



UNIVERSIDAD DE JAÉN
Escuela Politécnica Superior de Linares

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

Marzo, 2016

UNIVERSIDAD DE JAÉN
Escuela Politécnica Superior de Linares

Trabajo Fin de Máster

Curso 2015-2016

SEPARACIÓN Y BÚSQUEDA DE PATRONES REPETITIVOS PROCEDENTES DEL CORAZÓN HUMANO EN SEÑALES CARDIOPULMONARES

Alumno: *Pancorbo Rubio, Daniel*

Tutor: *Cañadas Quesada, Francisco Jesús*

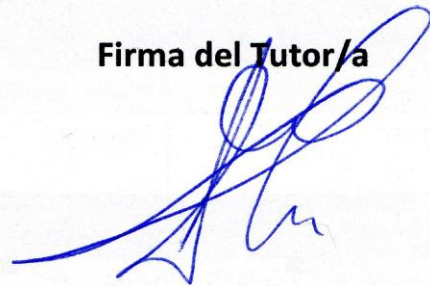
Depto.: *Ingeniería de Telecomunicación*

Visto Bueno a la defensa del TFM

Firma del autor /a



Firma del Tutor/a



A mi familia por su apoyo incondicional a lo largo de estos años y por sus continuos ánimos que han hecho que no me rindiese nunca. A mi pareja por estar siempre conmigo en las buenas y malas situaciones y por su apoyo incondicional. A mi tutor por su ayuda constante gracias a la cual la realización de este trabajo ha sido posible.

Índice General

PARTE I - MEMORIA.....	11
1. INTRODUCCIÓN	12
1.1. Contexto	12
1.2. Caracterización del sonido	13
1.2.1. Frecuencia fundamental (F_0), sonoridad, timbre y duración	15
1.2.2. Sonidos monocanal, estéreo y multicanal	17
1.2.3. Sonidos armónicos e inarmónicos	17
1.2.4. Sonidos de corazón y pulmón	18
2. DESCRIPCIÓN	26
2.1. Resumen general	26
2.2. Summary	28
3. OBJETIVOS	30
4. ESTADO DEL ARTE	32
4.1. Estado del arte de la separación de sonidos de corazón y pulmón	32
4.2. Introducción a Non-Negative Matrix Factorization (NMF)	41
4.2.1. Principios generales	41
4.2.2. Sparse NMF	45
4.2.3. Criterio de parada	47
5. DESARROLLO DEL TRABAJO FIN DE MÁSTER	48
5.1. Caracterización espectral de sonidos de corazón y pulmón	48
5.2. Metodología propuesta	54

5.3. Algoritmo de estimación del ritmo cardíaco	56
5.4. Sistema NMF Semisupervisado	62
5.4.1. Fase de entrenamiento	62
5.4.2. Modelo de separación	63
5.4.3. Función de coste, reglas de actualización multiplicativas y criterio de parada.	64
5.5. Sistema NMF No Supervisado	65
5.5.1. Modelo de Separación	65
5.5.2. Función de coste y reglas de actualización multiplicativas.	66
5.5.3. Clustering o Clasificación de bases	67
5.5.4. Modelo de factorización basado en el patrón rítmico	72
5.6. Reconstrucción y recuperación de las señales	74
5.7. Búsqueda de patrones repetitivos	75
6. RESULTADOS	82
6.1. Base de datos	82
6.2. Optimización de parámetros	83
6.2.1. Parámetros óptimos del sistema NMF no supervisado.	84
6.2.1.1. Parámetros óptimos del sistema NMF no supervisado utilizándose el modelo de factorización basado en el patrón rítmico	90
6.2.2. Parámetros óptimos del sistema NMF Semisupervisado	91
6.3. Testeo de sistemas	91
6.3.1. Testeo del sistema NMF No Supervisado.	92
6.3.1.1. Testeo del sistema NMF No supervisado usando el modelo de factorización basado en el patrón rítmico	99
6.3.2. Testeo del sistema NMF Semisupervisado.	101

6.4. Testeo del algoritmo de estimación del ritmo cardíaco	108
7. CONCLUSIONES	110
8. LÍNEAS DE INVESTIGACIÓN FUTUTAS	113
<i>ANEXO I - HERRAMIENTA CON INTERFAZ GRÁFICA.....</i>	<i>115</i>
9. HERRAMIENTA CON INTERFAZ GRÁFICA	116
<i>ANEXO II - COSTE COMPUTACIONAL.....</i>	<i>119</i>
10. COSTE COMPUTACIONAL	120
<i>ANEXO III - BIBLIOGRAFIA</i>	<i>121</i>
11. BIBLIOGRAFÍA	122

Índice de tablas

Tabla 1 - Posibles combinaciones de los técnicas de clustering.....	71
Tabla 2 - Posibles combinaciones de las técnicas de clustering para el nuevo modelo de factorización	73
Tabla 3 - Ejemplo del cálculo de parámetros óptimos 1.	85
Tabla 4 - Ejemplo del cálculo de parámetros óptimos 2.	86
Tabla 5 - Ejemplo del cálculo de parámetros óptimos 3.	87
Tabla 6 - Parámetros óptimos según el tipo de maximización para todas las combinaciones de clustering	89
Tabla 7 - Parámetros óptimos para la maximización de zonas de corazón o zonas de pulmón del nuevo modelo de factorización.....	90
Tabla 8 - Parámetros óptimos según el tipo de maximización de corazón en el sistema semisupervisado	91
Tabla 9 - Resultados finales sistema NMF No Supervisado	98
Tabla 10 - Resultados finales de maximización de corazón utilizando el nuevo método de factorización	99
Tabla 11 - Resultados finales de maximización de pulmón utilizando el nuevo método de factorización	100
Tabla 12 - Resultados finales sistema NMF Semisupervisado	107
Tabla 13 - Evaluación del algoritmo de estimación del ritmo cardíaco	108
Tabla 14 - Especificaciones ordenador.....	120

Índices de figuras

Figura 1 - Frecuencia fundamenta (F0) y armónicos	15
Figura 2 - Curvas isofónicas [52]	16
Figura 3 - Ilustración de corazón humano [50]	18
Figura 4 - Fases de bombeo de sangre en el corazón humano [51].....	19
Figura 5 - Espectrograma de sonido mezcla con ventana de 512 muestras	23
Figura 6 - Espectrograma de sonido mezcla con ventana de 1024 muestras	24
Figura 7 - Espectrograma de sonido mezcla con ventana de 2048 muestras	24
Figura 8 - Sistema basado en PCA	34
Figura 9 - Sistema basado en NMF 1	35
Figura 10 - Sistema basado en filtrado adaptativo	36
Figura 11 - Sistema basado en filtrado de modulación	38
Figura 12 - Sistema basado en NMF 2	39
Figura 13 - Sistema basado en NMF 3	40
Figura 14 - Forma de las distintas ecuaciones de coste [28]	41
Figura 15 - Matriz de activaciones H con λ cero	46
Figura 16 - Matriz de activaciones H con λ uno	46
Figura 17 - DEP de sonido de corazón S1.....	49
Figura 18 - DEP de sonido de corazón S2.....	49
Figura 19 - DEP de sonido de corazón S3.....	50
Figura 20 - DEP de sonido de corazón S4.....	50
Figura 21 - DEP de murmullo pulmonar.....	51

Figura 22 - DEP de sonido broncovesicular	52
Figura 23 - DEP de sonido roncus.....	52
Figura 24 - DEP de sibilancia monofónica.....	53
Figura 25 - DEP de sonido mezcla de corazón y pulmón	53
Figura 26 - Flujo espectral de señal mezcla	57
Figura 27 - Función de autocorrelación del flujo espectral R	60
Figura 28 - Patrón rítmico de señal mezcla	61
Figura 29 - Frecuencia de corte Roll-off para base de sonido de corazón	68
Figura 30 - Frecuencia de corte Roll-off para base de sonido de pulmón.....	69
Figura 31 - Función de búsqueda de patrones	76
Figura 32 - Búsqueda de patrones, instantes marcados	77
Figura 33 - Búsqueda de patrones, selección de la región.....	78
Figura 34 - Ejemplo de correlación cruzada normalizada 2-D.....	79
Figura 35 - Resultado de la búsqueda de patrones	80
Figura 36- Figura resumen de resultados SDR (dB) de corazón obtenidos por todos los métodos.....	92
Figura 37 - Figura resumen de resultados SIR (dB) de corazón obtenidos por todos los métodos.....	93
Figura 38 - Figura resumen de resultados SAR (dB) de corazón obtenidos por todos los métodos.....	93
Figura 39 - Figura resumen de resultados SDR (dB) de pulmón obtenidos por todos los métodos.....	94
Figura 40 - Figura resumen de resultados SIR (dB) de pulmón obtenidos por todos los métodos.....	95
Figura 41 - Figura resumen de resultados SAR (dB) de pulmón obtenidos por todos los métodos.....	95

Figura 42 - Figura resumen de resultados SDR (dB) equilibrado entre ambos sonidos obtenidos por todos los métodos.	96
Figura 43 - Figura resumen de resultados SIR (dB) equilibrado entre ambos sonidos obtenidos por todos los métodos.	97
Figura 44 - Figura resumen de resultados SAR (dB) equilibrado entre ambos sonidos obtenidos por todos los métodos.	97
Figura 45 - Figura resumen de resultados SDR (dB) de corazón para distintos valores de K.	101
Figura 46 - Figura resumen de resultados SIR (dB) de corazón para distintos valores de K.	102
Figura 47 - Figura resumen de resultados SAR (dB) de corazón para distintos valores de K.	102
Figura 48 - Figura resumen de resultados SDR (dB) de pulmón para distintos valores de K.	103
Figura 49 - Figura resumen de resultados SIR (dB) de pulmón para distintos valores de K.	104
Figura 50 - Figura resumen de resultados SAR (dB) de pulmón para distintos valores de K.	104
Figura 51 - Figura resumen de resultados SDR (dB) equilibrado para distintos valores de K.	105
Figura 52 - Figura resumen de resultados SIR (dB) equilibrado para distintos valores de K.	106
Figura 53 - Figura resumen de resultados SAR (dB) equilibrado para distintos valores de K.	106
Figura 54 – Interfaz de bienvenida a la herramienta gráfica	116
Figura 55 – Herramienta gráfica para el sistema NMF No Supervisado.....	117
Figura 56 – Herramienta gráfica para el sistema NMF Semisupervisado.....	118

PARTE I

MEMORIA

CAPÍTULO 1

1. Introducción

1.1. Contexto

Las enfermedades de corazón continúan siendo una de las principales causas de mortalidad en el mundo. Cada año mueren más personas por enfermedades cardiovasculares que por cualquier otra causa. Se calcula que en 2012, las muertes relacionadas por este tipo de enfermedades representaron un 31% de todas las muertes registradas en el mundo. Suponiendo 6,7 millones de muertes debidas a enfermedades relacionadas con el corazón [1].

La auscultación es un método de diagnóstico fundamental para las enfermedades del corazón, es no invasivo y barato. Sin embargo, este es altamente dependiente del entrenamiento y la experiencia del doctor en cuestión. Debido a esto, este tipo de diagnóstico se acompaña del uso de otras herramientas como el electrocardiograma, resonancia magnética y escáner tomográfico computarizado, suponiendo el uso de estas, una gran inversión económica solo accesible para los grandes hospitales [2]. Por tanto, el uso de herramientas computarizadas de bajo coste se torna fundamental como ayuda al diagnóstico de las posibles afecciones cardíacas que afecten al paciente.

Al tratarse la auscultación de un análisis acústico, se produce interferencia entre los sonidos de corazón y los sonidos procedentes de los pulmones, dificultándose así dicho diagnóstico y pudiendo llevar a una confusión del estado patológico del corazón, dado que los sonidos cardíacos pueden quedar enmascarados por aquellos procedentes de los pulmones, o la capacidad cognitiva del doctor puede verse mermada, dificultando así la detección de posibles síntomas relacionados con diversas enfermedades. Por estos motivos, dicha interferencia se ha convertido en uno de los mayores problemas en las investigaciones biomédicas [3];

Es objeto de este Trabajo Fin de Máster desarrollar una herramienta que permita la separación entre sonidos de corazón y pulmón, solventando en mayor medida dicha problemática, facilitando así al doctor el diagnóstico de enfermedades del corazón.

1.2. Caracterización del sonido

En esta sección, se realiza una introducción al concepto de sonido; este se define como la sensación producida en el oído humano debido a la vibración o perturbación de las partículas que se desplazan en forma de onda a través de un medio elástico que las propaga [13].

Para producir sonido se necesita, en primer lugar, la existencia de un cuerpo vibrante o foco y, en segundo lugar, un medio elástico que transmita dichas vibraciones en forma de ondas sonoras.

El sonido se caracteriza matemáticamente como onda a través de una función dependiente de la posición y el tiempo. Dicha función se conoce como ecuación de onda y se define como:

$$\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \frac{\partial^2 \Psi}{\partial z^2} = \frac{1}{v^2} \frac{\partial^2 \Psi}{\partial t^2} \quad (1)$$

siendo x, y, z las coordenadas cartesianas espaciales, v la velocidad de la onda, t la variable temporal y Ψ una magnitud que define el comportamiento de la onda dependiente de la posición y el tiempo.

Asimismo, las ondas sonoras pueden clasificarse en función del medio de propagación, la dirección en que vibran las partículas del medio y la geometría de los frentes de onda [14] y [15]:

Según el medio de propagación pueden ser:

- **Ondas mecánicas:** Se define como onda mecánica aquella onda cuya energía no puede transmitirse a través del vacío, necesitando así un medio físico y elástico para propagarse.
- **Ondas electromagnéticas:** Se define como onda electromagnética aquella onda cuya energía puede transmitirse a través el vacío, no necesitando un medio físico para dicha propagación.

Según la dirección de vibración de las partículas:

- **Ondas transversales:** La vibración de las partículas del medio se produce de forma perpendicular a la dirección de propagación de la energía.
- **Ondas longitudinales:** La vibración de las partículas del medio se produce de forma paralela a la dirección de propagación de la energía.

Según la geometría de los frentes de onda, lugar geométrico del espacio en el que se encuentran los puntos en los que la onda llega en un mismo instante, pueden clasificarse principalmente en:

- **Ondas planas:** Los frentes de onda forman un plano.
- **Ondas circulares:** Los frentes de onda son circunferencias.
- **Ondas esféricas:** Los frentes de onda son esferas radiales.

Por tanto, el sonido se describe como onda mecánica, ya que necesita de un medio material y elástico para desplazarse, como onda longitudinal, dado que el movimiento de dichas partículas se propagan en la misma dirección de la onda, y como onda esférica, ya que sus ondas son tridimensionales desplazándose en tres direcciones y sus frentes de onda son esferas radiales que se propagan en todas direcciones siendo su foco la fuente de la perturbación, [13].

1.2.1. Frecuencia fundamental (F0), sonoridad, timbre y duración

En este apartado, se revisarán brevemente las principales cualidades del sonido, las cuales nos permiten caracterizarlo. Dichas cualidades son:

- **La frecuencia fundamental (F0), altura o tono:** Con la vibración de un cuerpo, a menudo, se producen sonidos que presentan componentes con distintas frecuencias. Se considera frecuencia fundamental (F0) a la frecuencia de la componente con frecuencia más baja, siendo la frecuencia del resto de componentes o armónicos, múltiplos enteros de la frecuencia fundamental [13].

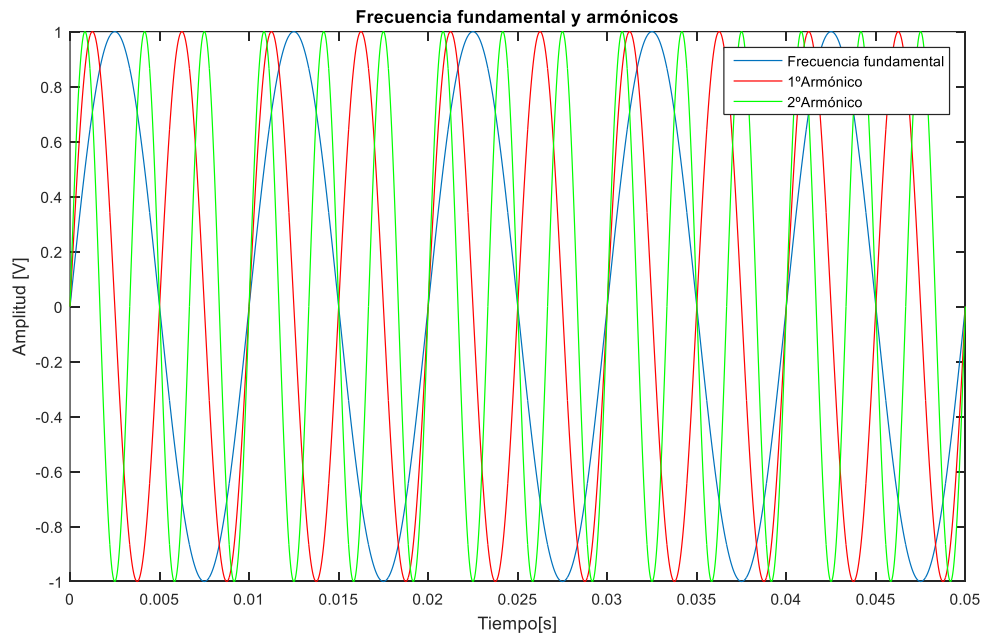


Figura 1 – Frecuencia fundamental (F0) y armónicos

- **Intensidad y sonoridad:** La intensidad sonora se define como el valor promedio de energía que atraviesa unidad de superficie en la unidad de tiempo, siendo esta dependiente de la presión acústica y la velocidad de la onda sonora. Por otro lado, la sonoridad se define como una medida subjetiva de la intensidad con la que el oído humano percibe un sonido, permitiendo la sonoridad clasificar los sonidos entre sonidos fuertes y débiles. Así mismo, la sonoridad se mide en fonios, siendo el fonio la sonoridad de un sonido senoidal de 1Khz con una intensidad de 0 dB_{SPL} [52].

Las primeras curvas de igual sonoridad, curvas isofónicas, ilustradas por Munson y Fletcher, muestran la relación entre la frecuencia y la intensidad en dB de dos sonidos para que se perciban igual de fuertes [52]:

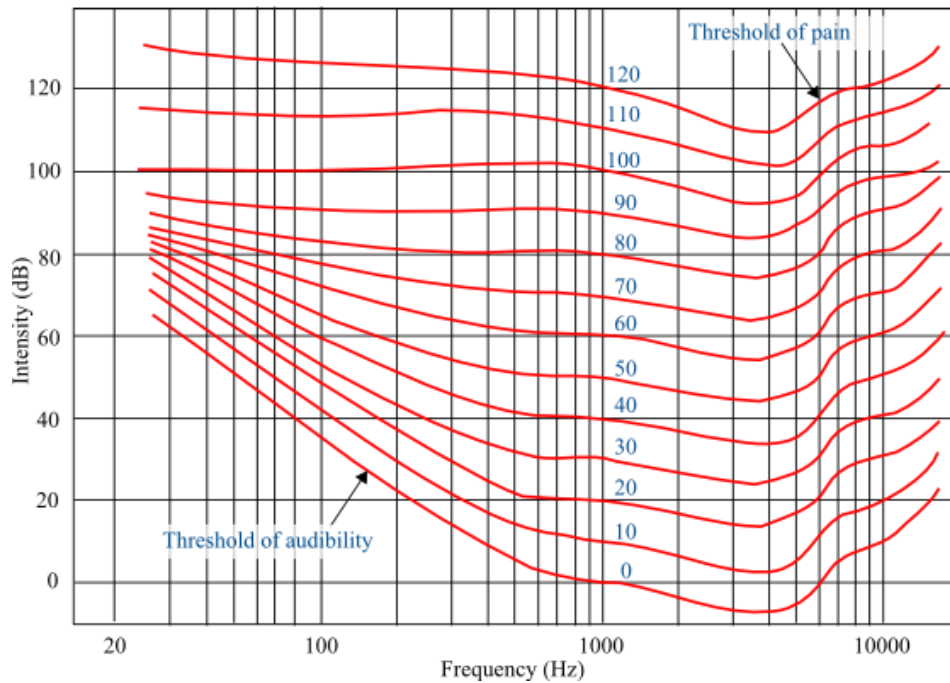


Figura 2 - Curvas isofónicas [52]

- **Timbre:** Es la cualidad del sonido por la cual somos capaces de distinguir entre dos sonidos de igual intensidad y frecuencia fundamental, dependiendo el timbre de la cantidad de armónicos que un sonido tenga y de la intensidad de cada uno de ellos [52].
- **Duración:** Es la cualidad del sonido que indica el tiempo que dura su vibración [53].

1.2.2. Sonidos monocanal, estéreo y multicanal

En otros tipos de clasificación del sonido se atiende, por ejemplo, a si este contiene o no información espacial. Así, se puede distinguir entre sonidos monocanal, estéreo o multicanal, en [16].

- **Sonidos monocanal:** Para estos sonidos, se ha usado un único micrófono durante su grabación, por esto, no contiene información espacial y resultan de la suma de todas las fuentes sonoras activas en dicho momento. Nuestro trabajo se centrará en este tipo de sonidos.
- **Sonidos estéreo:** Para estos sonidos, se han usado dos micrófonos independientes durante su grabación, aportando así información espacial de las fuentes sonoras activas.
- **Sonidos multicanal:** Para estos sonidos, se han usado más de dos micrófonos independientes, obteniéndose así sonidos con un número n de canales que aportan información espacial de las fuentes sonoras activas durante la grabación.

1.2.3. Sonidos armónicos e inarmónicos

Atendiendo a la periodicidad del sonido, podemos distinguir entre sonidos armónicos e inarmónicos:

- **Sonidos armónicos:** Son aquellos sonidos periódicos donde la frecuencia de todas sus componentes son múltiplos enteros de la frecuencia fundamental (F_0). Además, dichas componentes o armónicos permiten generar el timbre característicos de una determinada fuente de sonido, que lo diferenciará del resto de fuentes [54].

- **Sonidos inarmónicos:** Son aquellos sonidos no periódicos, donde la frecuencia de sus distintas componentes no son múltiplos de una frecuencia fundamental, siendo dicha frecuencia F_0 difícil de apreciar de forma precisa [54].

1.2.4. Sonidos de corazón y pulmón

En este apartado, se realizará en primer lugar un estudio de la anatomía del corazón humano, tras el cual se describirán los principales sonidos que en este se producen. En la siguiente figura, se representan las distintas partes que forman dicho órgano:

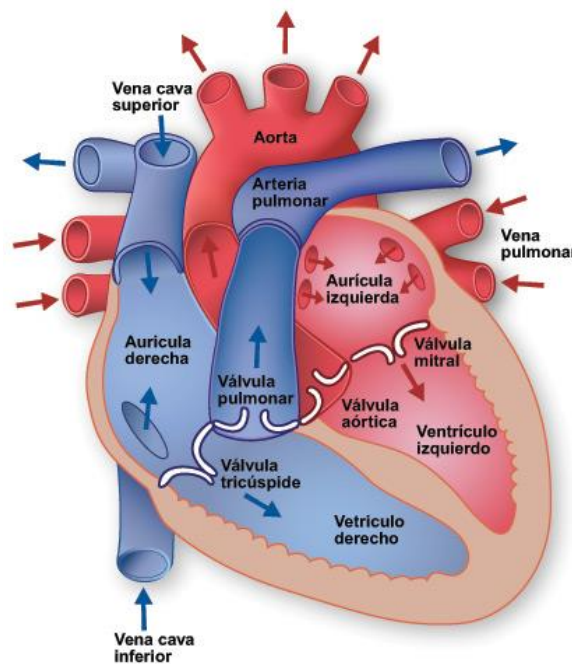


Figura 3 - Ilustración de Corazón humano [50]

El corazón humano es el órgano principal del sistema circulatorio humano [50]. Dicho órgano está compuesto por cuatro cavidades, dos superiores llamadas ‘aurícula izquierda’ y ‘aurícula derecha’ y dos inferiores llamadas ‘ventrículo izquierdo’ y ‘ventrículo derecho’. Además, se encuentra el ‘tabique’, una pared muscular cuya función es separar ambos ventrículos y ambas aurículas. Por otro lado, en el corazón humano se encuentran cuatro válvulas cardíacas, cada una con una función específica:

- **La válvula tricúspide:** Se encarga de regular el flujo sanguíneo entre la aurícula derecha y el ventrículo derecho.
- **La válvula pulmonar:** Se encarga de controlar el flujo sanguíneo entre el ventrículo derecho a las arterias pulmonares.
- **La válvula mitral:** Controla el paso de la sangre proveniente de los pulmones desde la aurícula izquierda al ventrículo izquierdo.
- **La válvula aórtica:** Controla el paso de la sangre desde el ventrículo izquierdo a la arteria aorta.

Una vez se conocen las distintas cavidades y válvulas de las que se compone el corazón humano, se describirá brevemente el proceso por el cual se produce el 'latido cardíaco' [51], el cual consiste en el bombeo de sangre en el corazón humano a través de dos fases llamadas Diástole y Sístole. En la siguiente figura se ilustran ambas fases:

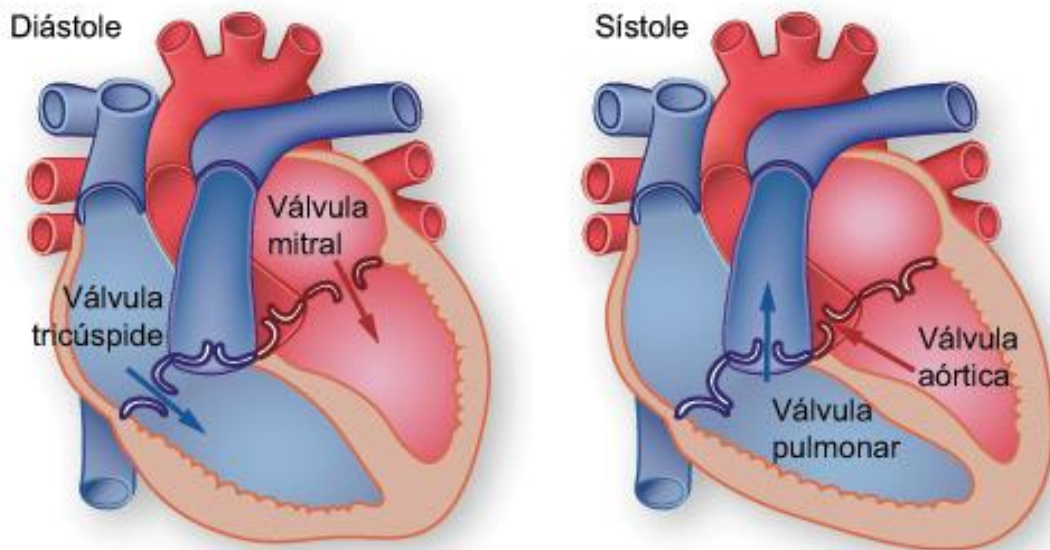


Figura 4 - Fases de bombeo de sangre en el corazón humano [51].

Durante la fase de Diástole, la más larga, consiste en el bombeo de sangre que, a través de las válvulas tricúspide y mitral, se bombea desde las aurículas derecha e izquierda, con la contracción de ambas, hacia los ventrículos derecho e izquierdo; fase durante la cual las válvulas aórtica y pulmonar permanecen cerradas. Durante la fase de Sístole, una vez los ventrículos están llenos de sangre, ambos se contraen bombeando la sangre hacia la arteria aorta, a través de la válvula aórtica, y la arteria pulmonar, a través de la válvula pulmonar; fase durante la cual las válvulas tricúspide y mitral permanecen cerradas. La apertura y cierre de todas las válvulas se realiza a través de impulsos eléctricos provenientes del marcapasos natural del corazón o nódulo SA.

Tras el estudio anatómico del corazón y del proceso por el cual se produce el latido humano, se realizará una breve introducción a los principales sonidos producidos por el corazón, S1 y S2, así como de los principales sonidos anómalos S3 y S4 [8].

- Sonidos normales producidos por el corazón:
 - a) **Primer sonido, S1:** Este sonido se produce por una desaceleración brusca de la sangre contra las válvulas mitral y tricúspide cuando están cerradas.
 - b) **Segundo sonido, S2:** Este sonido se produce por una desaceleración brusca de la sangre en la aorta y arteria pulmonar contra las válvulas semilunares.
- Sonidos anómalos producidos por el corazón:
 - a) **Tercer sonido, S3:** Sonido de baja intensidad que aparece en la fase inicial de la diástole, se produce durante el llenado ventricular una vez que se han abierto las válvulas tricúspide y mitral. Se considera como síntoma anómalo, cuando aparece en pacientes con una edad superior a 30 años.
 - b) **Cuarto sonido, S4:** Sonido producido por la relajación brusca y la vibración de los ventrículos cuando se produce la contracción auricular. Este sonido indica una

disminución de la compliancia ventricular o la incapacidad del ventrículo a volver a su estado original tras su compresión.

A continuación, se realizará una breve descripción de los sonidos pulmonares normales y sonidos anómalos de una disfunción típica respiratoria como es la bronquitis en [10] y [11]

- Sonidos normales producidos por los pulmones:
 - a) **Murmullos vesiculares**: Sonidos de tonalidad baja, auscultados en el tórax y producidos por las vibraciones originadas por el movimiento del aire al atravesar las vías aéreas.
 - b) **Respiración bronquial**: Sonidos auscultados en la tráquea y producidos por el movimiento del aire al atravesar la misma.
 - c) **Ruido broncovesicular**: Sonidos respiratorios que se auscultan entre los sonidos de los conductos bronquiales y los alveolares, o son una combinación de ambos sonidos.
- Sonidos anómalos producidos por los pulmones debido a la bronquitis:
 - a) **Roncus (estertores secos)**: Sonidos similares a los ronquidos, producidos cuando se obstruye o se vuelve áspero el aire a través de las vías respiratorias. Aparecen cuando se produce una bronquitis aguda.
 - b) **Sibilancia monofónica**: Sonido chillón que se produce con el paso del aire a través de vías respiratorias estrechas en los pulmones.

En cuanto a la representación de las señales utilizadas a lo largo de este TFM, se ha optado por el módulo de la Transformada discreta de Fourier de tiempo corto o STFT. Dicha

transformada proporciona una representación lineal espectro-temporal de las fuentes a separar. De este modo, la matriz de entrada a los sistemas implementados en este TFM, V , es:

$$V = |X_{STFT}| \quad (2)$$

Asimismo, la Transformada discreta de Fourier en tiempo corto se define matemáticamente como, [30]:

$$X_m(w) = \sum_{n=-\infty}^{\infty} x(n) * w(n - mR) * e^{-jwn} \quad (3)$$

donde $x(n)$ se corresponde con la señal de entrada en el instante n , $w(n)$ con una función ventana de longitud fija L , $X_m(w)$ con la Transformada discreta de Fourier de la señal enventanada en el instante mR y R con el solapamiento en muestras entre ventanas.

Con dicha transformada, se obtiene una resolución temporal y espectral constantes, calculadas ambas como:

$$Resolución_{temporal} = \frac{N}{fs} \quad (4)$$

$$Espaciado_{espectral} = \frac{fs}{N} \quad (5)$$

En nuestro trabajo, se ha utilizado un solapamiento espectral entre ventanas del cincuenta por ciento y una ventana Hanning de longitud 512 muestras utilizada en [31] que junto a una frecuencia de muestreo de 8000Hz de las señales mezclas monocanal generadas, proporciona una resolución temporal de 64 ms y un espaciado espectral de 15.65 Hz. Se ha utilizado dicha ventana, ya que la energía de los fuentes a descomponer se concentra en la zona de baja frecuencia, el uso de dicha ventana resulta adecuado dado que permite eliminar los artefactos de alta frecuencia que no forman parte de los fuentes a extraer [32].

Finalmente, se ha calculado el espectrograma tanto de las señales mezcla como de las fuentes extraídas. El espectrograma se define como:

$$S = |X_{STFT}|^2 \quad (6)$$

A continuación, se muestran espectrogramas correspondientes a un mismo sonido mezcla de sonidos de corazón, marcados con un cuadro rojo, y sonidos correspondientes a bronquitis conocidos como sibilancias polifónicas, marcados con un cuadro amarillo, utilizándose diferentes longitudes de ventana, calculándose la STFT para la banda de frecuencia completa; sin embargo, en estas figuras se realizara un zoom para centrarnos en la banda comprendida entre 0-1000Hz:

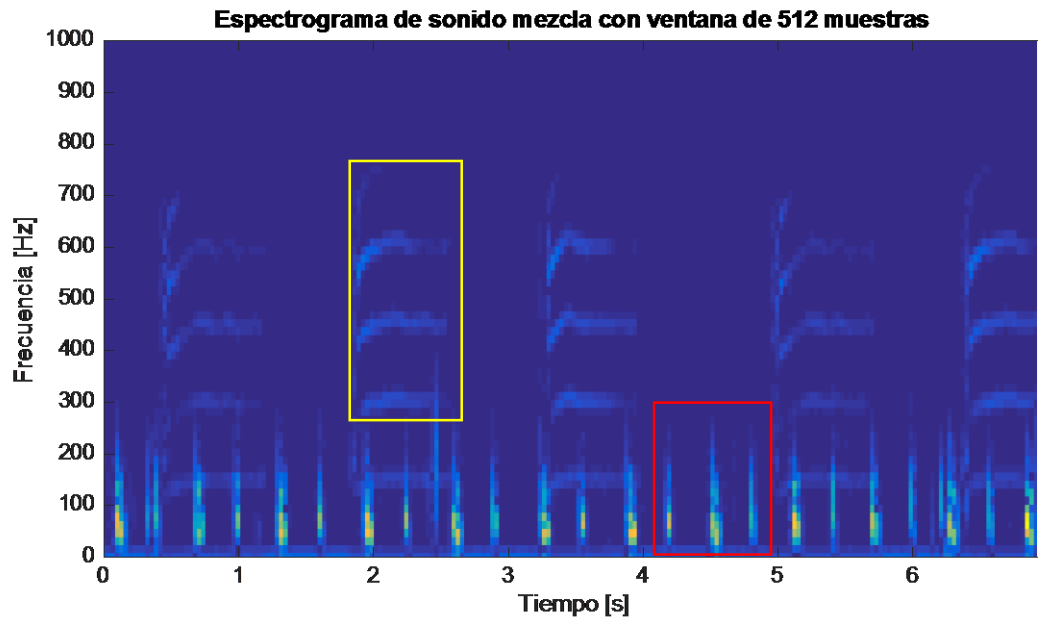


Figura 5 - Espectrograma de sonido mezcla con ventana de 512 muestras

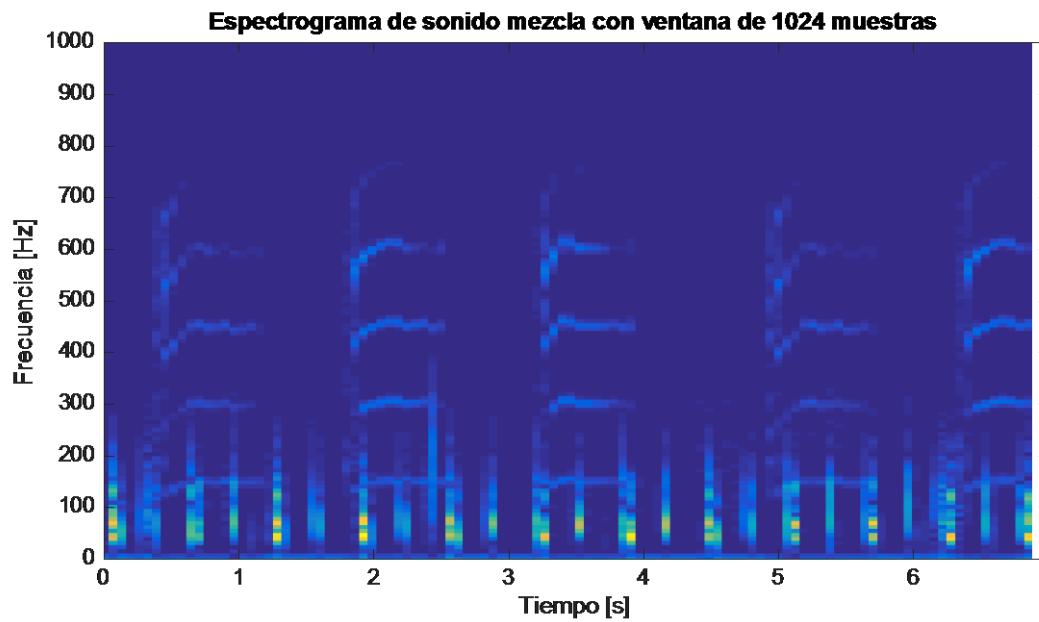


Figura 6 - Espectrograma de sonido mezcla con ventana de 1024 muestras

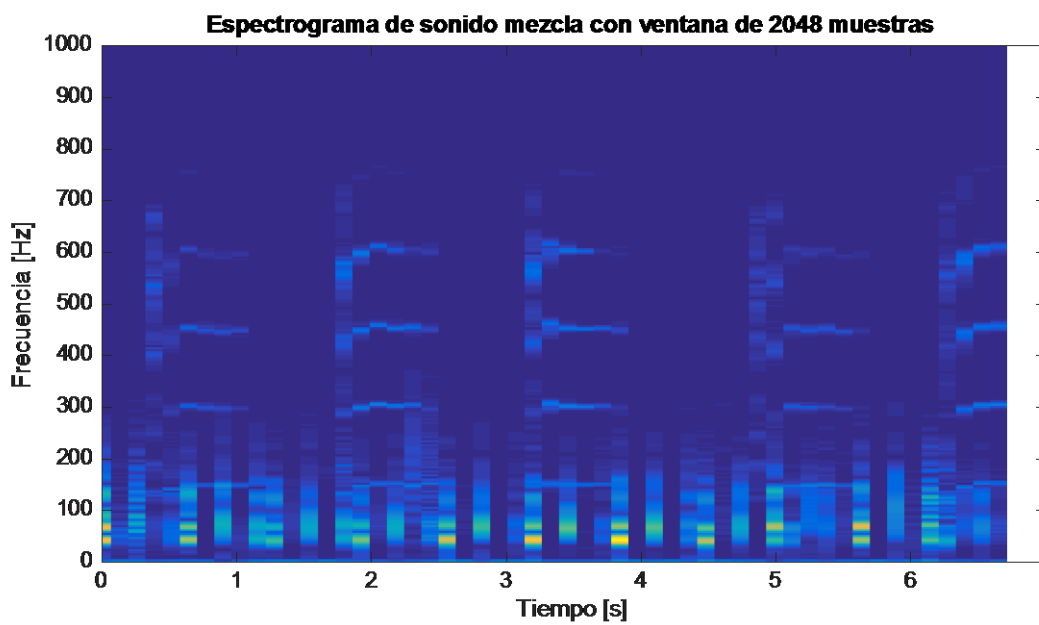


Figura 7 - Espectrograma de sonido mezcla con ventana de 2048 muestras

CAPÍTULO 1 - INTRODUCCIÓN

CAPÍTULO 2

2. Descripción

2.1. Resumen general

Dada la problemática anteriormente expuesta, se produce un solapamiento en tiempo-frecuencia entre los sonidos de corazón y pulmón, apareciendo una zona de solapamiento espectral en la banda de frecuencia comprendida entre 60-320 Hz [3] y [10].

En nuestro trabajo, se propone el uso de una herramienta gráfica sencilla basada en dos métodos distintos de separación de fuentes sonoras, ambos basados en las técnicas de factorización de matrices no negativas o NMF, con la finalidad de ayudar al doctor durante la realización del diagnóstico del paciente; tratándose de solventar en mayor medida la problemática que en este TFM se expone.

Así pues, el primer método de separación se basa en un sistema de separación NMF Semisupervisado donde, en una fase previa de entrenamiento, se han entrenado los patrones espectrales de sonidos cardíacos procedentes del corazón, caracterizando espectralmente dichos sonidos y creando una serie de diccionarios que son usados posteriormente en la descomposición.

El segundo método de separación se basa en un sistema NMF No Supervisado, el cual se acompaña de una fase posterior de Clustering o Clasificación de bases, donde se han introducido una serie de técnicas novedosas para este propósito, no usadas con anterioridad en otros trabajos, como son la medida del Roll-Off o distribución de energía en frecuencia, la correlación de patrones espectrales y la correlación de activaciones temporales de sonidos del corazón.

Además, se ha creado un algoritmo de estimación del ritmo cardíaco con dos propósitos principales:

- Realizar un tercer proceso de factorización NMF No Supervisado de la matriz de entrada en bloques, distinguiendo entre aquellos correspondientes a sonidos de corazón y sonidos de pulmón, utilizando un número de bases adaptado para la factorización de cada bloque.
- Generar un patrón de activaciones temporales adaptado a cada sonido mezcla, que permita realizar la correlación de activaciones temporales y la segmentación de los sonidos de corazón para la fase de entrenamiento del sistema Semisupervisado.

Asimismo, en dicha herramienta se ha implementado un algoritmo de búsqueda de patrones repetitivos en las señales de corazón y pulmón generadas tras la descomposición, con el fin de facilitar al doctor la búsqueda de aquellos sonidos que pudieran ser anómalos a juicio de este.

Finalmente, este TFM se acompaña de una herramienta gráfica donde el usuario tiene acceso a ambos sistemas, permitiéndose la configuración de parámetros para la separación y la visualización de resultados, y a dicha función de búsqueda de patrones, con la cual se permite la selección de sonidos de corazón o pulmón y se permite encontrar aquellos con un nivel mínimo de similitud con el sonido seleccionado. Además, se pueden escuchar los sonidos procedentes de la separación y los sonidos procedentes de la búsqueda de patrones, permitiéndose posteriormente su almacenamiento.

2.2. Summary

Due to the aforementioned problem, an overlap region in time and frequency domains appears between both sounds, producing an overlap area in the frequency band between 60-320 Hz [3] y [10].

In our work, we propose a simple graphical tool based on two different methods of source separation, both based on Non Negative Matrix Factorization, NMF, techniques in order to help the doctor make an accurate diagnose, trying to solve the discussed problems in this work.

Thus, the first separation method is based on a semi-supervised NMF separation system, in which, we use dictionaries that contain trained spectral patterns from heart sounds that we generated previously. In this way, we characterize spectral characteristics from heart sounds.

Furthermore, the second method is based on an Unsupervised NMF system and it is associated with a later Clustering stage in which we have introduced a selection of novel techniques for this purpose, such as the Roll-Off measuring, spectral patterns' correlation and temporary activations' correlation of heart sounds.

Also, we have created a heart rate estimation algorithm with two main purposes; firstly, performing a third process NMF unsupervised factorization in which the input matrix is divided into blocks, distinguishing between those corresponding to heart sounds and lung sounds and using an adapted number of basis for the factorization process for each block; secondly, estimating a temporary activations' pattern to each mixed sound, doing possible the temporary activation's correlation and the segmentation' process in the training stage for the semisupervised system.

In addition, it has been developed a repetitive patterns' searching algorithm to look for abnormal sounds from heart and lung sounds. Finally, this work is accompanied by a simple graphical interface where users can access to both systems and use the searching function. Users can configure the parameters freely and the results are shown clearly.

CAPÍTULO 2 – DESCRIPCIÓN

CAPÍTULO 3

3. Objetivos

El objetivo principal de este Trabajo Fin de Máster es la creación de una herramienta que permita la extracción por separado de aquellos sonidos procedentes del corazón humano que han sido interferidos por sonidos procedentes del sistema respiratorio durante el proceso de auscultación y que, además, permita la búsqueda de patrones repetitivos dentro de corazón o pulmón.

A continuación, se enumeran una serie de objetivos secundarios que se pretenden alcanzar con el presente TFM; así como nuevos objetivos que no eran objeto del mismo pero los cuáles han supuesto un desarrollo más extenso y enriquecedor del este.

- Estudiar los métodos NMF de separación de fuentes sonoras.
- Creación de un algoritmo de estimación del ritmo cardíaco.
- Creación de diccionarios de patrones espectrales de sonidos de corazón.
- Implementación en Matlab de un sistema NMF Semisupervisado.
- Implementación de distintas técnicas de Clustering que permiten explotar las características espectrales y temporales de los sonidos de corazón.
- Implementación en Matlab de un sistema NMF No Supervisado.
- Implementación de una funcionalidad de la búsqueda de patrones repetitivos tanto en señales de corazón como de pulmón.

CAPÍTULO 3 – OBJETIVOS

CAPÍTULO 4

4. Estado del arte

En este apartado se va a realizar una breve revisión de las diversas técnicas relacionadas con la separación de fuentes sonoras, así como de la diversidad técnicas tomadas para solventar la problemática que en este TFM se trata.

4.1. Estado del arte de la separación de sonidos de corazón y pulmón

En este apartado se realizará una breve descripción de los distintos métodos de separación de fuentes sonoras que podemos encontrar en la actualidad, ICA, SCA, NMF y PCA [4]. Cada uno de estos métodos se basa en explotar diferentes características de las señales como la independencia mutua, la no negatividad, dispersión, suavidad o combinaciones de las mismas para obtener las fuentes y reducir la influencia del ruido e interferencias.

a) El Análisis de Componentes Independientes (ICA)

Las técnicas ICA se basan en la independencia estadística de las fuentes a extraer. Así, las filas de la matriz S que representan a cada una de las fuentes a extraer deberán ser estadísticamente independientes. Encontramos este método en trabajos como [17] y [18].

En [17] se desarrolla un algoritmo para la separación de sonidos mezcla monocanal aplicando la técnica ICA de análisis de componentes independientes a la representación espectro-temporal de la señal a analizar. Se aplica la técnica ICA a espacios temporales contiguos del módulo de la Transformada de Fourier de tiempo corto o STFT de las señales a descomponer. Así, en primer lugar, se obtiene una envolvente lo suficientemente diferente para cada una de las fuentes sonoras, recuperando un espectro de potencia independiente para cada una de ellas. A

continuación, se hace uso de un sistema de puntuación basado en el análisis de señales auditivas para la asignación de cada uno de los espectros independientes a la fuente correcta. Después, se aplica una etapa de filtrado adaptativo a cada uno de los espectros para forzar su independencia mutua y, finalmente, se reconstruyen las fuentes haciendo uso de la Transformada inversa de Fourier de tiempo corto o iSTFT y el método de solapamiento y suma.

Además en [18], se desarrolla un algoritmo para la extracción de sonidos de tambor de pistas musicales polifónicas haciendo uso de las técnicas ICA. También se realiza una fase de particionado de las componentes obtenidas en subespacios para la clasificación de estas, en instrumentos percusivos y armónicos. En dicha fase de particionado se usan medidas como el cálculo de estadísticas de alto orden donde se mide la distancia entre componentes, el cálculo de Spectral Flatness donde se mide la tonalidad, el cálculo de la disonancia espectral o el cálculo de correlaciones haciendo uso de modelos de sonidos percusivos y ruido.

b) El Análisis de Componentes Dispersos (SCA)

En las técnicas SCA se considera que las fuentes son del tipo Sparse. Una señal es del tipo Sparse cuando la mayoría de sus muestras son cero o próximas a este valor, en cuyo caso se dice que son muestras inactivas y solo un porcentaje de las mismas toman valores significativos, en cuyo caso se dice que son muestras activas. Encontramos este tipo de técnicas en trabajos como [21] y [22].

En [21] se estudia la incorporación del análisis de componentes dispersos al algoritmo DUET de reconocimiento del habla en entornos con múltiples hablantes. A los componentes del habla extraídos mediante dicho algoritmo se añade la idea de la dispersión, de forma que la potencia dichos coeficientes debería tener un rápido decaimiento. Se demuestra en este trabajo que añadiendo el concepto de dispersión al habla, dichas componentes preservan la información necesaria para que la correcta transcripción del habla.

En [22] también se aborda el problema de separación de fuentes sonoras haciendo del análisis de componentes dispersas. En primer lugar, se realizan un conjunto de transformaciones ortogonales, por ejemplo la transformada Wavelet, en las cuales sólo unos pocos coeficientes resultantes tienen valores distintos de cero. Una vez se aplican dichas transformaciones a la señal a separar, se estima la matriz mezcla dentro de dicha transformada a partir de coeficientes de Clustering; después, la energía de dichas componentes se asocia a cada fuente estimada y, finalmente, la señal se reconstruye usando la transformada inversa.

c) *El Análisis de Componentes Principales (PCA)*

Las técnicas PCA consisten en un procedimiento matemático para la transformación de un número de posibles variables correladas en una pequeña cantidad de variables incorreladas llamadas ‘Componentes Principales’, teniendo estas en cuenta la variabilidad total de los datos en la mayor medida posible. Encontramos este tipo de análisis en trabajos como [23] y [24].

En [23] se propone un sistema no supervisado donde se aplica el análisis robusto de componentes principales, RPCA, donde se plantea un problema de optimización convexo. Dicho análisis se usará para la separación de la voz cantada del acompañamiento instrumental en grabaciones monocanal. En primer lugar, se calcula el espectrograma de dicha grabación mediante el cálculo de la STFT. Después, se usa el algoritmo ALM para la resolución de dicho problema, obteniéndose una matriz de bajo nivel, donde una fracción de sus componentes ha sido deteriorada por errores, correspondiente al acompañamiento instrumental y una matriz de dispersión correspondiente a la voz cantada. Finalmente, ambas señales se recuperan mediante el uso de máscaras espectro-temporales y la transformada inversa iSTFT.

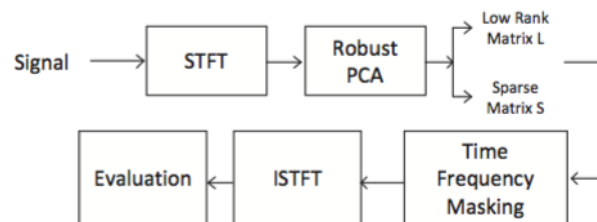


Figura 8 – Sistema basado en PCA

En [24], también se usa el análisis RPCA; sin embargo, este trabajo difiere del anterior en la metodología utilizada para la resolución del problema de optimización que plantea PCA. En este caso, se ha resuelto haciendo uso de la divergencia Itakura-Saito y estableciendo la restricción de no negatividad a las matrices de bajo rango y dispersión en el cálculo de dicha divergencia.

d) La Factorización No Negativa de Matrices (NMF)

Las técnicas NMF se basan en la no negatividad de las muestras de las fuentes a extraer. En el mundo real existen infinidad de escenarios en los que los datos son no negativos y sus componentes subyacentes solo tienen sentido físico cuando se da esta no negatividad. En este método, la matriz de entrada se descompone como el producto de una matriz de patrones espectrales (W) y una matriz de activaciones temporales (H), dependiendo de que matrices se actualicen durante la factorización, encontramos tres tipos de sistemas: Sistemas supervisados, semisupervisados o no supervisados. Así, nuestro trabajo se basará en el desarrollo de un sistema semisupervisado y un sistema no supervisado. Además, encontramos implementaciones de este tipo de sistemas en [19], [20].

En [19] se implementa un sistema NMF Semisupervisado donde se utiliza información extraída de la representación espectro-temporal de determinadas fuentes sonoras para la separación de fuentes en señales monocanal.

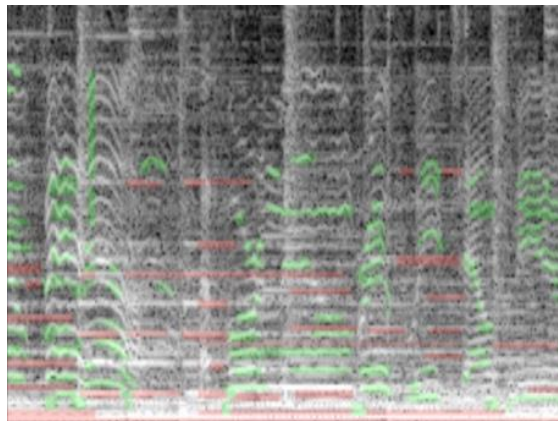


Figura 9 – Sistema basado en NMF 1

Con dicho propósito, se desarrolla una herramienta gráfica que permite al usuario extraer anotaciones para una determinada fuente, en este caso voz o instrumentos, de su representación espectro-temporal calculada en forma de espectrograma. Dichas anotaciones se usan posteriormente en la fase de separación para la extracción de fuentes.

Por otro lado en [20], también se desarrolla un sistema NMF Semisupervisado donde se toman restricciones de Sparse, cuyo objetivo es la reducción de ruido procedente del viento en señales habladas monocanal de forma que la extracción de la voz sea independiente del hablante. Para ello, se estima un diccionario de patrones espectrales de sonidos procedentes del viento; esto conlleva, a que la matriz de patrones espectrales correspondientes al viento se fija durante la separación y se actualizan el resto de matrices durante dicho proceso.

A continuación, se va a realizar una breve revisión de las diversas técnicas tomadas ante la problemática que en este TFM se trata sobre la separación entre sonidos de corazón y pulmón, entre las que se encuentran el uso de técnicas de filtrado adaptativo [5] eliminación de ruido basadas en metodología Wavelet [6] y [25], filtrado de modulación [7] y [25] o Factorización de matrices no negativas (NMF) en [26] y [27].

En [5] se usa metodología de filtrado adaptativo para minimizar la interferencia entre sonidos de corazón y sonidos pulmonares. El filtrado adaptativo, se basa en la adaptación de los parámetros de filtrado usados para este de la señal actual en el sistema, teniendo en cuenta los parámetros utilizados para la entrada anterior a este. En [5], se utiliza el siguiente esquema:

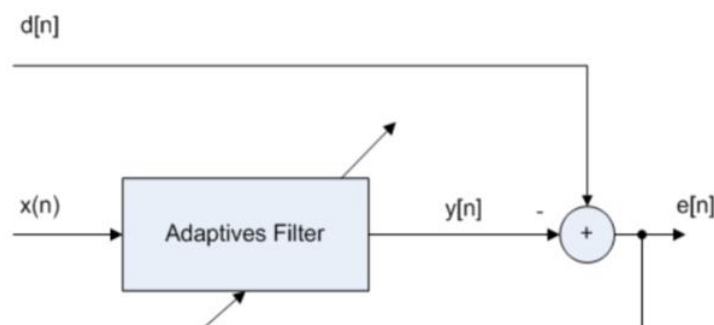


Figura 10 – Sistema basado en filtrado adaptativo

En este sistema, se considera como entradas principales $d[n]$, la señal de pulmón contaminada por sonidos de corazón, y $x[n]$, la señal de corazón tomada como referencia para la separación. Se considera además que la señal de entrada $d[n]$ se define como $d[n] = s[n] + x1[n]$, siendo $x1[n]$ el sonido de corazón que interfiere al sonido de pulmón y $s[n]$ el sonido de pulmón en el instante actual, el cual se quiere extraer. De esta forma, la señal $x1[n]$ de sonido de corazón interferente estará incorrelada con la señal de entrada $d[n]$, la cual contiene sonidos de pulmón, y mantendrá un cierto nivel de correlación con la señal de referencia $x[n]$, de sonido de corazón. Finalmente, se calcula la salida adaptativa del filtro $y[n]$ y el error, $e[n]$, que será la salida del sistema.

La señal de referencia $x[n]$ es filtrada de modo que se genera la salida $y[n]$ obteniendo una señal con el mayor nivel de correlación posible con $x1[n]$. Finalmente, se calcula la salida del sistema $e[n]$ como $e[n]=d[n]-y[n]$, donde esta señal de error se utilizará posteriormente para adaptar la respuesta al impulso del filtro para su estado siguiente. Por último, los siguientes algoritmos se han usado para la implementación del filtro adaptado: LMS, NLMS y RLS.

En [6] se utiliza una metodología de eliminación de ruido mediante técnicas Wavelet. En primer lugar, se encuentra una fase de adquisición de sonidos de pulmón haciendo uso de un acelerómetro de contacto piezoeléctrico en el tercer espacio intercostal.

A continuación, se encuentra una fase de detección de sonidos de corazón para la cual se hace uso de un análisis multiresolución Wavelet, detectándose los sonidos de corazón mediante el producto multiescala de los coeficientes aproximados Wavelet. Tres escalas se usan en la descomposición Wavelet, tomando como Wavelet madre la Symlet Wavelet. De esta manera, los coeficientes de la Transformada Discreta Wavelet, DWT, permiten que las señales de corazón se localicen como singularidades dentro de la señal de pulmón.

Posteriormente, se procede a la cancelación de los sonidos de corazón, una vez localizados, mediante umbralización. Finalmente, se realiza una fase de predicción y modelado de la información eliminada, estimando mediante el modelo de predicción lineal

ARMA la información eliminada, con el fin de seleccionar la mejor configuración para el sistema.

En [7] se aplican filtros de modulación para la separación entre ambos tipos de sonidos. El filtrado de modulación se basa en el filtrado de las trayectorias temporales de tiempo corto de las componentes espectrales. Para ello, se han usado dos filtros de modulación con respuesta al impulso finita, un filtro paso banda y un filtro elimina banda, los cuales se muestran a continuación:

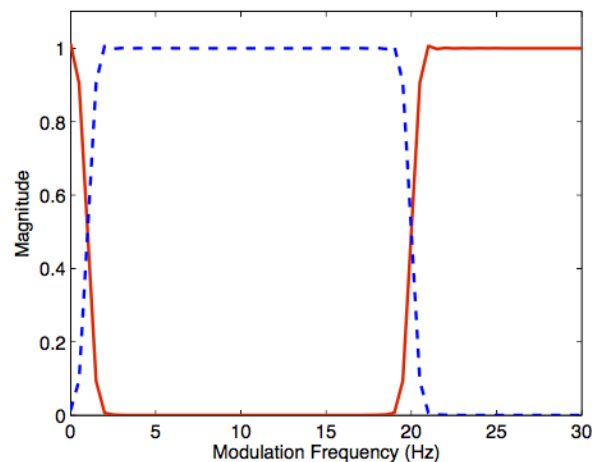


Figura 11 – Sistema basado en filtrado de modulación

En los experimentos realizados en [7], se ha utilizado la Transformada de Gabor. En primer lugar, se aplican técnicas de enventanado a la señal de pulmón interferida por sonidos de corazón. Posteriormente, se calcula la DFT de dicha señal enventanada $s(m)$, tomando por separado las componentes de magnitud y fase de cada ‘bin’ de frecuencia, siendo ambas entradas a una fase llamada “Procesado de modulación”, donde se aplican los filtros de modulación anteriormente expuestos y una compensación de retardo a la fase. Tras dicho proceso, se obtienen como salida las señales filtradas por los filtros de modulación y las componentes de fase con retardo. Finalmente, las señales de corazón y pulmón se obtienen por separado haciendo uso de la IDFT, técnicas de enventanado y el método de solapamiento y suma.

Por otro lado en [25], se comparan ambas técnicas, filtros de modulación y metodología Wavelet, para la separación entre ambos tipos de sonidos. Para el uso de los filtros de

modulación, en primer lugar se calcula la representación espectro-temporal de la señal de entrada, STFT. Después, se aplican filtros paso banda y elimina banda, cuyos resultados son los espectrogramas de sonidos de corazón y pulmón, respectivamente. Finalmente, se reconstruye la señal usando la iSTFT. Para el uso de metodología Wavelet, se aplican en primer lugar un filtro paso bajo y paso alto a la señal de entrada. Posteriormente, se establece un criterio por el cual se usa una familia de bases ortogonales para el cálculo de coeficientes y se aplica la Transformada Discreta Wavelet. Tras esto, los coeficientes, mediante umbralización, se clasifican en sonidos de corazón o pulmón, reiterándose tras esto en el proceso, hasta llegar a la condición establecida por el criterio de parada. Finalmente, se calculan los sonidos mediante el cálculo de la Transformada Inversa Discreta Wavelet de los coeficientes correspondiente a cada sonido.

En [26] se usan las técnicas de Factorización de matrices no negativa (NMF) para la separación entre ambos tipos de sonidos. En primer lugar, se propone un sistema NMF Semisupervisado donde se utiliza un algoritmo NMF modificado llamado NMF-SF o Factorización de matrices no negativas con factores compartidos; se busca con este método reforzar la bases espectrales de los sonidos principales, corazón y pulmón, durante el proceso de factorización. Se puede ver el modelo de separación utilizado en el proceso de factorización en la siguiente figura:

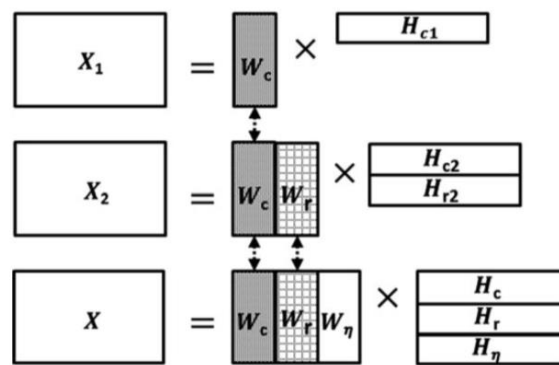


Figura 12 – Sistema basado en NMF 2

En segundo lugar, se presenta un sistema NMF No Supervisado, donde no se realiza ningún tipo de entrenamiento para la factorización y se acompaña de una serie de etapas de

Clustering para la clasificación de bases. En la primera fase, se particiona la señal de entrada de forma que se dividida en las señales correspondientes a un ciclo completo de señal de pulmón con la menor frecuencia cíclica. Se calculan las matrices de patrones espectrales y activaciones temporales de dichas señales usando el modelo estándar de NMF. En la segunda fase, se realiza un agrupamiento de dichas bases asignándose estas a cada una de las fuentes. Para ello, se utilizan medidas como el cálculo de la similitud entre bases y el cálculo de la correlación espectral donde se aprovecha el hecho de que por debajo de 100 Hz el corazón es dominante en la mezcla y por encima de 300 Hz el sonido dominante es el pulmón; seleccionándose una matriz W_i de referencia y calculando su correlación con el resto de matrices de patrones espectrales.

Por otro lado en [27], se usa un sistema NMF No Supervisado; sin embargo, se introduce el uso de la Transformada CQT o representación logarítmica del eje de frecuencias de la Transformada STFT. Dicha transformada se realiza como fase de preprocesado de la señal antes de la etapa de factorización NMF; con ello se busca reducir el rango completo de datos de la señal de entrada a factorizar, centrándose en los datos de baja frecuencia donde se encuentra la mayor parte de potencia de sonidos de corazón y pulmón, reduciendo así el tiempo de procesamiento del sistema.

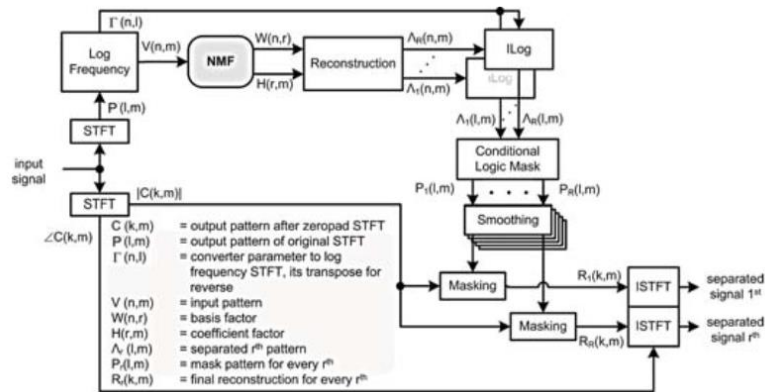


Figura 13 – Sistema basado en NMF 3

En la anterior figura, se muestra el diagrama de bloques del sistema propuesto en [27]. En primer lugar, se calcula la Transformada STFT de la señal de entrada añadiendo ceros a la señal de entrada obteniéndose la señal C y sin añadir ceros obteniéndose la señal P . Se aplica la Transformada CQT a la señal P , obteniéndose la señal V , reduciéndose el rango de datos de

dicha señal que serán factorizados con NMF. Una vez se realiza la factorización NMF para un número de R componentes, dos en este caso, se lleva a cabo una fase de reconstrucción donde se obtiene una señal Λ_r para componente. A continuación, se invierte la Transformada CQT para cada señal Λ_r , recuperando la representación lineal en tiempo-frecuencia de la Transformada STFT. Posteriormente, se aplican máscaras lógicas condicionales para filtrar cada ‘frame’ temporal y ‘bin’ de frecuencia obteniéndose $\mathbf{P}_r$ y se lleva a cabo una fase de suavizado columna por columna. Finalmente, se aplica la Transformada de Fourier a dichas señales $\mathbf{P}_r$ hasta que se alcanzan, en ellas, las dimensiones de la señal $\mathbf{C}$. Así se enmascaran las señales $\mathbf{P}_r$ a través del producto Hadamard de dichas señales con la señal $\mathbf{C}$ y se aplica la Transformada inversa iSTFT para obtener las señales de corazón y pulmón separadas.

4.2. Introducción a Non-Negative Matrix Factorization (NMF)

Dado que nuestro trabajo se basa en el uso de técnicas de Factorización de matrices no negativas NMF, se realizará un estudio de las mismas con objeto de introducir al lector en este tipo de técnicas. En primer lugar, se explicarán los principios generales que rigen estas. Después, se determinarán las ecuaciones necesarias para nuestro Trabajo Fin de Máster. A continuación, se introducirá al concepto de Sparse. Finalmente, se describirán los distintos criterios de parada usados en este tipo de algoritmos.

4.2.1. Principios generales

La Factorización de matrices no negativas o NMF es una técnica de factorización basada en la no negatividad de los datos a extraer [28]. Así, dada una matriz de entrada V de dimensiones $F \times N$ con datos no negativos, esta se obtiene como:

$$\mathbf{V} \approx \mathbf{W} * \mathbf{H} \quad (7)$$

De este modo, la matriz de entrada V se factoriza de forma aproximada como la multiplicación de dos matrices no negativas W y H con dimensiones $F \times K$ y $K \times N$, respectivamente, siendo K el número de bases usadas en el proceso de factorización.

Por otro lado, la matriz W contiene la información espectral para cada base K , mientras que la matriz H contiene las activaciones temporales de estas. Dado que la factorización realizada es un proceso aproximado, nos encontramos ante un problema de minimización del error en dicho proceso de factorización.

Dicho problema de minimización se representa matemáticamente [28] como:

$$\min_{W, H \geq 0} D(V|WH) \quad (8)$$

Donde $D(V|WH)$ es una función de coste definida [28] como:

$$D(V|WH) = \sum_{f=1}^F \sum_{n=1}^N d([V]_{fn}|[WH]_{fn}) \quad (9)$$

siendo $d(x|y)$ una función de coste escalar. Entre las funciones de coste escalares más populares para la separación de fuentes sonoras se encuentran:

- La distancia Euclídea definida [28] como:

$$d_{EU}(x|y) = \frac{1}{2} (x - y)^2 \quad (10)$$

- La divergencia Kullback-Leibler (KL) definida [28] como:

$$d_{KL}(x|y) = x \log \frac{x}{y} - x + y \quad (11)$$

- La divergencia Itakura Saito (IS) definida [28] como:

$$d_{IS}(x|y) = \frac{x}{y} - \log \frac{x}{y} - 1 \quad (12)$$

La distancia Euclidea, la divergencia KL y la divergencia IS [28] son funciones positivas y toman el valor cero únicamente cuando $x=y$, siendo las dos primeras funciones convexas entre $(0, \infty)$ y siendo la divergencia IS convexa en $(0, 2x]$ y cóncava en $[2x, \infty)$.

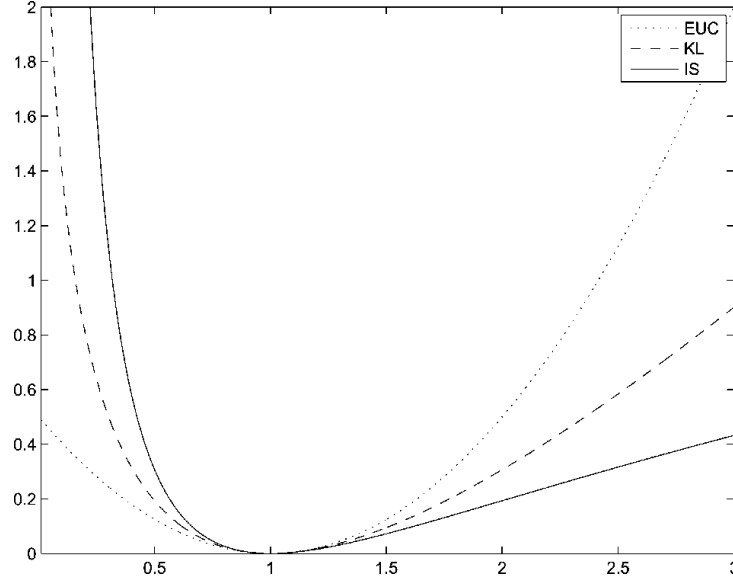


Figura 14 – Forma de las distintas ecuaciones de coste [28]

Para la minimización de dicha funciones de coste, se utiliza técnica del gradiente descendente a partir de la cual se calculan las reglas de actualización multiplicativas. Estas reglas garantizan la no negatividad de dichas funciones de coste mientras su valor no aumenta. Además, permiten encontrar un mínimo local, punto estacionario, dentro de dichas funciones que permita la convergencia del algoritmo tras un número ‘i’ de iteraciones.

Todas estas funciones de coste pertenecen a la familia de las funciones β -divergencia, definidas [28] como:

$$d_{\beta}(x|y) = \begin{cases} \frac{1}{\beta(\beta-1)}(x^{\beta} + (\beta-1)y^{\beta} - \beta xy^{\beta-1}) & \beta \in \mathbb{R} \setminus \{0,1\} \\ x \log \frac{x}{y} - x + y & \beta = 1 \\ \frac{x}{y} - \log \frac{x}{y} - 1 & \beta = 0 \end{cases} \quad (13)$$

Donde se encuentra que:

- La distancia Euclidea se corresponde con $\beta=2$
- La divergencia KL se corresponde con $\beta=1$
- La divergencia IS se corresponde con $\beta=0$.

Además, las reglas de actualización multiplicativas para esta familia de funciones, se determinan en función de β como:

$$H \leftarrow H \cdot \frac{W^T ((WH)^{[\beta-2]} \cdot V)}{W^T (WH)^{[\beta-1]}} \quad (14)$$

$$W \leftarrow W \cdot \frac{((WH)^{[\beta-2]} V) H^T}{(WH)^{[\beta-1]} H^T} \quad (15)$$

Por último, atendiendo a si actualizamos sólo las matrices W de patrones espectrales, sólo si actualizamos las matrices de activaciones temporales H o actualizamos ambas, se puede distinguir entre tres tipos de sistemas NMF:

- **Sistema supervisado:** Son aquellos sistemas donde la fase de separación de fuentes viene precedida por una fase de entrenamiento en la cual se entrenan los patrones espectrales de cada una de las fuentes sonoras que componen la señal mezcla, calculándose así sólo las matrices de activaciones temporales en la fase de separación.
- **Sistema semisupervisado:** Son aquellos sistemas donde la fase de separación de fuentes viene precedida por una fase de entrenamiento en la cual se entrenan los patrones espectrales únicamente de una de las fuentes sonoras que componen la señal mezcla, calculándose así las matrices de patrones espectrales del resto de fuentes durante la fase de separación así como las matrices de activaciones temporales para todos las fuentes que componen la mezcla.

- **Sistema no supervisado:** Son aquellos sistemas donde únicamente se realiza una fase de separación de fuentes sin calcular previamente ningún patrón espectral, lo que supone la imposición de una serie de restricciones para modelar determinados comportamientos de las fuentes a extraer, calculándose así en dicha fase las matrices de patrones espectrales y matrices de activaciones temporales para cada una de las fuentes de los sonidos mezcla.

4.2.2. Sparse NMF

Asumiendo que las fuentes a separar son dispersas, puede considerarse añadir a dichas funciones de coste un término de penalización Sparse a la matriz H de activaciones temporales con el fin de realizar la factorización de la matriz de entrada con el menor número de datos posible. De esta forma, el valor de la función de coste se incrementa para entradas de H distintas de cero, penalizándose así matrices densas [29].

La función de coste resultante usando la norma l_1 será:

$$D(V|WH) = \|d(V|WH)\|_F + \lambda \|H\|_1 \quad (16)$$

siendo $d(V|WH)$ cualquier de las funciones de coste escalares anteriores y λ un parámetro de medida del nivel de dispersión comprendido entre [0,1]. Además, la inclusión de este término de penalización resulta en la modificación de las reglas de actualización multiplicativas:

$$H \leftarrow H \cdot \frac{W^T((WH)^{[\beta-2]} \cdot V)}{W^T(WH)^{[\beta-1]} + \lambda} \quad (17)$$

$$W \leftarrow W \cdot \frac{((WH)^{[\beta-2]} \cdot V)H^T}{(WH)^{[\beta-1]}H^T} \quad (18)$$

A continuación, se muestran dos figuras donde se observan las matrices de activaciones temporales H utilizando y sin utilizar la restricción Sparse:

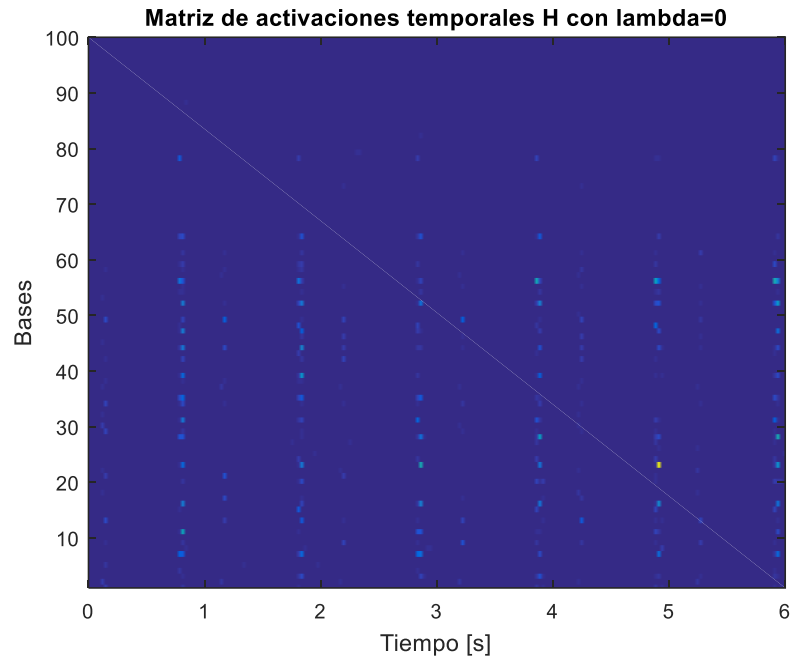


Figura 15 – Matriz de activaciones H con λ cero

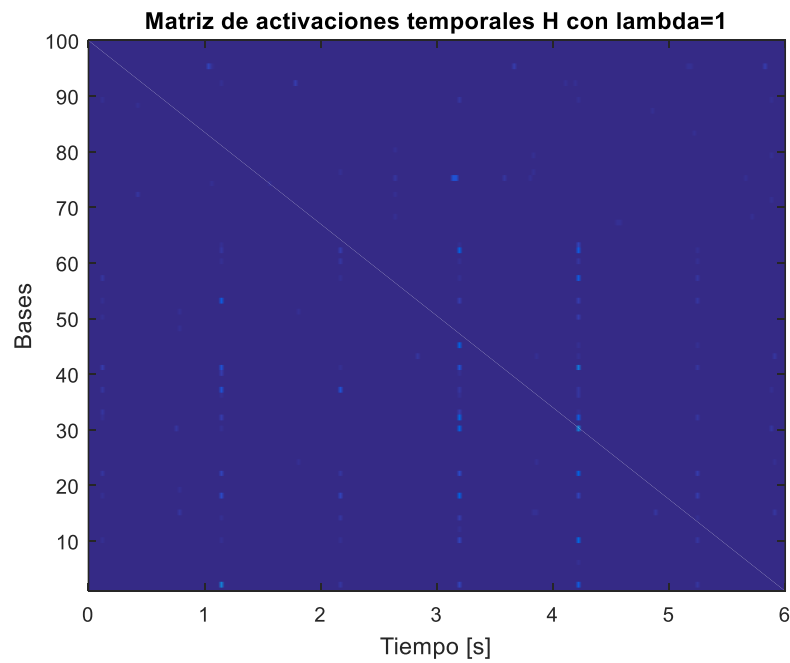


Figura 16 – Matriz de activaciones H con λ uno

Analizando ambas figuras, se observa que utilizándose la restricción Sparse se obtienen matrices temporales H con menor número de activaciones; Además, esta restricción permite la convergencia del algoritmo con un menor número de iteraciones.

4.2.3. Criterio de parada

Dado que el método del gradiente descendente proporciona un método de minimización de la función de coste basado en reglas iterativas de actualización, se vuelve lógico considerar un criterio de parada mediante el cual se considere que el algoritmo ha llevado a un punto estacionario, mínimo local, dentro de la función de coste. Así, como vemos en [4], existen múltiples criterios de parada para dicho fin:

- En el más simple, se considera que el algoritmo ha convergido tras llegar a un número máximo de iteraciones.
- La función de coste alcanza el valor cero o un valor por debajo de un umbral establecido ϵ , de esta forma:

$$D^i(V|(WH)^i) = \|V - (WH)^i\|_F^2 < \epsilon \quad (19)$$

- Hay muy poca mejora o ninguna entre sucesivas iteraciones en la minimización de la función de coste, de esta forma:

$$D^{i+1}((WH)^{i+1}|(WH)^i) = \|(WH)^i - (WH)^{i+1}\|_F^2 < \epsilon \quad (20)$$

Expresándose también como:

$$\frac{D^i(V|WH) - D^{i-1}(V|WH)}{D^i(V|WH)} < \epsilon \quad (21)$$

CAPÍTULO 5

5. Desarrollo del trabajo fin de Máster

En los sucesivos apartados se describirá el trabajo realizado en este Trabajo Fin de Máster que ha llevado a solventar la problemática que en él se expone. Este Trabajo Fin de Máster se distribuye como sigue: En primer lugar, se realizará una caracterización espectral de sonidos de corazón y pulmón, donde se obtendrán una serie de suposiciones utilizadas posteriormente. Después, se realizará un estudio de la representación adoptada para las señales mezcla a separar. A continuación, se describirán los sistemas de separación NMF implementados en este TFM, así como nuestras aportaciones a dichas técnicas de separación. También, se describirán las técnicas adoptadas para la reconstrucción de las señales obtenidas. Finalmente, se expondrá la metodología utilizada para la búsqueda de patrones repetitivos en ambas señales, la herramienta gráfica implementada, el proceso de evaluación de ambos sistemas y los resultados obtenidos por ellos.

5.1. Caracterización espectral de sonidos de corazón y pulmón

En este apartado, se ha realizado un estudio atendiendo a la caracterización espectral de los sonidos de corazón, considerados como normales, S1 y S2. Atendiendo a la información recogida en [3] y [9], la potencia de un sonido S1 se concentra en la banda comprendida mayormente en el rango de frecuencia 10 – 130 Hz, pudiendo aparecer componentes de sonidos S1 hasta los 300 Hz, y la potencia de un sonido S2 se concentra en la banda 24 – 144 Hz, pudiendo aparecer componentes de S2 hasta los 320Hz. Se ha calculado la densidad espectral de potencia (PSD) de dichas señales haciendo uso del programa de cálculo matemático MATLAB:

- Densidad espectral de potencia de sonido S1 [49]:

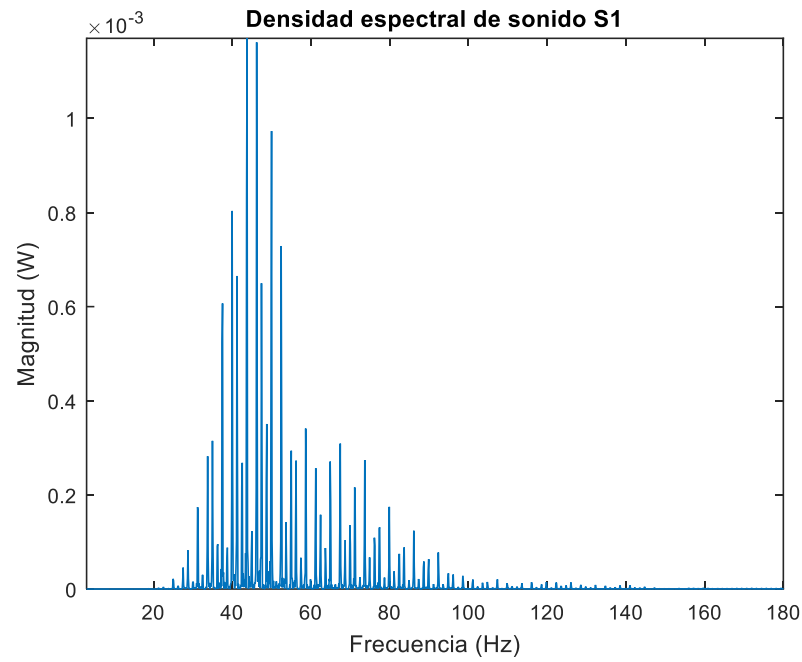


Figura 17 – DEP de sonido de corazón S1

- Densidad espectral de potencia de sonido S2 [49]:

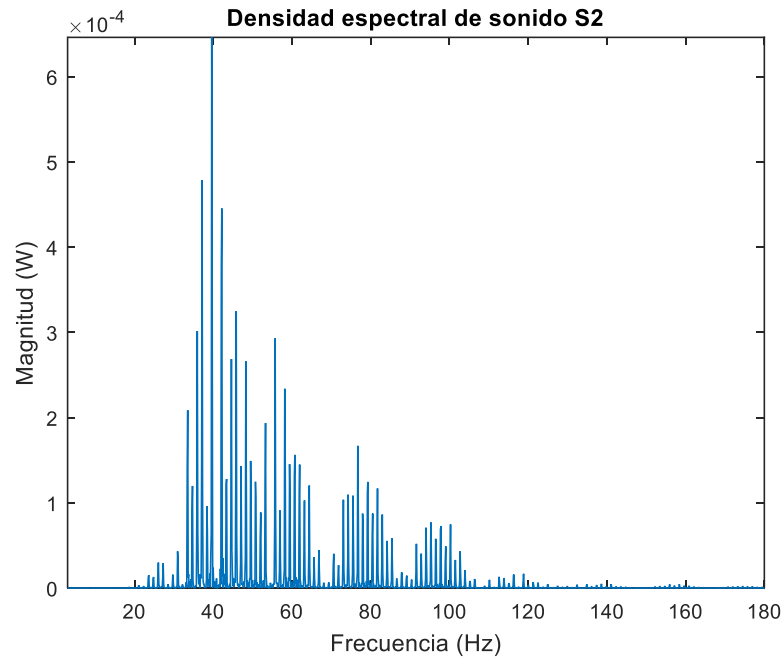


Figura 18 – DEP de sonido de corazón S2

Así mismo, también se ha caracterizado los sonidos considerados como anómalos S3 y S4. El resultado obtenido para estos es:

- Densidad espectral de potencia de sonido S3 [49]:

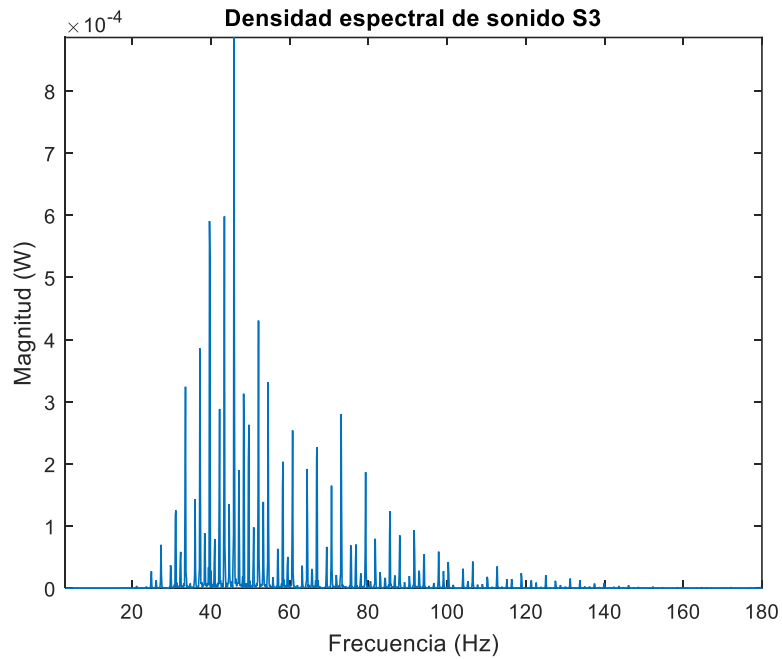


Figura 19 – DEP de sonido de corazón S3

- Densidad espectral de potencia de sonido S4 [49]:

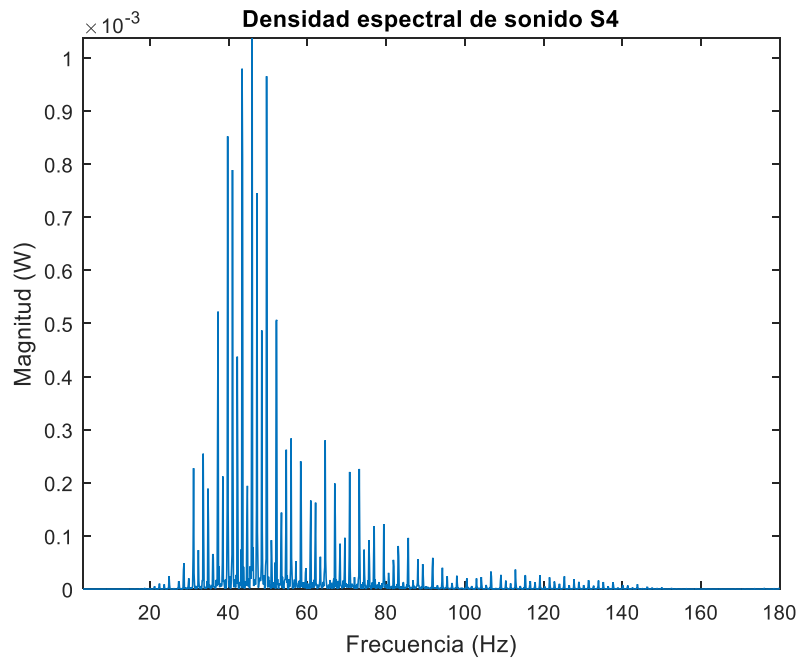


Figura 20 – DEP de sonido de corazón S4

A partir de los resultados obtenidos y los datos en [3] y [9], se asumen los sonidos de corazón como sonidos de banda estrecha con un ancho de banda de aproximadamente 300Hz, cuyas componentes aparecen en la banda 10-320Hz.

Tras el análisis de los sonidos de corazón, se ha realizado una caracterización espectral de los sonidos de pulmón. Atendiendo a la información recogida en [3] y [12], el ancho de banda espectral de dichos sonidos se extiende hasta 1Khz, concentrándose la mayor parte de la potencia entre los 60 – 600Hz.

- Densidad espectral de potencia de murmullo pulmonar [49]:

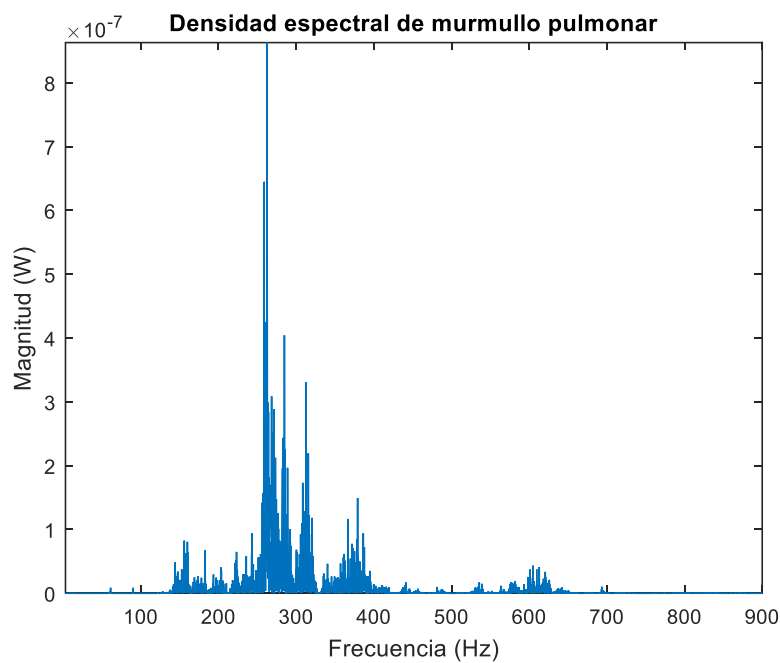


Figura 21 – DEP de murmullo pulmonar

- Densidad espectral de potencia de sonido Broncovesicular [49]:

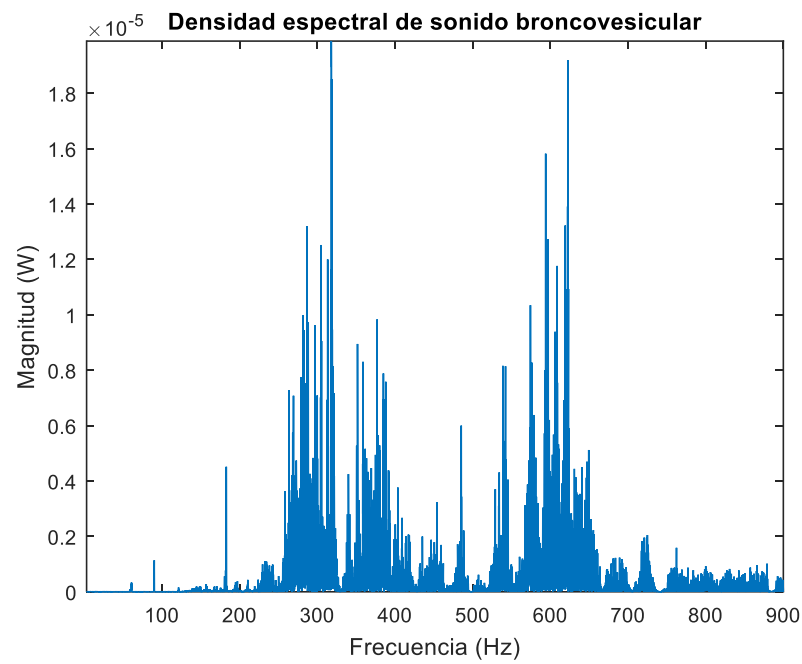


Figura 22 – DEP de sonido broncovesicular

- Densidad espectral de potencia de Roncus [49]:

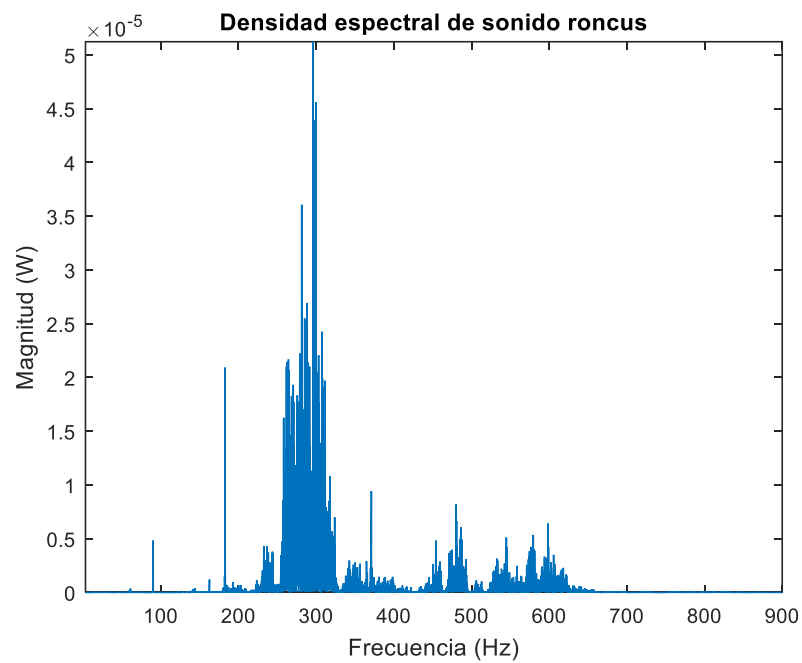


Figura 23 – DEP de sonido Roncus

- Densidad espectral de potencia de sibilancia monofónica [49]:

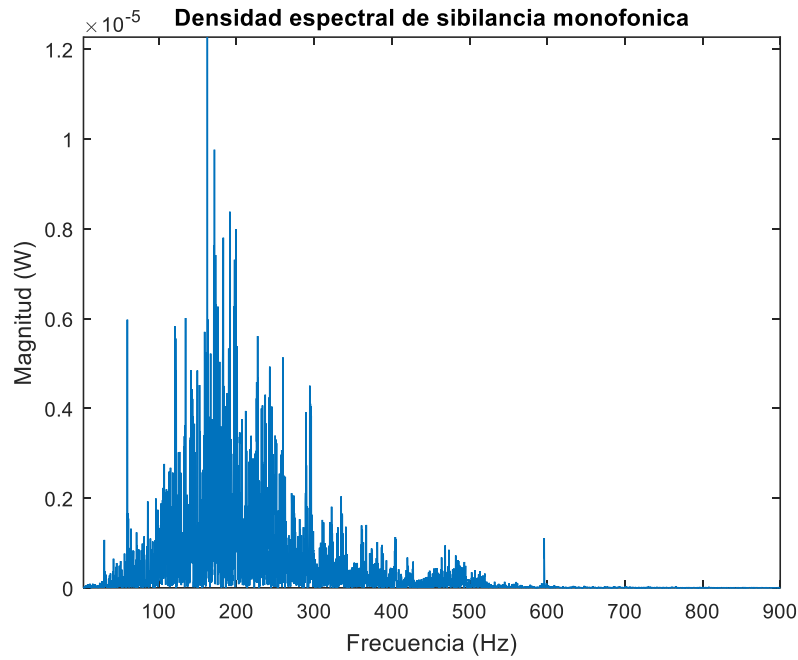


Figura 24 – DEP de sibilancia monofónica

En este caso, se asume que los sonidos pulmonares son sonidos de banda ancha con un ancho de banda de aproximadamente 1KHz. Así mismo, se corrobora el aparecimiento de una zona de solapamiento espectral [3] y [10] aproximadamente en la banda 60-320 Hz cuando ambos sonidos se interfieren mutuamente, como puede verse en la siguiente figura:

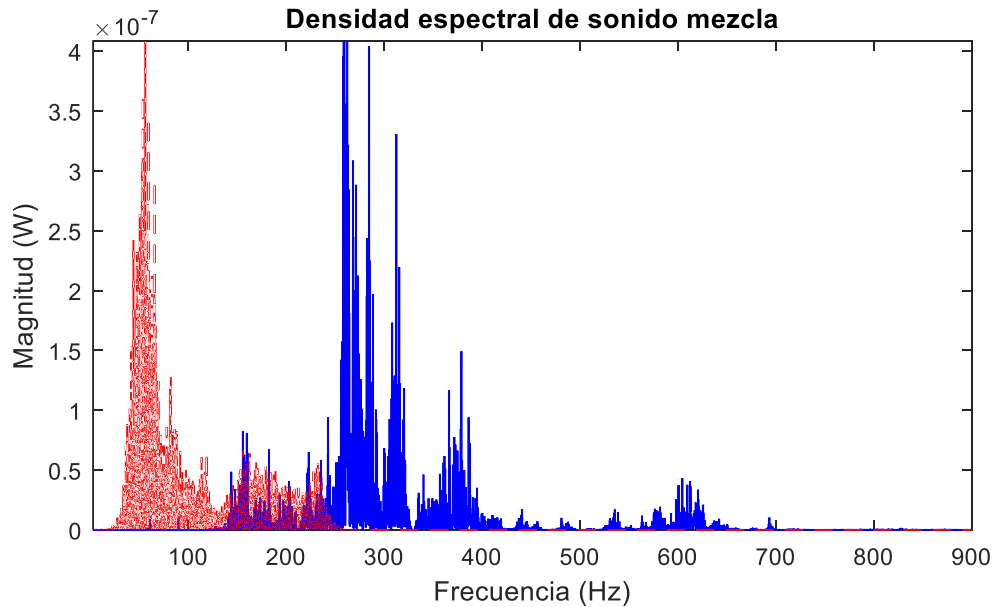


Figura 25 – DEP de sonido mezcla de corazón y pulmón

5.2. Metodología propuesta

Como se ha mencionado en apartados anteriores, se ha implementado para este Trabajo Fin de Máster dos sistemas diferentes de separación NMF para señales mezcla monocanal; por un lado, se ha desarrollado un sistema Semisupervisado; por otro, un sistema No Supervisado con dos posibles métodos de realización del proceso de factorización.

En primer lugar, se ha creado un algoritmo de estimación del ritmo cardiaco que permite la localización temporal de sonidos de corazón, generando así un patrón de sonidos de corazón adaptado a cada señal mezcla. Dicho patrón será usado durante la fase de entrenamiento del sistema Semisupervisado y durante la fase de Clustering del sistema No Supervisado. Asimismo, la generación de dicho patrón, permite la introducción de un segundo proceso de factorización en el sistema NMF no supervisado, basado en la factorización en bloques.

En segundo lugar, dado lo expuesto en el apartado 6.1, se observa como los sonidos de corazón y sonidos de pulmón presentan perfiles espectrales diferentes, siendo los primeros sonidos de banda estrecha y los segundos de banda ancha. Así, en el apartado 6.1, se demuestra la justificación para la implementación de un sistema Semisupervisado [32], donde se pone de manifiesto que dichos sistemas se comportan de forma adecuada cuando los perfiles espectrales de las fuentes a separar son lo suficientemente diferentes, siendo este nuestro caso. Posteriormente, dichos patrones espectrales son utilizados para la descomposición, utilizando un modelo de descomposición específico para esta, donde se distingue entre sonidos de corazón y pulmón.

En tercer lugar, se ha implementado un sistema No Supervisado utilizando el modelo de descomposición estándar de NMF. En este sistema, no se realiza ningún tipo de distinción entre ambos sonidos, por lo que se vuelve necesaria una fase posterior de Clustering o Clasificación de bases que permita su clasificación en sonidos de pulmón o sonidos de corazón. Para dicha fase, se han realizado una serie de aportaciones novedosas en cuanto a la utilización de técnicas de clasificación se refiere, utilizándose técnicas no vistas con anterioridad en otros trabajos semejantes. Dichas técnicas son el cálculo de la medida Roll-off y el cálculo de la correlación

de patrones espectrales y de activaciones temporales para cada una de las bases. En este sistema el patrón rítmico de sonidos de corazón es utilizado con dos finalidades; en primer lugar, se utiliza para el cálculo de la correlación de activaciones temporales; por otro lado, permite introducir un segundo proceso de factorización donde el espectrograma de entrada de la señal mezcla se divide en bloques sucesivos que son factorizados independientemente, encontrándose dos tipos de bloques: aquellos correspondientes a sonidos de corazón, donde pueden también encontrarse sonidos de pulmón, y aquellos correspondientes únicamente a sonidos de pulmón. De este modo, aquellos bloques correspondientes a sonidos de corazón son factorizados usando un número K_c de bases adaptado a dichos sonidos con el fin de obtener mejores resultados ya que NMF no establece un número de bases óptimo para el proceso de factorización.

Una vez ambos sistemas han sido implementados, se ha desarrollado un algoritmo de búsqueda de patrones repetitivos con un cierto porcentaje de similitud en las señales de corazón y pulmón obtenidas de dichos sistemas. Finalmente, se ha implementado una herramienta gráfica sencilla donde se integran los sistemas anteriormente expuestos y se permite la libre configuración de parámetros al usuario, así como la posibilidad de utilizar configuraciones predefinidas.

5.3. Algoritmo de estimación del ritmo cardíaco

Para la realización de dicho algoritmo, se parte de lo expuesto en la sección 6.1, donde se pone de manifiesto la existencia de una zona de solapamiento espectral entre los sonidos de corazón y pulmón en la banda 60-320 Hz. También se expone que la mayor parte de la energía de los sonidos de corazón se concentra en la banda 10-144Hz. Dado esto, se asume que en la banda comprendida entre los 10-60 Hz, donde no se produce solapamiento entre sonidos de corazón y pulmón, encontrándose sólo sonidos de corazón, habrá suficiente energía de estos.

De esta forma, el algoritmo de estimación del ritmo cardíaco se basa en los cambios de energía en dicha banda de frecuencia del espectrograma de la señal mezcla entre espacios temporales sucesivos. Para el cálculo de dichos cambios de energía, se utiliza una función Onset basada en el cálculo del flujo espectral o Spectral Flux [36].

En primer lugar, se calcula el espectrograma de la señal mezcla. Para ello, se calcula la Transformada STFT utilizando una longitud de ventana M específica para maximizar la eficiencia en la detección, en [35]:

$$M = \frac{fs}{8} \quad (22)$$

Una vez se calcula la transformada STFT, la matriz donde se realizará la búsqueda de sonidos de corazón se define como:

$$S = X_{STFT}(1:F_{60Hz}, 1:T) \quad (23)$$

En segundo lugar, se calcula el flujo espectral de dicha matriz S matemáticamente como, [36]:

$$H(i) = \sum_{f=1}^{F_{60Hz}} |S(f, i)| - |S(f, i - 1)| \quad (24)$$

donde $i = 2 \dots Ml$, siendo Ml el número de ventanas temporales usadas para el cálculo de la STFT, y f el bin de frecuencia actual.

Posteriormente, R se calcula aplicando a H un rectificador de media onda, de forma que se obtenga valores iguales o mayores que cero.

$$R(i) = \frac{H(i) + |H(i)|}{2} \quad (25)$$

A continuación, se muestra un ejemplo del resultado obtenido tras realizar dicha cálculo:

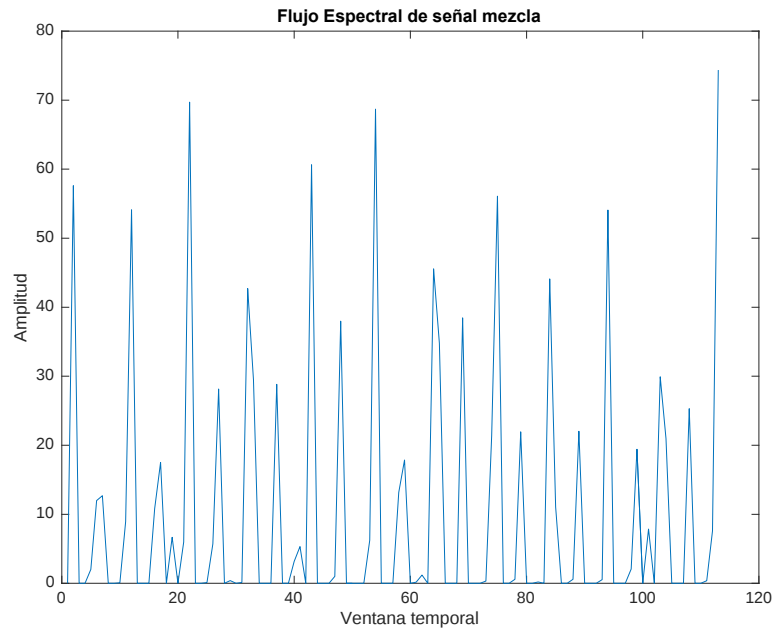


Figura 26 – Flujo Espectral de señal mezcla

Así, cada uno de los picos de R se corresponderá con un posible sonido de corazón. Para localizar la posición de dichos picos se utiliza el algoritmo de detección de picos multiescala desarrollado [37], conocido como AMPD. Dicho algoritmo se divide en las siguientes etapas:

Primero, se calcula el escalograma de máximos locales M . Para ello, se aplica a la señal de flujo espectral R un análisis de regresión lineal, con el fin de calcular la línea que mejor se adapta a la tendencia de los datos de dicha señal, eliminando posteriormente dicha tendencia, de forma que el valor de todos los datos varíe de la misma forma.

Teniendo la matriz M dimensiones $L \times N$ y siendo N la longitud de la señal y calculándose como $L = \left\lfloor \frac{N}{2} \right\rfloor - 1$, en cada posición de dicha matriz se tiene un valor $r + 1$, siendo r un valor aleatorio de una distribución uniforme entre 0 y 1. Después, se calculan los máximos locales para una determinada escala k utilizando una ventana con longitud variable, de forma que su longitud para dicha escala k se calcula como:

$$W_k = 2k, \quad k = 1, 2, \dots, L \quad (26)$$

Recorriéndose cada escala k , desde $i = k + 2, \dots, N - k - 1$, asignándose a cada $m_{k,i}$ de la matriz M los siguientes valores:

$$m_{k,i} = \begin{cases} 0 & R_{i-1} > R_{i-k-1} \wedge R_{i-1} > R_{i+k-1} \\ \text{No se modifica en otro caso} & \end{cases} \quad (27)$$

Por tanto, se asigna para cada escala k un valor 0 en aquellas posiciones donde se encuentre un máximo local, manteniéndose el resto. Se tiene así una matriz con valores en el rango $[0, 1 + r]$.

En segundo lugar, se calcula el vector φ como la suma del valor de los elementos de cada fila o escala k de la matriz M :

$$\varphi_k = \sum_{i=1}^N m_{k,i} \quad (28)$$

Obteniéndose $\varphi = [\varphi_1, \varphi_2, \dots, \varphi_L]$. Así el mínimo global de φ , $\lambda = \operatorname{argmin}(\varphi)$, se corresponde con la escala con mayor número de máximo locales localizados.

En tercer lugar, se genera una nueva matriz M_r donde se copian los valores de la anterior matriz M desde $k = 1 \dots \lambda$ e $i = 1 \dots N$.

En cuarto lugar, para detectar la posición de los picos de la señal R , se calcula la desviación estándar de cada una de las columnas de la matriz M_r . Finalmente, las posiciones i donde se obtenga una desviación estándar igual a cero, $\sigma_i = 0$, se corresponderán con las posiciones de los picos de nuestra señal R .

$$posiciones = i|_{\sigma_i=0} \quad (29)$$

Una vez obtenidas las posiciones de los picos de la señal de flujo espectral R , el instante temporal donde se localizan los sonidos de corazón se calcula como el valor central de la ventana temporal i , es decir:

$$times = T_{posiciones} \quad (30)$$

Dado que en ciertas ocasiones, el primer pico no es detectado por el algoritmo, se aplica una corrección a partir del cálculo del periodo cardíaco estimado de corazón. Dicho periodo se calcula a partir de la autocorrelación de la señal R , correspondiéndose dicho periodo al instante temporal donde se encuentra el siguiente máximo sucesivo al máximo principal de dicha función de autocorrelación. Dichos picos se calculan utilizando el algoritmo AMPD anterior expuesto.

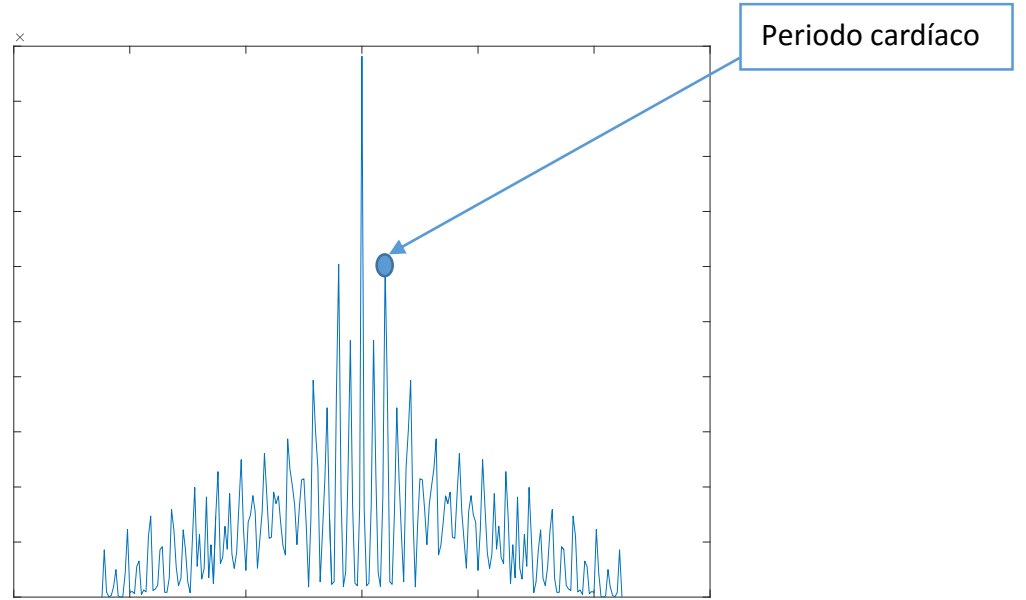


Figura 27 – Función de autocorrelación del flujo espectral R

Una vez calculado dicho periodo, la corrección se realiza como:

$$times_{primerpico} \begin{cases} times(1) - periodo, & \text{si } times(1) - periodo > 0 \\ times(2) - periodo, & \text{si } times(2) - periodo > 0 \\ \text{no se añade en cualquier otro caso} \end{cases} \quad (31)$$

Así se calculan todas las posiciones temporales como:

$$times = [times_{primerpico}, times] \quad (32)$$

Finalmente, el patrón rítmico de sonidos de corazón se calcula en función de la longitud de la ventana temporal y el instante central de esta utilizados en el sistema NMF. Para ello, en el algoritmo se considera dicha longitud N_{NMF} , y los instantes temporales centrales de las ventanas utilizadas en dichos sistemas, T_{NMF} , siendo Ml el número total de ventanas usadas.

Así para cada localización temporal $times$, se buscan las ventanas T_{NMF} más cercana a dicho valor temporal, obteniéndose las posiciones en forma de índices i e $i - 1$ del vector T_{NMF} .

Una vez se localizan dichos índices, en función de N_{NMF} , se calcula el patrón con dimensiones $1 \times Ml$ como:

$$P_{Ml} \begin{cases} 1 & \begin{cases} (i-1)-1:(i)+2 & \text{si } N_{NMF} = 512 \\ (i-1):(i)+1 & \text{si } N_{NMF} = 1024 \\ (i-1):(i) & \text{si } N_{NMF} = 2048 \end{cases} \\ 0 & \text{en cualquier otro caso} \end{cases} \quad (33)$$

De [38] se conoce que los sonidos de corazón S1 tienen una duración aproximada de 140ms y los sonidos S2 tienen una duración de 110ms. De esta forma, como justificación de la configuración anterior, se selecciona el número suficiente de ventanas temporales necesarias para cubrir cada sonido de corazón por completo; dicho número se determina teniendo en cuenta N_{NMF} y el solapamiento utilizado entre ventanas del 50%, obteniéndose con estos parámetros una separación específica entre espacios temporales según su valor, siendo esta:

$$T_{NMF}(i) - T_{NMF}(i+1) \begin{cases} 32ms & \text{si } N_{NMF} = 512 \\ 64ms & \text{si } N_{NMF} = 1024 \\ 128ms & \text{si } N_{NMF} = 2048 \end{cases} \quad (34)$$

De esta forma, cuando $N_{NMF} = 512$ se utilizan cinco ventanas temporales por sonido de corazón en P_{Ml} , para $N_{NMF} = 1024$ se utilizan tres y para $N_{NMF} = 2048$ se utilizan dos. Finalmente, un ejemplo del patrón generado para una señal mezcla se asemejaría a la siguiente figura:

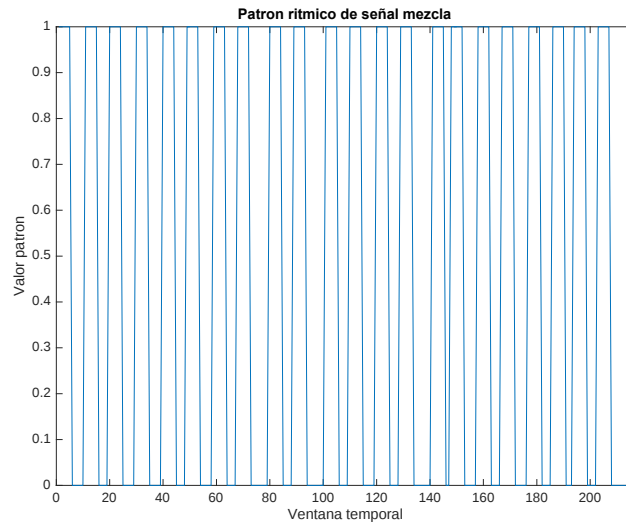


Figura 28 – Patrón rítmico de señal mezcla

5.4. Sistema NMF Semisupervisado

5.4.1. Fase de entrenamiento

Durante esta fase, se han entrenado los patrones espectrales para los sonidos de corazón dada su simplicidad y poca variabilidad respecto a los sonidos de pulmón. Para ello, se ha utilizado el modelo estándar de descomposición para NMF. Además, se han usado diferentes números de bases en el proceso de descomposición, así como diferentes longitudes de ventana para el cálculo de la matriz de entrada V ; generándose un conjunto de diccionarios atendiendo a la configuración establecida.

En primer lugar, se han seleccionado, para dicho propósito, dos tercios de una selección de sonidos de corazón obtenidos [33] y [34] dejándose el resto de sonidos para una posterior evaluación del sistema Semisupervisado. Después, utilizando el algoritmo de estimación del ritmo cardiaco, se han segmentado dichos sonidos de forma independiente para su posterior caracterización espectral, obteniéndose M sonidos. A continuación, cada sonido segmentado se ha factorizado usando un sistema NMF, donde se usa el modelo de descomposición estándar definido como:

$$V \approx W * H \quad (35)$$

Las ecuaciones de actualización multiplicativas utilizadas han sido:

$$W \leftarrow W \cdot \frac{((WH)^{[\beta-2]} V) H^T}{(WH)^{[\beta-1]} H^T} \quad (36)$$

$$H \leftarrow H \cdot \frac{W^T ((WH)^{[\beta-2]} V)}{W^T (WH)^{[\beta-1]}} \quad (37)$$

En dicho sistema NMF, se ha utilizado como función de coste la función Kullback-Leibler, la cual ofrece muy buenos resultados en la separación de fuentes sonoras, estableciéndose así un valor de $\beta = 1$. El número de bases K para la factorización utilizado ha

sido de 1, 2, 4 y 8 bases y las longitudes de ventana L utilizadas de 512, 1024 y 2048 muestras.

Para cada sonido segmentado de corazón, que puede incluir ruido, se ha obtenido una matriz de patrones W_i , descartándose su correspondiente matriz de activaciones H_i . Los diccionarios se obtienen con la concatenación de cada matriz W_i , resultando estos como:

$$W_{Corazón+ruido}^{K,L} = [W_1^{K,L}, W_2^{K,L} \dots W_M^{K,L}] \quad (38)$$

Finalmente, se han generado un total de doce diccionarios de patrones espectrales de sonidos de corazón que pueden o no estar acompañados por ruido, dependiendo de la configuración establecida para los parámetros K y L.

5.4.2. Modelo de separación

Para el modelo de separación utilizado en este sistema, se asume que los sonidos de corazón y pulmón son aditivos entre sí, es decir, la señal mezcla se obtiene de la suma de ambos sonidos. El modelo utilizado se define como:

$$V \approx Wc * Hc + Wp * Hp \quad (39)$$

Las matrices Wc, Hc se corresponden con las matrices de patrones espectrales y activaciones temporales para sonidos de corazón, con dimensiones $F \times b_c$ y $b_c \times T$, respectivamente; siendo b_c el número de bases utilizadas para la factorización de sonidos de corazón. Las matrices Wp, Hp se corresponden con las matrices de patrones espectrales y activaciones temporales para sonidos de pulmón, con dimensiones $F \times b_p$ y $b_p \times T$, respectivamente; siendo b_p el número de bases utilizadas para la factorización de sonidos de pulmón.

Al tratarse de un sistema Semisupervisado, la matriz Wc de patrones espectrales de sonidos de corazón, quedará fija durante el proceso de actualización de dichas matrices,

actualizándose el resto. La matriz Wc utilizada en el proceso de factorización se corresponderá con alguno de los diccionarios creados en la fase de entrenamiento.

Finalmente, una vez realizado el proceso de factorización, el espectrograma estimado de sonidos de corazón y sonidos de pulmón, se construye como:

$$X_c = Wc * Hc \quad (40)$$

$$X_p = Wp * Hp \quad (41)$$

5.4.3. Función de coste, reglas de actualización multiplicativas y criterio de parada.

Como función de coste para este sistema, se ha utilizado la función Kullback-Leibler a la cual se ha añadido un término de penalización Sparse a las matrices de activaciones temporales para cada tipo sonido, quedando la función de coste tal que:

$$D(V|WH) = V * \log_{10} \left(\frac{V}{Wc * Hc + Wp * Hp} \right) - V + (Wc * Hc + Wp * Hp) + \lambda_c \|Hc\|_1 + \lambda_p \|Hp\|_1 \quad (38)$$

Dada esta modificación en la función de coste KL original, donde se han añadido los términos de restricción λ_c y λ_p , se obtienen nuevas reglas de actualización multiplicativas definidas como:

- Ecuaciones para sonidos de corazón

$$Hc \leftarrow Hc \cdot \frac{Wc^T ((Wc * Hc)^{[\beta-2]} Y)}{Wc^T (Wc * Hc)^{[\beta-1]} + C_c * \lambda_c} \quad (42)$$

- Ecuaciones para sonidos de pulmón

$$Wp \leftarrow Wp \cdot \frac{((Wp * Hp)^{[\beta-2]} Y) Hp^T}{(Wp * Hp)^{[\beta-1]} Hp^T} \quad (43)$$

$$Hp \leftarrow Hp \cdot \frac{Wp^T((Wp * Hp)^{[\beta-2]}) \cdot Y}{Wp^T(Wp * Hp)^{[\beta-1]} + C_p * \lambda_p} \quad (44)$$

Siendo $\beta = 1$, dado el uso de la función KL, y C_c y C_p son valores de peso que aseguran que ambas restricciones tienen la misma importancia, siendo el valor máximo a sumar $b_c * T$ para la matriz de activaciones temporales H_c de corazón y $b_p * T$ para la matriz de activaciones temporales H_p de pulmón, definidas como:

$$C_c = \frac{F}{b_c} \quad (45)$$

$$C_p = \frac{F}{b_p} \quad (46)$$

Por último, se ha utilizado el siguiente criterio de parada, donde una vez se cumple la condición establecida con la cual se tiene en cuenta la variación en valor de la función de coste entre iteraciones sucesivas [27], se considera que el algoritmo ha convergido:

$$\frac{D^i(V|(Wc * Hc + Wp * Hp)) - D^{i-1}(V|(Wc * Hc + Wp * Hp))}{D^i(V|(Wc * Hc + Wp * Hp))} < 10^{-3} \quad (47)$$

5.5. Sistema NMF No Supervisado

5.5.1. Modelo de Separación

Para este sistema, se ha utilizado el modelo de descomposición estándar NMF definido como:

$$V \approx W * H \quad (48)$$

Las matrices $\mathbf{W}, \mathbf{H}$ se corresponden con las matrices de patrones espectrales y activaciones temporales de las fuentes a extraer, con dimensiones $F \times b$ y $b \times T$, respectivamente; siendo b el número de bases utilizadas para el proceso de factorización, que se obtiene como $b = b_{cor} + b_{pul}$, correspondiéndose dichos parámetros al número de bases para corazón y pulmón, respectivamente.

Al tratarse de un sistema no supervisado, ninguna matriz queda fija durante el proceso de actualización, actualizándose ambas a costa de añadir restricciones sobre los comportamientos de las fuentes a extraer.

5.5.2. Función de coste y reglas de actualización multiplicativas.

Como función de coste para este sistema, se ha utilizado la función Kullback-Leibler a la cual se ha añadido un término de penalización Sparse a la matriz de activaciones temporales, quedando la función de coste tal que:

$$D(V|WH) = V * \log_{10} \left(\frac{V}{W * H} \right) - V + (W * H) + \lambda * \|\mathbf{H}\|_1 \quad (49)$$

Dada esta modificación en la función de coste KL original, donde se ha añadido el término de restricción λ , se obtienen nuevas reglas de actualización multiplicativas definidas como:

$$\mathbf{W} \leftarrow \mathbf{W} \cdot \frac{((\mathbf{W} * \mathbf{H})^{[\beta-2]} \mathbf{Y}) \mathbf{H}^T}{(\mathbf{W} * \mathbf{H})^{[\beta-1]} \mathbf{H}^T} \quad (50)$$

$$\mathbf{H} \leftarrow \mathbf{H} \cdot \frac{\mathbf{W} \mathbf{C}^T ((\mathbf{W} * \mathbf{H})^{[\beta-2]} \mathbf{Y})}{\mathbf{W}^T (\mathbf{W} * \mathbf{H})^{[\beta-1]} + \mathbf{C} * \lambda} \quad (51)$$

Siendo $\beta = 1$, dado el uso de la función KL, y C un valor de peso definido como:

$$C = \frac{F}{b} \quad (52)$$

Por último, se ha utilizado el siguiente criterio de parada, donde una vez se cumple la condición establecida, se considera que el algoritmo ha convergido:

$$\frac{D^i(V|WH) - D^{i-1}(V|WH)}{D^i(V|WH)} < 10^{-3} \quad (53)$$

5.5.3. Clustering o Clasificación de bases

Dado que no se realiza ningún tipo de distinción entre bases, es necesaria una fase de Clustering, posterior al proceso de factorización, donde se permita clasificar cada una de ellas entre sonidos de corazón y sonidos de pulmón.

Así, nuestras aportaciones a este tipo de técnicas se basan en la implementación de una serie de técnicas para dicho propósito, no usadas con anterioridad en otros trabajos. Para cada una de las bases generadas durante la factorización, se calculan tres medidas diferentes:

1. Medida del Roll-off, con dicha medida se calcula cómo la energía de cada una de las bases se distribuye a través de la banda de frecuencia, [39]. Así centrándonos en la banda del corazón, la mayor parte de la energía de las bases correspondientes a sonidos de corazón se encontrará por debajo de una frecuencia $fc < 320Hz$. Por otro lado, la energía de aquellas bases correspondientes a sonidos de pulmón se concentrará mayoritariamente por encima de la frecuencia de corte fc establecida para sonidos de corazón, siendo $fc > 320Hz$ para los sonidos de pulmón.

Así, para el cálculo de esta medida se ha seguido el siguiente procedimiento: En primer lugar, se calcula el espectrograma correspondiente a cada una de dichas bases como:

$$X_{aux} = W(:, b_i) * H(b_i, :) \quad (54)$$

siendo b_i la base i -ésima. A continuación, utilizando la Transformada discreta inversa de Fourier en tiempo corto, iSTFT, se reconstruye la señal temporal $x_{b_i}(t)$, correspondiente a

dicha base, con el fin de obtener una mayor resolución frecuencial de dicha señal. Posteriormente, se calcula la Densidad Espectral de Potencia, PSD, de dicha señal como:

$$PSD_{b_i}(k) = |X(k)|^2 \quad (55)$$

siendo $X(k)$ la Transformada de Fourier de la señal $x_{b_i}(t)$. Finalmente, para dicha base se calcula como:

$$\sum_{K=0}^{fc(b_i)} PSD(K) = 0.85 * \sum_{K=0}^{sr/2} PSD(K) \quad (56)$$

siendo $sr/2$ la frecuencia de Nyquist y $fc(b_i)$, la frecuencia de corte Roll-off para la base i-ésima, donde dicha frecuencia es el punto por debajo del cual se encuentra el 85% de la energía de la señal $x_{b_i}(t)$.

A continuación, se muestra un ejemplo de la frecuencia Roll-off calculada para una base perteneciente a sonido de corazón:

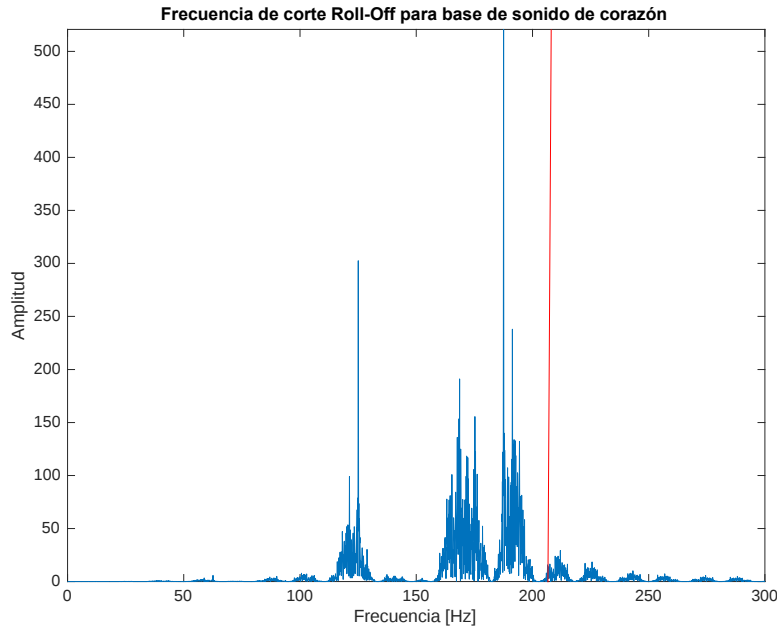


Figura 29 – Frecuencia de corte Roll-Off para base de sonido de corazón

Siendo dicha frecuencia de corte en este caso de 206.15Hz, siendo este valor menor a los 320Hz establecidos para sonidos de corazón.

En la siguiente, se muestra un ejemplo de la frecuencia Roll-off calculada para una base perteneciente a sonido de pulmón:

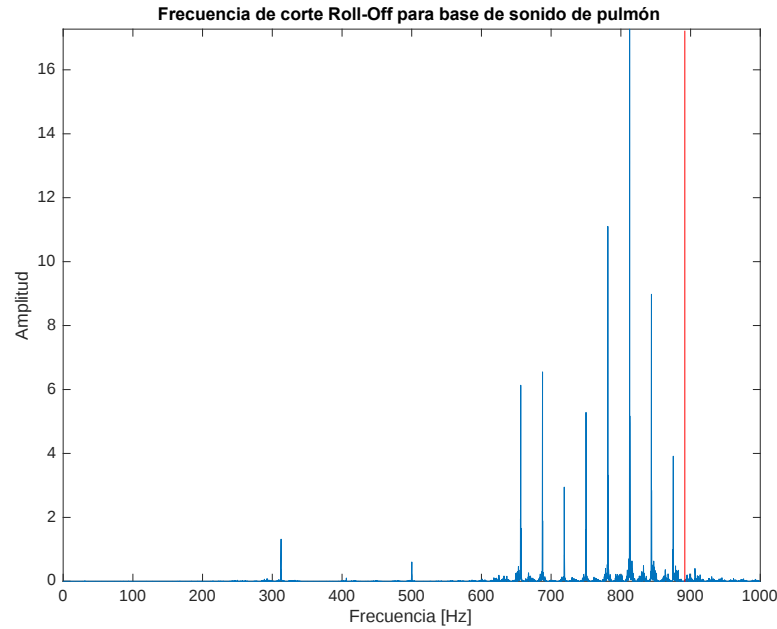


Figura 30 – Frecuencia de corte Roll-Off para base de sonido de pulmón

Siendo dicha frecuencia de corte en este caso de 875.98 Hz, mayor que los 320 Hz establecidos para sonidos de corazón, correspondiéndose por tanto a sonido de pulmón.

2. Medida de la correlación espectral: Con esta medida se mide el nivel de similitud entre la matriz de patrones espectrales generada en la descomposición, donde la columna $W(:, i)$ correspondiente a los patrones espectrales para la base ‘i’ se compara con cada una de las columnas ‘s’ del diccionario de patrones espectrales seleccionado $W_c(:, s)$, anteriormente almacenado para su uso en el sistema NMF Semisupervisado. La medida de similitud espectral para la base i se establece como el máximo valor obtenido tras comparar la matriz $W(:, i)$ con cada matriz $W_c(:, s)$, con $s = 1 \dots KM$. La selección del diccionario con el que comparar dependerá de la longitud de ventana usada para el cálculo de la STFT en el sistema NMF No Supervisado. Para la medida de dicha similitud entre matrices, se calcula la distancia del coseno entre ellas [40] obteniéndose una medida en el rango [0,1]:

$$cor(W(:, i), Wc(:, s)) = SIM(b_i) = \max \left(\frac{Wc(:, s) \cdot W(:, i)}{\|Wc(:, s)\| * \|W(:, i)\|} \right) \quad (57)$$

3. Medida de la correlación temporal: Para dicha medida, se calcula el coeficiente de correlación entre el patrón rítmico generado por el algoritmo del ritmo cardíaco con la matriz de activaciones temporales, $H(i, :)$, correspondiente a la base ‘i’ utilizada en el proceso de descomposición. Con este motivo, cada matriz $H(i, :)$, es preprocesada de forma que:

$$H(i, :) \begin{cases} 1 \text{ si } H(i, :) > media(H_{b_i}) \\ 0 \text{ si } H(i, :) < media(H_{b_i}) \end{cases} \quad (58)$$

De esta forma, todos los valores de cada matriz $H(i, :)$ se transforman al mismo rango de valores establecidos en el patrón rítmico. Así, una vez se ha aplicado dicho preprocesado, el coeficiente de correlación ρ se calcula como [41]:

$$\begin{aligned} cor(H(i, :), P_{Ml}) &= \rho(H(i, :), P_{Ml}) \\ &= \frac{1}{N-1} \sum_{j=1}^N \left(\frac{H(i, j) - \mu_{H(i, :)}}{\sigma_{H(i, :)}} \right) \left(\frac{P_{Ml}(1, j) - \mu_{P_{Ml}}}{\sigma_{P_{Ml}}} \right) \end{aligned} \quad (59)$$

Siendo N la longitud total de ambos vectores, en ambos casos Ml, j la posición actual en ambos vectores, $\mu_{H(i, :)}$ y $\mu_{P_{Ml}}$ la media correspondiente a $H(i, :)$ y P_{Ml} , respectivamente, y $\sigma_{H(i, :)}$ y $\sigma_{P_{Ml}}$ la desviación estándar correspondiente a $H(i, :)$ y P_{Ml} , respectivamente.

Se calcula así el coeficiente de correlación ρ entre el patrón rítmico y cada una de las matrices de activaciones temporales comparándose el valor en la misma posición ‘j’ en ambos vectores, obteniendo una medida en el rango [-1,1] para cada base ‘i’.

Dado que se usan tres medidas de clasificación, es necesario el uso de tres parámetros de discriminación entre sonidos de pulmón y corazón; por ello, se crean tres matrices diferentes: una correspondiente a sonidos de corazón, X_c , otra correspondiente a sonidos de pulmón, X_p , y

otra correspondiente al espectrograma para la base i-ésima, $X_{aux} = W(:, b_i) * H(b_i, :)$; de este modo, se establecen los siguientes criterios de clasificación:

$$X_{aux} = \begin{cases} X_c, & fc_{b_i} < fc_{aux} \\ X_p, & fc_{b_i} \geq fc_{aux} \end{cases} \quad (60)$$

$$X_{aux} = \begin{cases} X_c, & SIM(b_i) < SIM_{aux} \\ X_p, & SIM(b_i) \geq SIM_{aux} \end{cases} \quad (61)$$

$$X_{aux} = \begin{cases} X_c, & \rho(H_{b_i}, P_{ML}) < \rho_{aux} \\ X_p, & \rho(H_{b_i}, P_{ML}) \geq \rho_{aux} \end{cases} \quad (62)$$

Siendo fc_{aux} el parámetro de discriminación para la medida del Roll-off, SIM_{aux} el parámetro de discriminación para la medida de la correlación espectral y ρ_{aux} el parámetro de discriminación para la medida de la correlación temporal.

Por otro lado, al utilizarse tres tipos de medidas de clasificación, la fase de Clustering se puede realizar en función de una serie de combinaciones entre ellas, encontrándose así siete combinaciones posibles para la clasificación de bases que encontramos en la siguiente tabla. En adelante, se indicara la medida del Roll-off como ‘FC’, la correlación espectral como ‘SIM’ y la temporal como ‘CORR-TEMP’.

Métodos de Clasificación de bases
Sólo FC
Sólo SIM
Sólo CORR-TEMP
CORR-TEMP+SIM
CORR-TEMP+FC
SIM+FC
CORR-TEMP+FC+SIM

Tabla 1 – Posibles combinaciones de los técnicas de Clustering

Dada la gran variedad de combinaciones de técnicas posibles para la clasificación de bases; posteriormente, será necesaria una fase de evaluación donde se calculen los parámetros óptimos de discriminación $f_{c_{aux}}$, SIM_{aux} y ρ_{aux} , para dichas combinaciones.

5.5.4. Modelo de factorización basado en el patrón rítmico

Junto al sistema NMF No Supervisado, se ha añadido la posibilidad de realizar la descomposición del espectrograma de la señal mezcla de entrada al sistema utilizando la información proporcionada por el patrón rítmico. Este proceso de descomposición se realiza distinguiendo entre dos posibles zonas en dicho espectrograma y utilizándose un número de bases en la factorización adaptado a cada sonido en particular, siendo este un proceso novedoso con el que se intenta solventar uno de los principales problemas en las técnicas NMF, en las cuales no se establece el número necesario de bases necesarias para la realización del proceso de factorización. Se distingue así entre:

- **Zonas con presencia de sonidos de corazón:** Zonas del espectrograma de entrada en las que se encuentran sonidos procedentes del corazón humano con o sin solapamiento espectro-temporal de sonidos procedentes del sistema respiratorio.
- **Zonas con presencia de sonidos de pulmón:** Zonas del espectrograma de entrada donde únicamente se encuentran sonidos procedentes del sistema respiratorio o pulmones.

Así, el patrón rítmico permite determinar el número de espacios temporales o ‘frames temporales’ que conforman cada zona. Una vez dichas zonas se encuentran delimitadas, son factorizadas utilizando el modelo estándar NMF del sistema No Supervisado, utilizándose un número óptimo de bases para la factorización de los sonidos presentes en estas.

Dado que en dicho modelo de factorización se pierde la información espacial de los sonidos de corazón proporcionada por el patrón rítmico, la clasificación de las matrices

obtenidas W y H para cada base ‘i’ se puede realizar aplicando las siguientes metodologías de Clustering:

Métodos de Clasificación de bases para el nuevo modelo de factorización
Sólo FC
Solo SIM
SIM+FC

Tabla 2 – Posibles combinaciones de las técnicas de Clustering para el nuevo modelo de factorización

Una vez finalizada la fase de Clustering para las bases de cada zona, su resultado se coloca en la zona correspondiente dentro del espectrograma de corazón o pulmón a reconstruir. Finalmente, el proceso de reconstrucción y recuperación de la señales se mantiene inalterado respecto al procedimiento original.

5.6. Reconstrucción y recuperación de las señales

Una vez se obtiene el espectrograma estimado de sonidos de corazón X_c y pulmón X_p , se utiliza el mismo método de reconstrucción de los espectrogramas correspondientes a dichos sonidos y recuperación de señales al dominio temporal en ambos sistemas NMF.

El método de reconstrucción de dichos espectrogramas se basa en el uso de máscaras Wiener, en [41]. El uso de máscaras Wiener permite escalar la magnitud de la señal mezcla para cada componente de frecuencia de forma que se obtengan valores comprendidos entre 0 y 1, a partir de los cuales se encuentra el correspondiente valor del componente en frecuencia en el espectrograma estimado. De esta forma, se consigue que la suma de ambos espectrogramas se corresponda con el espectrograma de la señal original, evitándose el error producido por la factorización aproximada de las técnicas NMF. Así, se construyen dos máscaras correspondientes a cada uno de los sonidos:

$$M_p = \frac{X_p^2}{X_p^2 + X_c^2} \quad (63)$$

$$M_c = \frac{X_c^2}{X_p^2 + X_c^2} \quad (64)$$

Siendo M_p y M_c , las máscaras correspondientes a sonidos de pulmón y corazón respectivamente. En dicho cálculo, cada elemento de cada una de las matrices se eleva al cuadrado, siendo por tanto, una operación elemento a elemento. Finalmente, los espectrogramas se reconstruyen a partir del espectrograma original de la señal mezcla, manteniendo de esta forma la fase de la señal de entrada compleja, como:

$$X_{c_r} = M_c \otimes V \quad (65)$$

$$X_{p_r} = M_p \otimes V \quad (66)$$

Recuperándose de esta forma la fase de la señal original y cometiéndose cero error en la construcción de los espectrogramas ya que:

$$V = X_{c_r} + X_{p_r} \quad (67)$$

Una vez se reconstruyen los espectrogramas, para recuperar las señales de corazón y pulmón en el dominio temporal se utiliza la Transformada Discreta Inversa de Fourier en tiempo corto o iSTFT, de modo que:

$$x_c(t) = iSTFT(X_{c_r}) \quad (68)$$

$$x_p(t) = iSTFT(X_{p_r}) \quad (69)$$

5.7. Búsqueda de patrones repetitivos

En esta sección, se describirá el método desarrollado para la búsqueda de patrones repetitivos en los espectrogramas reconstruidos de corazón X_{c_r} y de pulmón X_{p_r} . En primer lugar, se ha implementado una fase de selección en dichos espectrogramas donde se permite marcar aquellas localizaciones temporales, a modo de recordatorio, donde se han encontrado sonidos para los cuales se estima oportuna su búsqueda a lo largo de la señal de corazón o pulmón, de forma que se conozca si dichos sonidos se producen de forma puntual o se producen de forma repetitiva a lo largo de la señal.

En esta fase de localización, se genera un nuevo espectrograma en el cual una barra avanza desde el comienzo del espectrograma al final de este, al mismo tiempo que la señal de corazón o pulmón, que se pretende analizar, se escucha de forma sincronizada al movimiento de dicha barra. Siendo este proceso de búsqueda muy cómodo en su utilización, a la vez que ha resultado costosa su implementación dado el problema presente en MATLAB en cuanto a los hilos de ejecución, permitiendo este programa un único hilo de ejecución simultáneo.

Así, una vez se localiza un determinado sonido que se estima oportuno analizar, se selecciona la posición temporal de la barra utilizando la ‘BARRA ESPACIADORA’ del teclado, marcándose tantas localizaciones como el usuario estime oportuno. Dicho proceso de recorrido del espectrograma se repite de manera cíclica, hasta que el usuario presione la tecla ‘ESC’ de su teclado.

A continuación, se muestra un ejemplo del funcionamiento de la herramienta en dicha fase:

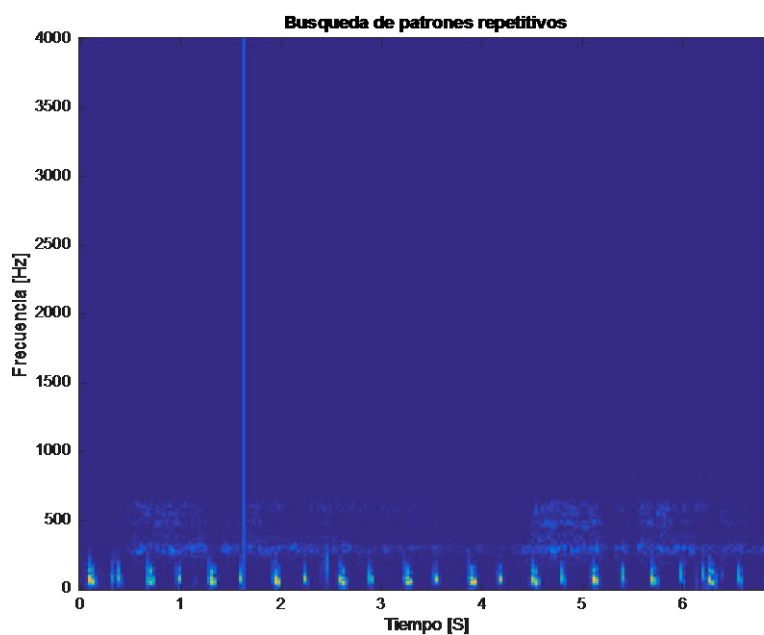


Figura 31 – Función de búsqueda de patrones

Dado que el programa MATLAB es un programa que utiliza un único hilo de ejecución, el movimiento constante de la barra y la utilización del teclado como método de selección de las localizaciones temporales, se ha resuelto utilizando una función secundaria ‘teclado’ [43] a la función principal ‘reproductor’ que constantemente espera la pulsación de alguna de las teclas del teclado y, a partir de dicha pulsación, se actúa en consecuencia en la función ‘reproductor’, a través de la propiedad ‘KeyPressFcn’ de la función ‘Figure’ y la función ‘drawnow’, necesaria para poder actuar cuando una tecla se presiona. Para la creación de la barra, se utiliza el valor máximo del espectrograma en el cual se realiza la búsqueda, sustituyendo de forma sucesiva, desde principio a fin, el valor de todos los bins de frecuencia por dicho valor máximo en cada

uno de los espacios temporales, dejando el resto sin modificar. Para sincronizar el movimiento de la barra con el sonido de corazón o pulmón, se utiliza la función ‘pause’ en cada iteración tras establecer la barra, pausándose el hilo de ejecución principal un tiempo determinado de forma que el sonido y el movimiento de la barra estén sincronizados. Finalmente, los sonidos se reproducen utilizando la función ‘play’ y se paran con la función ‘stop’.

A continuación, se muestra un ejemplo de espectrograma con varias posiciones temporales marcadas:

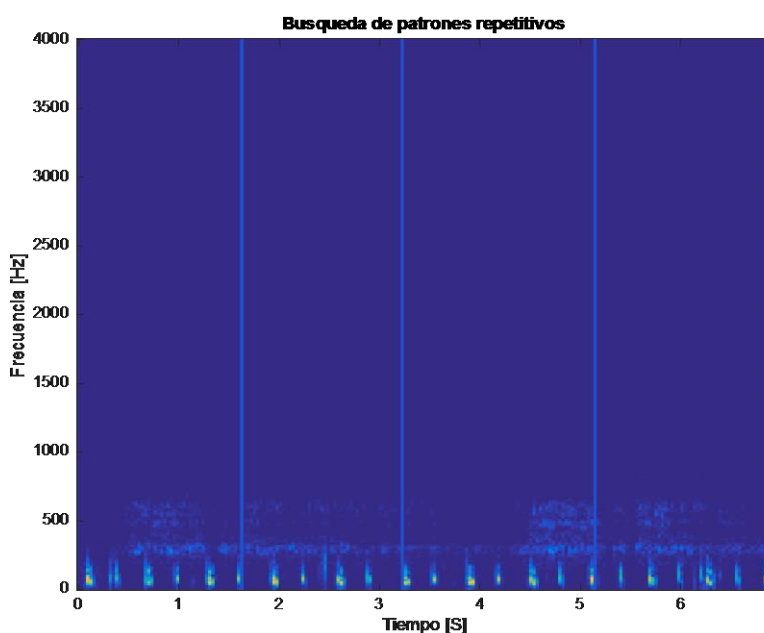


Figura 32 – Búsqueda de patrones, instantes marcados

Tras finalizar la selección de las localizaciones temporales oportunas, se encuentra una pantalla como la anterior. El siguiente paso, en el proceso de búsqueda de patrones, consiste en seleccionar una región en dicha pantalla. Para ello, se utiliza la función ‘spectro_extraction’, que a través del uso de la función ‘getrect’, permite dicha selección y se devuelve a la función principal ‘reproductor’ como la matriz ‘ X_{sel} ’.

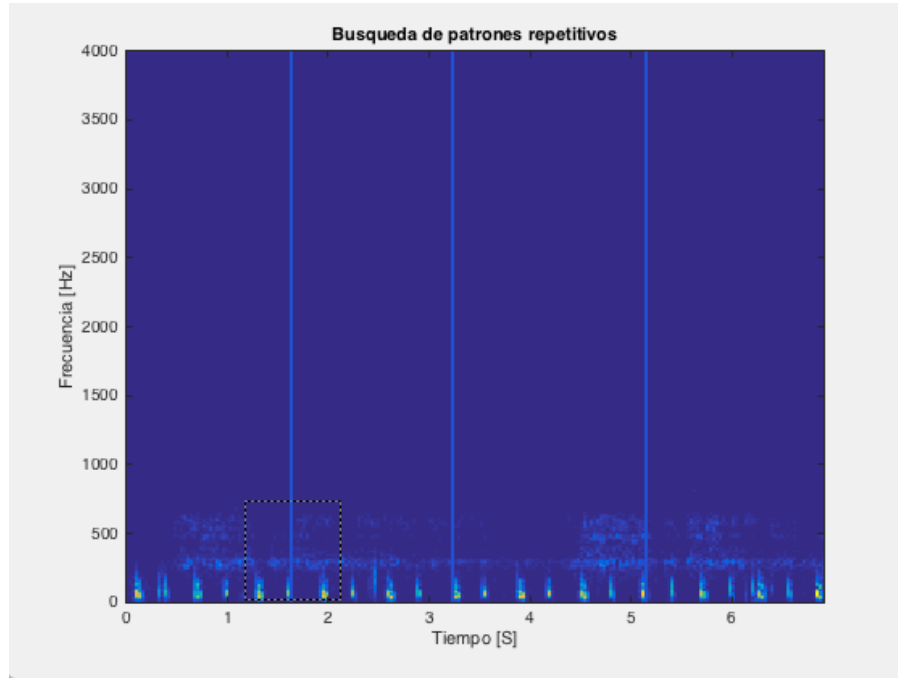


Figura 33 – Búsqueda de patrones, selección de la región

Una vez se obtiene dicha región, como medida de la similitud entre esta y el espectrograma reconstruido de corazón o de pulmón, se calcula la correlación cruzada normalizada en dos dimensiones entre dichas matrices. Dicha medida se define como [44]:

$$C(i, j) = \frac{\sum_{x,y} (S(i-1+x, j-1+y) - \overline{S_{i,j}}) * (X_{sel}(x, y) - \overline{X_{sel}})}{\left(\sum_{x,y} (S(i-1+x, j-1+y) - \overline{S_{i,j}})^2 \sum_{x,y} (X_{sel}(x, y) - \overline{X_{sel}})^2 \right)^{\frac{1}{2}}} \quad (70)$$

Siendo S el espectrograma reconstruido de corazón X_{c_r} o de pulmón X_{p_r} . De esta forma se obtiene una matriz de similitud con coeficientes en el rango $[-1,1]$ que miden el nivel de semejanza entre cada punto de dicha región seleccionada y cada bin de frecuencia y espacio temporal del espectrograma reconstruido.

Un ejemplo de dicha matriz se muestra en la siguiente figura:

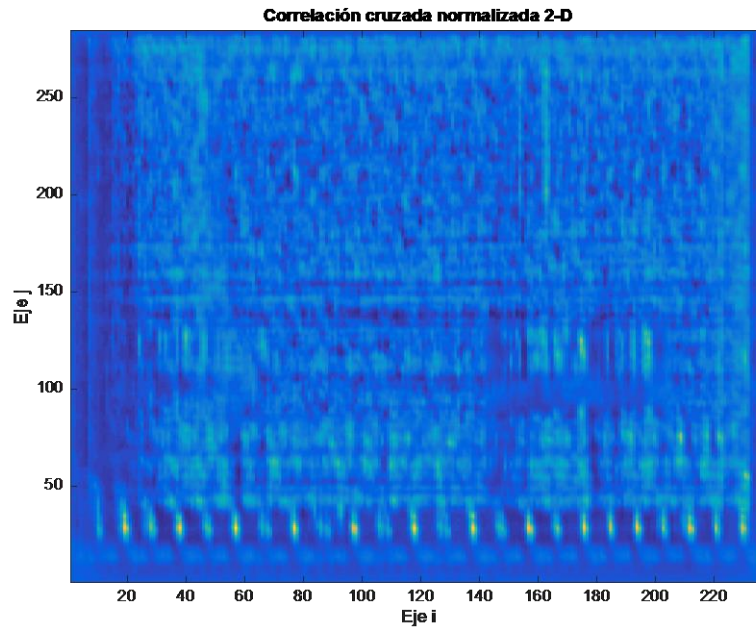


Figura 34 - Ejemplo de Correlación cruzada normalizada 2-D

Finalmente, el proceso de búsqueda de patrones repetitivos se realiza de forma iterativa a través de dicha matriz hasta que se deja de cumplir una condición establecida por el usuario. Dicha condición es un parámetro de similitud mínima en tanto por ciento entre ambas matrices que se extrapola al rango $[0,1]$. El proceso, por tanto se desarrolla como sigue:

- En primer lugar, se calcula el máximo de la matriz C , obteniéndose las posiciones (i,j) en dicha matriz de similitud.
- Después, si dicho máximo cumple la condición de similitud mínima, se asumen que en esa posición se encuentra un sonido similar al buscado y se calculan las posiciones (x,y) de este en el espectrograma Xc_r o Xp_r ; para ello, se elimina a las posiciones (i,j) de la matriz de similitud el desplazamiento en filas y columna introducido como resultado del cálculo de dicha correlación, así:

$$x = i - n_{col} \quad (71)$$

$$y = j - n_{fil} \quad (72)$$

Siendo n_{col} y n_{fil} el número total de columnas y filas de la región seleccionada.

- Una vez se calculan dichas posiciones, se elimina de la matriz de similitud una región con centro en el máximo localizado en (i,j) y con dimensiones $n_{col} \times n_{fil}$. Por otro lado, la matriz donde se almacenan los sonidos encontrados se calcula como:

$$S_1 = S(:, x'); \quad (73)$$

Siendo x' , un vector calculado a través de la posición x y n_{col} , que se adecua en cada caso a la región del espectrograma donde se ha encontrado dicho sonido, evitándose así posibles problemas en la selección.

- Finalmente, una vez realizado este proceso, se repite de forma iterativa hasta que se deja de cumplir la condición de similitud preestablecida, obteniéndose la matriz S_1 con todos los sonidos encontrados.

Como ejemplo a lo anteriormente expuesto, si se realiza una búsqueda con un porcentaje de similitud del 90%, se obtiene el siguiente espectrograma como resultado:

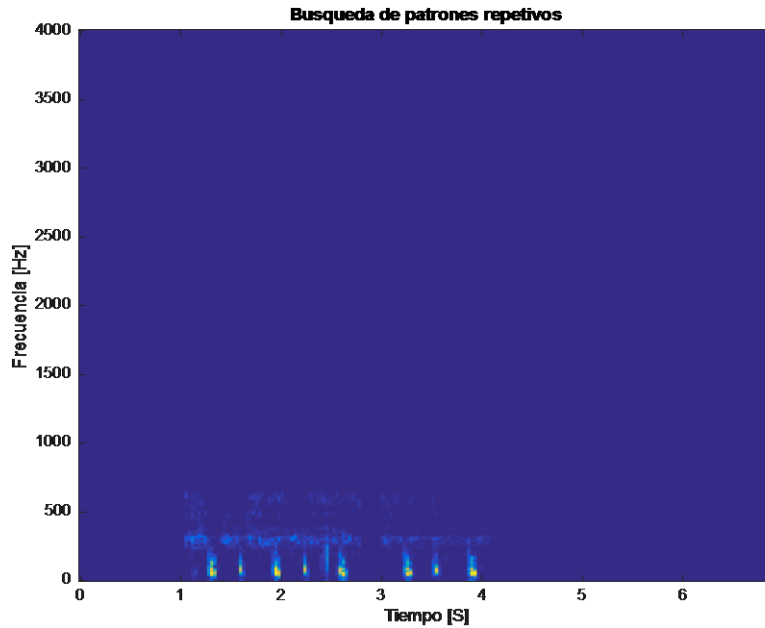


Figura 35 – Resultado de la búsqueda de patrones

CAPÍTULO 5 – DESARROLLO DEL TFM

CAPÍTULO 6

6. Resultados

En este capítulo, se describirá el método utilizado para la generación de la bases de datos de señales mezclas a separar y la metodología utilizada para la evaluación de parámetros de ambos sistemas, así como para su testeo. Finalmente, se realizará un análisis de la eficiencia del algoritmo de ritmo cardíaco.

6.1. Base de datos

Los sonidos de corazón utilizados durante este TFM, se han obtenido de [33] y [34], y los sonidos de pulmón se han obtenido de [45] y [46]. Dado que los sonidos de corazón obtenidos no son sonidos completamente puros, se ha realizado una selección subjetiva de aquellos sonidos de corazón con menor ruido añadido. Así, se han obtenido un total de trece sonidos de corazón. Por otro lado, se han seleccionado nueve sonidos de pulmón.

A partir de dichos sonidos de corazón, se han creado dos bases de datos diferentes, una base de datos para el sistema Semisupervisado y otra base de datos para el sistema No supervisado.

En cuanto a la base de datos para el sistema Semisupervisado, se ha seleccionado aproximadamente dos tercios del total de sonidos corazón, para la creación de diccionarios de patrones espectrales en la fase de entrenamiento. Con el tercio restante y junto a los nueve sonidos de pulmón se han creado cincuenta y cuatro mezclas de siete segundos de duración cada una.

Se han obtenido dichas señales mezcla, seleccionando cada sonido de corazón y mezclándolo con los nueve sonidos de pulmón a diferentes relaciones señal-ruido, SNR: -5 dB, 0dB y 5dB. Anterior a dicha mezcla, se ha igualado la potencia de ambas señales, siempre

seleccionando la menor potencia entre ambas, reduciendo la de la señal con mayor potencia y evitándose, de este modo, la saturación de las señales.

Finalmente, de las cincuenta y cuatro señales mezcla, en dicha base de datos, se ha seleccionado aproximadamente dos tercios para la fase de optimización de parámetros y un tercio para la fase de testeo, usándose para ambos la señal mezcla con SNR de 5dB.

En cuanto a la base de datos para el sistema No Supervisado, de los trece sonidos de corazón se han seleccionado ocho de ellos. Se ha seguido la misma metodología del caso anterior para generar las señales mezcla obteniéndose, en este caso, un total de setenta y dos mezclas de siete segundos cada una. Se ha utilizado, del mismo modo, dos tercios para la fase de optimización de parámetros y un tercio para la fase de testeo, usándose para ambos la señal mezcla con SNR de 5dB.

6.2. Optimización de parámetros

Para la fase de optimización de parámetros, se ha optado por un análisis hiperparamétrico, dada la gran variedad de parámetros a optimizar y métodos de Clustering a evaluar. De esta forma, mientras se varía el valor de un parámetro, el resto de parámetros se mantienen fijos. Además, dado que los sistemas NMF proporcionan un proceso de factorización aproximado, se ha evaluado cada señal mezcla tres veces obteniendo así un valor medio de cada uno de los parámetros obtenidos para cada señal, siendo estas SDR, SIR y SAR, definidas como:

- **SDR (Source-to-distorsion ratio):** Proporciona información de la calidad global de separación del sistema.

$$SDR(dB) = 10 * \log_{10} \frac{\|s_{true}\|^2}{\|e_{interf} + e_{noise} + e_{artif}\|^2} \quad (74)$$

- **SIR (Source-to-interferences ratio):** Proporciona una medida de la presencia de sonidos de corazón en la señal de pulmón y viceversa:

$$SIR(dB) = 10 * \log_{10} \frac{\|s_{true}\|^2}{\|e_{interf}\|^2} \quad (75)$$

- **SAR (Source-to-artifacts ratio):** Proporciona información de los artefactos en la señal separada debido al proceso de separación y reconstrucción.

$$SAR(dB) = 10 * \log_{10} \frac{\|s_{true} + e_{interf} + e_{noise}\|^2}{\|e_{artif}\|^2} \quad (76)$$

Siendo s_{true} una matriz que contiene la fuente original, e_{noise} es una matriz que contiene la componente espacial de distorsión, e_{interf} es una matriz que contiene la componente de interferencia y e_{artif} es una matriz que contiene la componente de artefactos [47].

6.2.1. Parámetros óptimos del sistema NMF no supervisado.

Para el cálculo óptimo de parámetros, se ha tenido en cuenta el valor SDR obtenido de la separación; de esta forma, se han obtenido tres conjuntos de parámetros en función de dicha medida: parámetros para la maximización de SDR de sonidos de corazón, para la maximización de SDR de sonidos de pulmón y para la maximización de SDR media entre ambos sonidos o maximización equilibrada.

Los parámetros de discriminación de cada método Clustering se han calculado utilizando una longitud de ventana de 512 muestras, que proporciona una resolución temporal de 64ms, 64 bases para los sonidos de corazón, 64 bases para los sonidos de pulmón y Sparse cero.

A modo de ejemplo para la comprensión del procedimiento seguido, se indica a continuación como se calcularían los parámetros de discriminación cuando se hace uso del método SIM+FC+CORR-TEMP para la maximización de corazón. Se indicará con color azul el parámetro que se varía en cada momento. También, en adelante, cuando cualquier parámetro de discriminación no sea utilizado, se marcará en la tabla con el símbolo ‘-’.

- En primer lugar, se variará el parámetro de discriminación correspondiente a la correlación SIM_{aux} entre 0 y 1, estableciéndose fijos los valores de los parámetros de discriminación fc_{aux} de la técnica Roll-Off como 320 Hz, límite en frecuencia establecido para los sonidos de corazón en el apartado 5.1, y ρ_{aux} del cálculo de la correlación temporal como 0, ya que para esta correlación se obtenían medidas en el rango [-1,1], donde supondremos que cuando se obtienen valores mayores a 0, las bases se corresponden a sonidos de corazón y cuando se obtienen valores menores a 0 las bases se corresponden a sonidos de pulmón. El valor de la restricción Sparse se mantiene a 0.

Parámetros para la maximización de corazón			
	Parámetros de discriminación		
Método de Clustering	SIM_{aux}	fc_{aux}	ρ_{aux}
SIM+FC+CORR-TEMP	0.42	320	0
Parámetro Sparse, λ	0		
Número de bases de corazón, b_{cor}	64		
Número de bases de pulmón, b_{pul}	64		

Tabla 3 – Ejemplo del cálculo de parámetros óptimos 1.

Así, se observa que cuando se varía el valor de SIM_{aux} , se obtiene para cada sonido mezcla evaluado de la base de datos para la optimización, un valor determinado de

SIM_{aux} con el cual se maximiza la medida SDR de corazón. Finalmente, se obtiene el valor final de SIM_{aux} como la mediana entre todos los valores de SIM_{aux} obtenidos para cada sonido mezcla, estableciéndose que $SIM_{aux} = 0.42$.

- En segundo lugar, se fija dicho valor para SIM_{aux} y se varía el valor del parámetro fc_{aux} entre los valores 140Hz, límite superior en frecuencia en el que aparecen las componentes principales de sonidos de corazón, y 320Hz, límite superior para el resto de componentes de sonidos de corazón. El resto de parámetros se mantienen igual.

Parámetros para la maximización de corazón			
	Parámetros de discriminación		
Método de Clustering	SIM_{aux}	fc_{aux}	ρ_{aux}
SIM+FC+CORR-TEMP	0.42	277	0
Parámetro Sparse, λ	0		
Número de bases de corazón, b_{cor}	64		
Número de bases de pulmón, b_{pul}	64		

Tabla 4 – Ejemplo del cálculo de parámetros óptimos 2.

Así, se observa que cuando se varía el valor de fc_{aux} , se obtiene para cada sonido mezcla evaluado de la base de datos para la optimización, un valor determinado de fc_{aux} con el cual se maximiza la medida SDR de corazón. Finalmente, se obtiene el valor final de fc_{aux} como la mediana entre todos los valores de fc_{aux} obtenidos para cada sonido mezcla, estableciéndose que $fc_{aux} = 277$.

- En tercer lugar, se fijan los valores para SIM_{aux} y fc_{aux} y se varía el valor del parámetro ρ_{aux} en el rango [-1,1]. El resto de parámetros se mantienen igual.

Parámetros para la maximización de corazón			
	Parámetros de discriminación		
Método de Clustering	SIM_{aux}	fc_{aux}	ρ_{aux}
SIM+FC+CORR-TEMP	0.42	277	0.05
Parámetro Sparse, λ	0		
Número de bases de corazón, b_{cor}	64		
Número de bases de pulmón, b_{pul}	64		

Tabla 5 – Ejemplo del cálculo de parámetros óptimos 3.

Así, se observa que cuando se varía el valor de ρ_{aux} , se obtiene para cada sonido mezcla evaluado de la base de datos para la optimización, un valor determinado de ρ_{aux} con el cual se maximiza la medida SDR de corazón. Finalmente, se obtiene el valor final de ρ_{aux} como la mediana entre todos los valores de ρ_{aux} obtenidos para cada sonido mezcla, estableciéndose que $\rho_{aux} = -0.05$.

Se sigue el mismo procedimiento para el cálculo de los parámetros de discriminación para el resto de métodos de Clustering. Una vez se calculan los parámetros para cada método, se calcula el valor correspondiente a la restricción Sparse λ , el número de bases de corazón b_{cor} y el número de bases de pulmón, b_{pul} . Para ello, se ha utilizado la metodología SIM+FC+CORR-TEMP, extrapolando los resultados al resto de metodologías.

- Para el cálculo del valor del parámetro correspondiente a la restricción Sparse λ , se varía su valor en el rango $[0,1]$, fijando el valor de los tres parámetros de discriminación y fijando el número de bases de pulmón y corazón a 64. El valor final del parámetro λ , se calcula como la mediana entre todos los posibles valores, de Sparse, obtenidos para cada una de las señales mezcla evaluadas, estableciéndose finalmente como $\lambda = 0.55$.

- Para el cálculo del valor del número de bases de corazón b_{cor} , se varía su valor en el rango $[1,64]$, fijándose el valor de Sparse λ a 0, el número de bases de pulmón a 64 y el valor de los tres parámetros de discriminación. El valor final del parámetro b_{cor} , se calcula como la mediana entre todos los posibles valores, de b_{cor} , obtenidos para cada una de las señales mezcla evaluadas, estableciéndose finalmente como $b_{cor} = 40$.
- Por último, para el cálculo del valor del número de bases de pulmón b_{pul} , se varía su valor en el rango $[1,64]$, fijándose el valor de Sparse λ a 0, el número de bases de corazón a 64 y el valor de los tres parámetros de discriminación. El valor final del parámetro b_{pul} , se calcula como la mediana entre todos los posibles valores, de b_{pul} , obtenidos para cada una de las señales mezcla evaluadas, estableciéndose finalmente como $b_{pul} = 39$.

Siguiendo el procedimiento anterior, se obtiene la siguiente tabla de parámetros óptimos según el tipo de maximización de resultados deseada para todas las combinaciones de Clustering posibles:

	Parámetros para la maximización de corazón			Parámetros para la maximización de pulmón			Parámetros para maximización equilibrada		
	<i>Parámetros de discriminación</i>			<i>Parámetros de discriminación</i>			<i>Parámetros de discriminación</i>		
<i>Método de Clustering</i>	SIM_{aux}	fc_{aux}	ρ_{aux}	SIM_{aux}	fc_{aux}	ρ_{aux}	SIM_{aux}	fc_{aux}	ρ_{aux}
Sólo SIM	0.59	-	-	0.4	-	-	0.56	-	-
Sólo FC	-	235	-	-	259	-	-	259	-
Sólo CORR-TEMP	-	-	0.08	-	-	0.05	-	-	0.06
CORR-TEMP + SIM	0.49	-	-0.03	0.31	-	-0.26	0.4	-	-0.11
CORR-TEMP+FC	-	247	0.04	-	292	-0.02	-	288	0.04
SIM+FC	0.6	261	-	0.36	293	-	0.51	294	-
SIM+FC+CORR-TEMP	0.42	277	0.05	0.16	284	-0.04	0.28	283	0.02
<i>Parámetro Sparse, λ</i>	0.55			0.73			0.64		
<i>Número de bases de corazón, b_{cor}</i>	40			40			20		
<i>Número de bases de pulmón, b_{pul}</i>	39			29			21		

Tabla 6 – Parámetros óptimos según el tipo de maximización para todas las combinaciones de Clustering

6.2.1.1. *Parámetros óptimos del sistema NMF no supervisado utilizándose el modelo de factorización basado en el patrón rítmico*

Los parámetros óptimos para dicho sistema se toman de los obtenidos en el apartado anterior, obviando aquellos correspondientes a los métodos de Clustering donde se tiene en cuenta la técnica CORR-TEMP. Así mismo, se ha optimizado el número de bases necesario para la factorización en zonas con presencia de sonidos de corazón; para ello se ha utilizado la metodología SIM+FC y se han fijado los valores de discriminación para esta. El valor final del número de bases de corazón para dichas zonas se obtiene como la mediana entre todos los valores obtenidos para todas las señales mezclas evaluadas, estableciéndose finalmente un valor de 18 bases para dichas zonas. De esta forma:

	Parámetros para la maximización de zonas de corazón		Parámetros para la maximización de zonas de pulmón	
	<i>Parámetros de discriminación</i>		<i>Parámetros de discriminación</i>	
Método de Clustering	SIM_{aux}	fc_{aux}	SIM_{aux}	fc_{aux}
Sólo SIM	0.59	-	0.4	-
Sólo FC	-	235	-	259
SIM+FC	0.6	261	0.36	293
Parámetro Sparse, λ	0.55		0.73	
Número de bases de corazón	18		-	
Número de bases de pulmón	-		29	

Tabla 7 - Parámetros óptimos para la maximización de zonas de corazón o zonas de pulmón del nuevo modelo de factorización

6.2.2. Parámetros óptimos del sistema NMF Semisupervisado

Al igual que en el caso del sistema NMF No Supervisado, se calculan tres conjuntos de parámetros: parámetros para la maximización de SDR de sonidos de corazón, para la maximización de SDR de sonidos de pulmón y para la maximización de SDR media entre ambos sonidos o maximización equilibrada. Los parámetros optimizados han sido: el número de bases de pulmón requeridas para la factorización, el valor Sparse λ_c para sonidos de corazón y el valor Sparse λ_p para sonidos de pulmón. Para dicha optimización se ha seguido el mismo análisis hiperparamétrico utilizado en el sistema No Supervisado. Así mismo, no se optimiza el número de bases de corazón empleadas durante la factorización, ya que dicho valor se determinará en función del diccionario W_c utilizado y vendrá establecido por el factor $K * M$.

Parámetros óptimos para la maximización de corazón			Parámetros óptimos para la maximización de pulmón			Parámetros óptimos para la maximización equilibrada		
<i>Parámetros optimizados</i>			<i>Parámetros optimizados</i>			<i>Parámetros optimizados</i>		
λ_c	λ_p	b_p	λ_c	λ_p	b_p	λ_c	λ_p	b_p
0.51	0.47	64	0.44	0.51	64	0.49	0.54	64

Tabla 8 - Parámetros óptimos según el tipo de maximización de corazón en el sistema Semisupervisado

6.3. Testeo de sistemas

Una vez obtenidos los parámetros óptimos para ambos sistemas, se realiza una fase de testeo de estos. Esta fase, nos permitirá determinar de forma objetiva la eficiencia de ambos sistemas en la separación de sonidos de corazón y pulmón. Con dicho fin, se utilizan los parámetros óptimos anteriormente calculados y se obtienen los resultados SDR, SIR y SAR para ambos sistemas como la media de cada uno de estos tras ejecutar tres veces el sistema, dado que el proceso de factorización es un procedimiento aproximado.

6.3.1. Testeo del sistema NMF No Supervisado.

Para el testeo del sistema No Supervisado, se calculan las medidas de SDR, SIR y SAR utilizando cada uno de los posibles métodos de Clustering con y sin la restricción Sparse, cuando se quiere maximizar los resultados SDR de corazón, de pulmón y ambos a la vez, utilizándose los parámetros óptimos establecidos en el apartado 6.2.1 para cada tipo de maximización. Para la selección de la metodología que en cada caso proporciona mejores resultados, se analizarán los valores de SDR y SIR en (dB) obtenidos de forma conjunta.

Así en las siguientes gráficas, se muestran los resultados de SDR, SIR y SAR en dB medios para sonidos de corazón obtenidos con todas las posibles configuraciones de parámetros:

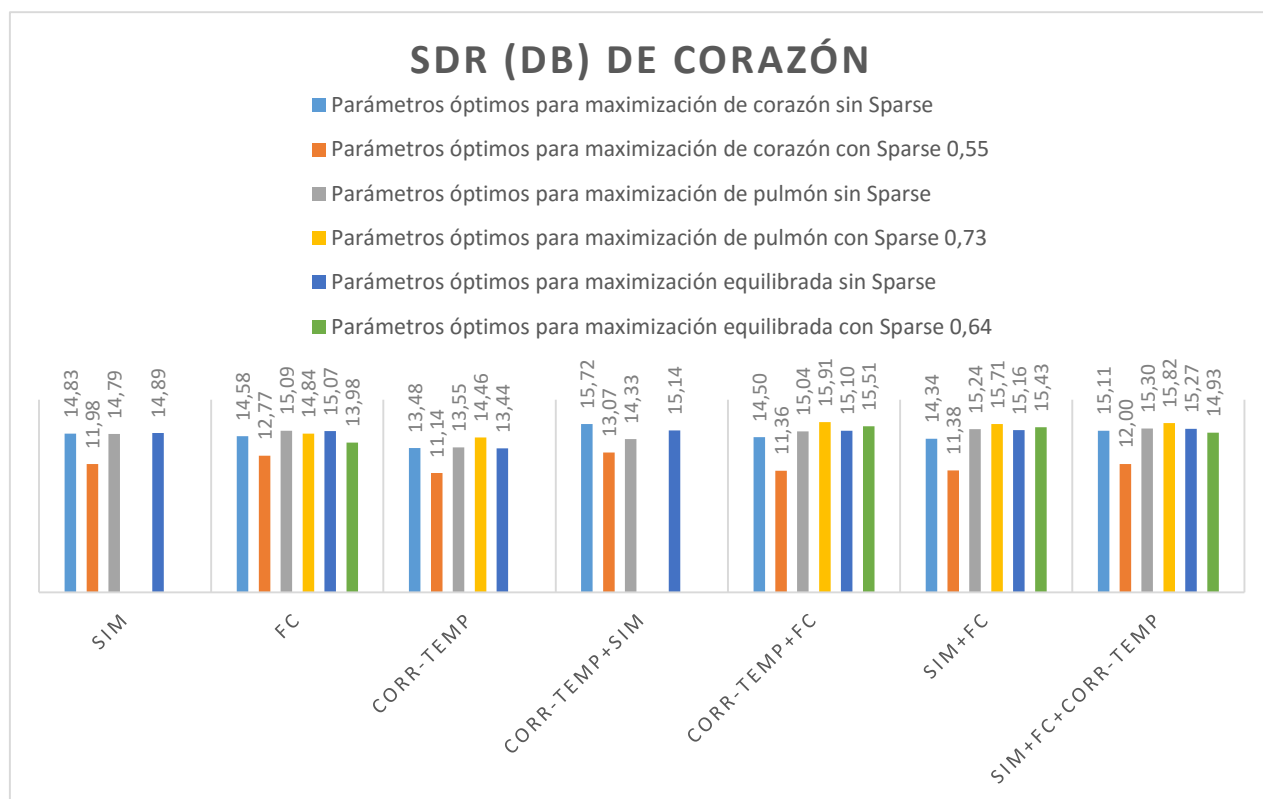


Figura 36- Figura resumen de resultados SDR(dB) de corazón obtenidos por todos los métodos.

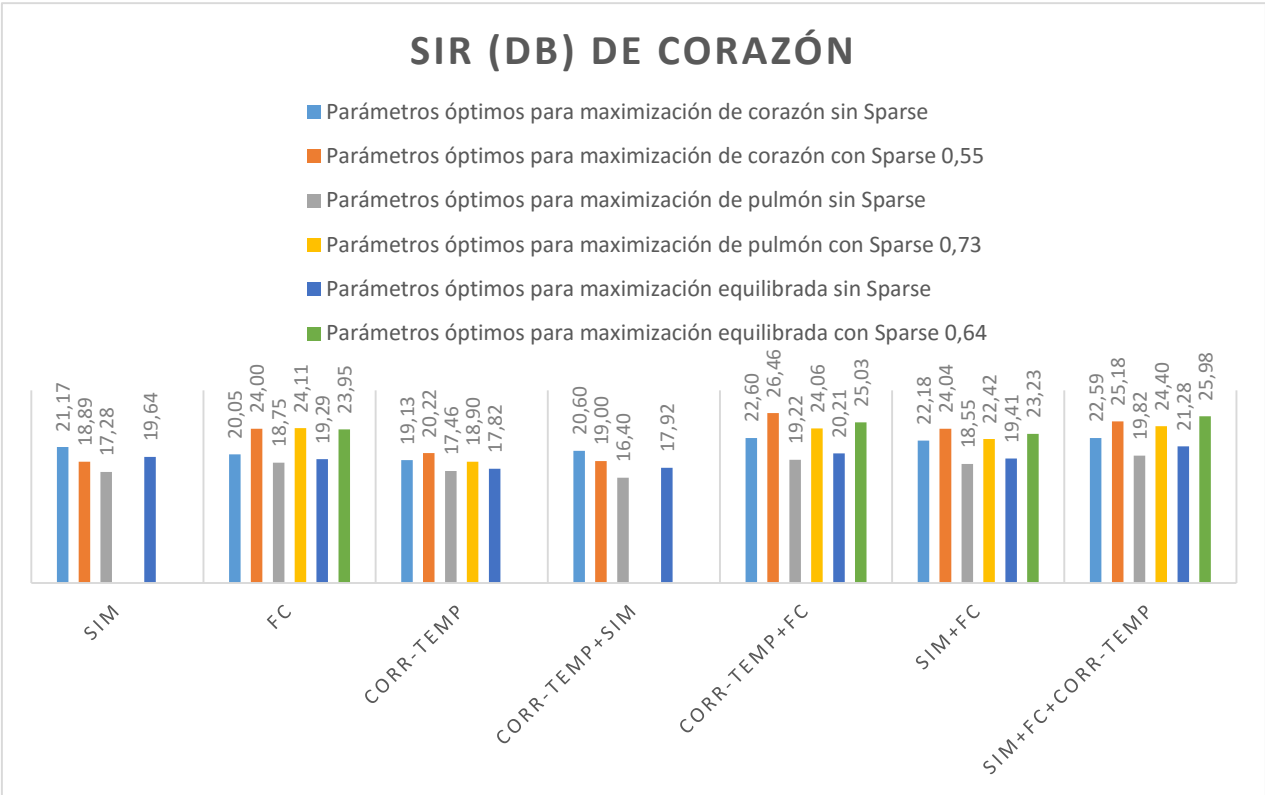


Figura 37 - Figura resumen de resultados SIR(dB) de corazón obtenidos por todos los métodos.

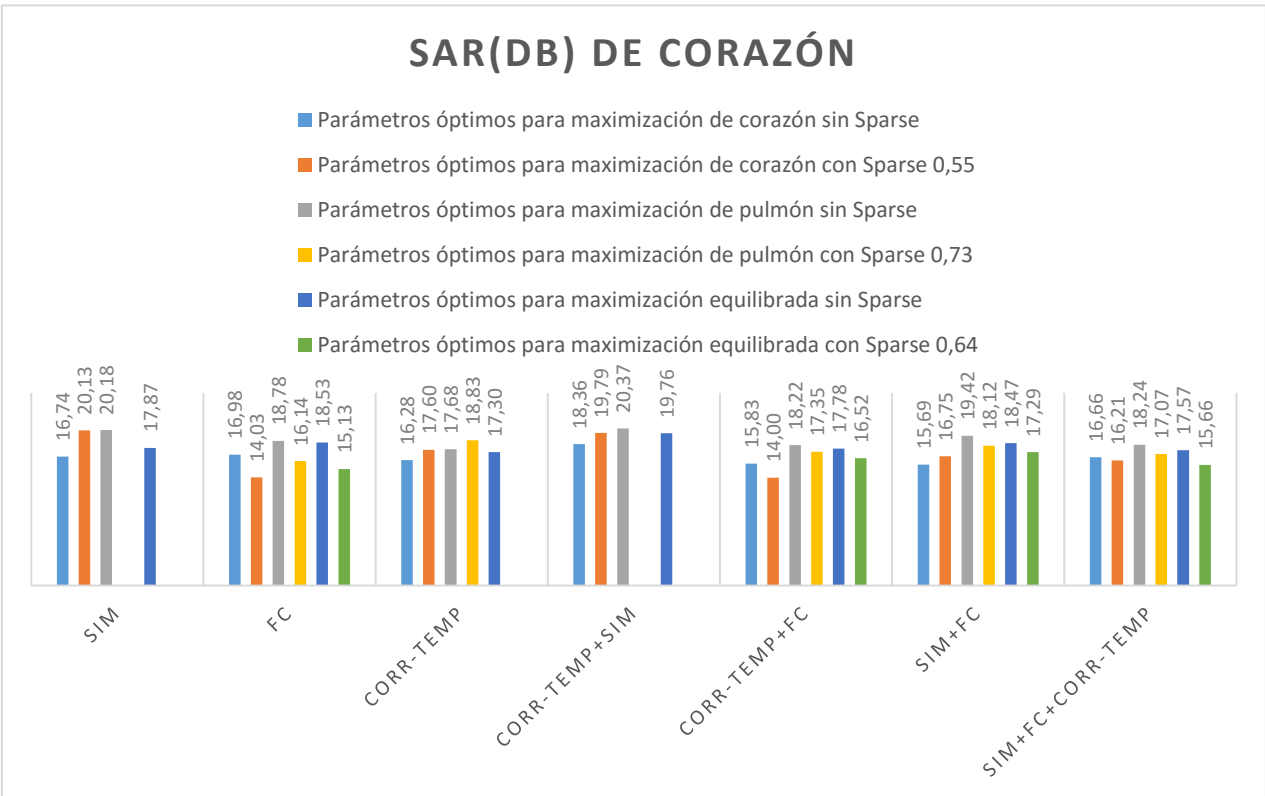


Figura 38 - Figura resumen de resultados SAR(dB) de corazón obtenidos por todos los métodos.

Atendiendo a los datos recogidos en las figuras anteriores, se observa como conjuntamente los resultados de SDR y SIR en dB se maximizan para sonidos de corazón cuando se utiliza la metodología **SIM+FC+CORR-TEMP** utilizando los parámetros óptimos que se habían establecido para la maximización de pulmón con Sparse de 0.73, columna amarilla.

A continuación, en las siguientes gráficas, se muestran los resultados de SDR, SIR y SAR en dB medios para sonidos de pulmón obtenidos con todas las posibles configuraciones de parámetros:

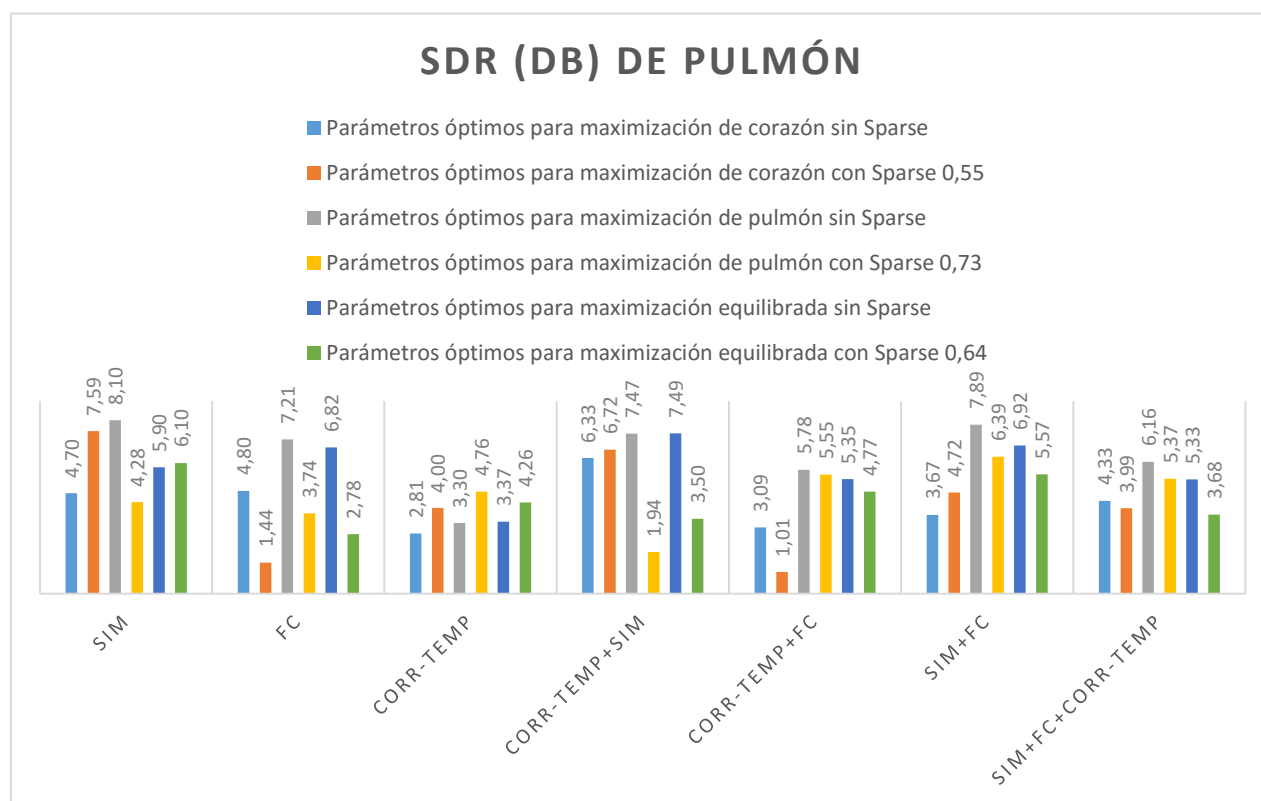


Figura 39 - Figura resumen de resultados SDR(dB) de pulmón obtenidos por todos los métodos.

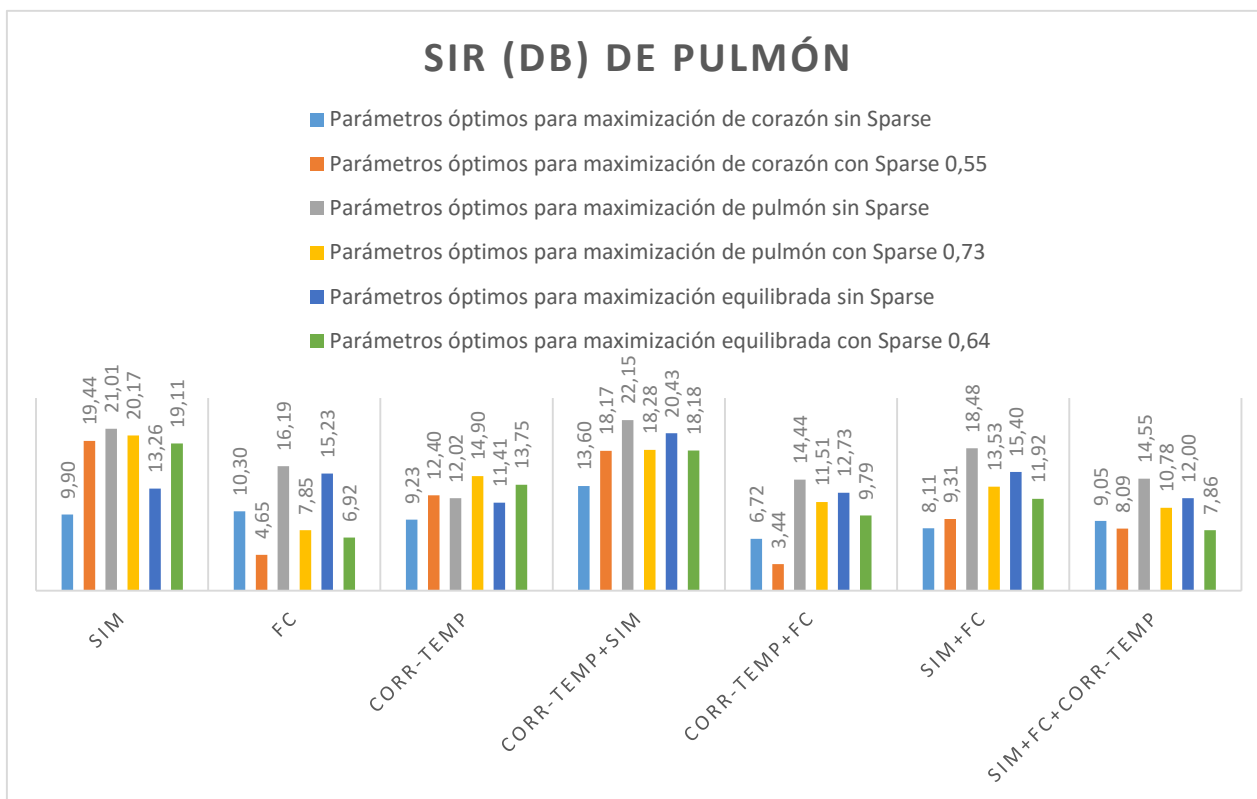


Figura 40 - Figura resumen de resultados SIR(db) de pulmón obtenidos por todos los métodos.

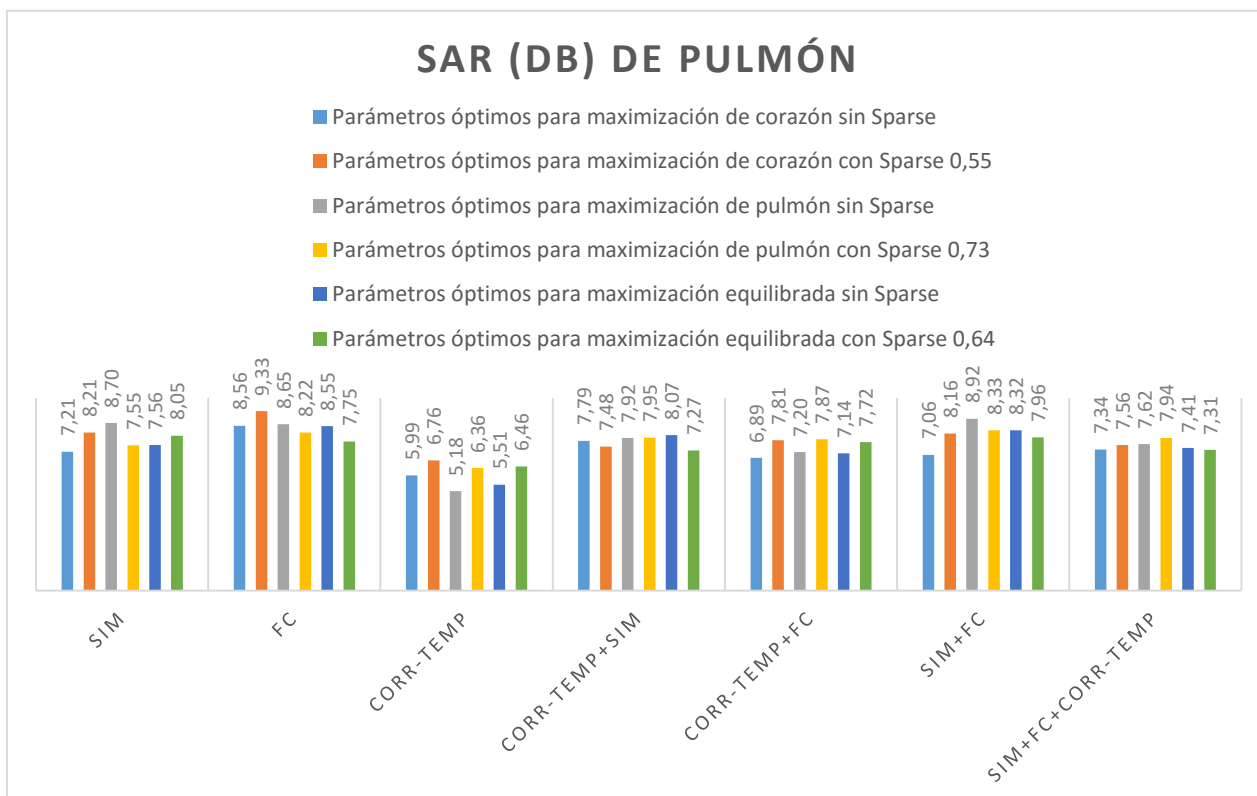


Figura 41 - Figura resumen de resultados SAR(db) de pulmón obtenidos por todos los métodos.

Atendiendo a los datos recogidos en las figuras anteriores, se observa como conjuntamente los resultados de SDR y SIR en dB se maximizan para sonidos de pulmón cuando se utiliza la metodología **SOLO SIM** utilizando los parámetros óptimos que se habían establecido para la maximización de pulmón sin Sparse, columna gris.

A continuación, en las siguientes gráficas, se muestran los resultados de SDR, SIR y SAR en dB medios entre ambos sonidos obtenidos, para la maximización equilibrada entre sonidos de corazón y pulmón, con todas las posibles configuraciones de parámetros:

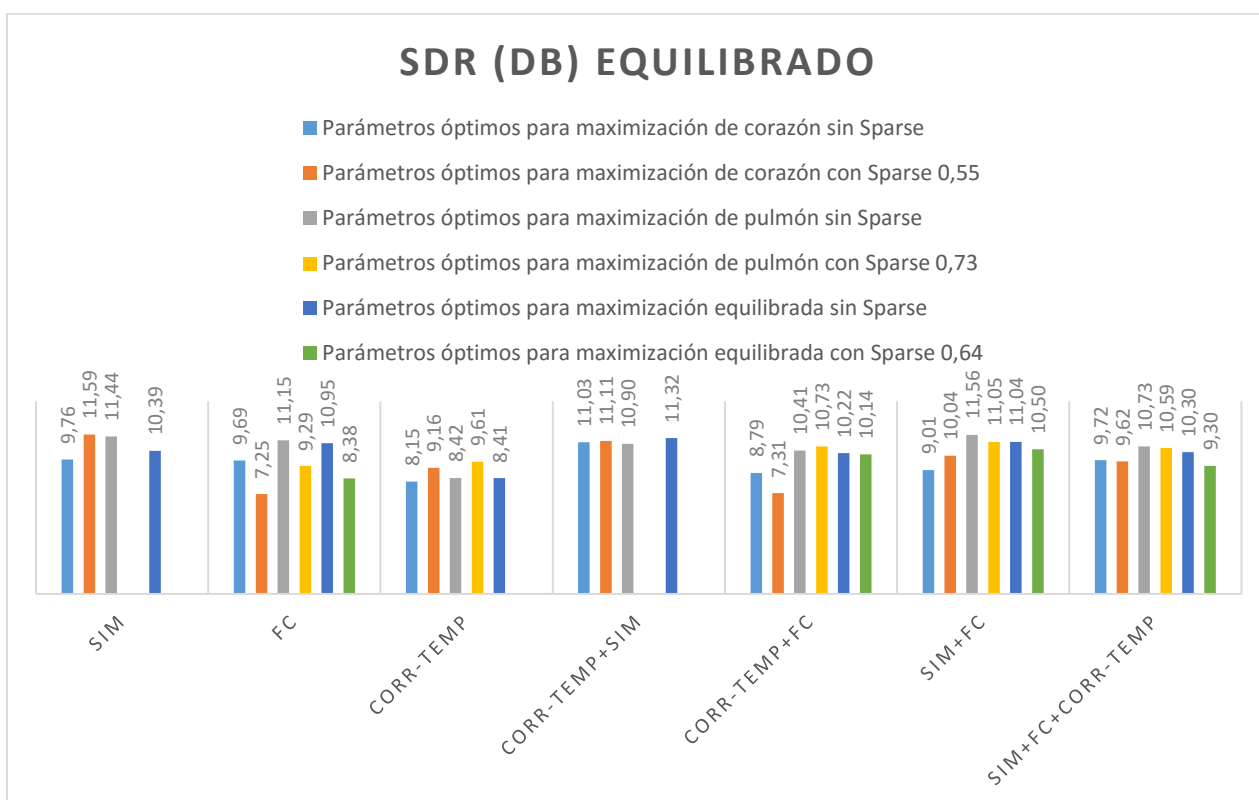


Figura 42 - Figura resumen de resultados SDR(dB) equilibrado entre ambos sonidos obtenidos por todos los métodos.

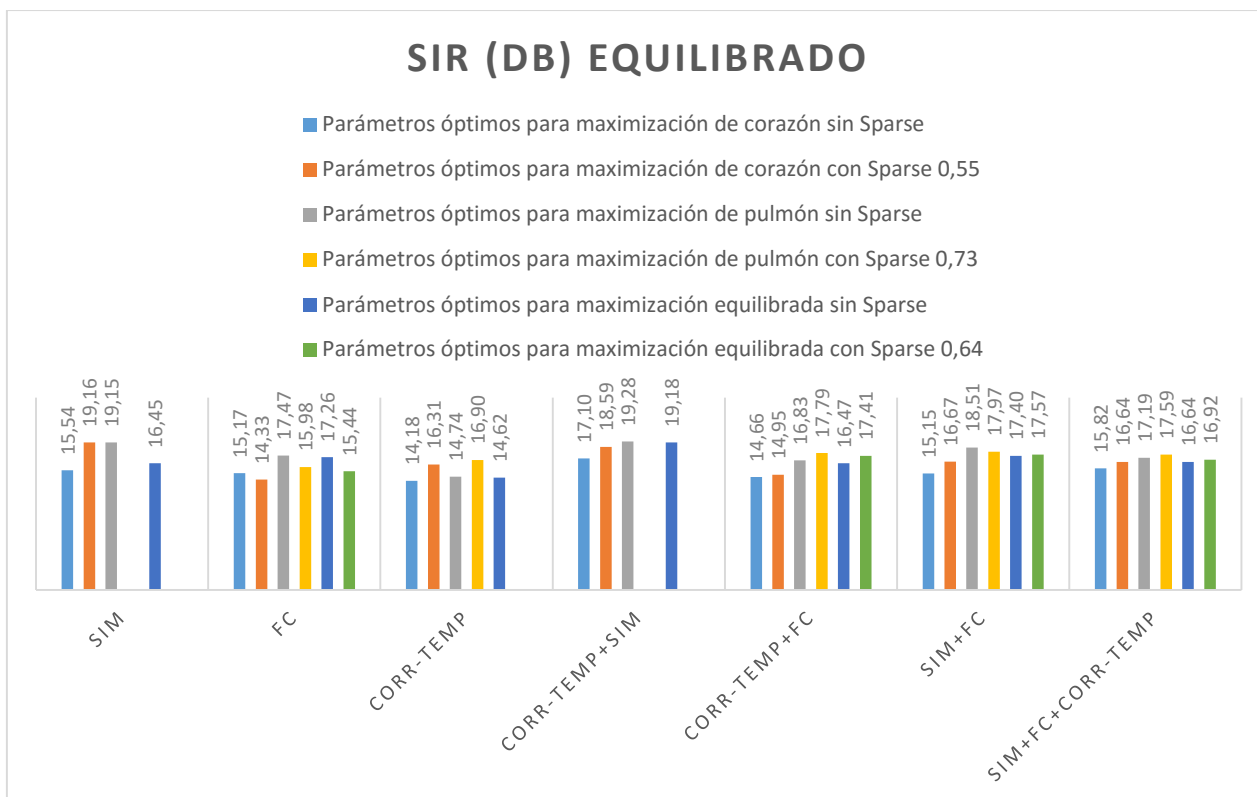


Figura 43 - Figura resumen de resultados SIR(dB) equilibrado entre ambos sonidos obtenidos por todos los métodos.

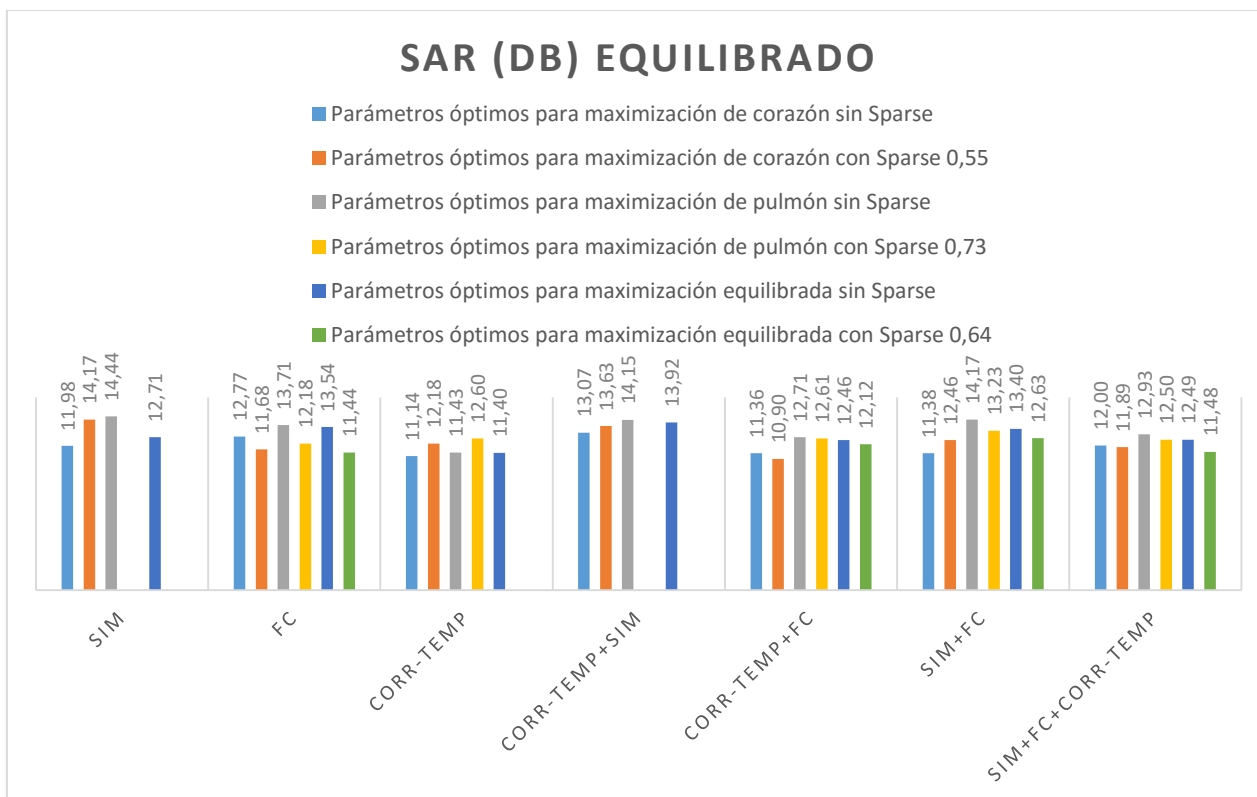


Figura 44 - Figura resumen de resultados SAR(dB) equilibrado entre ambos sonidos obtenidos por todos los métodos.

Atendiendo a los datos recogidos en las figuras anteriores, se observa como conjuntamente los resultados de SDR y SIR en dB se maximizan ambos sonidos simultáneamente cuando se utiliza la metodología **SOLO SIM** utilizando los parámetros óptimos que se habían establecido para la maximización de corazón con Sparse 0.55, columna naranja.

Finalmente, a modo de resumen, los resultados óptimos del sistema NMF No Supervisado se obtienen cuando:

- **Maximizar resultados corazón:** Se utiliza la metodología SIM+FC+CORR-TEMP con la configuración de parámetros óptimos establecida para maximizar pulmón con Sparse de 0.73, Tabla X.
- **Maximizar resultados pulmón:** Se utiliza la metodología SOLO SIM con la configuración de parámetros óptimos establecida para maximizar pulmón sin Sparse, Tabla X.
- **Maximizar resultados de ambos sonidos simultáneamente:** Se utiliza la metodología SOLO SIM con la configuración de parámetros óptimos establecida para maximizar corazón con Sparse 0.55, Tabla X.

Los resultados medios finales para las medidas SDR, SIR y SAR para cada maximización, usando el método de Clustering apropiado, son:

Resultados finales Sistema NMF No Supervisado			
Maximización/Tipo de medida	SDR (dB)	SIR(dB)	SAR(dB)
Maximizar resultados corazón	15,82	24,40	17,07
Maximizar resultados pulmón	8,10	21,00	8,70
Maximizar resultados de ambos sonidos simultáneamente	11,59	19,16	14,17

Tabla 9 – Resultados finales Sistema NMF No Supervisado

Como podemos comprobar, el sistema NMF No Supervisado proporciona resultados muy buenos en la extracción de sonidos de corazón y pulmón. Además, puede afirmarse que el método más restrictivo en la extracción de sonidos de pulmón ha resultado ser el cálculo de la correlación espectral o SIM, pudiéndose deber a la falta de información espacial de dichos sonidos, aprovechándose únicamente su características espectrales; por tanto, el uso de la metodología ‘SOLO SIM’ permite maximizar los resultados SDR, SIR y SAR obtenidos para los sonidos de pulmón. Por otro lado, se comprueba como el uso conjunto de las tres técnicas de Clustering implementadas en este TFM, metodología llamada ‘SIM+FC+CORR-TEMP’, permite obtener los mejores resultados en la extracción de sonidos de corazón, debiéndose principalmente a que de este modo se explotan las características espectrales y espaciales de forma conjunta para dichos sonidos. Por último, se comprueba como la restricción Sparse tomada sobre la matriz de activaciones temporales en este sistema supone un aumento de los resultados obtenidos para sonidos de corazón, no suponiendo mejora para los resultados de sonidos de pulmón.

6.3.1.1. Testeo del sistema NMF No supervisado usando el modelo de factorización basado en el patrón rítmico

En base al análisis de resultados realizado en el apartado anterior, se ha decidido utilizar las siguientes metodologías para la fase de Clustering: uso conjunto de **la correlación espectral y cálculo del Roll-off** para los sonidos de corazón, siendo esta combinación la siguiente mejor a la elegida en el apartado anterior para dichos sonidos, y uso de **la correlación espectral** para sonidos de pulmón, siendo esta técnica la siguiente mejor a la elegida en el apartado anterior para dichos sonidos. Para esta fase de testeo se han utilizado los parámetros óptimos establecidos en las tablas X y X.

Así los resultados medios obtenidos, usando este nuevo modelo de factorización, han sido:

Maximización corazón		
<i>SDR (dB)</i>	<i>SIR(dB)</i>	<i>SAR(dB)</i>
7.299	14.8209	8.6438

Tabla 10 – Resultados finales de maximización de corazón utilizando el nuevo método de factorización

En cuanto a los resultados de sonidos de pulmón se han obtenido los siguientes valores medios de SDR, SIR y SAR:

Maximización pulmón		
<i>SDR (dB)</i>	<i>SIR(dB)</i>	<i>SAR(dB)</i>
-5.4991	-0.3792	1.5693

Tabla 11 – Resultados finales de maximización de pulmón utilizando el nuevo método de factorización

Como puede comprobarse, en base a los resultados, el método de factorización basado en bloques no ha supuesto una mejora en el sistema de factorización NMF No Supervisado como se esperaba, probablemente debido a que no puede predecirse de forma exacta el número de bases correcto para las zonas con presencia de sonidos de corazón, debido principalmente a que:

- En dichas zonas podrían aparecer solamente sonidos de corazón. En este caso, sería necesaria una única base en la factorización o dos bases si consideramos que podría aparecer algún tipo de ruido, ya que las señales utilizadas durante este TFM no son señales de corazón totalmente puras. En otro caso, el uso de un número elevado de bases, cuando encontramos este caso, podría suponer descomponer el patrón espectral de un sonido de corazón en muchos subpatrones espectrales diferentes, lo que llevaría a un empeoramiento sustancial de los resultados.
- En dichas zonas podrían aparecer sonidos de pulmón solapados a los sonidos de corazón. En este caso, el uso de una única base en la factorización no podría garantizar la correcta extracción de ambos sonidos.

Por tanto, como no puede predecirse a priori el tipo de escenario, no se puede establecer con seguridad el número correcto de bases para necesarias para la factorización de zonas con presencia de corazón.

6.3.2. Testeo del sistema NMF Semisupervisado.

Para el testeo del sistema NMF Semisupervisado, se ha seguido un procedimiento similar al tomado para la evaluación de sistema No Supervisado. Por esto, una vez se han establecido los valores óptimos para cada tipo de maximización en este sistema, se calculan los resultados obtenidos en los casos de maximización de SDR de corazón, pulmón y ambos a la vez, para cada una de las señales mezcla a testear. Así mismo, se ha determinado el número de bases K necesarias en la creación de diccionarios W_c de patrones de corazón que maximiza cada caso comparando los resultados obtenidos cuando $K = 2,4,8$.

Así en las siguientes gráficas, se muestran los resultados de SDR, SIR y SAR medios en dB para sonidos de corazón obtenidos con todas las posibles configuraciones de parámetros:

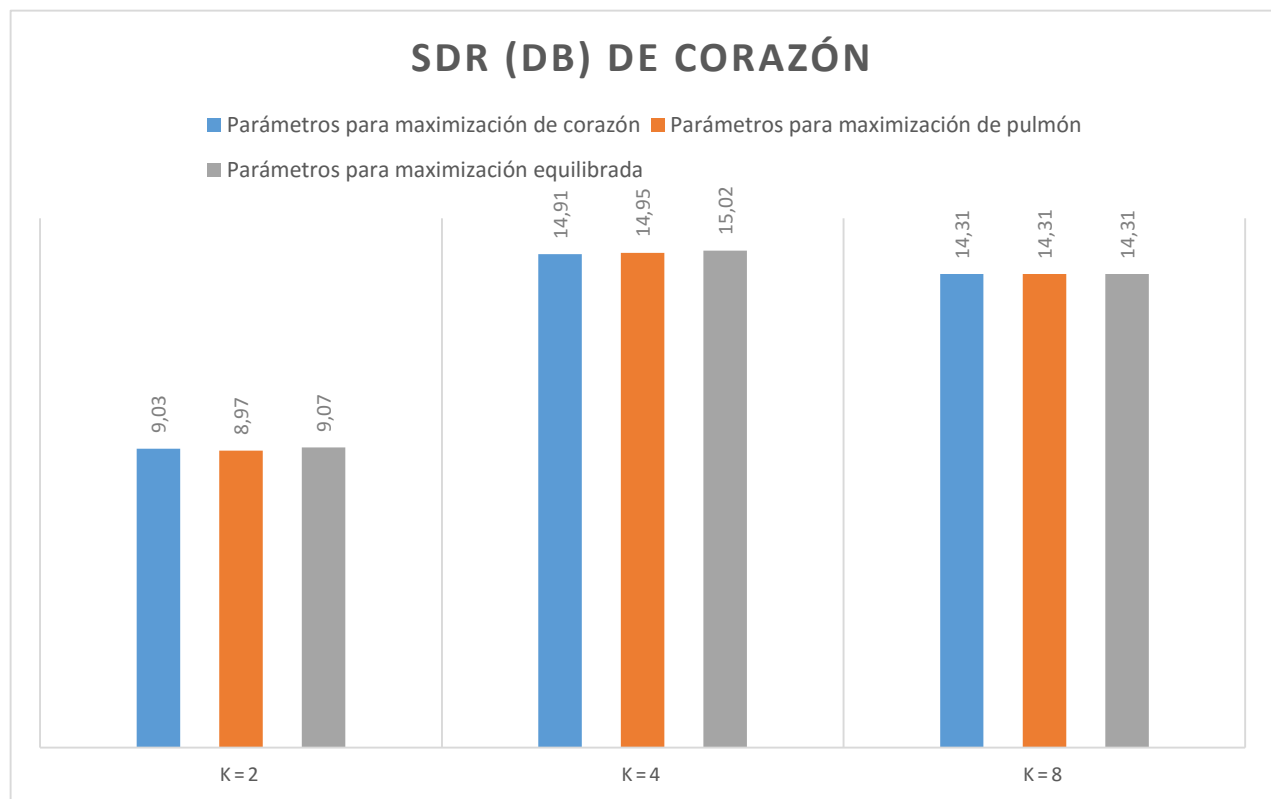


Figura 45 - Figura resumen de resultados SDR(dB) de corazón para distintos valores de K .

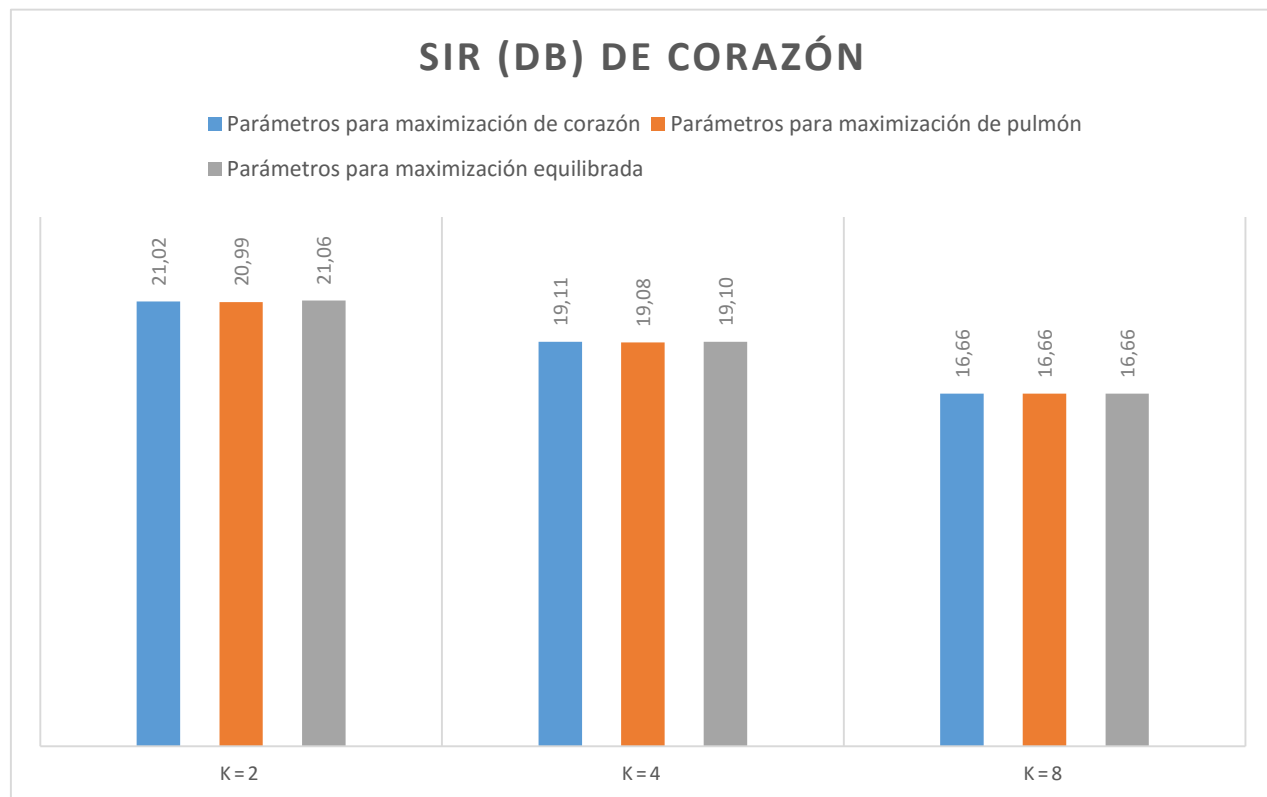


Figura 46 - Figura resumen de resultados SIR(dB) de corazón para distintos valores de K.

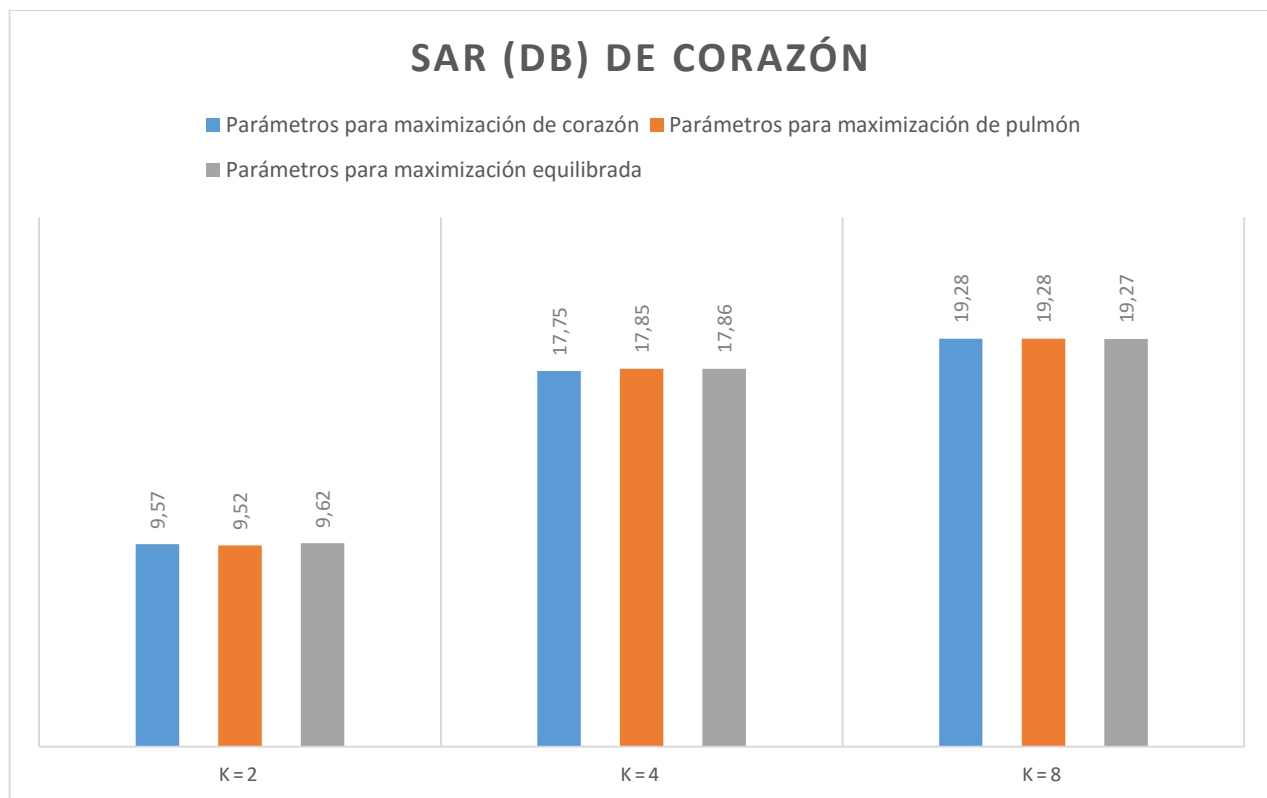


Figura 47 - Figura resumen de resultados SAR(dB) de corazón para distintos valores de K.

Atendiendo a los datos recogidos en las figuras anteriores, se observa como los resultados de SDR en dB se maximizan para sonidos de corazón cuando se usa un número K de bases igual a 4. Dado que para los tres tipos de configuraciones en dicho caso proporcionan resultados similares, se determinará que la configuración de parámetros óptimos para maximización de corazón es aquella que ofrece mejores resultados en este caso, azul.

A continuación, en las siguientes gráficas, se muestran los resultados de SDR, SIR y SAR medios en dB para sonidos de pulmón obtenidos con todas las posibles configuraciones de parámetros:

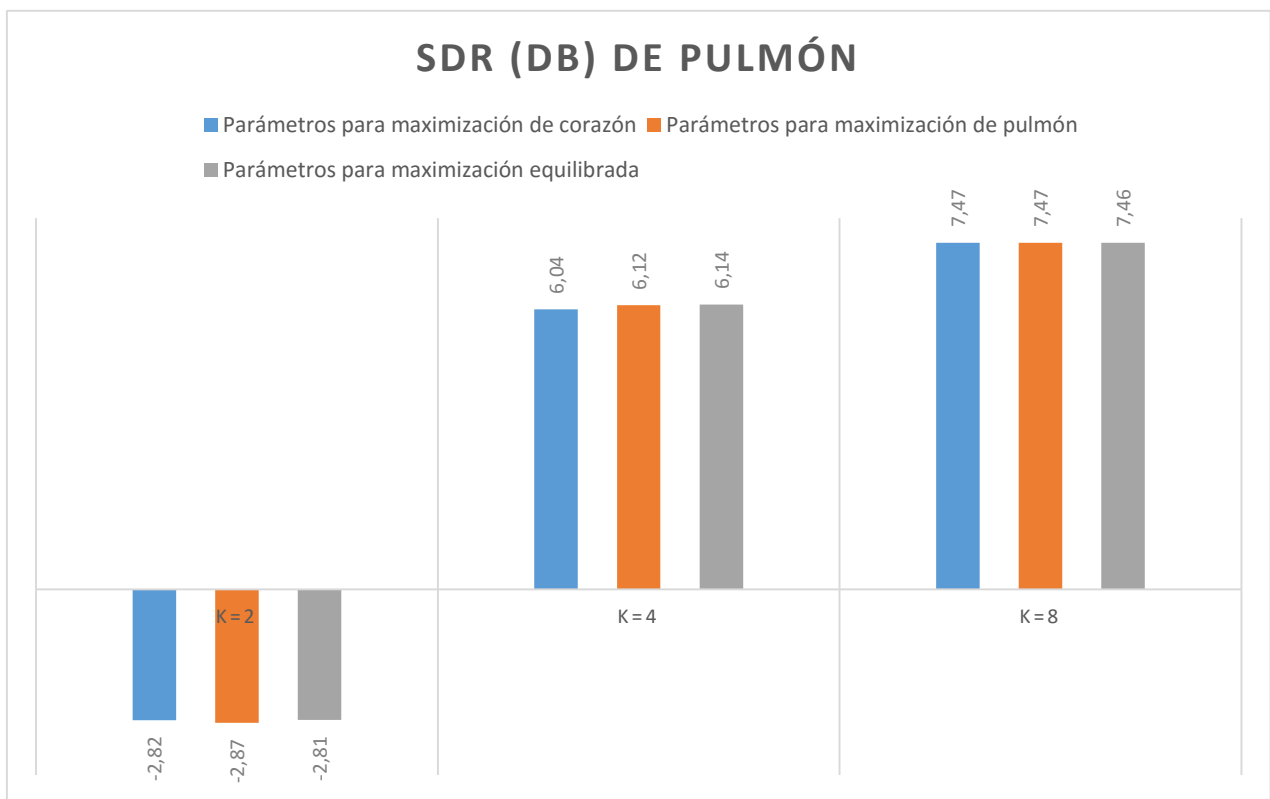


Figura 48 - - Figura resumen de resultados SDR(dB) de pulmón para distintos valores de K.

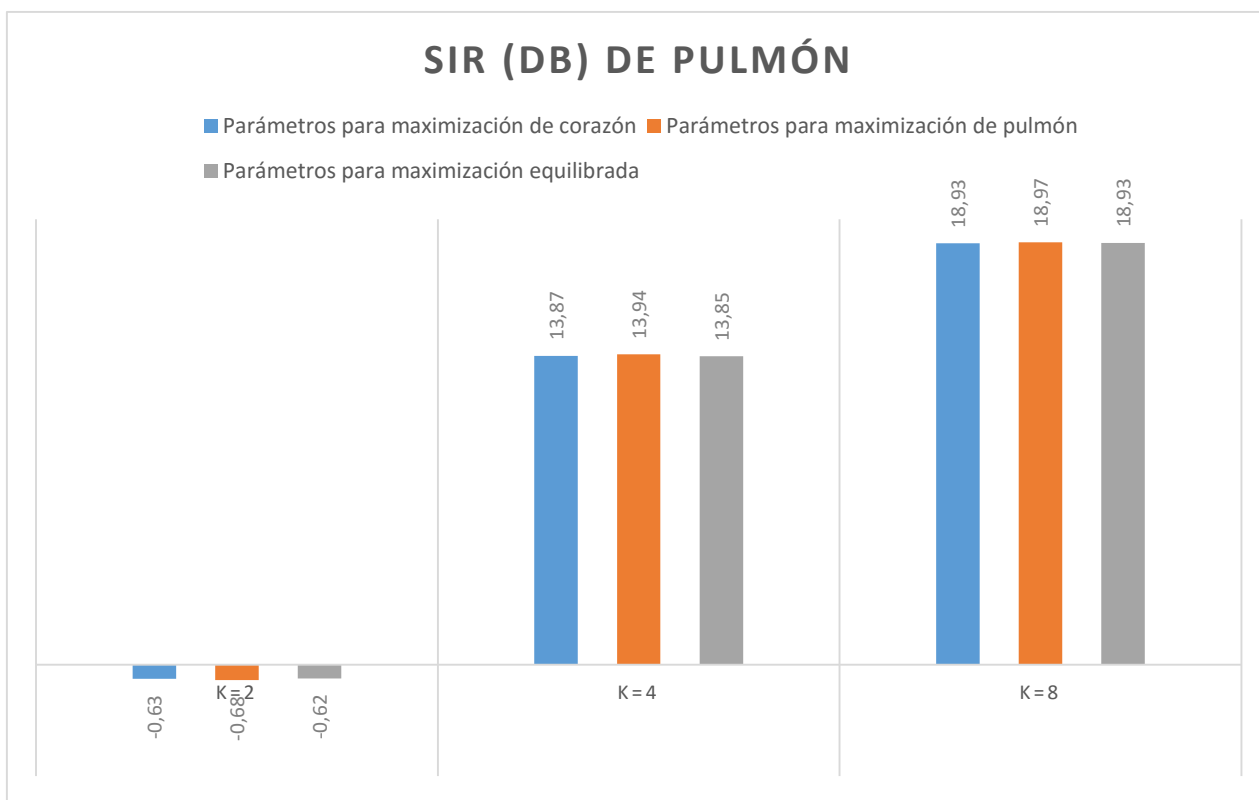


Figura 49 - Figura resumen de resultados SIR(dB) de pulmón para distintos valores de K.

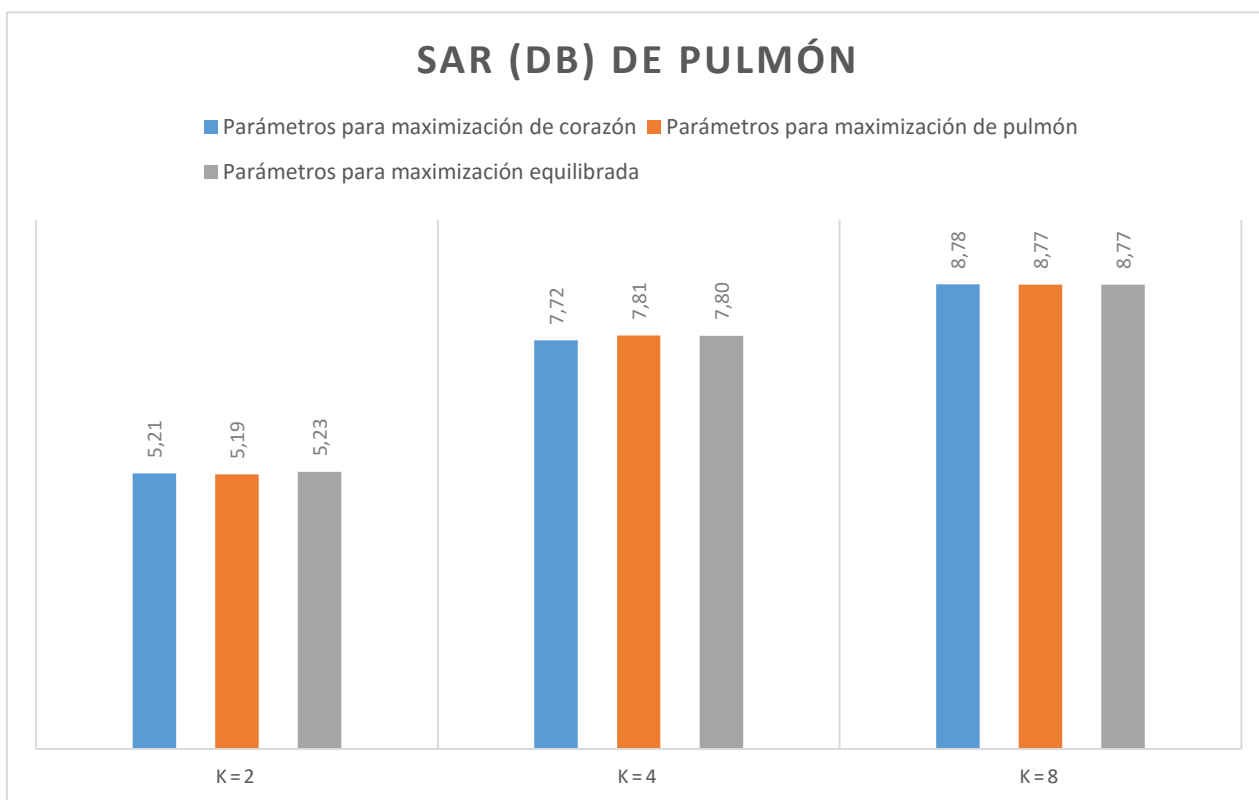


Figura 50 - - Figura resumen de resultados SAR(dB) de pulmón para distintos valores de K.

Atendiendo a los datos recogidos en las figuras anteriores, se observa como los resultados de SDR en dB se maximizan para sonidos de pulmón cuando se usa un número K de bases igual a 8. Dado que para los tres tipos de configuraciones en dicho caso proporcionan resultados similares, se determinará que la configuración de parámetros óptimos para maximización de pulmón es aquella que ofrece mejores resultados en este caso, naranja.

A continuación, en las siguientes gráficas, se muestran los resultados de SDR, SIR y SAR medios en dB para sonidos de pulmón obtenidos con todas las posibles configuraciones de parámetros:

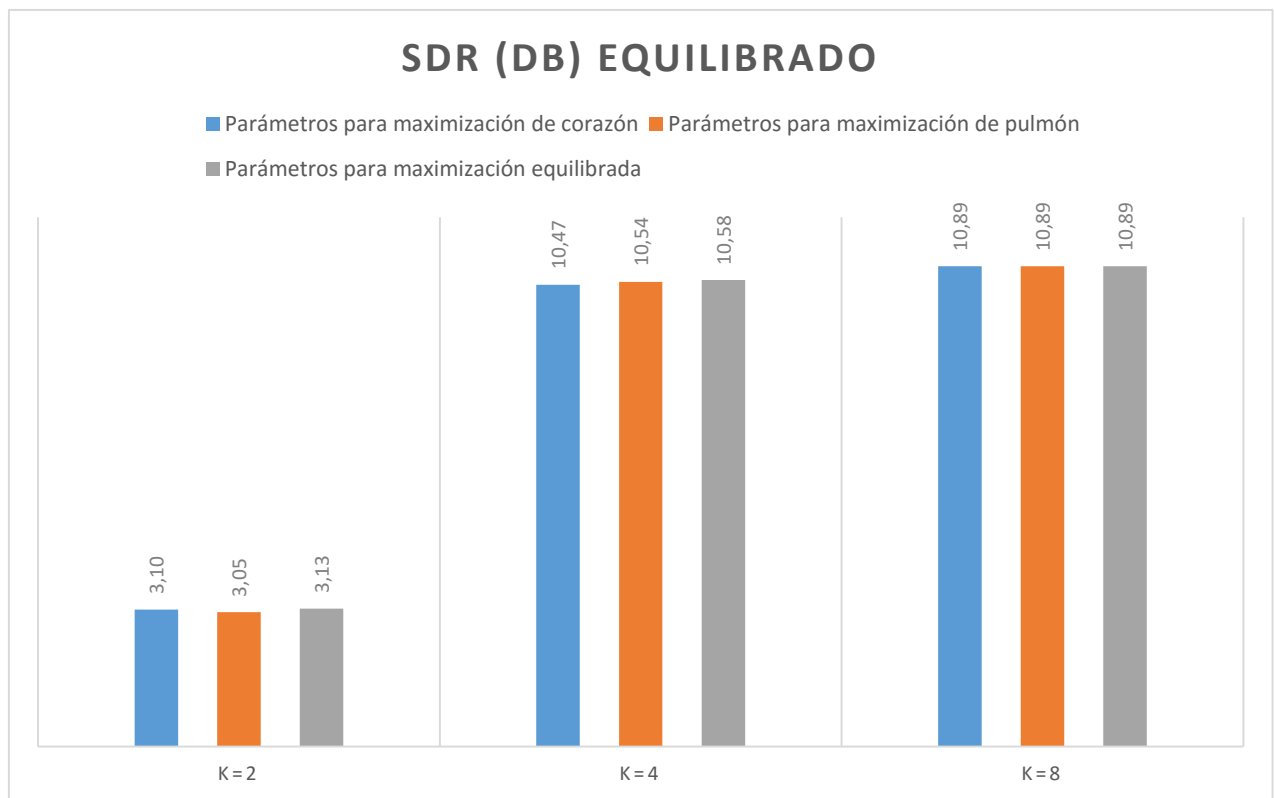


Figura 51 - Figura resumen de resultados SDR(dB) equilibrado para distintos valores de K.

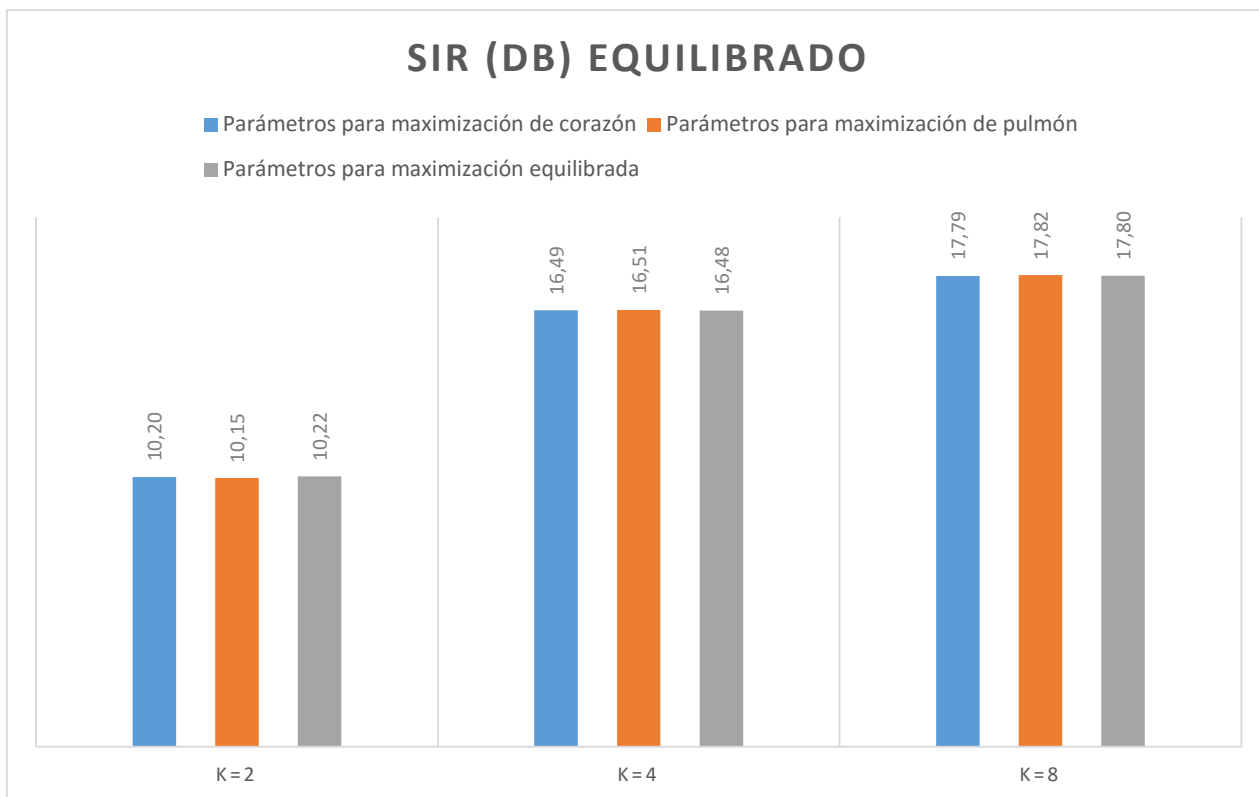


Figura 52 - Figura resumen de resultados SIR(dB) equilibrado para distintos valores de K.

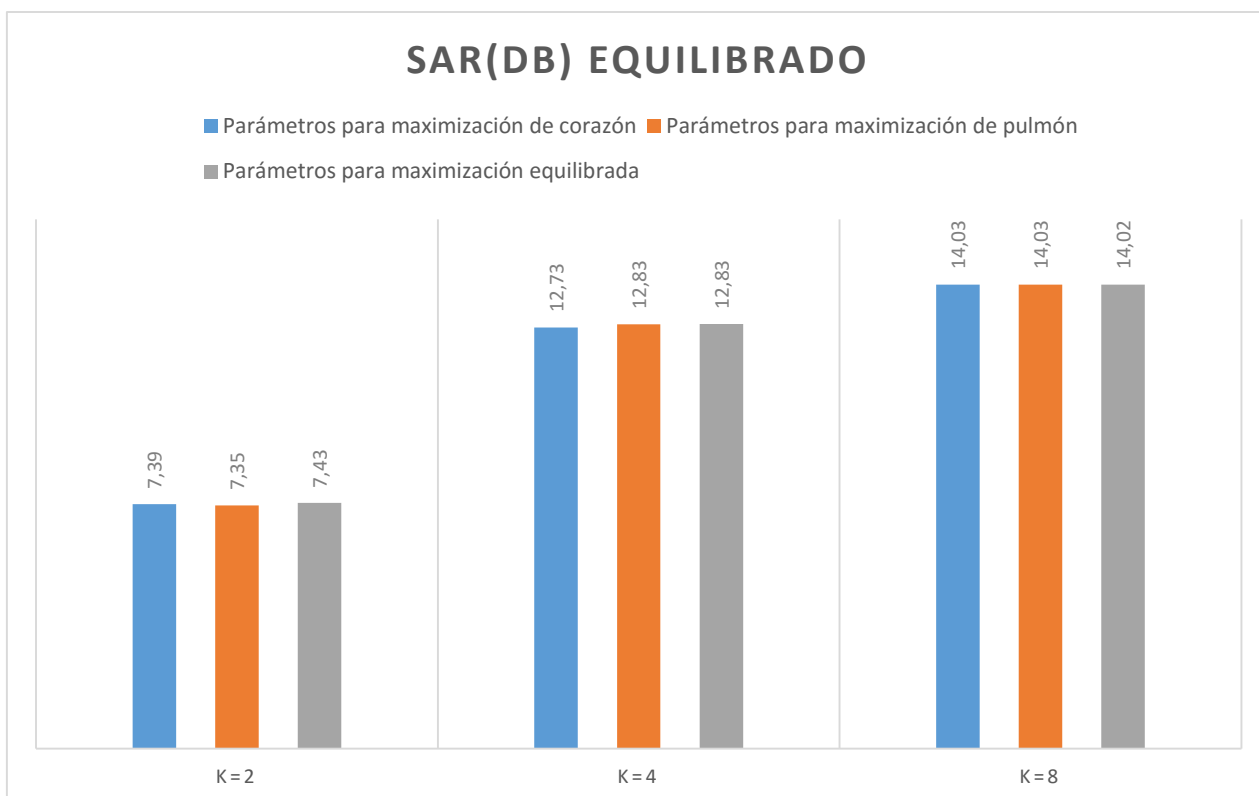


Figura 53 - Figura resumen de resultados SAR(dB) equilibrado para distintos valores de K.

Atendiendo a los datos recogidos en las figuras anteriores, se observa como los resultados de SDR en dB se maximizan para ambos sonidos simultáneamente cuando se usa un número K de bases igual a 8. Dado que para los tres tipos de configuraciones en dicho caso proporcionan resultados similares, se determinará que la configuración de parámetros óptimos para maximización equilibrada es aquella que ofrece mejores resultados en este caso, gris.

Los resultados medios finales para las medidas SDR, SIR y SAR para cada maximización, usando el método de Clustering apropiado, son:

Resultados finales Sistema NMF Semisupervisado			
<i>Maximización/Tipo de medida</i>	<i>SDR (dB)</i>	<i>SIR(dB)</i>	<i>SAR(dB)</i>
Maximizar resultados corazón	14,91	19,11	17,75
Maximizar resultados pulmón	7,47	18,97	8,77
Maximizar resultados de ambos sonidos simultáneamente	10,89	17,80	14,02

Tabla 12 - Resultados finales Sistema NMF Semisupervisado

Finalmente, atendiendo a los resultados obtenidos en este apartado, puede afirmarse que el sistema NMF Semisupervisado, ofrece muy buenos resultados en la extracción de sonidos de corazón y pulmón. Así mismo, se concluye que son necesarias al menos 4 bases K para la correcta caracterización espectral de sonidos de corazón.

6.4. Testeo del algoritmo de estimación del ritmo cardíaco

También, se ha calculado la eficiencia del algoritmo de estimación del ritmo cardíaco. Para ello, se han tenido en cuenta tres posibles condiciones en el proceso de detección de sonidos de corazón: True Positives, Falses Positives y Falses Negatives. Así la eficiencia del algoritmo se calcula como:

$$Eficiencia(\%) = \frac{TP - FP - FN}{TS} * 100 \quad (77)$$

Siendo TP el número total de sonidos correctamente detectados, FP el número de falsos positivos, FN el número de falsos negativos y TS el número total de sonidos de corazón a detectar. En la siguiente tabla se muestran los resultados obtenidos para los sonidos analizados:

Señales mezcla	TS	TP	FP	FN
Mezcla 1	19	17	3	0
Mezcla 14	23	22	0	1
Mezcla 26	19	18	6	1
Mezcla 35	21	21	3	0
Mezcla 44	17	16	0	1
Mezcla 57	21	20	1	1
Mezcla 69	15	15	7	0
TOTAL	135	129	20	4

Tabla 13 - Evaluación del algoritmo de estimación del ritmo cardíaco

Aplicando los datos de la tabla anterior se calcula la eficiencia como:

$$Eficiencia(\%) = \frac{129 - 20 - 4}{135} * 100 \approx 78\%$$

Finalmente, se obtiene que dicho algoritmo tiene una eficiencia del **78%**.

CAPÍTULO 6- RESULTADOS

CAPÍTULO 7

7. Conclusiones

Finalmente, tras comprobar el correcto funcionamiento de los sistemas NMF Semisupervisado y No Supervisado desarrollados en este TFM, así como la correcta integración de ambos junto al algoritmo para la estimación del ritmo cardíaco y la funcionalidad de búsqueda de patrones en una herramienta gráfica sencilla, se ha alcanzado con éxito el objetivo principal de este TFM, la creación de una herramienta que permita la obtención por separado de aquellos sonidos procedentes del corazón humano que han sido interferidos por sonidos procedentes del sistema respiratorio durante el proceso de auscultación y que, además, permita la búsqueda de patrones repetitivos dentro de dicha señal.

Atendiendo a los resultados obtenidos durante la fase de testeo de ambos sistemas, se ha llegado a las siguientes conclusiones:

En primer lugar, se observa como ambos sistemas de separación ofrecen resultados excepcionales en la separación entre sonidos de corazón y pulmón, siendo los resultados obtenidos por el sistema NMF No Supervisado ligeramente superiores a los obtenidos con el sistema NMF Semisupervisado.

En cuanto al sistema NMF No Supervisado, se comprueba como la utilización de las técnicas de clasificación de bases implementadas junto a él, permiten la correcta clasificación de bases entre ambos tipos de sonidos explotando tanto sus características espectrales como temporales. Además, se demuestra también como la correlación temporal, implementada con la información obtenida del algoritmo de estimación del ritmo cardíaco, mejora los resultados obtenidos para este sistema. Por otro lado, se comprueba como el modelo estándar de factorización ofrece mayores resultados a los obtenidos por el modelo de factorización basado en bloques, posiblemente debido a la presencia simultánea entre sonidos de corazón y pulmón

dentro de un mismo bloque en la mayoría de los casos, lo que conlleva a que los sonidos de corazón y pulmón no se separen correctamente y, por tanto, los resultados con este sistema se vean degradados. Por último, en este sistema, la utilización de la restricción Sparse ha resultado ser eficaz cuando se utiliza en la extracción de sonidos de corazón, debido probablemente a la simplicidad de dichos sonidos, mientras resulta inapropiada para la extracción de sonidos de pulmón.

En cuanto al sistema NMF Semisupervisado, se comprueba como la correcta caracterización espectral de los sonidos de corazón permite su correcta extracción, necesitándose al menos cuatro bases por sonido de corazón para su correcta caracterización durante la fase de entrenamiento. Además, se corrobora la asunción tomada de [32], demostrándose el correcto funcionamiento del sistema Semisupervisado cuando las fuentes a extraer presentan características espectrales suficientemente diferentes. Por último, para este sistema, se observa como la utilización de la restricción Sparse resulta eficaz la para la extracción tanto de sonidos de corazón como de pulmón.

Por otro lado, se ha demostrado la eficacia del algoritmo de estimación del ritmo cardíaco, siendo su eficiencia de un 78%. Sin embargo, aunque dicho algoritmo permite la detección de todos los sonidos de corazón en la mayoría de casos, su eficiencia se ve reducida en situaciones donde se encuentran sonidos con suficiente energía como para ser detectados. Los sonidos de corazón, que han sido utilizados durante este TFM, no son sonidos puros y cuentan con cierto nivel de ruido de baja frecuencia, hecho que ha contribuido a disminuir la eficiencia de este algoritmo.

Finalmente, la implementación de la funcionalidad de búsqueda de patrones repetitivos en las señales de corazón y pulmón se ha implementado con éxito, permitiendo de forma exitosa la búsqueda de aquellos sonidos detectados por el usuario con el grado de semejanza establecido para esta.

CAPÍTULO 7 - CONCLUSIONES

CAPÍTULO 8

8. Líneas de investigación futuras

En el siguiente apartado, se enumeran las líneas de futuro a seguir en una futura continuación del trabajo realizado en este TFM:

- Implementación en tiempo real de los sistemas desarrollados en este TFM.
- Desarrollo de un estetoscopio electrónico que permita almacenar sonidos resultantes del proceso de auscultación.
- Mejora del algoritmo del ritmo cardíaco que permita modelar en tiempo real el ritmo cardíaco.
- Creación de un servicio web que permita la separación entre ambos sonidos de forma remota.
- Mejora del sistema NMF No Supervisado basado en el modelo de factorización en bloques.
- Implementación de una funcionalidad que permita la detección de determinadas enfermedades de manera probabilística.
- Integración de nuevas técnicas de clasificación de bases que complementen a las ya integradas.

CAPÍTULO 8 – LINEAS DE INVESTIGACIÓN FUTURAS

ANEXO I

HERRAMIENTA CON INTERFAZ GRÁFICA

9. Herramienta con interfaz gráfica

Como parte final de este Trabajo Fin de Máster, se ha creado una herramienta gráfica donde se integran el sistema de separación NMF Semisupervisado y No Supervisado, así como la funcionalidad de búsqueda de patrones no repetitivos.

Dicha herramienta, se ejecuta a través del archivo ‘Separacion_sonidos_cardiopulmonares.exe’, donde se muestra una pantalla de bienvenida que dará acceso a dichos sistemas:



Figura 54 – Interfaz de bienvenida a la herramienta gráfica

En esta pantalla, se muestra cierta información vinculada al proyecto. Además, se han implementado dos botones para dar acceso a cada uno de los sistemas.

Si se accede al sistema NMF No Supervisado, encontramos la siguiente pantalla:

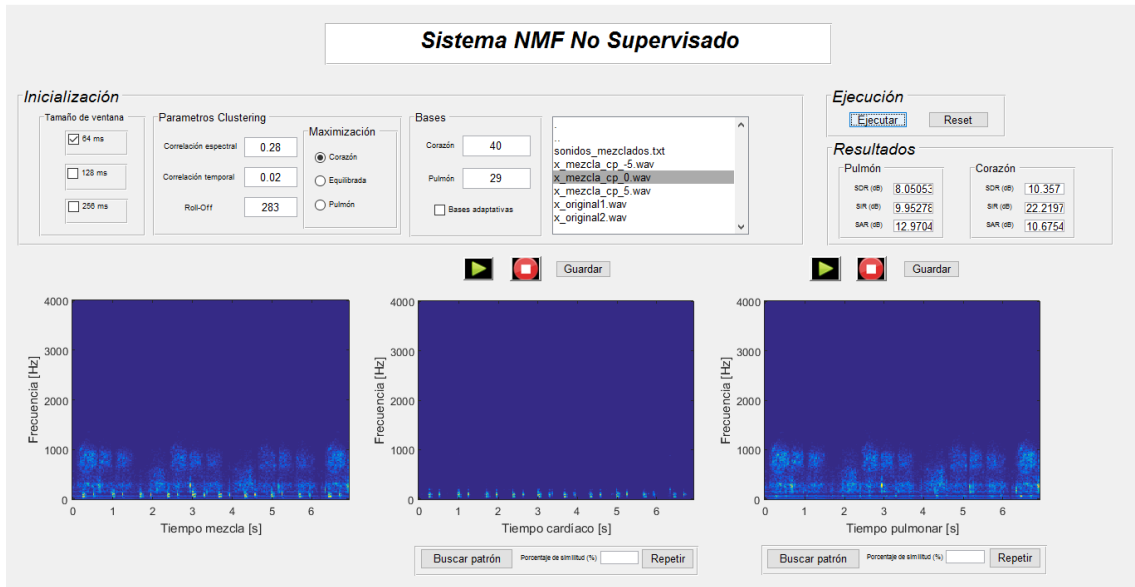


Figura 55 – Herramienta gráfica para el sistema NMF No Supervisado

En esta interfaz, encontramos tres secciones. Una primera sección donde se inician los parámetros necesarios para el proceso de descomposición NMF No Supervisado como son: la longitud de la ventana, los parámetros de discriminación utilizados para las técnicas de Clustering y una serie de opciones según se quiera maximizar los resultados para corazón, pulmón o de forma equilibrada, el número de bases utilizadas para el proceso de factorización, la opción de utilización del modelo de factorización basado en el patrón rítmico y un cuadro de navegación para la selección de la señal mezcla a descomponer.

Después, se encuentra una segunda sección para la ejecución del sistema o la inicialización predeterminada de parámetros. También, se encuentra una tercera sección donde se muestran los resultados SDR, SIR y SAR para los sonidos de corazón y pulmón.

En la primera figura de esta pantalla, se muestra el espectrograma de la señal mezcla; en las siguientes, se muestra el espectrograma reconstruido de corazón y de pulmón obtenidos tras el proceso de factorización. Así mismo, por encima de estas dos figuras, se han colocado los

botones de reproducción, parada o guardar, con el fin de reproducir, parar o guardar, la señal que se muestra en ese momento en dichas figuras. Finalmente, por debajo de dichas figuras, se encuentra el acceso a la funcionalidad de búsqueda de patrones, donde se introduce el nivel de similitud deseado para la búsqueda. Dicha búsqueda se ejecuta utilizando el botón ‘Buscar patrón’; dado que pueden seleccionarse tantas localizaciones temporales como el usuario desee, se ha añadido el botón “Repetir”, que permite realizar la búsqueda en el resto de localizaciones temporales, tantas veces como el usuario desee.

Si se accede, por el contrario, al sistema NMF Semisupervisado, se muestra la siguiente interfaz:

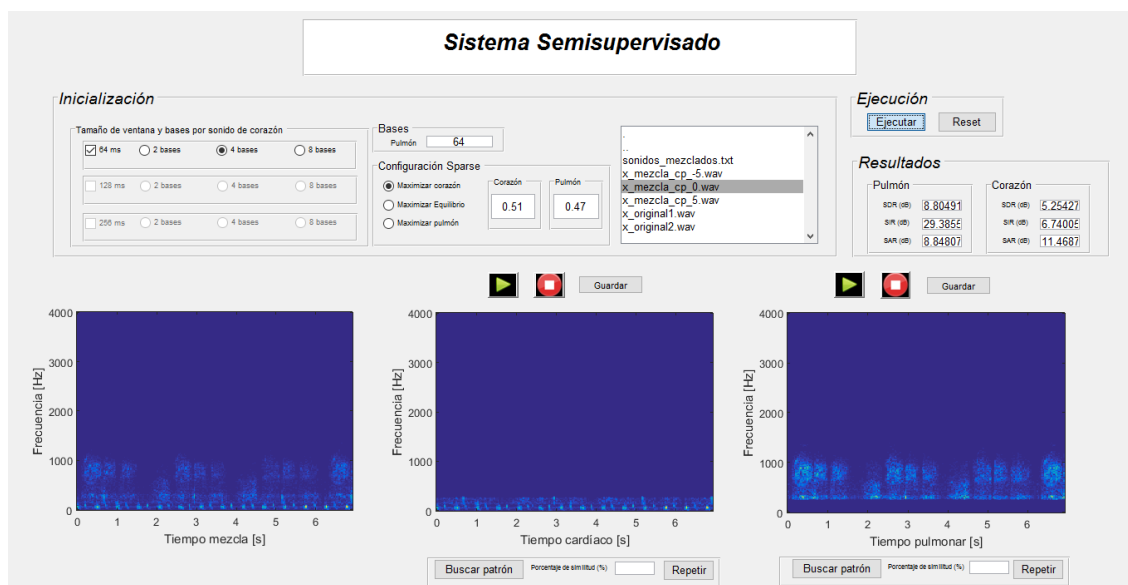


Figura 56 – Herramienta gráfica para el sistema NMF Semisupervisado

Como se observa, la distribución de este sistema es muy similar a la del sistema anterior, cambiando únicamente la sección de ‘Inicialización’, donde se permite además de la selección de la longitud de la ventana a utilizar, la selección del número de bases utilizadas durante la fase de entrenamiento, dando así acceso a cualquiera de los diccionarios generados. También, en dicho sistema, la configuración para maximización se centra en el uso de la configuración de la restricción Sparse.

ANEXO II

COSTE COMPUTACIONAL

10. Coste computacional

Para el cálculo del coste computacional de los sistemas implementados, se ha usado un ordenador Toshiba L855-11K con especificaciones comunes en el mercado de ordenadores, teniendo este las siguientes especificaciones principales:

<i>Componentes</i>	<i>Modelo</i>
<i>Procesador</i>	Dual Core i5-2450m
<i>Memoria RAM</i>	8Gb
<i>Capacidad disco duro</i>	1Tb

Tabla 14 - Especificaciones ordenador

Así, utilizándose un ordenador con las anteriores características:

- El sistema NMF No Supervisado, utilizando el modelo estándar de separación, necesita 0.45s de procesamiento por cada segundo de señal mezcla a separar.
- El sistema NMF No Supervisado, utilizando el modelo de factorización en bloques con bases adaptativas, necesita 1.7s de procesamiento por cada segundo de señal mezcla a separar.
- El sistema NMF Semisupervisado necesita 0.19s de procesamiento por cada segundo de señal mezcla a separar.

ANEXO III
BIBLIOGRAFÍA

11. Bibliografía

- [1] OMS. Enfermedades cardiovasculares. Obtenida el 19 de Enero de 2016, de <http://www.who.int/mediacentre/factsheets/fs317/es/>
- [2] Finahari, N. (2012). Existence of interference between the heart and respiratory sounds: preliminary report. Critical Care 2012, Volum 16 Suppl 1, S70.
- [3] ChingShun,, L., Erwin H. (2013).Blind Source Separation of Heart and Lung Sounds based on Nonnegative Matrix Factorization. 2013 International Symposium on Intelligent Signal Processing and Communication Systems, 731-736.
- [4] Andrzej, C., Rafael, Z., Ahn, H. P. y Shun-Ichi, A. (2009). Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis And Blind Source Separation. Wiley, noviembre 2009.
- [5] R. M. Potdar et.al., “Performance evaluation of different adaptive filtering algorithms for reduction of heart sound from lung sound”, International Journal of Engineering and Advanced Technology, Vol. 1,No. 3, pp. 61–67, 2012.
- [6] Siby, T. y Nishi, S. H. Detection and elimination of Heart sound from Lung sound based on wavelet multi resolution analysis technique and linear prediction. The International Journal Research Publication’s, Vol 1, No 10 (2012).
- [7] T. Falk, C. Wai. Modulation Filtering for heart and lung sound separation from breath sound recordings. in Proc. Ann. Int. Conf. IEEE EMBS. Vancouvert, Canada. 2008. pp. 1859-1862.
- [8] J.D. Echeverry, A.F. López, J.F. López. Reconocimiento de valvulopatías cardíacas en señales de fonocardiografía empleando la transformada Gabor. Scientia et Technica Año XIII, No 34, Mayo de 2007. Universidad Tecnológica de Pereira. ISSN 0122-1701.
- [9] S.M.DEBBAL, F.BEREKSI REGUIG. Spectral analysis of the PCG signals. The Internet Journal of Medical technology 2007; Vol4; N°1

- [10] H. Pasterkamp, R. Fenton, A. Tal, and V. Chernick, "Interference of cardiovascular sounds with phonopneumography in children," *Am. Rev. Respir. Dis.*, vol. 131, no. 1, pp. 61–64, Jan. 1985.
- [11] Gastro Mérida. Auscultación. Obtenida el 19 de Enero de 2016 de : <http://www.gastromerida.com/secciones/semiologia/pulmon/pulmoausc.html>
- [12] V. Gross, A. Dittmar, T. Penzel, F. Schuttler, y P. V. Wichert. The relationship between normal lung sounds, age, and gender. *Am. J. Respir. Critic. Care Med.*, vol. 162, no. 3, pp. 905-909, 2000.
- [13] Amir Al-Majdalawi. Conceptos básicos sobre el sonido. Obtenida el 25 de Enero de 2016 de: https://www.lpi.tel.uva.es/~nacho/docencia/ing_ond_1/trabajos_05_06/io2/public_html/sonido.html
- [14] Milka Villayandre Llamazares. Tipos de ondas sonoras. Obtenida el 25 de Enero de 2016 de: <http://fhyc.unileon.es/Milka/FyF/33.pdf>
- [15] David Blanco. Tema 5. Ondas. Obtenida el 25 de Enero de 2016 de: http://www.ugr.es/~dblanc0/FFI_antiguo/Apuntes_Tema5.pdf
- [16] Raúl, Z. A. y Miroslav, Z. (2009). Estudio de pitch en señales polifónicas. Proyecto Fin de Carrera en la Universidad Pública de Navarra.
- [17] Dan Barry, Derry Fitzgerald, Eugene Coyle and Bob Lawlor 'Single Channel Source Separation using Short-time Independent Component Analysis' Audio Engineering Society 119th Convention (2005).
- [18] C. Uhle, C. Dittmar and T. Sporer. Extraction of Drum Tracks from polyphonic Music using Independent Subspace Analysis. in *Proc. of the Fourth International Symposium on*

Independent Component Analysis, Nara, Japan, 2003

[19] A. Lefèvre, F. Bach, and C. Févotte. Semi-supervised NMF with time-frequency annotations for single-channel source separation. Proc. International Conference on Music Information Retrieval (ISMIR), 2012.

[20] Mikkel N Schmidt et al. Wind Noise Reduction using Non-Negative Sparse Coding. MACHINE LEARNING FOR SIGNAL PROCESSING, 2007 IEEE WORKSHOP ON, IEEE, PI, 1 August 2007 (2007-08-01), pag. 431-436, XP031199125, ISBN: 978-1-4244-1565-6

[21] Asaei, Afsaneh, Boulard, H. and Garner, P. N. Sparse Component Analysis for Speech Recognition in Multi-Speaker Environment. Proceedings of Interspeech, Makuhari, Japan, 2010.

[22] Andrew Nesbit, Maria Jafari, Emmanuel Vincent, Mark Plumbley. Audio source separation using sparse representations. W. Wang. Machine Audition: Principles, Algorithms and Systems, IGI Global, pp.246–265, 2010, <10.4018/978-1-61520-919-4.ch010>. <inria-00544030>

[23] Li, Y. / Wang, D. L.. Singing-Voice Separation From Monoaural Recordings Using Robust Principal Component Analysis. IEEE TRANSACTIONS ON AUDIO SPEECH AND LANGUAGE PROCESSING; 15, 4; 1475-1487; IEEE TRANSACTIONS ON AUDIO SPEECH AND LANGUAGE PROCESSING; 2007

[24] Madhuri A. Patil. Perceptually Motivated Robust Principal Component Analysis Based Separation of Singing Voice from Music. International Research Journal of Engineering and Technology (IRJET), Volume 2, Issue 03, Junio 2005.

[25] Lozano, Fernando Enrique, Salazar, Antonio José, Alvarado, Catalina, System of heart and lung sounds separation for store-and-forward telemedicine applications. Revista Facultad de Ingeniería Universidad de Antioquia 2012.

[26] Shah, Ghafoor; Koch, Peter; Papadias, Constantinos B. On the Blind Recovery of Cardiac and Respiratory Sounds. IEEE J. Biomedical and Health Informatics 19, N0. 1, 151-157(2015)

- [27] ChingShun, L., Erwin H. (2013). Blind Source Separation of Heart and Lung Sounds based on Nonnegative Matrix Factorization. 2013 International Symposium on Intelligent Signal Processing and Communication Systems, 731-736.
- [28] Cédric Févotte. Nonnegative matrix factorization with the Itakura-Saito divergence: With application to music analysis. C Févotte, N Bertin, JL Durrieu. Neural computation 21 (3), 793-830, 2009.
- [29] Felix Weninger. openBliSSART. User Manual Version 1.2, Pag 29, Mayo 2010.
- [30] J. B. Allen and L. R. Rabiner. A unified approach to short-time Fourier analysis and synthesis.' Proc. IEEE, vol. 65, pp. 1558-1564, Nov. 1977.
- [31] Ghafoor Shah, Constantinos B. Papadias. (2013). Blind recovery of cardiac and respiratory sounds using non-negative matrix factorization & time-frequency masking. BIBE 2013: 1-5
- [32] P. Smaragdis, B. Raj, and M. Shashanka, Supervised and semi-supervised separation of sounds from single-channel mixtures. Proceedings of the 7th International Conference on Independent Component Analysis and Blind Signal Separation (ICA '07), pp. 414-421, Londres, UK, Septiembre 2007.
- [33] Peter Bentley. Classifying Heart Sounds Challenge. Obtenida el 29 de Enero de 2016 de: <http://www.peterjbentley.com/heartchallenge/>
- [34] Andrea Spadaccini. Performance Evaluation of Heart Sounds Biometric Systems on an Open Dataset. Base de datos obtenida el 20 de Noviembre de 2015 de: <http://www.diit.unict.it/hsct11/>
- [35] Constantine Barabasa, Mohsen Jafari y Mark D Plumbey. A robust method for S1/S2 heart sounds detection without ECG reference based on music beat tracking. Electronics and Telecommunications (ISETC), 2012 10th International Symposium, Pages 307-310, IEEE, 15/11/2012.
- [36] L. Su and Y.-H. Yang. Power-scaled spectral flux and peak-valley group-delay methods for robust musical onset detection. Proc. Sound Music Comput. Conf., 2014.

- [37] Felix Scholkmann, Jens Boss, Martin Wolf. An Efficient Algorithm for Automatic Peak Detection in Noisy Periodic and Quasi-Periodic Signals. *Algorithms* 5(4): 588-603 (2012).
- [38] Eugenijus Kaniusas, *Biomedical Signals and Sensors II: Linking Acoustic and Optic Biosignals and Biomedical Sensors [Versión Electrónica]*. Pag 6.
- [39] G. Peeters. A large set of audio features for sound description (similarity and classification) in the CUIDADO Project. CUIDADO I.S.T. Project Report, Pag 15, 2004.
- [40] Zu,Q. “Human Centered Computing, First International Conference, HCC 2014, Phnom Penh, Cambodia, November 27-29, 2014, Revised Selected Papers”, Pag 712
- [41] Documentación Matlab. Correlation Coefficients. Obtenida el 1 de Febrero de 2016 de: <http://es.mathworks.com/help/matlab/ref/corrcoef.html>
- [42] Emad M. Grais and Hakan Erdogan. Single channel speech music separation using nonnegative matrix factorization and spectral mask. *Digital Signal Processing (DSP)*, 17th International Conference on, 2011, 1 - 6
- [43] Matlab fórum. Matlab Answers. Obtenida el 2 de Febrero de 2016 de: <http://www.mathworks.com/matlabcentral/answers/100980-how-do-i-write-a-loop-in-matlab-that-continues-until-the-user-presses-any-key>
- [44] Zafar Rafii, Antoine Liutkus, Bryan Pardo. A simple user interface system for recovering patterns repeating in time and frequency in mixtures of sounds. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Abril de 2015, Brisbane, Francia. 2015.

- [45] Littman ®. Tutorial sobre sonidos respiratorio básicos. Obtenido el 15 de Octubre de 2015 de: http://solutions.productos3m.es/wps/portal/3M/es_ES/3M-Littmann-EMEA/fonendoscopios/area-educativa/ruidos-cardiacos-y-pulmonares/tutorial-de-ruidos-pulmonares/
- [46] University of Maryland, Baltimore. Spring DPTE 516 Medical Issues I – Lung Sounds. Itunes U.
- [47] Emmanuel Vincent, R´emi Gribonval, C´edric F´evotte. Performance measurement in blind audio source separation. IEEE Transactions on Audio, Speech and Language Processing, Institute of Electrical and Electronics Engineers, 2006, 14 (4), pp.1462–1469.
- [49] Littman ®. Sonidos de corazón y pulmón. Obtenido el 15 de Octubre de 2015 de: http://solutions.productos3m.es/wps/portal/3M/es_ES/Littmann/stethoscope/education/heart-lung-sounds/
- [50] Texas Heart ® Institute. Anatomía del corazón. Obtenido el 27 de Febrero de 2016 de: http://www.texasheart.org/HIC/Anatomy_Esp/anato_sp.cfm
- [51] Texas Heart ® Institute. El latido cardíaco. Obtenido el 27 de Febrero de 2016 de: http://www.texasheart.org/HIC/Anatomy_Esp/systole_sp.cfm
- [52] Juan Ignacio Arribas, Ph.D. Definiciones y Conceptos. Obtenido el 27 de Febrero de 2016 de: https://www.lpi.tel.uva.es/~nacho/docencia/ing_ond_1/trabajos_06_07/io1/public_html/definini.htm
- [53] Instituto Nacional de Tecnologías Educativas y de Formación del Profesorado. Cualidades del sonido. Sonido y música, Módulo 2 – El sonido y su representación. Obtenido el 27 de Febrero de 2016 de: http://www.ite.educacion.es/formacion/materiales/60/cd/02_elsonido/2_cualidades_del_sonido.html