$T(\%)$	M
0.007	10	100	10

Tabla 5.8 - Valor óptimo del parámetro h en el método NLM

Tras realizar la optimización, se obtiene un valor óptimo de $h = 0.007$, el cual se fija para llevar a cabo, a continuación, la optimización del parámetro P , cuyo valor variará dentro del intervalo de 1 a 100. Se muestran los resultados a continuación:

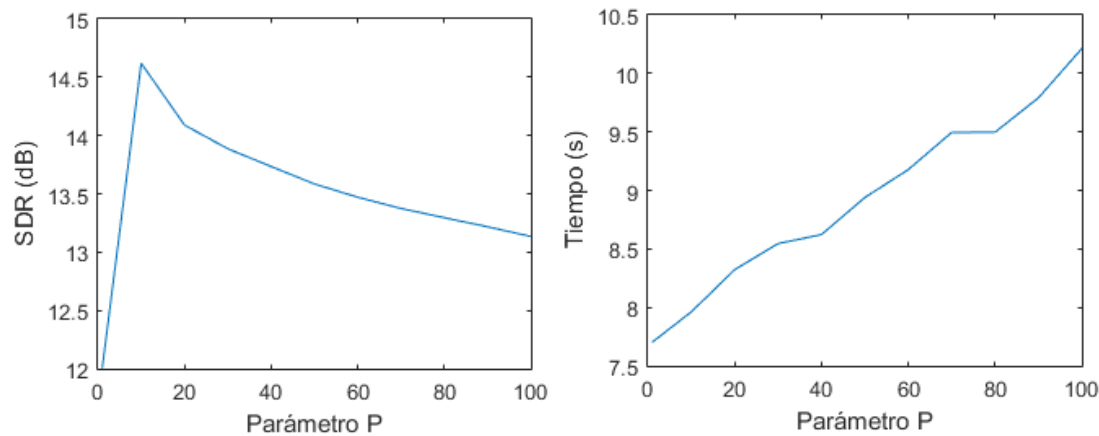


Figura 5.8 - Optimización del parámetro P en el método NLM (en la derecha, el eje de ordenadas se corresponde con el tiempo que tarda en procesar un segundo de la señal de entrada)

NLM			
Optimización parámetro P			
h	P	$T(\%)$	M
0.007	10	100	10

Tabla 5.9 - Valor óptimo del parámetro P en el método NLM

Como se puede observar en las figuras anteriores, incrementar el valor P aumenta considerablemente el tiempo de cálculo. Además, una vez realizada la optimización, se obtiene un valor óptimo de $P = 10$. Posteriormente se fija el valor obtenido y se realiza la optimización del parámetro T , en el cual, los valores varían dentro del rango de 1% a 100% como se muestra a continuación:

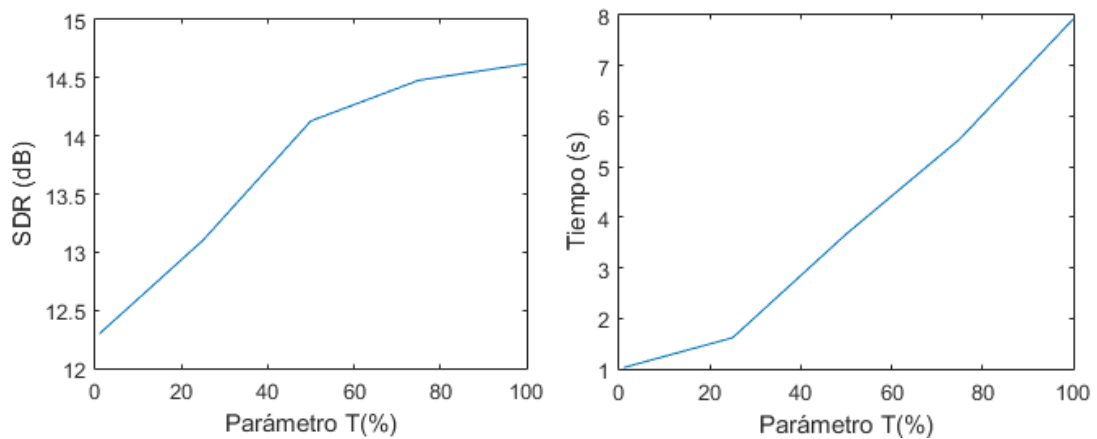


Figura 5.9 - Optimización del parámetro T en el método NLM (en la derecha, el eje de ordenadas se corresponde con el tiempo que tarda en procesar un segundo de la señal de entrada)

NLM			
Optimización parámetro T			
h	P	$T(\%)$	M
0.007	10	100	10

Tabla 5.9 - Valor óptimo del parámetro P en el método NLM

Como se puede apreciar, a medida que se aumenta el valor del parámetro T , es decir, se tengan en cuenta más periodos de la señal de entrada, los resultados de SDR aumentaran, pero, sin embargo, también ocurrirá lo mismo con el coste computacional debido a que se incrementará considerablemente el número de muestras a comparar. Se obtiene, por tanto, un valor óptimo de $T = 100\%$ el cual se fija para continuar realizando la optimización del parámetro M .

Una vez realizadas algunas pruebas, se determinó que aumentar el valor de M incrementaba ligeramente los resultados obtenidos mientras que el coste computacional crecía enormemente. Debido a que el algoritmo ya ofrecía unos buenos resultados, se decidió dejar el valor de $M = 10$ como estaba fijado por defecto.

5.4. Resultados

Una vez realizada la optimización y obtenidos los valores de cada uno de los parámetros, se procede a continuación a realizar el testeo utilizando el tercio de señales restantes de la base de datos detallada al inicio del capítulo. Se obtendrá, para cada una de las veintiuna señales tres valores correspondientes a la SDR, SIR y SAR del sonido cardiaco a los cuales se le realiza la media. A continuación, se muestran los resultados obtenidos:

Resultados finales ALE LMS, ALE VSS-LMS y NLM				
		SDR (dB)	SIR (dB)	SAR (dB)
ALE	LMS	7,81	11,6	10,31
	VSS-LMS	8,95	13,05	12,77
NLM		12,98	21,31	14,21

Tabla 5.10 - Resultados finales de los métodos ALE LMS, ALE VSS-LMS y NLM (SDR, SIR y SAR)

Se realizará también, la comparación de los métodos para la separación de sonidos cardiacos y pulmonares desarrollados en este trabajo con el método NMF desarrollado en [39]. A la hora de realizar la comparación entre dos métodos, lo principal es comenzar comparando el valor de SDR, esta medida será la que aporte el valor más fiable respecto a la calidad de separación. El hecho de que un método presente una SIR igual o más elevada que otro no necesariamente indica que sea mejor, esto puede significar que ese método ha conseguido eliminar más restos del sonido pulmonar que el otro método, pero quizá también ha perdido una mayor cantidad de sonido cardiaco y por tanto tener un valor de SDR más bajo. Sin embargo, si dos métodos presentan aproximadamente el mismo nivel de SDR y uno de ellos tiene un valor de SIR mucho más elevado que el otro esto si indica que se realiza una mejor separación de ambos sonidos. Se muestra a continuación la comparación:

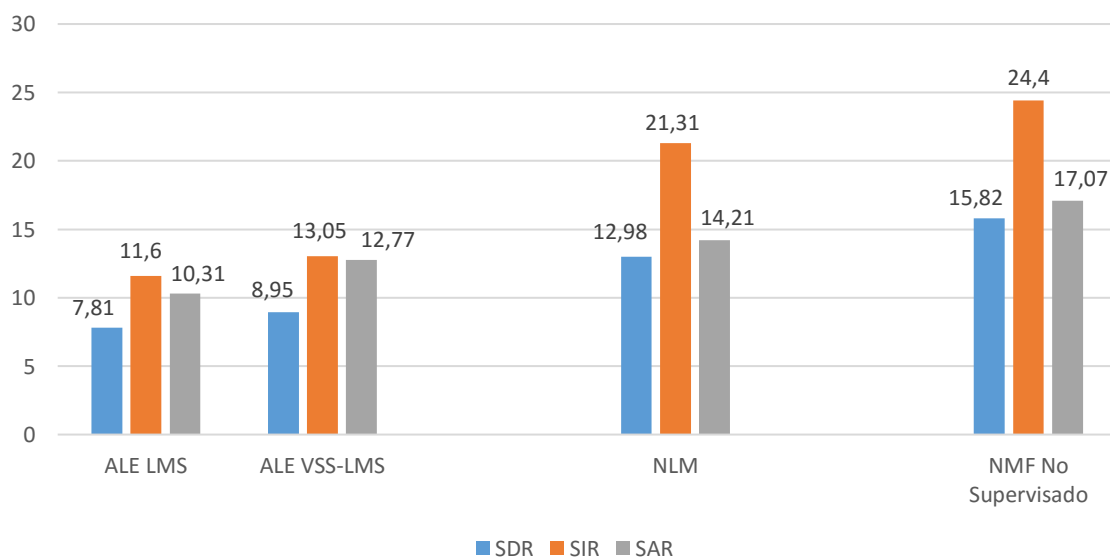


Figura 5.10 - Resultados finales de los métodos ALE LMS, ALE VSS-LMS, NLM y comparación con NMF No Supervisado

Tras analizar y comparar los resultados obtenidos en base a los valores de SDR, SIR y SAR, se puede determinar que el método ALE VSS-LMS, a pesar de mejorar los resultados del esquema ALE LMS ofrece unos resultados más bajos en cuanto a la calidad de la separación de sonidos cardiacos y pulmonares si se le compara con otros métodos como NLM y NMF.

Como se pudo ver en las pruebas realizadas con sonidos cardiacos y pulmonares en el capítulo correspondiente a la implementación, el filtro ALE VSS-LMS realizaba una buena separación, pero, sin embargo, dejaba unos pequeños restos de la señal de pulmón en la señal de corazón recuperada y viceversa (ver **Figura 4.13**). Se podría pensar que la solución para reducir esos restos de señal de pulmón en la señal de corazón, consiste únicamente en aumentar el valor del retardo, y efectivamente ocurre así, pero, sin embargo, a medida que se incrementa este valor de retardo también va resultando más complicado recuperar la señal de corazón (se va recortando), lo que provoca que se pierda una mayor información de estos sonidos, reduciendo de igual manera la calidad de la separación. Todo esto lleva a que sea realmente difícil recuperar el sonido cardiaco sin restos del sonido pulmonar y viceversa utilizando un esquema ALE VSS-LMS.

Por otro lado, en el algoritmo NLM, se tuvo en cuenta las variaciones en el ritmo cardiaco, lo que provocó que fuese necesario alinear los distintos periodos de la señal de entrada, pero, sin embargo, no se ha tenido en cuenta que los sonidos cardiacos que fueron explicados en la introducción (S1, S2, S3, S4, ...) y que forman parte de los distintos periodos, también presentan variaciones en el tiempo que los separa, lo que provoca que no se encuentren completamente alineados con sus equivalentes en otros periodos y, por tanto, que se esté reduciendo la calidad del filtrado. No obstante, a pesar de esto, el filtro funciona realmente bien en la separación de sonidos cardiacos y pulmonares, siendo capaz de realizar la separación de ambos sonidos sin perder prácticamente información de ninguno de ellos, permitiéndole ofrecer unos resultados muy elevados en comparación con el método ALE-VSS-LMS y bastante similares a los obtenidos con el método NMF.

Una vez mostrados los resultados obtenidos en función de las métricas objetivas de calidad, y tras ser comparados con otro método para la separación de sonidos cardiacos y pulmonares basado en NMF, se muestra a continuación los resultados en función del coste computacional de los métodos desarrollados en este trabajo para los parámetros óptimos obtenidos:

Resultados finales ALE LMS, ALE VSS-LMS y NLM		
		Tiempo (s)
ALE	LMS	0,07
	VSS-LMS	0,07
NLM		9,72

Tabla 5.11 - Resultados finales de los métodos ALE VSS-LMS y NLM (tiempo que se tarda en procesar un segundo de la señal de entrada)

A partir de los datos mostrados en la figura anterior, se puede observar cómo se ha mejorado considerablemente el coste computacional del método NLM. Trabajando con señales de siete segundos de duración, se ha pasado de necesitar una hora para filtrar la señal de entrada reduciendo el radio de búsqueda a los dos pulsos cardiacos más próximos, a necesitar prácticamente un minuto en filtrar la señal completa y teniendo en cuenta todos los pulsos cardiacos. Sin embargo, aun así, el coste computacional del método NLM es muy superior al del esquema ALE VSS-LMS.

CAPÍTULO 6

6. Conclusiones

En este trabajo, se han implementado dos de la multitud de métodos utilizados para la separación de sonidos cardiacos y pulmonares. Estos métodos están basados, principalmente, en filtrado adaptativo y media no local. Una vez han sido implementados y evaluados ambos métodos, se pueden sacar las siguientes conclusiones.

- La aplicación de una máscara Wiener a la salida de ambos métodos ha hecho posible la eliminación de algunos restos de sonidos de pulmón presentes en el sonido de corazón y viceversa que el propio método no era capaz de eliminar. Esto ha permitido mejorar en gran medida la calidad de la separación de ambos sonidos.
- El método ALE VSS-LMS ha conseguido, con una ligera modificación en el algoritmo haciendo el tamaño de paso variable, mejorar ligeramente los resultados en cuanto a la calidad de la separación respecto al esquema ALE LMS, pero, sin embargo, su principal inconveniente es que sigue dejando algunos restos de la señal de pulmón en la señal de corazón recuperada y viceversa. Se debe mantener un equilibrio entre cuanta señal de pulmón se quiere eliminar de la señal de corazón y cuanta señal de corazón se desea conservar a la salida. Esto lo convierte en un método que no presenta unos buenos resultados en la separación de sonidos cardiacos y pulmonares si se le compara con otros métodos como NLM y NMF.
- La principal ventaja del método ALE VSS-LMS es su escaso coste computacional, esto lo convierte en una opción a tener en cuenta para su adaptación en aplicaciones de separación de sonidos cardiacos y pulmonares en tiempo real.
- El inconveniente del método NLM es que el coste computacional sigue siendo grande. La mejora aplicada en este trabajo aprovechando la división del sonido cardiaco en periodos repetitivos, ha conseguido que el método sea realizable en la

práctica al reducir el tiempo de cálculo de horas a apenas minutos manteniendo la “no localidad” pero, sin embargo, el tiempo sigue siendo algo elevado.

- La principal ventaja del método NLM es su precisión al llevar a cabo la separación de ambos sonidos a pesar de que, como se ha comentado en capítulos anteriores, no se estén teniendo en cuenta algunos factores como la variación en la separación entre sí de los distintos sonidos que forman cada pulso cardiaco, por lo que, además, este método podría constar de un gran margen de mejora. Por otro lado, esta mejora aplicada en el algoritmo NLM, también puede ser aplicada en la eliminación de ruido de cualquier otra señal periódica como es el caso de las señales ECG en medicina.

CAPÍTULO 7

7. Líneas futuras

En este capítulo se determinarán algunas de las líneas futuras que pueden ser abarcadas para continuar con el desarrollo de este trabajo basado en la separación de sonidos cardiacos y pulmonares:

- Mejorar el alineamiento entre los distintos sonidos que forman cada periodo del sonido cardiaco con sus equivalentes en cualquier otro periodo, esto podría incrementar considerablemente los resultados en cuanto a la calidad de la separación obtenidos con el método NLM. Una posible solución podría basarse en el uso de algunos algoritmos capaces de segmentar el sonido cardiaco en los distintos sonidos que lo forman, de esta forma se podría mejorar la alineación de los distintos sonidos con sus equivalentes.
- Intentar reducir aún más el coste computacional del algoritmo NLM. Una posible solución podría basarse, por ejemplo, en evitar repetir el cálculo de los pesos entre las mismas muestras. A media que se van procesando más muestras del algoritmo existen la posibilidad de que estas muestras ya hayan sido comparadas con periodos anteriores, por lo que evitar nuevamente el cálculo de los pesos podría ahorrar un gran coste computacional.
- Implementar los algoritmos en dispositivos móviles. Además, en la actualidad, existen algunos accesorios capaces de convertir el dispositivo móvil en un estetoscopio. Por lo que se podrían utilizar las señales recogidas por el sensor y llevar a cabo la separación de ambos sonidos con el propio Smartphone.

CAPÍTULO 8

8. Interfaz gráfica

Con el objetivo de facilitar la tarea al usuario, se han desarrollado dos interfaces gráficas en MATLAB de manera que resulte mucho más fácil e intuitivo realizar la separación de señales de corazón y de pulmón a partir de los dos métodos desarrollados anteriormente, así como evaluar la calidad de dicha separación. Este software se adjunta en la carpeta de nombre “Software y Ficheros de Audio” anexa a la memoria.

8.1. Separacion_HSLS

Esta interfaz gráfica permite realizar la separación de una señal formada por la mezcla de sonidos cardiacos y pulmonares mediante la utilización de uno de los dos métodos que han sido desarrollados en este trabajo. Para ejecutarla, únicamente hay que situar el directorio actual de Matlab donde se encuentre guardado el archivo *Separacion_HSLS.m*. Se muestra a continuación una imagen de la interfaz:

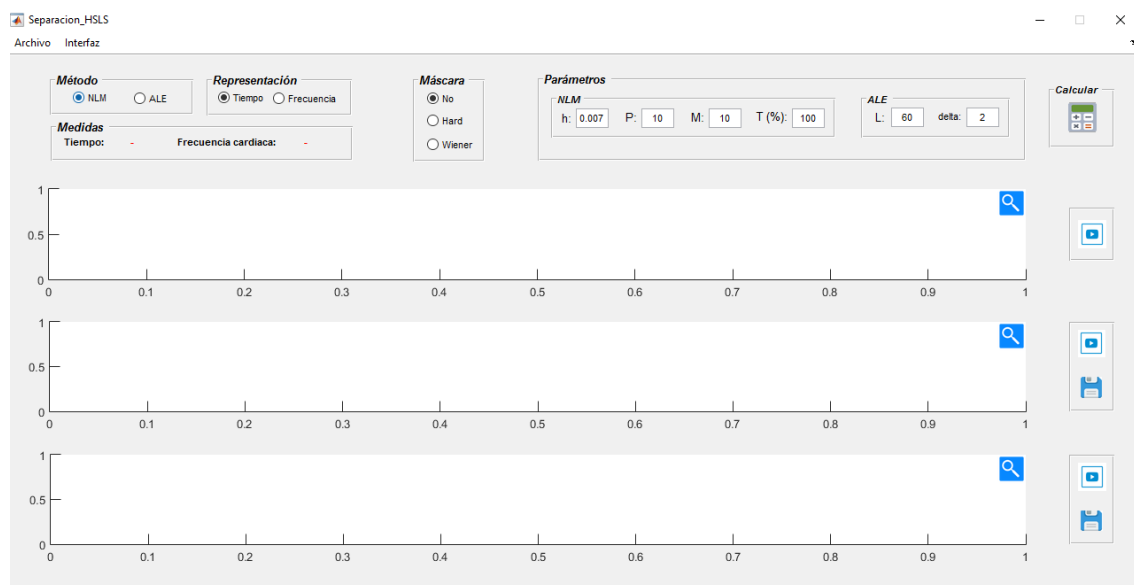


Figura 8.1 - Pantalla principal de la interfaz Separacion_HSLS

Como se puede observar en la figura anterior, la interfaz gráfica está dividida en distintas partes que permiten configurar desde el método y parámetros a utilizar para llevar a cabo la optimización, así como la máscara e incluso ver información sobre el tiempo que ha tardado en realizar la separación y una estimación de la frecuencia cardiaca de la señal de entrada. Se explica a continuación el proceso a seguir para realizar la separación de una señal de entrada:

- 1) En primer lugar, en el menú superior, se hace clic en *Archivo* → *Abrir archivo* → *Señal mezcla*. Posteriormente, se abrirá una ventana en la cual se debe cargar el archivo de sonido formado por la mezcla del sonido cardiaco y pulmonar. Una vez hecho esto, aparecerá la señal cargada, así como la estimación de la frecuencia cardiaca. También se puede reproducir la señal introducida haciendo clic en el botón situado a la derecha de la señal.

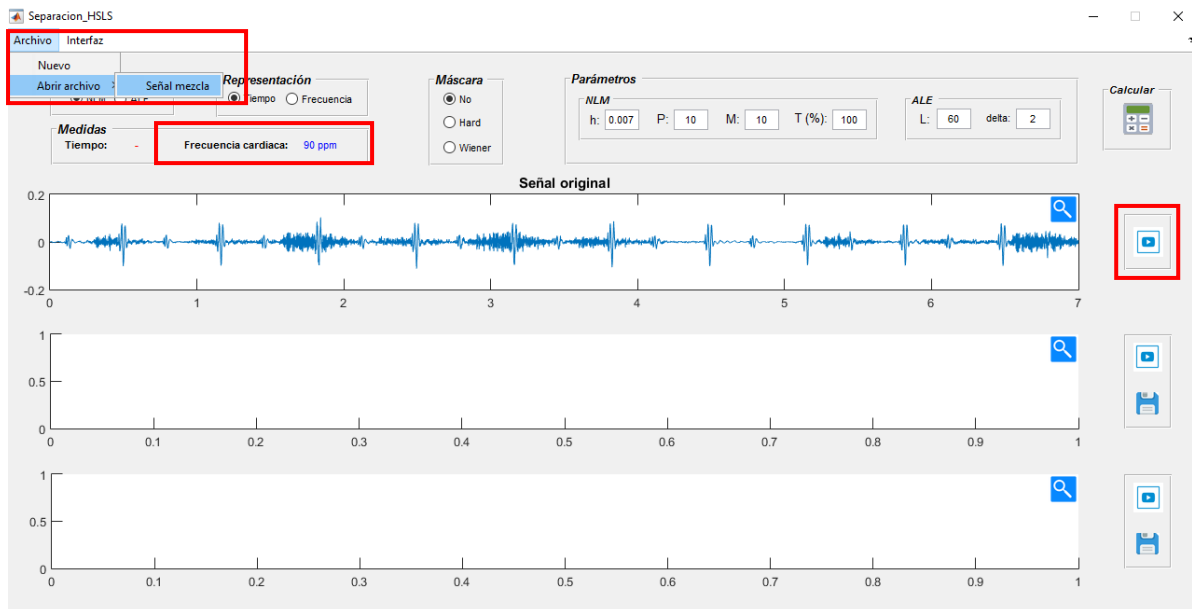


Figura 8.2 - Cargar archivo de sonido en la interfaz Separacion_HSLs

- 2) Se selecciona el método para llevar a cabo el filtrado, así como el tipo de máscara y los valores de los parámetros del método seleccionado, los cuales estarán fijados por defecto en su valor óptimo.

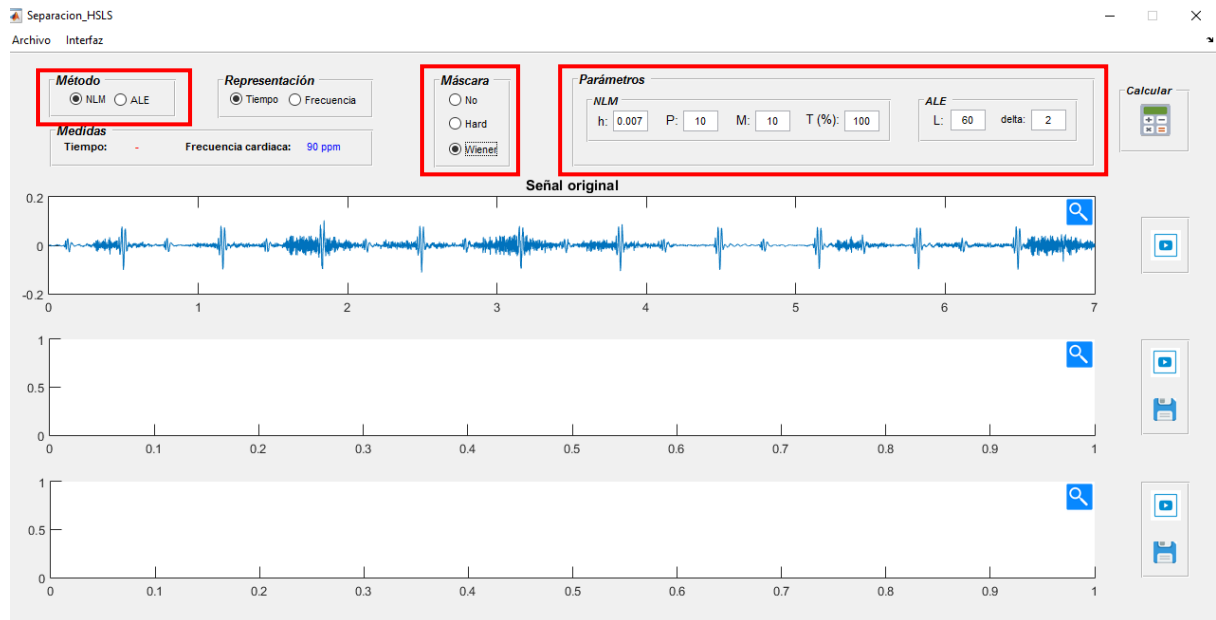


Figura 8.3 - Selección método, máscara y parámetros en la interfaz Separacion_HSLs

- 3) Una vez rellenados los campos anteriores, se hace clic en el botón *Calcular* y aparecerá un cuadro con el porcentaje y el tiempo restante para realizar el filtrado de la señal de entrada:

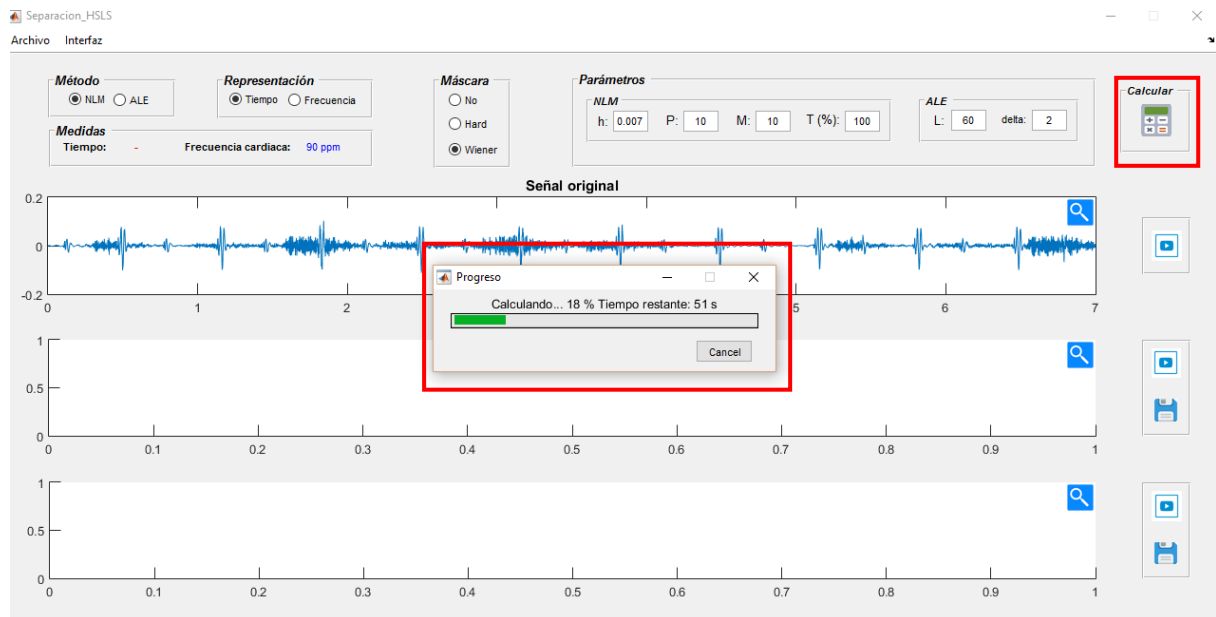


Figura 8.4 - Realizar filtrado en la interfaz Separacion_HSLs

- 4) Una vez realizado el filtrado, aparecerá las señales de corazón y pulmón recuperadas, así como el tiempo invertido en realizar el cálculo. Estas señales podrán ser guardadas (en dos archivos con el nombre *c_recuperado.wav*, correspondiente al sonido de corazón recuperado y *p_recuperado.wav*, correspondiente al sonido de pulmón recuperado) o reproducidas. También se puede modificar la representación del dominio temporal al dominio de la frecuencia.

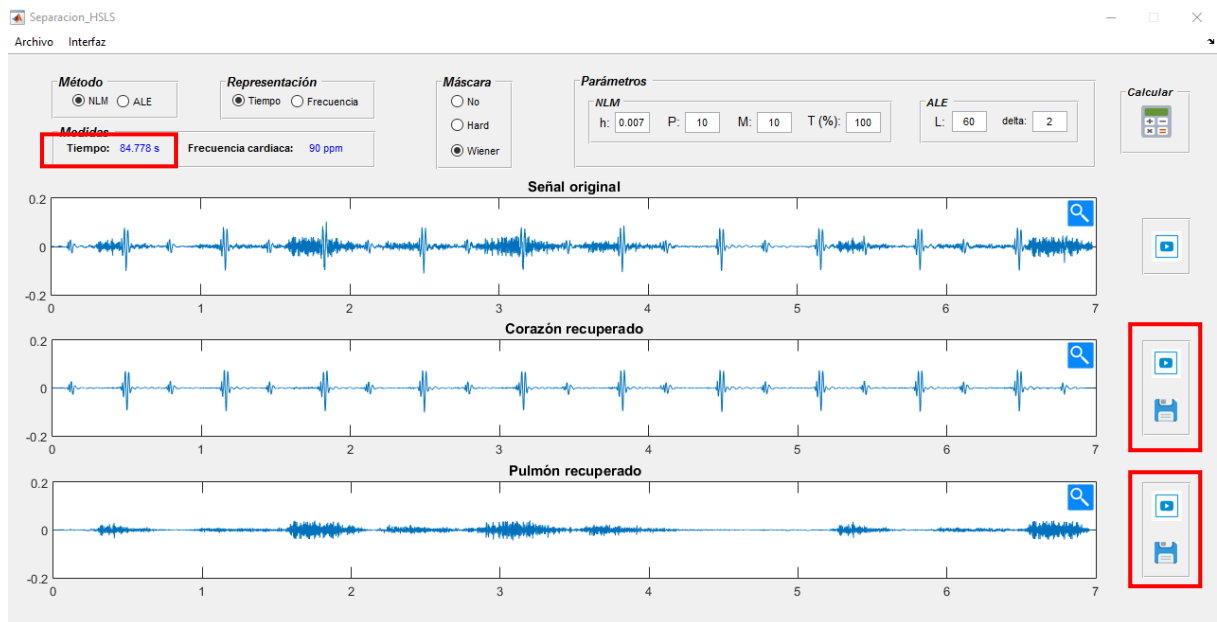


Figura 8.5 - Sonidos de corazón y de pulmón recuperados en la interfaz Separacion_HSLs

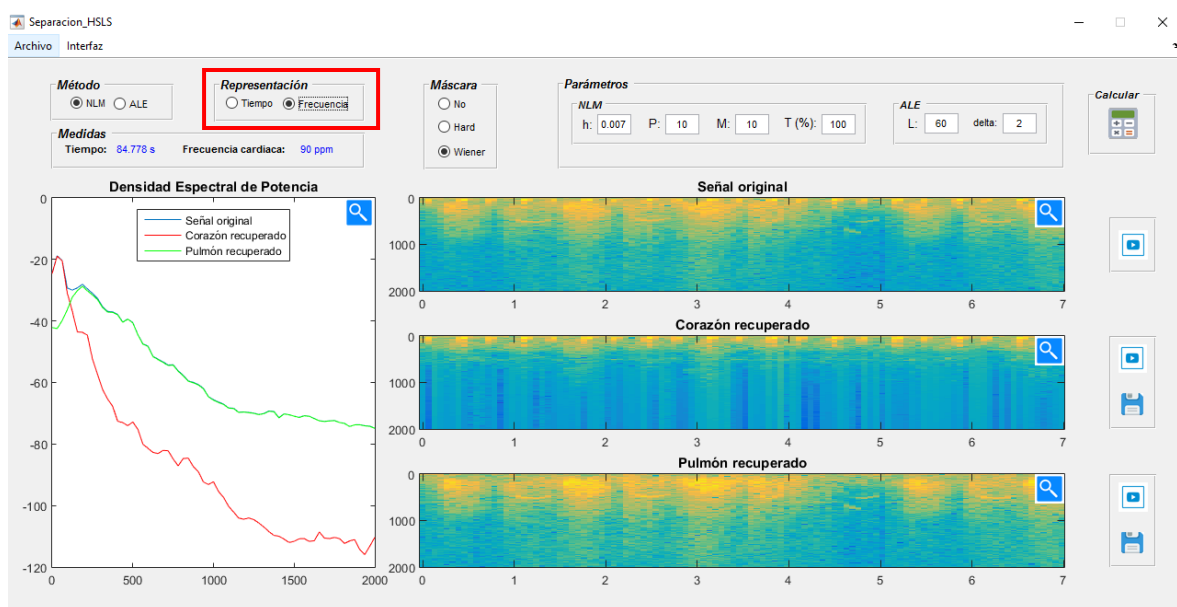


Figura 8.6 - Representación en el dominio de la frecuencia en la interfaz Separacion_HSLs

8.2. Calidad_Separacion_HSLS

Junto a *Separacion_HSLS*, se ha creado también la interfaz *Calidad_Separacion_HSLS* que permite obtener los valores de SDR, SIR y SAR de los sonidos de corazón y pulmón recuperados en el caso de que se disponga de las señales originales, así como la SNR con la que fueron mezclados ambos sonidos. Esta interfaz puede ser abierta directamente desde la interfaz *Separacion_HSLS*, haciendo clic en *Interfaz* → *Calidad_Separacion_HSLS*. De este modo, se puede evaluar la calidad de la separación realizada. Se muestra a continuación la pantalla principal de esta interfaz:

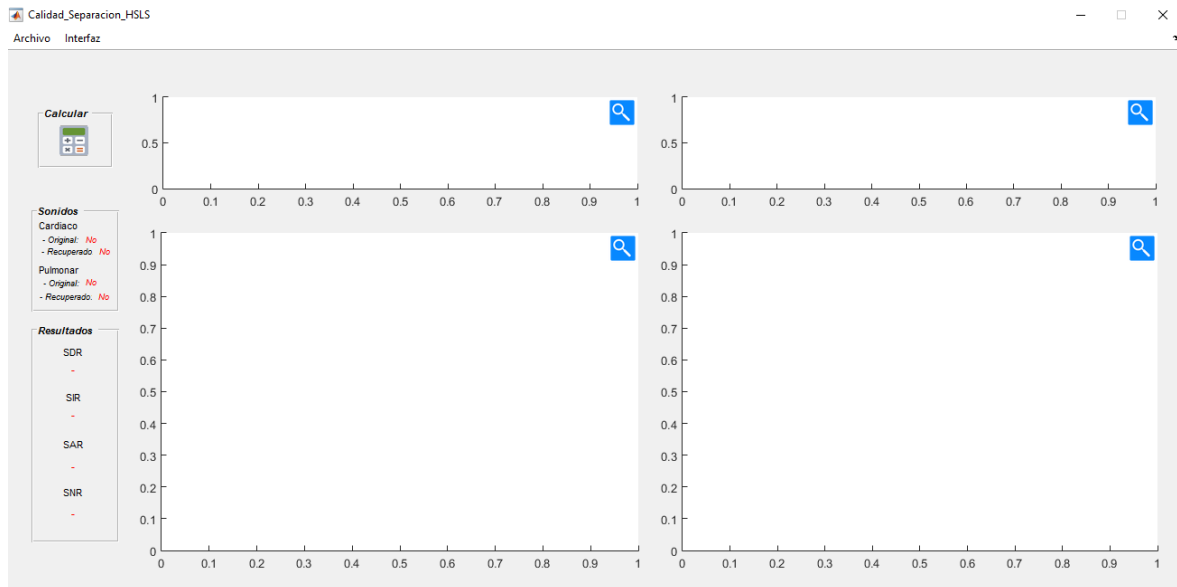


Figura 8.7 - Pantalla principal de la interfaz Calidad_Separacion_HSLS

Para mostrar los resultados, es necesario, al igual que en la interfaz anterior, cargar las señales necesarias que en este caso serán cuatro: el sonido de corazón original, el sonido de pulmón original, el sonido de corazón recuperado y el sonido de pulmón recuperado. Se muestran a continuación los pasos a seguir:

- 1) En primer lugar, se introducen las cuatro señales mencionadas anteriormente haciendo clic en *Archivo* → *Abrir archivo*.

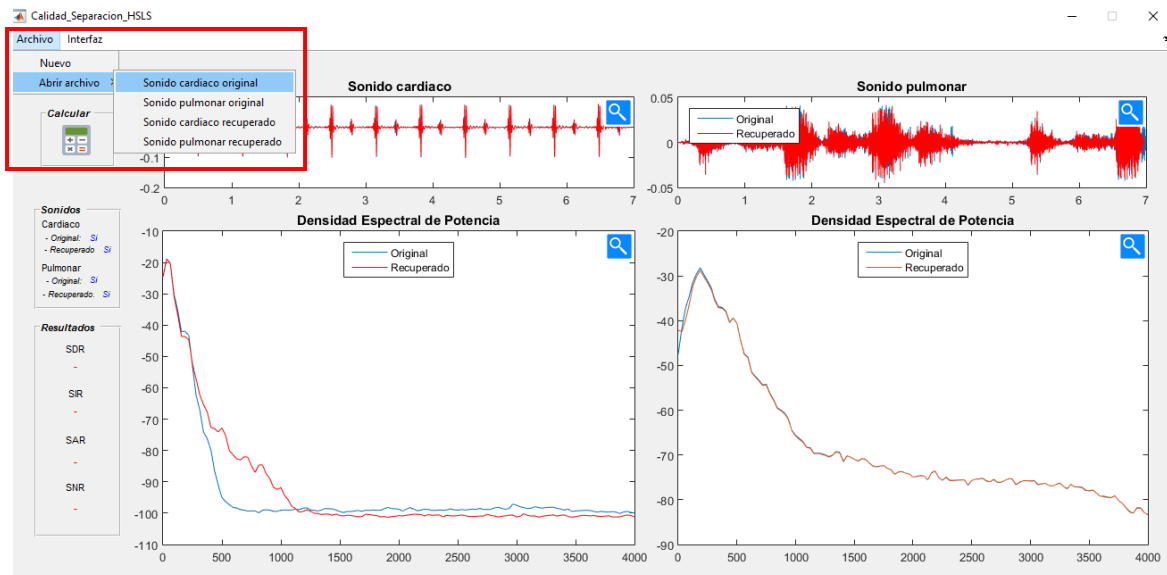


Figura 8.8 - Cargar archivos en la interfaz Calidad_Separacion_HSLs

- 2) Una vez cargados los cuatro archivos de sonido, únicamente faltaría pulsar en el botón calcular para obtener de esta forma los resultados.

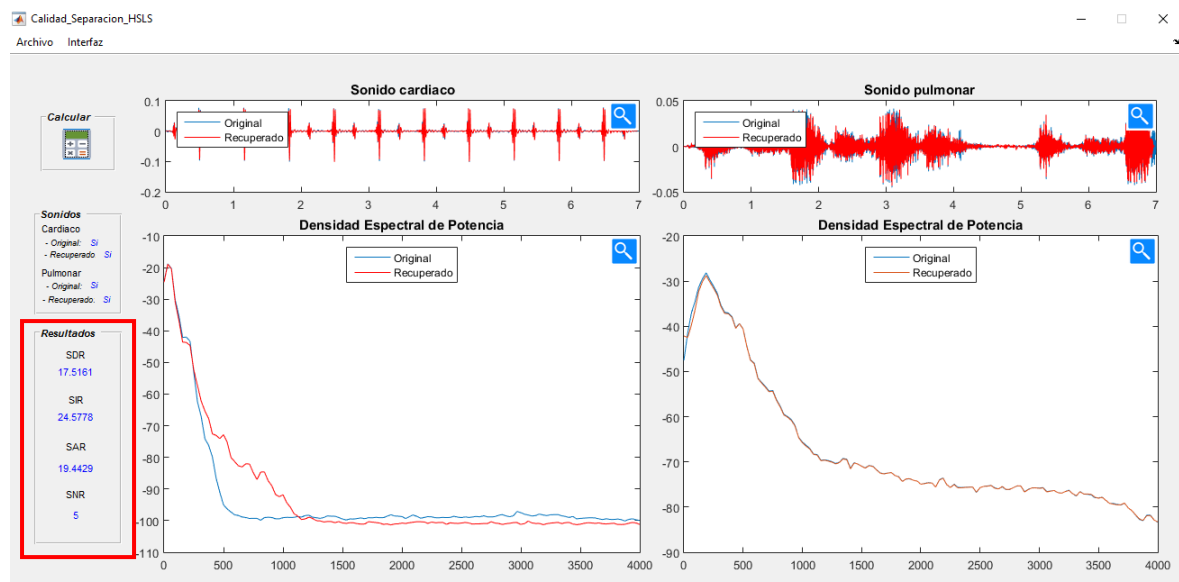


Figura 8.9 - Resultados en la interfaz Calidad_Separacion_HSLs

CAPÍTULO 9

9. Bibliografía

- [1] 3M™ Littmann® Stethoscopes: http://www.3m.com/3M/en_US/littmann-stethoscopes-us/ (último acceso: Julio de 2016)
- [2] Nurse Theory: <http://www.nursetheory.com/advantages-disadvantages-of-electronic-stethoscopes/> (último acceso: Julio de 2016)
- [3] Grais, Emad M.; Erdogan, Hakan. Single channel speech music separation using nonnegative matrix factorization and spectral masks, 2011 17th International Conference on Digital Signal Processing (DSP). IEEE, 2011. p. 1-6
- [4] Ministerio de Educación, Cultura y Deporte | Instituto Nacional de Tecnologías Educativas y de Formación del Profesorado: http://www.ite.educacion.es/formacion/materiales/60/cd/02_elsonido/index.html (último acceso: Julio de 2016)
- [5] Universidad de León: <http://fhyc.unileon.es/Milka/FyF/35.pdf> (último acceso: Julio de 2016)
- [6] Fundación Española del Corazón: <http://www.fundaciondelcorazon.com/prevencion/riesgo-cardiovascular/frecuencia-cardiaca.html> (último acceso: Julio de 2016)
- [7] Texas Heart Institute: http://www.texasheart.org/HIC/Anatomy_Esp/valve_sp.cfm (último acceso: Julio de 2016)
- [8] J. Gnitecki and Z. Moussavi. Separating heart sounds from lung sounds-Accurate diagnosis of respiratory disease depends on understanding noises", 2007, IEEE Eng. Med. Biol. Mag., vol. 26, no. 1, pp. 20-29
- [9] Medscape Education: <http://emedicine.medscape.com/article/1894036> (último acceso: Junio de 2016)

- [10] 3M™ Littmann® Stethoscopes: http://www.littmann.ca/wps/portal/3M/en_CA/3M-Littmann-CA/stethoscope/littmann-learning-institute/heart-lung-sounds/heart-sounds/ (último acceso: Junio de 2016)
- [11] Rudnitskii, A. G. Using nonlocal means to separate cardiac and respiration sounds. *Acoustical Physics*, 2014, vol. 60, no 6, p. 719-726.
- [12] Kids Health: <http://kidshealth.org/es/teens/lungs-esp.html> (último acceso: Julio de 2016)
- [13] National Library of Medicine: <https://www.nlm.nih.gov/medlineplus/ency/images/ency/fullsize/1103.jpg> (último acceso: Junio de 2016)
- [14] Practical Clinical Skills: <http://www.practicalclinicalskills.com/auscultation-course-contents.aspx?courseid=201> (último acceso: Junio de 2016)
- [15] Sapiens Medicus: <http://sapiensmedicus.org/blog/2015/01/13/ruidos-respiratorios/> (último acceso: Junio de 2016)
- [16] The electronic stethoscope: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4496820/> (último acceso: Julio de 2016)
- [17] Gavriely, Noam; Palti, Yoram; Alroy, Gideon. Spectral characteristics of normal breath sounds. *Journal of Applied Physiology*, 1981, vol. 50, no 2, p. 307-314.
- [18] Charbonneau, G., et al. An accurate recording system and its use in breath sounds spectral analysis. *Journal of Applied Physiology*, 1983, vol. 55, no 4, p. 1120-1127.
- [19] Iyer, Vijay K., et al. Reduction of heart sounds from lung sounds by adaptive filtering. *IEEE transactions on biomedical engineering*, 1986, no 12, p. 1141-1148.
- [20] Kompis, Martin; Russi, Erich. Adaptive heart-noise reduction of lung sounds recorded by a single microphone. *Engineering in Medicine and Biology Society, 1992 14th Annual International Conference of the IEEE. IEEE*, 1992. p. 691-692.
- [21] Lee, Y. J., et al. Single input ANC for suppression of breath sound. *World Academy of Science, Engineering and Technology*, 2010, vol. 44, p. 1160-1162.

- [22] Yip, L.; Zhang, Y. T. Reduction of heart sounds from lung sound recordings by automated gain control and adaptive filtering techniques. Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE. IEEE, 2001. p. 2154-2156.
- [23] Tsalaile, Thato; Sanei, Saeid. Separation of heart sound signal from lung sound signal by adaptive line enhancement. Signal Processing Conference, 2007 15th European. IEEE, 2007. p. 1231-1235.
- [24] Sathesh, K.; Muniraj, N. J. R. Real time heart and lung sound separation using adaptive line enhancer with NLMS. Journal of Theoretical & Applied Information Technology, 2014, vol. 65, no 2.
- [25] Hadjileontiadis, Leontios J.; Panas, Stavros M. Adaptive reduction of heart sounds from lung sounds using fourth-order statistics. IEEE Transactions on Biomedical Engineering, 1997, vol. 44, no 7, p. 642-648.
- [26] Gnitecki, J.; Moussavi, Z.; Pasterkamp, H. Recursive least squares adaptive noise cancellation filtering for heart sound reduction in lung sounds recordings. Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE. IEEE, 2003. p. 2416-2419.
- [27] Basak, Kausik, et al. Phonocardiogram signal analysis using adaptive line enhancer methods on mixed signal processor. 2010 International Conference on Signal Processing and Communications (SPCOM). IEEE, 2010. p. 1-5.
- [28] Charleston, S.; Azimi-Sadjadi, Mahmood R. Reduced order Kalman filtering for the enhancement of respiratory sounds. IEEE Transactions on Biomedical Engineering, 1996, vol. 43, no 4, p. 421-424.
- [29] Buades, Antoni; Coll, Bartomeu; Morel, J.-M. A non-local algorithm for image denoising. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). IEEE, 2005. p. 60-65.
- [30] Rudnitskii, A. G. Separating heart sound from lung sound using Non-Local Means filter. Electronics and Nanotechnology (ELNANO), 2014 IEEE 34th International Conference on. IEEE, 2014. p. 361-364.

- [31] Tracey, Brian H.; Miller, Eric L. Nonlocal means denoising of ECG signals. *IEEE Transactions on Biomedical Engineering*, 2012, vol. 59, no 9, p. 2383-2386.
- [32] Pourazad, M. T.; Mousavi, Z. K.; Thomas, G. Heart sound cancellation from lung sound recordings using adaptive threshold and 2D interpolation in time-frequency domain. *Engineering in Medicine and Biology Society*, 2003. *Proceedings of the 25th Annual International Conference of the IEEE*. IEEE, 2003. p. 2586-2589.
- [33] Pourazad, M. T., et al. Heart sounds separation from lung sounds using independent component analysis. *Proc. Ann. Int. Conf. IEEE EMBS*. 2005. p. 2736-2739.
- [34] Lin, ChingShun; Hasting, Erwin. Blind source separation of heart and lung sounds based on nonnegative matrix factorization. *Intelligent Signal Processing and Communications Systems (ISPACS)*, 2013 *International Symposium on*. IEEE, 2013. p. 731-736.
- [35] Flores-Tapia, Daniel; Moussavi, Zahra MK; Thomas, Gabriel. Heart sound cancellation based on multiscale products and linear prediction. *IEEE transactions on biomedical engineering*, 2007, vol. 54, no 2, p. 234-243.
- [36] Haykin, Simon S. *Adaptive filter theory*. Pearson Education India, 2008.
- [37] Widrow, Bernard, et al. Adaptive noise cancelling: Principles and applications. *Proceedings of the IEEE*, 1975, vol. 63, no 12, p. 1692-1716.
- [38] Kwong, Raymond H.; Johnston, Edward W. A variable step size LMS algorithm. *IEEE Transactions on signal processing*, 1992, vol. 40, no 7, p. 1633-1642.
- [39] Trabajo Fin de Máster, “Separación y búsqueda de patrones repetitivos procedentes del corazón humano en señales cardiopulmonares”, Alumno: Daniel Pancorbo Rubio, Tutor: Prof. D. Francisco Jesús Cañadas Quesada, Depto.: Ingeniería de Telecomunicación Universidad de Jaén, Marzo, 2016
- [40] E. Vincent, R. Gribonval and C. Févotte, Performance measurement in blind audio source separation, *IEEE Trans. Audio, Speech and Language Processing*, 2006, 14(4) pp 1462-1469.