



**Universidad Autónoma de San Luis
Potosí**



Facultad de Ciencias

**Estimación de la Frecuencia Respiratoria Instantánea
Mediante Representaciones Tiempo-Frecuencia de
Señales de Fotopletismografía de Imagen**

Tesis para obtener el grado de:

Maestría en Ciencias de la Vida con Orientación en Bioingeniería

Presenta:

Ing. Susana Morales Lugo

Director de tesis:

Dr. Bersaín Alexander Reyes, Facultad de Ciencias, UASLP

San Luis Potosí, Mexico, Febrero 2025

Agradecimientos

A Ludwig, quien ha estado conmigo en momentos difíciles cuidándome con amor. Gracias por motivarme para seguir adelante y sacarme una sonrisa cuando lo necesito.

También, le agradezco a mi director de tesis, el Dr. Bersaín Reyes, quien me ha acompañado durante este proceso con comprensión y paciencia. Sin él, este proyecto no sería posible.

Quisiera agradecer a mi familia, que siempre confiaron en que yo podría completar esta etapa.

Por último, extendiendo un agradecimiento a CONAHCYT por ser beneficiaria de una de sus becas (CVU: 1240309).

Estimación de la Frecuencia Respiratoria mediante Representaciones Tiempo-Frecuencia de Señales de Fotopletismografía de Imagen © 2025 por Susana Morales Lugo está licenciada bajo CC BY-NC-ND 4.0. Para ver una copia de esta licencia, visite <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Índice de contenido

Agradecimientos	I
Índice de contenido.....	IV
Índice de tablas	VII
Índice de figuras.....	IX
Lista de abreviaturas	XII
Resumen	XIV
Abstract.....	XV
Capítulo 1. Introducción	1
1.1 Motivación	1
1.2 Antecedentes.....	3
1.3 Planteamiento del problema.....	6
1.4 Objetivo general	7
1.5 Objetivos específicos.....	8
1.6 Estructura de la tesis	8
Capítulo 2. Marco teórico.....	10
2.1 Fisiología	10
2.1.1 Sistema respiratorio.....	10
2.1.2 Arritmia sinusal respiratoria	11
2.1.3 Respiraciones controladas	12
2.1.4 Monitoreo de la frecuencia respiratoria	13
2.2 Fotopleletismografía de imagen.....	14
2.2.1 Propiedades ópticas de la piel.....	15
2.2.2 Fotopleletismografía de contacto	16

2.2.3	Estimación de la frecuencia respiratoria mediante iPPG	16
2.3	Análisis tiempo-frecuencia	20
2.3.1	Espectrograma	20
2.3.2	Escalograma	24
2.3.3	Transformada <i>Synchrosqueezing</i> basada en Fourier	26
2.3.4	Espectro de Hilbert-Huang	27
Capítulo 3.	Metodología	31
3.1	Base de datos	32
3.1.1	Adquisición de datos.....	32
3.1.2	Maniobras respiratorias	34
3.2	Extracción de señales de iPPG	36
3.2.1	Elección del canal de color	36
3.2.2	Región de interés.....	38
3.3	Preprocesamiento de señales de iPPG.....	39
3.3.1	Remuestreo	39
3.3.2	Filtrado digital	40
3.3.3	Normalización	40
3.3.4	Alineación de la señal de iPPG y la señal de respiración	40
3.3.5	Detección de picos de iPPG	42
3.4	Simulaciones	44
3.4.1	Maniobras respiratorias simuladas	44
3.4.2	Ruido aditivo	47
3.5	Selección de parámetros de la TFR	49
3.5.1	Obtención de la iRR a partir de las TFR.....	51
3.5.2	Índices de desempeño	54
3.6	Estimación de RR instantánea en datos reales.....	55
3.6.1	Estimación de iRR de referencia	55
3.6.2	Estimación de iRR con las TFR seleccionadas.....	56
3.7	Análisis de desempeño en señales reales	56
Capítulo 4.	Resultados y discusión	57
4.1	Preprocesamiento de señales de iPPG.....	57

4.2	Simulaciones	58
4.2.1	Maniobras respiratorias	59
4.2.2	Ruido aditivo	61
4.3	Selección de parámetros de la TFR	61
		62
4.3.1	Marginal temporal de la TFR	63
4.3.2	Índices de desempeño	76
4.4	Estimación de RR instantánea en datos reales.....	95
4.4.1	Estimación de iRR con las TFR seleccionadas	95
4.4.2	Análisis de desempeño	97
4.5	Comparación con estudios reportados en la literatura	107
	Capítulo 5. Conclusiones.....	111
5.1	Limitaciones	113
5.2	Trabajo a futuro	113
	Referencias.....	115
	Anexo A.....	127
	Anexo B.....	131

Índice de tablas

Tabla 3.1 Duración de los registros de cada maniobra respiratoria.	35
Tabla 3.2 Señales simuladas. A cada maniobra respiratoria se le adicionaron cuatro cantidades de relación señal-a-ruido (SNR).....	49
Tabla 3.3 Parámetros utilizados en cada técnica tiempo-frecuencia.....	51
Tabla 4.1 Ventanas con errores más pequeños con el espectrograma en la maniobra simulada después del ejercicio.	78
Tabla 4.2 Parámetros seleccionados en el espectrograma para todas las maniobras.	80
Tabla 4.3 Índices de desempeño de algunos resultados de las simulaciones constantes con el escalograma.....	82
Tabla 4.4 Índices de desempeño de dos wavelets de las simulaciones de la maniobra después del ejercicio.	83
Tabla 4.5 Parámetros seleccionados en el escalograma para cada una de las maniobras.	85
Tabla 4.6. Índices de desempeño de algunas ventanas utilizadas en el FSST con las simulaciones de cambios abruptos.	85
Tabla 4.7 Ventanas utilizadas en la transformada synchrosqueezing basada en Fourier con las simulaciones a 6 bpm.	86
Tabla 4.8 Índices de desempeño de las iRRs obtenidas de la transformada synchrosqueezing basada en Fourier de las simulaciones espontaneas utilizando un tamaño de 60 s en las ventanas.	89

Tabla 4.9 <i>Parámetros seleccionados en la transformada synchrosqueezing basada en Fourier.</i>	90
Tabla 4.10 <i>Índices de desempeño de cuatro espectros Hilbert-Huang de la maniobra constante a 6 bpm. La SD de ruido esta fija en 1.</i>	91
Tabla 4.11 <i>Índices de desempeño de las simulaciones de después del ejercicio utilizando el espectro Hilbert-Huang. Solo se muestran datos donde se usó el filtro. El SD de ruido es 1.</i>	93
Tabla 4.12 <i>Índices de desempeño de las maniobras constantes a 18 bpm y a 24 bpm. La SD de ruido esta fija en 1.</i>	92
Tabla 4.13 <i>Índices de desempeño de las simulaciones espontaneas utilizando el espectro Hilbert-Huang. Se muestran los datos con uso del filtro y SD de ruido unitario.</i>	94
Tabla 4.14 <i>Realizaciones seleccionadas en el HSS vía CEEMDAN.</i>	94
Tabla 4.15 <i>Índices de desempeño de la iRR estimada a partir de la señal de iPPG de la maniobra de cambios abruptos.</i>	98
Tabla 4.16 <i>p valores de las maniobras metrónomas para cada índice de desempeño. .</i>	100
Tabla 4.17 <i>Índices de desempeño de la iRR estimada a partir de las señales de iPPG de 6 y 18 bpm.</i>	102
Tabla 4.18 <i>Índices de desempeño de la iRR estimada a partir de las señales de iPPG de 6 y 18 bpm.</i>	101
Tabla 4.19 <i>Índices de desempeño de la iRR estimada a partir de la señal de iPPG después del ejercicio.</i>	103
Tabla 4.20 <i>Índices de desempeño de la maniobra espontanea de la base de datos.</i>	105

Índice de figuras

Figura 2.1. Arritmia sinusal respiratoria. La frecuencia cardíaca aumenta durante la inhalación y disminuye durante la exhalación	12
Figura 2.2. Modulaciones en la iPPG causadas por la respiración. Se incluye la modulación en línea de base, en amplitud y en frecuencia.	18
Figura 2.3. Representación tiempo-frecuencia	21
Figura 2.4. Ventanas de análisis utilizadas en el análisis tiempo-frecuencia en A) el dominio del tiempo y B) dominio de la frecuencia.....	23
Figura 2.5. Resolución TF en A) espectrograma y en B) escalograma.	25
Figura 2.6. Representación tiempo-frecuencia de una senoidal con frecuencia de 0.4 Hz (24 bpm), utilizando una ventana Hamming de 45 s	27
Figura 2.7. Representación tiempo-frecuencia de una senoidal con cambios en la frecuencia	29
Figura 3.1. Diagrama general de la metodología.....	31
Figura 3.2. Ejemplificación de cómo se obtiene el canal verde	37
Figura 3.3 Elección de la ROI en un 40 % de los píxeles (128 x 96) y en el centro del canal de los verdes	38
Figura 3.4 Espectrogramas con ventana Gaussiana de 45 s	48
Figura 3.5 Funcionamiento del algoritmo greedy de búsqueda heurística hacia adelante y hacia atrás. Se presenta un ejemplo de una representación tiempo-frecuencia (TFR).....	53
Figura 4.1 Preprocesamiento de señales	58

Figura 4.2 Ejemplo de la señal de respiración filtrada superpuesta en la serie de tiempo de la frecuencia de pulso obtenida a partir del iPPG.....	59
Figura 4.3 Señales y representaciones tiempo-frecuencia (TFR) ideales.....	60
Figura 4.4 Simulaciones del mismo segmento temporal, pero con diferentes SNR.....	62
Figura 4.5 Marginales (iRR) de los escalogramas de la simulación de cambios abruptos con 20 dB de SNR.....	63
Figura 4.6 Diagramas de cajas y bigotes del RMSE (bpm) de las iRRs obtenidas a partir de las simulaciones de la respiración con cambios abruptos.....	77
Figura 4.7 Diagrama de cajas y bigotes del RMSE (bpm) de las iRRs obtenidas a partir de las simulaciones de la respiración espontanea.....	79
Figura 4.8 Diagrama de cajas y bigotes del MAE (bpm) de las iRRs obtenidas a partir de las simulaciones con cambios abruptos.....	81
Figura 4.9 Diagrama de cajas y bigotes del RMSE (bpm) de los escalogramas de las simulaciones espontaneas.....	84
Figura 4.10 Diagrama de cajas y bigotes del MAE para las simulaciones constantes en 12 bpm de la transformada synchrosqueezing basada en Fourier.....	87
Figura 4.11 Diagrama de cajas y bigotes del RMSE (bpm) de las iRRs obtenidas a partir de las simulaciones después del ejercicio y la FSST.....	88
Figura 4.12 Diagramas de cajas y bigotes de las simulaciones con cambios abruptos de los resultados del espectro Hilbert-Huang.....	91
Figura 4.13 Diagrama de cajas y bigotes del MAE (bpm) de las iRRs obtenidas a partir de las simulaciones en 24 bpm.....	92
Figura 4.14 Ejemplos de TFR e iRR de la base de datos.....	96

Figura 4.15 <i>iRR de la señal respiratoria de referencia (en morado) y iRR estimada mediante el escalograma ($\gamma = 30$, $\beta = 2$) de la señal PR (en línea roja discontinua) del mismo voluntario al respirar a 6 bpm (0.1 Hz).</i>	97
Figura 4.16 <i>RMSE (bpm) de las técnicas TF y parámetros utilizados en la maniobra de cambios abruptos de la base de datos.</i>	99
Figura 4.17 <i>Índice de correlación cruzada de las técnicas y parámetros utilizados en la maniobra constante en 12 bpm de la base de datos.</i>	100
Figura 4.18 <i>Diagrama de cajas y bigotes del coeficiente de correlación cruzada de la maniobra respiratoria espontanea.</i>	104
Figura 4.19 <i>Diagrama de cajas y bigotes de los MAE (bpm) de la maniobra respiratoria después del ejercicio.</i>	104
Figura 4.20 <i>Ejemplo de estimación de la iRR con las técnicas tiempo-frecuencia aplicadas a los datos de un voluntario.</i>	106
Figura A.1 <i>iRRs de las simulaciones con cambios abruptos con 10 dB.</i>	127
Figura A.2 <i>iRRs de las simulaciones con Constante 12 bpm con 20 dB.</i>	128
Figura A.3 <i>iRRs de las simulaciones con después del ejercicio con 30 dB.</i>	129
Figura A.4 <i>iRRs de las simulaciones con espontanea con 40 dB.</i>	130

Lista de abreviaturas

A

ASR.....arritmia sinusal respiratoria

B

bpm respiraciones-por-minuto(*breaths-per-minute*)

C

CEEMDAN.....descomposición empírica de modos de conjunto completo con ruido adaptativo (*complete ensemble empirical mode decomposition with adaptive noise*)

CWT transformada *wavelet* continua (*continuous wavelet transform*)

E

ECG.....Electrocardiografía

EEMD.....descomposición empírica de modos por conjuntos (*ensemble empirical mode decomposition*)

EMD descomposición empírica de modos (*empirical mode decomposition*)

EPOC Enfermedad pulmonar obstructiva crónica

F

fps.....fotogramas-por-segundo (*frames-per-second*)

FSST transformada *synchrosqueezing* basada en Fourier (*Fourier-based synchrosqueezing transform*)

G

GMW*wavelet* Morse generalizadas (*generalized Morse wavelets*)

I

iPPG fotopletiografía de imagen (*image photoplethysmography*)

iRR frecuencia respiratoria instantánea (*instantaneous respiratory rate*)

P

PPG..... fotopleletismografía (*photoplethysmography*)

PR frecuencia de pulso (*pulse rate*)

PRV variabilidad de la frecuencia de pulso (*pulse rate variability*)

R

RR frecuencia respiratoria (*respiratory rate*)

S

SC.....escalograma (*scalogram*, SC)

SP.....espectrograma (*spectrogram*, SP)

STFT transformada de Fourier de tiempo corto (*short time Fourier transform*)

T

TF tiempo-frecuencia

TFR representaciones tiempo-frecuencia (*time-frequency representation*)

V

VFCDM.....demodulación compleja de frecuencia variable (*variable frequency complex demodulation*)

W

WT..... transformada *wavelet* (*wavelet transform*)

Estimación de la Frecuencia Respiratoria mediante Representaciones Tiempo-Frecuencia de Señales de Fotopletismografía de Imagen

Ing. Susana Morales Lugo

Universidad Autónoma de San Luis Potosí

Resumen

La frecuencia respiratoria (*respiratory rate*, RR) es un signo vital básico que refleja el estado de salud del sistema cardiorrespiratorio y que tiene importancia en eventos clínicos y enfermedades como el asma, hipoventilación, taquipnea, arrestos cardiopulmonares, entre otros. Con el objetivo de obtener una medición sencilla, económica y fiable de la RR, se han investigado diversas estrategias, entre ellas el uso de técnicas de análisis tiempo-frecuencia en series de tiempo de la variabilidad del pulso cardíaco, provenientes de señales de fotopletismografía de imagen (*image photoplethysmography*, iPPG). Algunas de estas técnicas son el espectrograma, la transformada *synchrosqueezing* basada en Fourier (*Fourier-based synchrosqueezing transform*, FSST), el escalograma y el espectro Hilbert-Huang. En todos los casos, los resultados de estas técnicas son imágenes, denominadas representaciones tiempo-frecuencia (*time-frequency representations*, TFR), las cuales dependen de diversas variables. Sin embargo, el procedimiento de selección de los parámetros que se usan no siempre es detallado o metodológico. En este trabajo de tesis se propone una metodología para que las TFR sean evaluadas utilizando señales respiratorias sintéticas de maniobras respiratorias, empleando diferentes parámetros para encontrar cuales de estos logran el mejor rendimiento para estimar la RR instantánea. Después, se aplican los mejores parámetros, resultantes para cada TFR, a series temporales de frecuencia de pulso derivadas de señales de iPPG, adquiridas previamente en voluntarios sanos por el grupo de trabajo, en la modalidad de contacto. Finalmente, se hace una comparación y análisis estadístico del desempeño de la estimación con las diversas TFR en términos de su error absoluto promedio (*mean absolute error*, MAE), raíz del error cuadrático medio (*root mean square error*, RMSE) e índice de correlación cruzada. Utilizando los parámetros seleccionados, se obtuvo un RMSE promedio de 0.127 respiraciones por minuto (*breaths-per-minute*, bpm) mediante la FSST, en la maniobra respiratoria constante a 6 bpm, de hasta 2.951 bpm en una maniobra con cambios repentinos en la frecuencia con el espectrograma, de 2.394 para una maniobra después del ejercicio con la FSST y de 1.621 bpm en la maniobra espontánea con el espectrograma. Los resultados resaltan la relevancia de una adecuada selección de los parámetros de las TFR cuando se emplean para la estimación de la RR instantánea a partir de los registros de iPPG.

Respiratory Rate Estimation by means of Time-Frequency Representations of Image Photoplethysmography Signals

B.M.E. Susana Morales Lugo

Universidad Autónoma de San Luis Potosí

Abstract

Respiratory rate (RR) is a basic vital sign reflecting the health status of the cardiorespiratory system and is relevant in clinical events and diseases such as asthma, hypoventilation, tachypnea, cardiopulmonary arrests, among others. In order to obtain a simple, economical and reliable measurement of RR, several strategies have been investigated, including the use of time-frequency analysis techniques to time series of heart rate variability extracted from image photoplethysmography (iPPG) signals. Some of these techniques include the spectrogram, Fourier-based synchrosqueezing transform (FSST), scalogram, and Hilbert-Huang spectrum. In any case, the results of these techniques are images, called time-frequency representations (TFRs), which depend on diverse variables. However, the procedure for selecting their parameters is not always detailed or methodological. In this thesis, a methodology is proposed to evaluate time-frequency representations using synthetic respiratory signals emulating respiratory maneuvers, while using different parameters to find which of these achieve the best performance to estimate instantaneous RR. Then, the TFRs computed using the selected parameters are applied to pulse rate time series derived from contact-based iPPG signals previously acquired from healthy volunteers by our research group. Finally, a comparison and statistical analysis of the performance of the estimation with the various TFRs is made in terms of their mean absolute error (MAE), root mean square error (RMSE) and cross-correlation index. Using the selected parameters, an average RMSE of 0.127 breaths-per-minute (bpm) was obtained by the FSST, in the constant breathing maneuver at 6 (breaths-per-minute) bpm, up to 2.951 bpm in a maneuver with sudden changes in frequency with the spectrogram, 2.394 bpm for a post-exercise maneuver with the FSST and 1.621 bpm in the spontaneous maneuver with the spectrogram. The obtained results highlight the relevance of an adequate selection of the parameters of the TFR when they are used for the estimation of the instantaneous RR from iPPG data.

Capítulo 1

Introducción

1.1 Motivación

En México, las enfermedades relacionadas con las vías respiratorias se encuentran entre las principales causas de muerte y ocupan del tercer al quinto lugar según la lista de la OMS de países de ingresos medianos altos, a la cual pertenece México de acuerdo con el Banco Mundial. En orden de principales causas de muerte, estas enfermedades respiratorias son la enfermedad pulmonar obstructiva crónica (EPOC), el cáncer de tráquea, bronquios y pulmón y las infecciones de las vías respiratorias inferiores [1]. Además, a nivel mundial las defunciones por la EPOC representan al 6% de las muertes totales y en 2019, antes de la pandemia de COVID-19, 2.6 millones de muerte fueron debido a infecciones de las vías respiratorias. En estos dos tipos de enfermedades, junto a otras como el asma, tuberculosis, COVID-19 y disnea, el estudio de la respiración resulta relevante [2]. La respiración es un proceso biológico con alcance en el seguimiento de enfermedades respiratorias y cardiorrespiratorias, puesto que nos puede indicar si existe una respiración anormal. Entonces, para un apropiado diagnóstico y seguimiento en las enfermedades respiratorias, la respiración es un indicador de apoyo para conocer el estado de salud del paciente. Añadido a esto, en estudios de meditación para conocer sus efectos cognitivos y emocionales, se ha encontrado que controlar la respiración en bajas frecuencias pareciera producir impactos benéficos en el sistema nervioso central y autónomo [3].

Un parámetro que se puede obtener de la respiración es la frecuencia respiratoria (*respiratory rate*, RR), la cual es un signo vital y clínicamente es definida como el número de respiraciones que se producen en un minuto o respiraciones-por-minuto (*breaths-per-minute*, bpm). Se ha reportado que el monitorear la frecuencia respiratoria en un área de cuidados intensivos puede evitar readmitir pacientes [4].

Asimismo, la estimación de la RR se puede llevar a cabo con la ayuda de diversas técnicas, cada una con sus propias ventajas y desventajas, incluyendo la auscultación, la cual depende de que alguien haga un conteo visual o con ayuda de un estetoscopio, lo que puede dar resultados no tan precisos. También, el uso de bandas respiratorias, capnografía, espirometría, electrocardiografía (ECG) o la fotopletismografía (PPG), las cuales necesitan de equipos especializados y de personal de salud entrenado, que sepa cómo usar y manejar los equipos como el capnógrafo, sensores de flujo o espirómetros. Otro aspecto de estos métodos es que algunos necesitan de consumibles como electrodos, bandas elásticas o cánulas nasales, los cuales suelen ser costosos y pueden llegar a ser incómodos, e.g., las cánulas nasales se introducen en las fosas nasales, lo cual puede distraer y resultar molesto para la persona [5].

Otra técnica que sirve para estimar la RR, es la fotopletismografía de imagen (*image photoplethysmography*, iPPG), la cual es un método que se ha propuesto en los últimos años debido a que, a diferencia de los demás métodos, no necesita de equipo biomédico especializado. En la iPPG se utilizan cámaras de dispositivos inteligentes, lo que permite que se lleve a cabo en cualquier lugar y en cualquier momento. Dicho método está basado en la PPG tradicional, una técnica óptica y no invasiva, para estimar la saturación de oxígeno y la frecuencia de pulso cardíaco (*pulse rate*, PR), aplicando luz de longitudes de onda específicas en una zona del cuerpo para captar el cambio en su reflexión o transmisión, dado por la modificación en el volumen de la sangre de vasos sanguíneos provocado con cada latido del corazón [6].

En la técnica de iPPG, se plantea que el principio de la PPG se aplique capturando secuencias de videos de una parte del cuerpo iluminada por luz natural o artificial [7], lo que ha permitido que se proponga el uso de sistemas móviles para adquirir la iPPG. Para el 2023, aproximadamente 78% de la población por arriba de los 10 años tiene a su disposición un teléfono móvil, según la ONU [8], lo que los hace una opción tecnológica accesible en la vida cotidiana para aplicarse en la salud. Para estimar la RR, se han utilizado los sensores de los dispositivos inteligentes, como los acelerómetros, al colocarlos encima de la caja torácica o utilizando el micrófono para capturar los sonidos respiratorios [9]. Sin embargo, la iPPG se ve afectada por artefactos de movimiento cuya banda de frecuencia esta alrededor de 0.2 Hz, frecuencia que está dentro del rango de las frecuencias respiratorias [5]. Además, la señal se ve afectada por la luz ambiental y los movimientos involuntarios del usuario.

Desde el punto de vista fisiológico, la RR puede ser obtenida de la señal de PPG debido a que la respiración modula a esta última señal en frecuencia, amplitud y línea de base [10], por consiguiente también puede ser obtenida de la señal de iPPG [11]. Además de la frecuencia respiratoria promedio, se puede obtener la serie de tiempo de la RR o RR instantánea (*instantaneous respiratory rate*, iRR), la cual brinda información sobre como varia la RR a través del tiempo, resultando útil para un monitoreo continuo de este signo vital. En particular, puede estimarse la iRR a partir de la serie de tiempo de PR o de su variabilidad (*pulse rate variability*, PRV) [11]. Entonces, la motivación de este trabajo consiste en mejorar la estimación de la iRR a partir de datos adquiridos directamente con un *smartphone*, en situaciones donde se requiere monitorizar los cambios en la frecuencia en cada instante de tiempo, siendo de utilidad después de alguna actividad física, en reposo o incluso para evaluar la rehabilitación pulmonar [12].

1.2 Antecedentes

Para estimar la iRR se han utilizado diversas técnicas de procesamiento digital de señales, incluido el análisis tiempo-frecuencia (TF), como el clásico espectrograma

(SP) [13], la transformada *wavelet* (*wavelet transform*, WT) [14], [15], la descomposición empírica de modos (*empirical mode decomposition*, EMD) [11], la demodulación compleja de frecuencia variable (*variable frequency complex demodulation*, VFCDM), entre otras [11], [14]. En particular, el procesamiento de las señales de PPG o de iPPG, se ha hecho de varias formas, incluyendo diversas técnicas de análisis TF para obtener las denominadas representaciones TF (*time-frequency representation*, TFR). A continuación, se presenta un breve resumen de algunos trabajos donde se ha usado el PPG tradicional o la iPPG de contacto obtenida de dispositivos móviles para la estimación de la RR.

La VFCDM se ha aplicado para extraer de la señal de PPG la modulación en frecuencia (FM) y la modulación en amplitud (AM) por Chon et al., a partir de señales de oximetría [16] y de *smartphones* [14], [17], donde ambas tienen buenos resultados en frecuencias de 0.2 a 0.4 Hz, sobre todo la FM. A la par, este análisis tiempo frecuencia ha sido comparado con la CWT y por algoritmos que involucran variaciones inducidas por la respiración, también aplicados para FM y AM y se ha reportado que la VFCDM y la CWT tienen mejor desempeño, pero la primera tiene un mejor tiempo computacional, aunque en [14] en un rango de frecuencias de 0.4 a 0.6 Hz la CWT-FM obtuvo los mejores resultados. Alafeef y Fraiwan también usaron la WT para obtener la frecuencia respiratoria de la grabación de un dedo tomada con un *smartphone*, donde adquirieron un error de 2.2 % [15].

Así mismo, la transformada *synchrosqueezing* fue utilizada por Dehkordi et al. para calcular la frecuencia respiratoria a partir de PPG y usaron 2 bases de datos abiertas, donde se aplicó la transformada con una ventana Gaussiana de 4 s a la AM de la señal de PPG. Una base de datos es *Capnabase* que contiene datos de PPG de 42 personas anestesiadas, con la que obtuvieron un RMSE de 1.8 rpm y 2 SD de 3.97 bpm. La otra base de datos es *Sleep Data Set* que tiene el registro de PPG de 43 niños dormidos con la que el RMSE fue de 2.3 bpm y 2 SD de 5.9 bpm [18]. Igualmente, la técnica se utilizó en la *VORTAL dataset* que incluye la señal de

PPG de 39 sujetos sanos en reposo, donde también se usó una ventana Gaussiana y se obtuvo un RMSE de 4.28 bpm con 3.42 2 SD [19].

Por otro lado, la HHS fue usada con el EMD por Garde et al. en [20], aplicada a señales de PPG, reportaron un RMSE de 3.5 bpm con SD de 1.1 bpm. También, se utilizó por Karlen et al. en [11] obtuvieron un RMSE de 7.17 bpm con un oxímetro conectado a una cámara, de 4.31 bpm y 7.23 bpm con oxímetros conectados a dos diferentes iPads, los autores también utilizaron *SmatFusion* y obtuvieron errores más pequeños con este algoritmo. En 2020 Sharma hizo uso del *Ensamble* EMD y menciona que el error en la estimación disminuye al hacer uso de ventanas grandes, donde comparo ventanas de 32 y 64 segundos [10]. Además, Shuzan et al. en 2021 con ayuda de la transformada de Hilbert obtuvieron un modelo con 3 señales del dominio del tiempo con diferentes características de la señal de PPG, para introducirlas en algoritmos de aprendizaje máquina. Con el ajuste a un modelo de regresión de proceso Gaussiano lograron resultados con un RMSE de 2.66 bpm y 2SD de 5.3 bpm [21].

Por otro lado, también se ha empleado la técnica de descomposición CEEMDAN (*complete ensemble empirical mode decomposition with adaptive noise*), técnica basada en la EMD, para identificar el complejo QRS, la frecuencia cardiaca, entre otros, en señales de ECG [22]. Al respecto, cabe mencionar que en la búsqueda bibliográfica no se encontró que la técnica CEEMDAN haya sido empleada para obtener la TFR correspondiente al espectro de Hilbert-Huang (*Hilbert-Huang spectrum*, HHS) con el propósito de estimar la iRR a partir de señales de iPPG.

En esfuerzos anteriores del grupo de trabajo, se utilizó la técnica TF del SP para obtener la iRR a partir de la señal de PRV obtenida del iPPG mediante la grabación de secuencias del dedo índice izquierdo con la cámara de un *smartphone* [13]. Sin embargo, solo la técnica TF del SP fue empleada y, además solo fue calculada con un único tamaño y tipo de ventana escogido de forma heurística.

En otro aspecto, las *apps* móviles de salud se han introducido en la vida cotidiana en los últimos años, e. g., hay personas que suelen estar al pendiente de *apps* que cuentan sus pasos. Por medio de los dispositivos inteligentes se ha incluido la estimación de la RR, haciendo uso de sus componentes, como la cámara o los micrófonos para procesar los sonidos respiratorios, o los acelerómetros, con los cuales se detectan los movimientos torácicos al colocar el equipo en el pecho [9]. En tiendas digitales, como la Google Play Store se encuentran disponibles *apps* como “Respi-Rate” o “Respiratory rate monitor”, las cuales brindan la frecuencia respiratoria promedio, i. e., no brindan la iRR.

1.3 Planteamiento del problema

En la práctica clínica, existen distintos métodos para estimar la RR, pero estos suelen ser incómodos o poco precisos, como la auscultación. Además, con la pandemia por COVID-19 hizo más evidente la necesidad de desarrollar tecnologías innovadoras que permitan la adquisición y transmisión remota de datos fisiológicos y clínicos, con el propósito de reducir el acceso a servicios de salud, evitando que los centros de atención médica se saturen y reducir la transmisión de enfermedades. Esta estrategia está orientada a ayudar a reducir costos en la atención médica y permitir que el acceso a estas tecnologías de adquisición y análisis de datos estén al alcance de las masas fuera de los ambientes clínico y de investigación, para llevar un control en deportistas, en el análisis de emociones o simplemente en la vida diaria [23], [24]. Entonces, debido a que en los últimos años la iPPG se ha estudiado en mayor medida para estimar la PR, se puede tomar ventaja de esto para también estimar la iRR.

Para sobrellevar las limitaciones anteriores, en este trabajo de tesis se plantea una metodología orientada a probar distintas TFR, que han reportado ventajas sobre el SP en otras señales biomédicas, con el propósito de seleccionar objetivamente la técnica TF y sus parámetros más relevantes, de tal forma que

permitan mejorar el desempeño en la estimación de la iRR con respecto a los resultados obtenidos anteriormente por el grupo de trabajo.

Así, el propósito de este trabajo de tesis es explorar diferentes técnicas TF para estimar la iRR a partir de señales de series de tiempo de PRV obtenidas a partir de señales de iPPG, adquiridas de sujetos voluntarios sanos mediante la modalidad de contacto empleando un *smartphone*. Inicialmente, se simularán datos de maniobras respiratorias, de las cuales se obtendrán sus TFR con diferentes conjuntos de parámetros. A decir, se exploran el espectrograma, el escalograma (*scalogram*, SC) obtenido a partir de la WT, la FSST y el HHS. Después, se comparan cuantitativamente los resultados empleando índices de desempeño, para buscar los parámetros más adecuados de cada TFR, utilizando su TFR ideal simulada. Posteriormente, se aplican las TFR con los parámetros seleccionados a los datos reales de iPPG, obtenidos previamente por el grupo de trabajo, y finalmente los resultados de la iRR estimada se comparan con los obtenidos de una señal respiratoria de referencia adquirida simultáneamente con las secuencias de video. Con este trabajo de tesis se aborda la estimación de la iRR aportando en la adecuada selección de los parámetros utilizados a la hora de hacer un análisis TF, a diferencia de lo realizado en el pasado por el grupo de trabajo, donde únicamente se utilizó el espectrograma con una única selección de parámetros, elegidos de forma heurística. Así, la búsqueda de parámetros de las técnicas TF se considera relevante para hacer una estimación adecuada de la iRR, i.e., de la serie de tiempo de la frecuencia respiratoria, a partir de señales de iPPG. En particular, se explora la selección de parámetros de las técnicas TF del SP, la FSST, el SC y el HHS como técnicas de análisis.

1.4 Objetivo general

Estimar la frecuencia respiratoria instantánea a partir de señales de variabilidad de la frecuencia de pulso extraídas de señales de fotopletismografía de imagen

mediante diversas representaciones tiempo-frecuencia calculadas con parámetros seleccionados de forma cuantitativa.

1.5 Objetivos específicos

- Preprocesar señales de iPPG y respiración (referencia) adquiridas previamente por el grupo de trabajo durante diferentes maniobras respiratorias
- Explorar diversas técnicas de análisis TF a través de la simulación de señales de respiración y sus TFR ideales.
- Seleccionar los parámetros adecuados de cada técnica de análisis TF.
- Estimar la iRR de los datos con las TFR calculadas a partir de sus parámetros seleccionados.
- Comparar las estimaciones de iRR considerando la señal de respiración de referencia.

1.6 Estructura de la tesis

El resto de este documento de tesis se divide en cuatro capítulos adicionales descritos a continuación.

En el capítulo 2 se explora el marco teórico, donde se describen los fundamentos y antecedentes en los que se basa el trabajo de tesis. Se comienza con la fisiología del sistema respiratorio y su relación con la PR. También se menciona el principio de funcionamiento de la técnica de iPPG y por qué con esta señal es posible estimar la iRR mediante técnicas TF.

El capítulo 3 presenta la metodología empleada para lograr los objetivos, comenzando por la descripción del preprocesamiento de la base de datos para ganar entendimiento sobre la simulación de señales de respiración para las diversas maniobras respiratorias de interés con el fin de obtener la serie de tiempo de PR, así como su iRR de referencia. Después se describe la selección de parámetros de

las técnicas TF a partir de las simulaciones, los cuales se utilizan en los datos reales adquiridos previamente por el grupo de trabajo.

El capítulo 4 presenta los resultados obtenidos y la discusión de cada etapa de la metodología. Se presentan los índices de desempeño calculados para las TFR a partir de las simulaciones y como estos se usaron para seleccionar los parámetros de las TFR que luego se aplicaron a los datos reales, de los cuales también se presenta sus resultados.

El capítulo 5 presenta las conclusiones, limitaciones y trabajo a futuro de este proyecto de investigación.

Finalmente, el anexo A tiene figuras de las iRRs de las simulaciones con diferentes técnicas TF, fijando un parámetro por técnica, mientras que en el anexo B se presenta la miniatura de un poster expuesto en Congreso Internacional sobre Investigación en STEM y Ciencias Sociales: Desafíos Transdisciplinarios Actuales.

Capítulo 2

Marco teórico

2.1 Fisiología

2.1.1 Sistema respiratorio

El cuerpo humano funciona gracias a la sinergia de todos sus sistemas, los cuales necesitan constantemente de energía, que para producirla necesita de oxígeno el cual ingresa del medio exterior al cuerpo cuando se respira. El sistema respiratorio humano es el encargado de introducir oxígeno y expulsar dióxido de carbono para mantener la homeostasis y regular el balance ácido-base del cuerpo [23]. Por este intercambio de gases las células de todos los tejidos son capaces de obtener oxígeno a través de la red de vasos sanguíneos, los cuales llevan oxígeno. Luego, las células expulsan el dióxido de carbono hacia las arterias las cuales regresan el flujo sanguíneo a los pulmones, donde el compuesto químico es extraído y finalmente exhalado.

La respiración o ventilación pulmonar ocurre todo el tiempo basalmente con los movimientos mecánicos del tórax y abdomen, además, esta cambia según las demandas del cuerpo [25], es por eso que cambios extremos son indicativos de un estado de salud en deterioro. La frecuencia respiratoria normal en adultos tiene un rango de 6 a 24 bpm [25], [26]. La respiración por debajo de 6 bpm se llama hipoventilación y la que asciende de 24 bpm es taquipnea [27].

El proceso de la respiración consta de dos fases, la inhalación y la exhalación. En la inhalación o inspiración los pulmones se llenan de aire y el tórax se expande para permitir el paso del aire fácilmente, hay una disminución en la presión alveolar e intrapleurar lo que facilita la expansión de los pulmones y el retorno venoso y los músculos intercostales externos y el diafragma se contraen [28]. En la exhalación el aire sale y la caja torácica se comprime y la presión alveolar e intrapleurar aumentan, aunque la última se mantiene siempre negativa.

El sistema circulatorio y el respiratorio están estrechamente vinculados por el intercambio de moléculas que hay entre ellos, por lo que no es raro que se sincronicen de diferentes formas. Una de ellas, es el acoplamiento cardio-ventilatorio que ocurre cuando hay concordancia entre los latidos del corazón y el inicio de la inspiración, otra es la sincronización del volumen sistólico respiratorio que es la diferencia de fase entre los volúmenes sistólicos derecho e izquierdo durante un ciclo respiratorio, y también está la arritmia sinusal respiratoria [29] la cual se aborda a continuación.

2.1.2 Arritmia sinusal respiratoria

En 1847 se demostró, por Carl Ludwig, que la fluctuación en la frecuencia cardíaca está relacionada con la respiración [30]. La arritmia sinusal respiratoria (ASR) es un proceso en el cual la respiración influye en el ciclo cardíaco. La ASR funciona acortando los intervalos entre las ondas R del corazón durante la inhalación, i. e., la frecuencia cardíaca aumenta, y alargándolos en la exhalación, i. e., la frecuencia cardíaca disminuye [31], como se observa en la **Figura 2.1** [32]. Intentar explicar la causa de este fenómeno ha llevado a teorizar que se debe a los cambios mecánicos inducidos en la presión intratorácica, durante la inhalación y la espiración. Este proceso también puede ser que la ASR ocurra para salvar energía en la exhalación durante los ciclos respiratorios, aprovechando la entrada de oxígeno [33]. Además, cuando el valor de RR está incrementada, la amplitud de las oscilaciones de la frecuencia cardíaca disminuye, y cuando la RR disminuye, la amplitud incrementa,

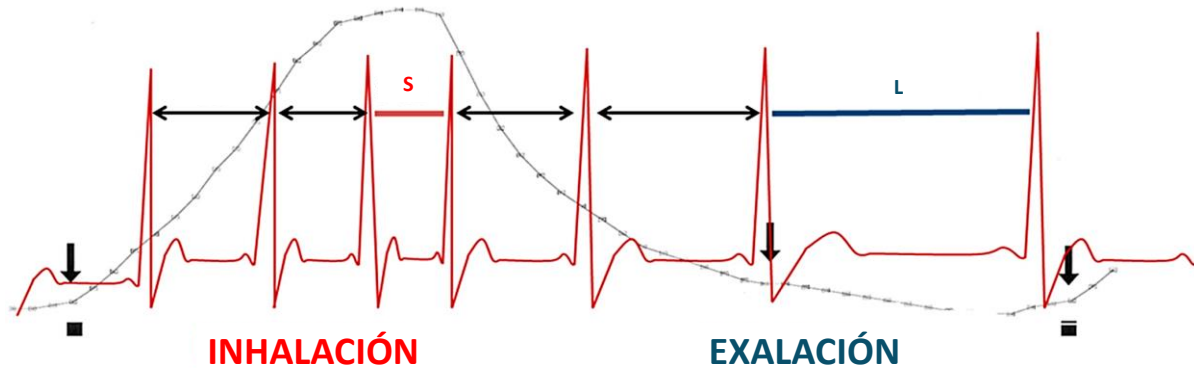


Figura 2.1. Arritmia sinusal respiratoria. La frecuencia cardíaca aumenta durante la inhalación y disminuye durante la exhalación, lo que resulta en el período cardíaco más largo durante la exhalación (L: *long*) y el período cardíaco más corto durante la inhalación (S: *short*). Modificada de [32].

sin importar si la respiración es en reposo o es controlada [34]. Debido a la ASR es que existe una modulación en señales como el ECG, PPG, iPPG, entre otras, causada por la respiración.

2.1.3 Respiraciones controladas

La respiración es un proceso que se puede contralar conscientemente y que también sucede inconscientemente. En el tronco encefálico hay dos áreas que controlan el ritmo de la respiración, el área inspiratoria y el área espiratoria. En un estado de reposo la respiración sucede automáticamente, porque el área inspiratoria genera impulsos nerviosos rítmicos que se propagan hacia los músculos intercostales externos y al diafragma. Pero cuando se quiere contralar la respiración, el área inspiratoria activa al área espiratoria, lo cual genera impulsos que contraen a los músculos intercostales internos y los abdominales, presionando el aire hacia afuera, además durante la inspiración forzada se contraen los músculos esternocleidomastoideo y escalenos [27]. Como se mencionó anteriormente, la ASR está presente en ambas respiraciones, espontánea y controlada.

El control consiente de la respiración puede tener beneficios en la salud física y mental de las personas. El entrenar la respiración puede darle bienestar a asmáticos y reducir el uso de broncodilatadores⁵. Asimismo, bajar la frecuencia respiratoria a propósito, ayuda a pacientes con hipertensión al disminuir la frecuencia cardíaca y la presión arterial, y reduce la respuesta del reflejo de los quimiorreceptores a la hipercapnia y la hipoxia [35]. Además, diferentes estados emocionales y cognitivos influyen en la frecuencia de la respiración y viceversa [3], [36].

2.1.4 Monitoreo de la frecuencia respiratoria

Como se ha mencionado, existen diferentes métodos para estimar la RR. Hay acercamientos por monitorearla por oximetría de pulso, por medio del flujo respiratorio, siguiendo los movimientos de la caja torácica y el abdomen, por los sonidos respiratorios, con métodos ópticos e incluso por uso de sensores térmicos, entre algunos otros métodos.

Por ejemplo, en el caso de los sonidos respiratorios se utiliza un micrófono cerca de la nariz o en la garganta y la señal rescatada es procesada para hacer un análisis de frecuencia [37]. Esto supone que se debe colocar un micrófono sobre la de la persona, lo cual pudiera ser menos incomodo que otros artefactos como los que se utilizan para medir el flujo respiratorio, donde se necesita de un sensor que pueda medir diferencias en la temperatura, la humedad o el CO₂ entre la fase inhalatoria y la exhaladora. Para que el sensor este bien colocado se necesita de mascarillas o termistores nasales [38].

En el supuesto de querer estimar la RR por los movimientos torácicos se hace con bandas elásticas, donde se pone un sensor piezoeléctrico que detecte la presión que se ejerce en cada inhalación por la expansión de los pulmones, también, se pueden usar sensores de deformación que miden la resistencia basándose en el área del sensor, lo cual aumenta al inhalar [39]. Además, hay sensores de conductividad, que miden el cambio en la impedancia del pecho, debido

al incremento y disminución de aire en cada respiración. Otra forma de estimar la RR es con sensores de movimiento como el acelerómetro, giroscopio o magnetómetros, los cuales son colocados en el pecho y detectan los movimientos producidos en este [40].

Asimismo, la RR se encuentra inmersa en señales cardiacas como el pulso o el ECG. El pulso cardiaco se mide por oximetría de pulso, técnica que aplica luz en frecuencias infrarrojas a una parte del cuerpo para obtener la cantidad de hemoglobina saturada con oxígeno. Esta técnica óptica se conoce como pletismografía, en la cual es necesario el uso de un oxímetro de pulso [41]. Además, mediante el ECG también se puede hacer un procesamiento para encontrar la señal de respiración por la ASR. La medición de la señal de ECG requiere de sensores, los cuales se deben colocar en sitios específicos. Una ventaja de este método es la reducción en el uso de algún otro tipo de sensor para obtener la respiración, ya que solo con los sensores de ECG se pueden adquirir ambos parámetros fisiológicos [42].

A distancia, también se han desarrollado algoritmos que estiman la RR por radar [43], cámaras de video [12], sensores de temperatura y por cámaras infrarrojas [44]. Estas alternativas le evitan al paciente cualquier artefacto invasivo en su cuerpo, molestias e incomodidad por la colocación de sensores o artefactos médicos. Sin embargo, hay una mayor exposición a artefactos de movimiento, como una mala iluminación o los movimientos del paciente, lo cual se podría resolver con un algoritmo robusto, pero esto supone mayor tiempo de procesamiento [39].

2.2 Fotopletismografía de imagen

La iPPG es una técnica óptica a partir de la cual se obtiene el pulso cardiaco. Esta es una técnica no invasiva, lo que facilita su implementación. Además, la tecnología necesaria para adquirirla está al alcance de la mayoría de la población, puesto que solo necesita de un dispositivo con cámara, e. g., teléfono inteligente, que pueda grabar videos. La iPPG extiende los principios de la PPG, técnica que utiliza la

reflexión y transmisión de la luz de una zona del cuerpo, para obtener el pulso. Así, dicho principio se extrapoló a videograbadoras, lo que originó a la iPPG, donde en lugar de realizar mediciones puntuales se obtienen mediciones en un arreglo rectangular de sensores que resultan en imágenes de fotoplethysmografía. La iPPG tiene dos modalidades, una por contacto y otra de no-contacto o remota, donde en la de contacto comúnmente se usa un dedo que toca a la cámara mientras graba, mientras que en la de no contacto se graba a distancia una parte del cuerpo, siendo generalmente el rostro [14].

2.2.1 Propiedades ópticas de la piel

La iPPG mide los cambios en la intensidad de la luz en los videos registrados de la piel, cambios que ocurren por que la piel es un tejido que puede absorber la luz y luego transmitirla. Los fotones rebotan con las estructuras del cuerpo y dependiendo de con cual choquen, se transmiten de regreso, con una intensidad diferente [45]. Es por esto que a partir de la luz recolectada se extrae información sobre los parámetros físicos y ópticos del tejido [46]. Las mediciones de la óptica de la piel, pueden medir melanina, hemoglobina oxigenada y reducida, así como bilirrubina [47]. Entre esos factores, resulta de particular interés la hemoglobina por su presencia en la sangre. Además, la luz que es remitida hacia la piel es importante según el objetivo, e. g., las ondas de luz de longitudes cortas se absorben en su mayoría por la melanina, en cambio las ondas más largas, como las que están en los rangos ultravioleta e infrarrojo se absorben mejor por los líquidos [48].

Tradicionalmente, las señales de PPG son tomadas en lugares como los dedos o los lóbulos de las orejas, porque son lugares donde la luz transmitida es captada con facilidad y donde los tejidos circundantes a los capilares tienen una absorbancia constante, dejando al flujo sanguíneo como el único factor cambiante. Los cambios en el volumen dependen del ciclo cardíaco. Durante la sístole ventricular el corazón eyecta sangre del corazón a la aorta y pulmones, lo que provoca que el volumen sanguíneo en capilares aumente [27], por lo tanto la

absorción de luz es mayor, luego el volumen disminuye y la absorción de luz es menor, i. e., la luz transmitida es menor en la sístole y aumenta en la fase diastólica.

2.2.2 Fotopletismografía de contacto

La PPG convencional necesita de una fuente de iluminación, para iluminar el tejido, y de un fotodetector, el cual detecta las variaciones en la luz reflejada desde el tejido causadas por el cambio de volumen. Además, debe existir un contacto directo y fijo con la piel, para que las mediciones sean concisas. Se ha encontrado que la fuerza con la que se efectúa el contacto entre la piel y el sensor afecta la amplitud de onda, por lo que la fuerza debe ser poca y controlada para obtener datos adecuados [48]. Los sitios del cuerpo donde se hace la medición de PPG son los dedos, la frente, las muñecas y los lóbulos de las orejas. Por su parte, la iPPG tiene dos modalidades:

- iPPG de contacto: la zona del cuerpo toca suavemente a la cámara que toma el video y se utiliza el *flash* del dispositivo inteligente para iluminar el área. Comúnmente se usa un dedo para hacer la medición.
- iPPG remota: la cámara se pone a una distancia del cuerpo y se analiza la grabación en la zona de interés.

En este trabajo de tesis, se empleó la técnica de fotopletismografía de imagen en la modalidad de contacto. A continuación, se brindan detalles acerca de esta técnica y su utilización para estimar la frecuencia respiratoria.

2.2.3 Estimación de la frecuencia respiratoria mediante iPPG

El significado de iPPG, se divide en tres partes: 1), la pletismografía se refiere al cambio en el volumen sanguíneo en una parte del cuerpo, 2) la palabra foto incluye la parte óptica y 3) imagen significa que la PPG se obtiene por medio de imágenes, i. e., un conjunto consecutivo de imágenes extraído de un video. La iPPG brinda información en tres dimensiones por cada instante en el que se captura una imagen, i. e., largo de la imagen, ancho de la imagen y el valor de cada píxel, sin embargo,

la señal de iPPG está en dos dimensiones, por lo que es necesario convertir cada imagen del video en un valor único. Así, la adquisición de la señal de iPPG consta de tres pasos:

- Grabación del video: el video es grabado con cualquier dispositivo inteligente que tenga una cámara de video y *flash* embebidos.
- Detección de la región de interés (*region of interest, ROI*): La ROI es un área de píxeles que se selecciona en cada imagen para extraer la señal. No es necesario tener toda la imagen, puesto que con una zona se tiene la información necesaria para reconstruir la señal, además de esta manera puede reducirse la carga computacional.
- Extracción de la señal iPPG: Una vez que se selecciona la ROI, se utiliza un algoritmo de extracción de iPPG, el cual consiste en un promediado de los píxeles en la ROI. Los valores que se obtienen se guardan para reconstruir la señal, la cual después pasara por una etapa de preprocesamiento, donde será remuestreada, filtrada y normalizada [11].

Después de llevar a cabo estos pasos, la señal resultante puede ser analizada. Estimar la RR a partir de la iPPG es posible, porque la respiración la modula en la línea de base, en amplitud y en frecuencia, como se ilustra en la **Figura 2.2** modificada de [49]. A continuación, se detallan las modulaciones [11], [18], [41]:

- Modulación en línea de base (*baseline wander, BW*): Hay cambios en el ciclo cardiaco debidos a cambios en la presión intratorácica causados por la respiración lo que conduce a una modulación en la línea de base. Durante la inspiración, la presión intratorácica disminuye, provocando que también haya una disminución en la presión venosa central que aumenta el retorno venoso. En la expiración ocurre lo contrario, disminuyendo el retorno venoso. La variación en la intensidad causada por la respiración en la iPPG se obtiene

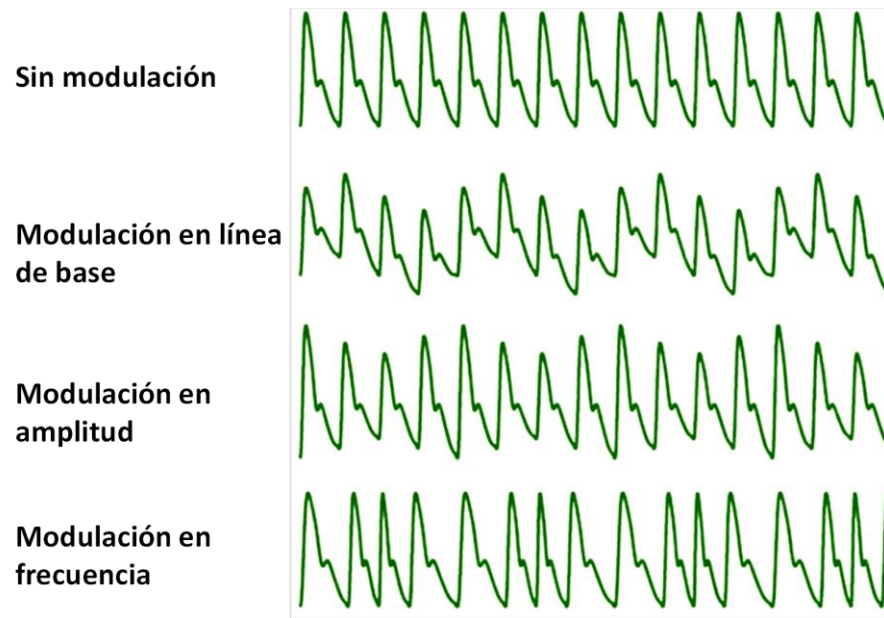


Figura 2.2. Modulaciones en la iPPG causadas por la respiración. Se incluye la modulación en línea de base, en amplitud y en frecuencia. Modificada de [49].

de la serie de tiempo del promedio de las amplitudes superior, i. e., las amplitudes de los picos, e inferior, i. e., las amplitudes de los valles de la señal de iPPG.

- Modulación en amplitud (*amplitud modulation*, AM): Esta modulación se debe a los cambios en el volumen de la caja torácica durante el ciclo respiratorio. Durante la inspiración, el volumen sistólico del ventrículo izquierdo disminuye lo que lleva a una disminución en la amplitud del pulso. Lo contrario pasa durante la expiración, donde la amplitud aumenta. Entonces, la variación en amplitud inducida por la respiración está dada por la serie de tiempo de las amplitudes de los picos de la iPPG.
- Modulación en frecuencia (*frequency modulation*, FM): Como se mencionó anteriormente, la ASR es la que causa este fenómeno, que primordialmente se debe a la regulación autonómica de la frecuencia cardíaca durante la

respiración. La serie de tiempo de la variación en la frecuencia inducida por la respiración proviene de la distancia entre los picos sucesivos de la iPPG.

Se han reportado trabajos que abordaron el problema de estimar la RR a partir de la iPPG, donde se usa uno, dos o incluso los tres tipos de modulación, e.g, se ha usado la FM y AM de señales de distintos dispositivos para aplicarles la VFCDM y a la CWT, donde esta última con FM obtuvo menores errores que las demás en bajas frecuencias [13], [16]. También, se han usado algoritmos de *smart fusion*, donde las tres variaciones son usadas para llegar a un promedio e introducirlo en una etapa de EMD, para llegar a una estimación de la RR a partir de señales de iPPG de dos diferentes modelos de *smartphone* con los datos de 19 adultos mientras respiraban espontáneamente, con un dispositivo se obtuvo un RMSE de 2.27 bpm y con el otro de 2.29 [11].

Asimismo la iPPG ha sido adquirida por luz visible y luz infrarroja para pasar una etapa donde se hace un cálculo de pesos y selección de pesos que proporcionan la señal de pulso con la SNR más alta, luego se aplicó el SP con una ventana hanning de 8 segundos, así, se obtuvo un error medio de 1.74 bpm con luz visible y de 2.27 bpm con luz infrarroja en registros de respiración guiada y de 4.72 bpm y 7.40 bpm, respectivamente en una respiración espontánea [50].

A su vez, con la FSST se ha utilizado la iPPG para la estimación de la iRR, usando las tres modulaciones, se usó dos veces la FSST y un algoritmo de fusión condicionado por picos para fusionar estimaciones de iRR simultáneas [18]. Además, en [19] a la PPG se le extrajo la AM, luego, se hizo una descomposición de modos y se escogió el pico espectral más largo por debajo de 1 Hz, para después aplicar un filtrado *particle* y por último se usó la FSST con una ventana Gaussiana, donde se obtuvo una RMSE de 4.28 bpm con una SD de 3.42 bpm.

Cambiando la perspectiva al SC, se ha usado el canal de los ojos para extraer la señal de iPPG de 15 señales de voluntarios, donde se usó la WT discreta que descompone la señal y se llegó a un nivel de descomposición nueve, debido a

que se menciona que ese nivel revela información de frecuencia sobre una frecuencia respiratoria tan baja como 4 bpm, para esto se usó la wavelet de Daubechies de orden 2. La frecuencia respiratoria se calculó con un buscador de picos aplicado a la señal respiratoria encontrada y obtuvieron una precisión promedio de 97.8% [15]. Ahora bien, con el HHS, para estimar la RR a partir de señales de iPPG de 19 voluntarios, obtenidas de una cámara conectada a un oxímetro, se usó la EMD para descomponer las señales en cuatro funciones de modo intrínseco (*intrinsic mode functions*, IMF) y hubo un RMSE de 7.17 bpm [11].

2.3 Análisis tiempo-frecuencia

Nuestro mundo está lleno de señales que varían en el tiempo como las ondas de radio, señales sísmicas, la luz, el sonido, señales biológicas, la temperatura, la presión, la velocidad, e incluso hay señales digitales que adquieren valores discretos como los interruptores, los lenguajes de computadora, los videos, entre otros. Estas señales se componen por su amplitud, fase y frecuencia, con los cuales es posible analizarlas. Comúnmente las señales se procesan para adquirir su contenido en frecuencia, lo cual es de utilidad cuando la señal es constante en su frecuencia, más esto no suele ocurrir con la mayoría de las señales biomédicas, donde su frecuencia cambia a través del tiempo, por lo que puede ser pertinente un análisis que mapee su contenido en la frecuencia a través del tiempo. Aquí, entra el análisis TF, donde se obtienen mapas de intensidades de colores que dependen del tiempo y la frecuencia [51], como en la **Figura 2.3** que muestra una TFR, donde en tonos amarillos se muestra las zonas donde hay mayor contenido en frecuencia.

2.3.1 Espectrograma

El SP fue utilizado anteriormente por el grupo de trabajo, por lo que se utilizó como punto de partida para saber si habría otros parámetros que pudieran mejorar la estimación de la iRR. El SP es la técnica tiempo-frecuencia (TF) por excelencia, debido a que se obtiene a partir de la transformada de Fourier (*Fourier transform*, FT) [52]. La FT de una señal $x(t)$, en el tiempo t , está definida por [53] :

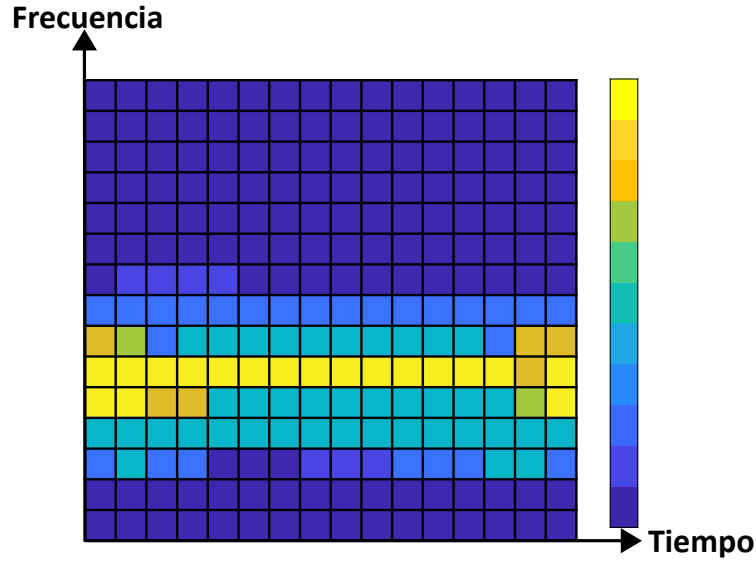


Figura 2.3. Representación tiempo-frecuencia

$$X(\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt \quad (2.3.1)$$

donde $X(\omega)$ es la FT dependiente de la frecuencia ω . En específico, para el SP se utiliza la FT de tiempo corto (*short time Fourier transform*, STFT), la cual es una versión de la FT que se le aplica a segmentos sucesivos de la señal, definidos por:

$$s_{\tau}(t) = s(t) g(t - \tau) \quad (2.3.2)$$

donde $s(t)$ es la señal, que al ser multiplicada por una ventana g centrada en el tiempo τ , se obtiene una versión de la señal $s_m(t)$ que solo tiene valores alrededor de τ . Entonces, al aplicar la TF se obtiene la información en el dominio de la frecuencia en este tiempo, i. e., la STFT [54].

$$S_{\tau}(t, \omega) = \int_{-\infty}^{\infty} s(\tau) g(\tau - t) e^{-j\omega \tau} d\tau \quad (2.3.3)$$

Finalmente, su densidad espectral está dada por el cuadrado de su magnitud

$$SP(t, \omega) = |S_\tau(t, \omega)|^2 \quad (2.3.4)$$

y al obtenerla en momentos sucesivos de la señal, se obtiene el SP [52]. Aunque, esta técnica es fácil de obtener presenta un compromiso en la resolución de la imagen TF, pues está limitada por el principio de incertidumbre de Heisenberg-Gabor, el cual dice que $\Delta f \Delta t \geq 1/4\pi$, donde Δf y Δt representan el cambio en la frecuencia y en el tiempo, respectivamente [55]. En consecuencia, cuando la ventana es larga, la resolución en el tiempo es pobre y en la frecuencia es buena, por el contrario, cuando la ventana es corta la resolución en el tiempo es buena y en la frecuencia es mala. Por lo que, si de inmediato se quiere conocer cuando ocurre un cambio en la frecuencia, la ventana debe ser muy corta, pero eso significara que los componentes en la frecuencia no serán óptimos. Ante esta situación, se debe encontrar un punto medio entre la resolución en el tiempo y la frecuencia [19].

Ahora bien, usar esta técnica para obtener la RR en contextos donde se quieren los cambios en la frecuencia rápidamente, podría no ser lo más adecuado, pero la señal puede ser guardada al ser obtenida por iPPG, situación que permite que el SP sea usado cuando no se requiere de una respuesta pronta [19].

Con respecto del tipo de ventanas que existes, hay algunas comunes como:

- La ventana Hamming dada por

$$W(n) = 0.54 - 0.46 \cos\left(2\pi \frac{n}{N}\right), 0 \leq n \leq N \quad (2.3.5)$$

- La ventana Hanning es

$$W(n) = 0.5 \left(1 - \cos\left(2\pi \frac{n}{N}\right)\right), 0 \leq n \leq N \quad (2.3.6)$$

- La ventana Blackman-Harris es:

$$W(n) = 0.36 - 0.49 \cos\left(\frac{2\pi n}{N}\right) + 0.14 \cos\left(\frac{4\pi n}{N}\right) - 0.01 \cos\left(\frac{6\pi n}{N}\right), \quad 0 \leq n \leq N \quad (2.3.7)$$

- La ventana rectangular es

$$W(n) = \begin{cases} 1, & 0 \leq n \leq N \\ 0, & \text{otro caso} \end{cases} \quad (2.3.8)$$

- La ventana Gaussiana, cuyos coeficientes están dados por

$$W(n) = e^{-n^2/2\sigma^2}, \quad 0 \leq n \leq N \quad (2.3.9)$$

donde en todos los casos $N + 1$ es el tamaño de la ventana. Para la ventana Gaussiana la desviación estándar es $\sigma = N/2 \propto$, con $\propto$ como factor de amplitud (en MATLAB el default tiene valor de 2.5) [56]. En la **Figura 2.4** se puede ver la representación de todas estas ventanas.

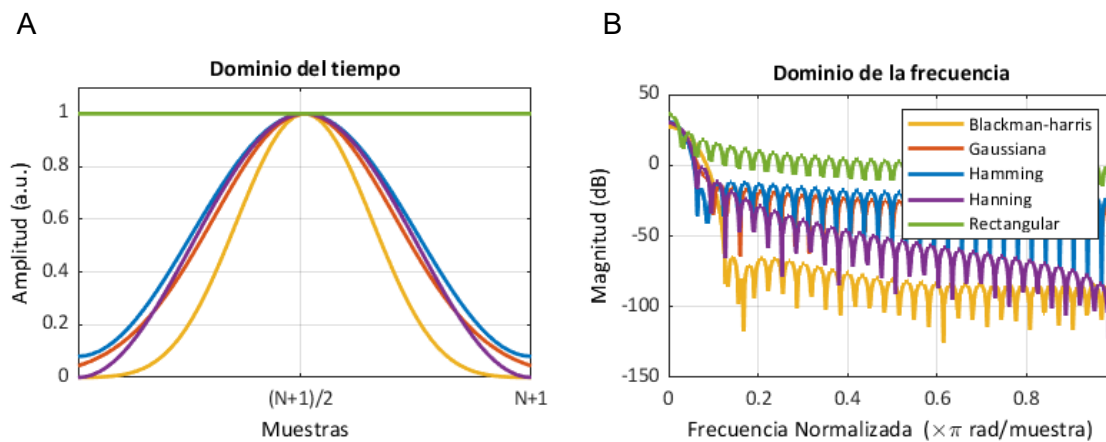


Figura 2.4. Ventanas de análisis utilizadas en el análisis tiempo-frecuencia en A) el dominio del tiempo y B) dominio de la frecuencia

2.3.2 Escalograma

Una alternativa al SP es la WT, para enfrentar la resolución TF con mayor flexibilidad. Además, la WT puede descomponer una señal según su frecuencia, utilizando una ondulatoria, llamada *wavelet* madre, la cual es la que descompone la señal y puede obtener su información TF, donde, el resultado depende del tipo de wavelet que se use [57]. La WT continua (*continuous wavelet transform*, CWT) se define como:

$$CWT(a, b; \psi) = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) dt \quad (2.3.10)$$

donde ψ es la *wavelet* madre, con un desplazamiento en el tiempo b y una dilatación en el tiempo a , estos parámetros encuentran secciones de la señal para analizarla en diferentes dilataciones [57].

Acerca de las *wavelet* madre, estas son funciones que deben cumplir con tener media cero y energía finita [58]. Por otro lado, cuando se tiene una señal con modulación oscilatoria, se deben usar las *wavelet* analíticas [59]. Entre sus características se encuentra que en el dominio de la frecuencia la onda es 0 para $w < 0$, i. e., no tienen respuesta para frecuencias negativas. Las funciones analíticas suelen ser complejas, pero la familia generalizada Morse de *wavelets* (*generalized Morse wavelets*, GMW) solo tiene parte real [60]. Las GMW fueron propuestas por Olhede y Walden en [61], que en el dominio de la frecuencia está definida por:

$$\Psi_{\beta,\gamma}(w) = \int_{-\infty}^{\infty} \psi_{\beta,\gamma}(t) e^{-i\omega t} dt = U(w) a_{\beta,\gamma} w^{\beta} e^{-w^{\gamma}} \quad (2.3.11)$$

donde $U(w)$ es el escalón unitario en el dominio de la frecuencia, β controla la oscilación, γ determina la simetría de la *wavelet* y $a_{\beta,\gamma} = 2(e\gamma/\beta)^{\beta/\gamma}$ es la constante

de normalización. La forma de la *wavelet* es controlada por γ y β , las cuales deben ser mayores a cero [60]. Adicionalmente, para que las GMW se mantengan analíticas deben estar dentro del área de Heisenberg, donde se localiza la energía de la *wavelet*. El límite inferior para entrar a esta área, está dado por $\gamma = 3$ [59]

A su vez, los parámetros de las GMW generalizan algunas familias de wavelets, por ejemplo, cuando $\gamma = 3$ las wavelets son las más gaussianas, y si además $\beta = 30$, su respuesta en la frecuencia es casi idéntica a la de las wavelet Morlet [62].

Para obtener la visualización de la información TF se calcula el cuadrado de la magnitud de la CWT la cual depende del tiempo t y la frecuencia ω , al resultado se le llama SC.

$$SC(t, \omega) = |CWT(a, b; \psi)|^2 \quad (2.3.12)$$

donde, la resolución en el tiempo y la frecuencia de la representación varía según la banda de frecuencia [60]. En la **Figura 2.5.B** se puede ver como el SC tiene múltiples resoluciones dependiendo de la localización a diferencia del SP (**Figura**

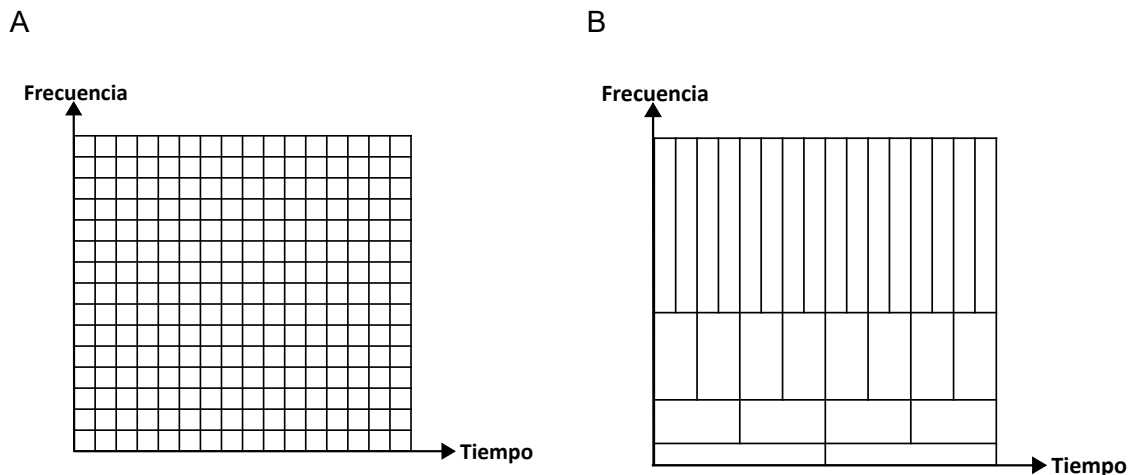


Figura 2.5. Resolución TF en A) espectrograma y en B) escalograma.

2.5.A), en el que la resolución es la misma independientemente del valor en la frecuencia. Esta técnica ha sido utilizada con el propósito de estimar la RR, donde en algunos trabajos se usa una única *wavelet* madre, como la *wavelet* Morlet [14], [16] o la Daubechies [15], aunque esta última ha sido utilizada con la WT discreta. La CWT ha sido un buen referente para usarse en señales no paramétricas y ha dado buenas aproximaciones en la estimación de la RR [14].

2.3.3 Transformada *Synchrosqueezing* basada en Fourier

Originalmente, la transformada *synchrosqueezing* fue propuesta por Daubechies y Maes en [63], la cual se aplica a la CWT, pero la transformada *synchrosqueezing* basada en Fourier (*Fourier-based synchrosqueezing transform*, FSST) reasigna la STFT. La FSST es un método no lineal que obtiene una versión más fina del SP [64], al hacer la reasignación solo en la frecuencia [65], por lo que esta técnica despertó interés para compararla con el SP y evaluar si hay una mejora debido a sus características. Además, en la búsqueda literaria que se hizo, esta técnica ha sido poco estudiada para estimar la RR, siendo Pirhonen et al. los únicos autores que se encontró que la hayan usado [19], aunque la FSST ha sido utilizada para procesar otras señales biomédicas como la frecuencia cardíaca a partir de señales de PPG [65].

Por su relación con la STFT, los parámetros que se varían en esta técnica también son el tipo de ventana y la duración. En la **Figura 2.6** se muestra un ejemplo de aplicar el SP y la FSST a la misma señal, donde en el SP se ve una definición menor alrededor de la frecuencia de la señal y al alejarse de esta se ve un degradado en la intensidad de las frecuencias, en cambio, en la FSST prácticamente solo hay dos intensidades de frecuencia, 1 y 0. Encima, la definición se realiza sin importar cual es la frecuencia o el tiempo, diferente a lo que ocurre en el escalograma, donde la resolución en las frecuencias grandes disminuye, pero aumenta en el tiempo.

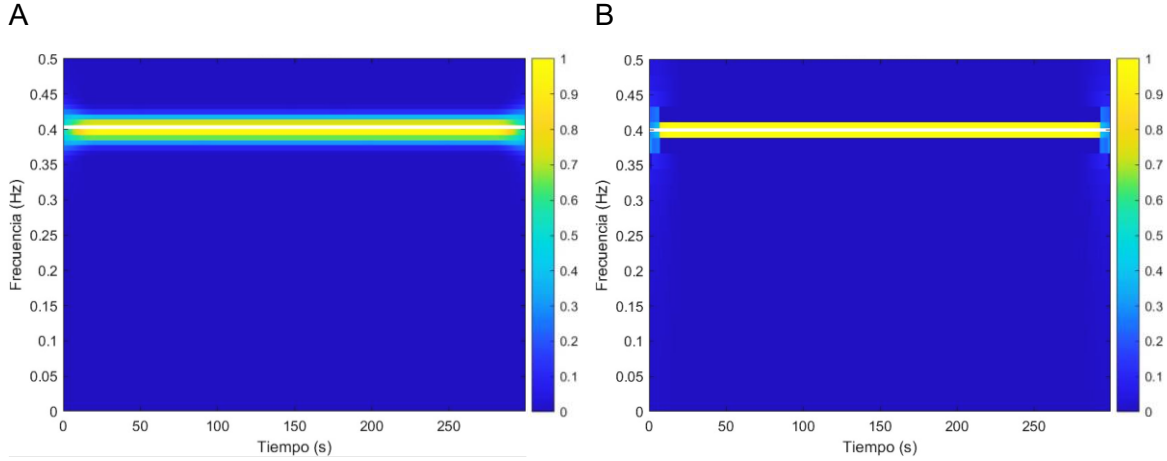


Figura 2.6. Representación tiempo-frecuencia de una senoidal con frecuencia de 0.4 Hz (24 bpm), utilizando una ventana Hamming de 45 s en el A) Espectrograma y en la B) transformada *synchrosqueezing* basada en Fourier.

El algoritmo de la FSST tiene 3 pasos, primero, se aplica la STFT a la señal bajo análisis (ver **ecuación (2.3.3)**). Segundo, se filtra la frecuencia instantánea de la siguiente forma:

$$\widehat{\omega}_f(t, \eta) = \text{Re} \left\{ \frac{1}{j2\pi} \frac{\partial}{\partial t} \frac{STFT(t, \eta)}{STFT(t, \eta)} \right\} \quad (2.3.13)$$

donde, $\widehat{\omega}_f(t, \eta)$ es la frecuencia instantánea local η en un tiempo t [66]. Tercero, se integra la STFT con el dominio de integración restringido al intervalo donde $\widehat{\omega}_f(t, \eta) = \omega$, entonces la FSST se define como:

$$FSST(t, \omega) = \frac{1}{g(0)} \int_{-\infty}^{\infty} STFT(t, \eta) \delta(\widehat{\omega}_f(t, \eta) - \omega) d\eta \quad (2.3.14)$$

donde g es la ventana [66], [67].

2.3.4 Espectro de Hilbert-Huang

El HHS es un método para analizar señales no estacionarias derivadas de sistemas no lineales y no estacionarios, a diferencia del análisis de Fourier, donde se asume

que las señales son estacionarias y la superposición de senoidales [68]. Adicionalmente, considera la energía regional en el plano local TF y proporciona una representación avanzada de las características de TF de señales multicomponentes [69]. El HHS se utiliza con algoritmos que descomponen a la señal en funciones moduladas en amplitud y frecuencia llamadas funciones de modo intrínseco (*intrinsic mode functions*, IMF), o simplemente modos de oscilación y la señal original se puede expresar como la sumatoria de estos IMF más una tendencia monótona final [22].

El método de descomposición básico es el EMD, el cual hace una separación local de la señal en oscilaciones rápidas y lentas. Con el EMD puede que las oscilaciones tengan escalas muy dispares en un IMF o escalas similares en diferentes modos, fenómeno que se conoce como “mezcla de modos” [69]. Es por eso por lo que se propuso la descomposición empírica de modos por conjuntos (*ensemble empirical mode decomposition*, EEMD), que genera un conjunto de versiones ruidosas de la señal original, obteniendo los resultados finales mediante promediado. Así, se obtienen modos más regulares y con escalas similares, pero la señal reconstruida contiene ruido residual y diferentes realizaciones de señal más ruido pueden producir un número diferente de modos [10].

Para enfrentar los problemas restantes, se propuso la descomposición empírica de modos de conjunto completo con ruido adaptativo (*complete ensemble empirical mode decomposition with adaptive noise*, CEEMDAN), la cual es una mejora importante al EEMD, por tener un error de reconstrucción despreciable y obteniendo el mismo número de modos para diferentes realizaciones de la señal. Cabe mencionar que EMD y EEMD han sido utilizadas para estimar la frecuencia respiratoria en varias ocasiones [11], [20], [21], [70], a diferencia del CEEMDAN, que solo se encontró que ha sido usado para procesar señales de electroglotografía, ECG y electroencefalografía [22].

En la **Figura 2.7** hay una comparativa de cómo se ven las TFR de una misma señal utilizando tanto el SP como el HHS vía CEEMDAN, en ambas imágenes se nota una distorsión alrededor de las frecuencias objetivo, aunque en el HHS se

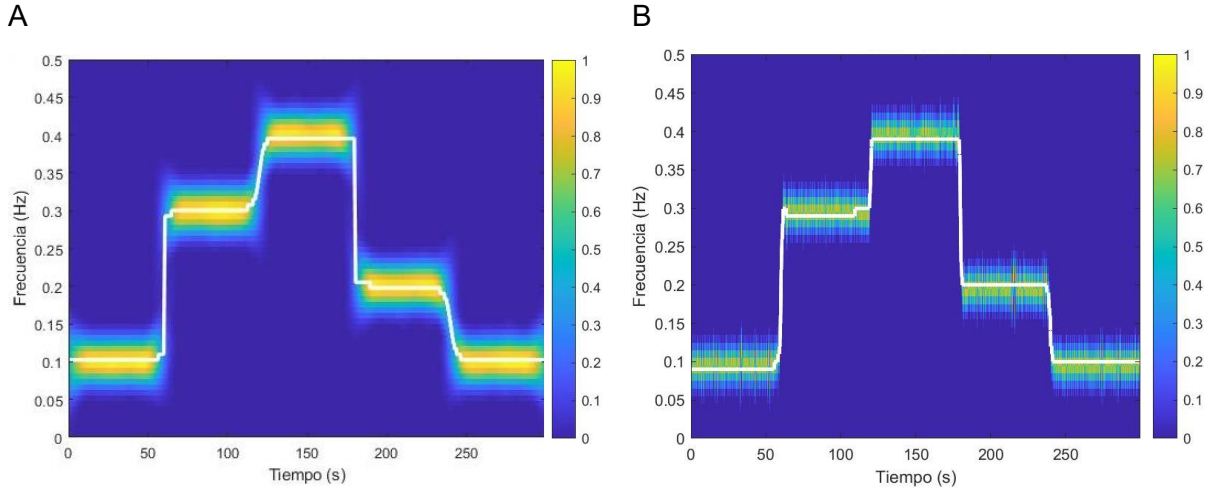


Figura 2.7. Representación tiempo-frecuencia de una senoidal con cambios en la frecuencia. A) Espectrograma utilizando una ventana Blakman-Harris de 30 s y B) Espectro Hilbert-Huang con ruido de 1 y 20 realizaciones.

observa una mayor definición en la intensidad de la frecuencia, sin embargo, las intensidades de frecuencia aledañas presentan mayor ruido.

El algoritmo comienza obteniendo los promedios locales por EMD y calculando el primer residuo, de la siguiente forma:

$$r_1(t) = M\left(s + \beta_0 E_1(w^{(i)})\right) \quad (2.3.15)$$

donde, M es el operador que produce el promedio local de la señal, β_0 es > 0 , $w^{(i)}$ es la realización número i de varianza unitaria con promedio cero de ruido blanco y E_k es el operador que produce el modo k obtenido por EMD y s es la señal original. Luego se obtiene el primer modo:

$$\tilde{d}_1 = s - r_1 \quad (2.3.16)$$

Después, se estima el segundo residuo como el promedio de las medias locales de las realizaciones $r_1 + \beta_1 E_2(w^{(i)})$ y entonces se define el segundo modo:

$$\tilde{d}_2 = r_1 - M \left(r_1 + \beta_1 E_2(w^{(i)}) \right) \quad (2.3.17)$$

Estos pasos se repiten hasta obtener K modos [22]. El siguiente paso es aplicar la transformada Hilbert en cada uno de los modos, la cual está dada por:

$$\mathcal{H}\{s(t)\} = s(t) * \frac{1}{\pi t} \quad (2.3.18)$$

donde $\mathcal{H}$ denota la transformada y $*$ al operador convolución. A partir del resultado la señal s , se transforma en una señal analítica z :

$$z(t) = s(t) + i\mathcal{H}\{s(t)\} \quad (2.3.19)$$

tal que, la amplitud instantánea $a(t)$ y la fase $\theta(t)$ están dadas por:

$$a(t) = [s^2(t) + \mathcal{H}^2\{s(t)\}]^{1/2}, \quad \theta(t) = \arctan\left(\frac{\mathcal{H}\{s(t)\}}{s(t)}\right) \quad (2.3.20)$$

Luego la frecuencia instantánea ω se obtiene derivando la fase de z con respecto al tiempo del siguiente modo [69]:

$$\omega = \frac{d\theta(t)}{dt} \quad (2.3.21)$$

Para cada instante de tiempo de la **Ec. (2.3.21)** existe un solo valor de frecuencia pudiendo representar solamente un componente y para encontrarle sentido a la frecuencia instantánea, a esto se le conoce como “banda estrecha”. El ancho de banda puede ser definido en términos de momentos espectrales y la frecuencia definida por la aproximación de fase estacionaria coincide con el mejor ajuste local de una función sinusoidal, i. e., no es forzoso contar con un periodo oscilatorio completo para poder definir un valor de frecuencia [69].

Capítulo 3

Metodología

En este capítulo se abordarán los pasos seguidos para completar los objetivos propuestos. En primer lugar, se incluye una sección con la descripción del conjunto de datos y de la instrumentación usada para adquirirlos. A continuación, se explica cómo se obtuvo la señal de iPPG a partir de los videos y la sección posterior está relacionada con la extracción y el acondicionamiento de la señal iPPG. Luego, se incluye una sección donde se expone como se realizaron las simulaciones y otra que describe cómo se seleccionaron los parámetros de las técnicas TF a partir de las simulaciones. También se incluye una sección que describe la estimación de la iRR a partir de datos reales utilizando los parámetros seleccionados y por último hay una sección sobre la comparación del desempeño del estimador considerando una señal respiratoria de referencia. En la **Figura 3.1** se muestra un diagrama de la metodología presentada.

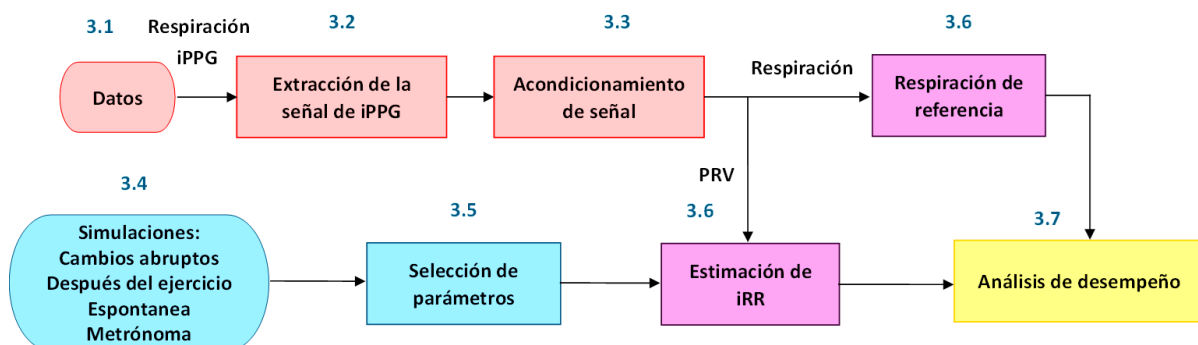


Figura 3.1. Diagrama general de la metodología. En azul se indica la sección donde se explica el bloque a detalle.

3.1 Base de datos

La base de datos, perteneciente al grupo de trabajo, almacena la información de 9 voluntarios sanos de entre 23 y 27 años, con un peso de 73.2 ± 11.96 kg y una altura de 1.68 ± 0.11 m. Cada uno de los voluntarios fue informado sobre el protocolo de adquisición y dieron su consentimiento informado. La adquisición de los datos se hizo mientras se llevaban a cabo diferentes maniobras respiratorias, como se describe en la **sección 3.1.2**. En cada maniobra se adquirieron las señales de iPPG, respiración y electrocardiografía (ECG). Para este proyecto, solo son fundamentales las primeras dos. Sin embargo, se hace uso de la señal de ECG con propósitos de alineación debido a la utilización de dos diferentes sistemas de adquisición.

Al momento de la adquisición, ninguno de los voluntarios padecía de enfermedades en las vías aéreas y pulmones como asma, bronquitis, enfisema o fibrosis quística. Incluso, la adquisición solo se realizó si no se encontraban enfermos de resfriado común, ni de ninguna otra infección respiratoria.

También, se les pidió a los voluntarios que no ingieran alimentos, no hicieran ninguna actividad física agotadora, no ingirieran alcohol o cafeína, ni que fumaran como mínimo dos horas antes de la adquisición, para evitar cambios en el flujo sanguíneo que pudieran afectar la señal de iPPG [71] .

3.1.1 Adquisición de datos

Se adquirieron simultáneamente las señales de respiración (referencia), ECG e iPPG, para las diversas maniobras respiratorias. La adquisición de la señal de respiración y ECG se realizó con el dispositivo inalámbrico BiosignalPlux (PLUX Wireless Biosignal S.A., Lisboa, Portugal) configurado a una frecuencia de muestreo de 1000 Hz. Cabe mencionar que esta elevada frecuencia de muestreo fue necesaria debido a que se adquirieron señales de ECG para propósitos ajenos a este trabajo de tesis. La señal de respiración se adquirió con ayuda de una banda flexible de respiración piezoeléctrica colocada alrededor de la caja torácica, a la

hora de la adquisición es recomendable eludir la colocación del sensor sobre vestimenta gruesa, holgada o que contribuya a reducir el estiramiento de la banda piezoeléctrica. El ECG se obtuvo con la derivación CM5, emulando la derivación estándar II, empleando tres electrodos adhesivos desechables de Ag/AgCl colocados sobre el tórax del voluntario.

Para la señal de iPPG, se utilizó la cámara de un *smartphone* de la marca Huawei modelo P9 Lite 2017, a una resolución de 320 x 240 píxeles (QVGA) de 8 bits con valores de 0 a 255 y a 30 cuadros-por-segundo (*frames-per-second*, fps). En archivos de imagen o video, una resolución elevada se traduce a un gran número de píxeles, lo que causa un mayor costo computacional durante su análisis. Si se tiene un archivo de 3 minutos a 30 fps a una resolución de 1920×1080 píxeles (full-HD), tomando una ROI del 100% se procesan aproximadamente la misma cantidad de píxeles que 27 archivos de video con las mismas características, pero a una resolución de 320 × 240 píxeles (QVGA) [13].

Además, el *smartphone* utilizado para adquirir secuencias de video, a partir de los cuales se extrae la señal de iPPG, cuenta con 12 megapíxeles con *flash* y un tamaño de píxel de 1.12 μm . Los datos se almacenaron en archivos de texto y videos compatibles para su posterior procesamiento en MATLAB (The Mathworks, MA, USA).

Todas las adquisiciones se realizaron en una oficina de la Facultad de Ciencias (UASLP), con iluminación ambiental común mediante ventanas y lámparas fluorescentes ubicadas en el techo. Se evitó la interrupción de otras personas durante la adquisición. Se empleó un calentador ambiental eléctrico en el área de medición para mantener una temperatura agradable, porque la temperatura del dedo durante la medición está vinculada a la perfusión sanguínea, lo que a su vez afecta la calidad de la señal de iPPG.

Durante la adquisición, se le pidió a cada voluntario que permaneciera sentado, callado y quieto. Además, se les proporcionó entre 1:00 y 1:30 minutos de

descanso si el voluntario así lo deseaba. A continuación, se describen las maniobras respiratorias realizadas por cada voluntario.

3.1.2 Maniobras respiratorias

Para obtener una diversidad de datos de respiración se emplearon diferentes maniobras respiratorias que se refieren a la forma en la que los voluntarios respiraron. La adquisición de la respiración espontánea es la más común en repositorios de bases de datos públicas y reportes de investigación, debido a que refleja la actividad natural del sujeto voluntario sin pedirle ninguna clase de esfuerzo [11]. Otro tipo de maniobra comúnmente utilizada es la guiada, donde el sujeto respira cada vez que se le pide para obtener registros con frecuencias deseadas, entre estas, la más común es la respiración constante, donde se han reportado rangos que van de 0.1 a 0.9 Hz, siendo lo más frecuente el intervalo de 0.2 a 0.6 Hz, con cambios de 0.1 Hz entre registros [14], [15], [72]. También, hay maniobras donde en un solo registro se hacen cambios en la frecuencia, e. g., Gastel *et al.* utilizaron maniobras de respiración constante con eventos pequeños de respiraciones a diferente frecuencia [18]. Las muestras en todas las referencias se tomaron de voluntarios sanos, sin enfermedades respiratorias o cardiovasculares.

En nuestro caso, la base de datos cuenta con cuatro tipos de maniobras adquiridas en una sola sesión y con pequeños descansos entre cada adquisición, debido a que algunas maniobras respiratorias podrían haber agitado al voluntario. A continuación, se detalla cada maniobra.

- **Espontanea:** Respiración en reposo, sin orientar de ninguna forma la respiración del voluntario.
- **Metronoma:** Se refiere a una respiración guiada, que es constante en la frecuencia. La base de datos tiene cuatro registros por voluntario con respiraciones de 6, 12, 18 y 24 bpm, que corresponden a 0.1, 0.2, 0.3 y 0.4 Hz, respectivamente.

- **Cambios abruptos:** Maniobra guiada donde la respiración tiene una frecuencia constante por un determinado intervalo de tiempo y de forma súbita la frecuencia cambia a otro valor constante y se mantiene en ella por otro intervalo de tiempo. La secuencia de cambios de la base de datos corresponde a 0.1, 0.3, 0.4, 0.2 y 0.1 Hz, con una duración de un minuto en cada valor.
- **Después del ejercicio:** Después de una actividad física que elevó la frecuencia respiratoria, se regresa a la respiración en reposo. Para hacer la adquisición, se conectó el sistema BiosignalPlux al voluntario, quien hizo una actividad física consistente en 20 a 25 sentadillas para aumentar su frecuencia respiratoria. Al terminar el ejercicio, de inmediato se comenzó la adquisición con el *smartphone*.

En total se adquirieron siete registros por voluntario, la mayoría con una duración de 3 minutos, excepto para el caso de la maniobra de respiración con cambios abruptos, como se muestra en la **Tabla 3.1** Cabe mencionar que algunos de los registros adquiridos se corrompieron y no estaban disponibles, mientras que otros no se grabaron correctamente. En ambos casos, se omitió su análisis. Por lo tanto, solo se contó con 44 registros del total de 63 (7 registros por cada uno de los 9 voluntarios), cada uno consistente en un video obtenido con el *smartphone* (para extraer la señal de iPPG) y un archivo con datos de ECG y respiración (señal de referencia en esta tesis).

Tabla 3.1 Duración de los registros de cada maniobra respiratoria.

Maniobra	Espontanea	Metrónoma	Cambios abruptos	Después del ejercicio
Duración	3 minutos	3 minutos	5 minutos, i. e., 1 minuto por cada frecuencia fija	3 minutos

3.2 Extracción de señales de iPPG

La adquisición de la señal de iPPG se realizó a partir de videos, donde para poder obtener la estimación de la RR es necesario transformarlos en una señal de pulso. En la **sección 3.1.1** se menciona que la resolución de los videos adquiridos fue de 320 x 240 pixeles (QVGA) a una frecuencia de 30 fps, puesto que se ha reportado una mayor amplitud pico a pico del iPPG a esta resolución [13], además de una menor carga computacional comparado con otras resoluciones mayores. La adquisición se hizo con el *flash* encendido, para obtener una amplitud mayor en los valores de intensidad [11], [14].

Para procesar los videos se utilizó MATLAB, donde los videos estaban en formato YCbCr '4:2:2' y se transformaron al formato de canales RGB (rojo: *Red*, verde: *Green*, azul: *Blue*). De ese modo, cada uno de los fotogramas de video pasaron a ser tres matrices con valores normalizados de tipo *double* de 0 a 1 de 8 bits y un tamaño de u renglones y v columnas.

Después, se fijó una región de interés (*region of interest*, ROI), i. e., solo se eligió un grupo de pixeles para extraer la señal de iPPG, con el propósito de mejorar el tiempo de procesamiento, además de que no son necesarios todos los pixeles para obtener la información con el fin de reconstruir esta señal [73]. Finalmente, la ROI fue promediada y de ese modo se extrajo un solo valor por fotograma. Es preciso mencionar que este proceso de extracción fue propuesto en otra tesis del grupo de trabajo. A continuación, se brindan mayores detalles respecto a la elección del canal de color, la ROI y el cómputo de la serie de tiempo de iPPG.

3.2.1 Elección del canal de color

En la literatura se ha mencionado que el mejor canal de color para extraer las señales de PR y RR corresponde al verde (G), porque es el menos afectado por artefactos de movimiento y cambios en la luz, al compararlo con el canal rojo (R) y azul (B) [74], [75], [72]. El canal B suele presentar picos que son falsos positivos,

ruido y tiene una amplitud pico a pico pequeña. En el caso del canal R, este funciona bien cuando se tiene el *flash* apagado, de lo contrario, el valor de todos sus pixeles se satura en 255 [13]. En algún momento, este canal se reportaba como el más adecuado para extraer la señal de iPPG [11], [76], aunque en la actualidad las cámaras de los *smartphones* tienen una mayor resolución, lo que ha colocado al canal G como el predilecto.

Asimismo, el uso de la luz artificial del *flash* de los dispositivos inteligentes, cuando se está haciendo la adquisición, ayuda a obtener una mayor intensidad en los pixeles y mejor amplitud pico a pico, como se ha mencionado en diversas investigaciones [72], [13], [76], [77].

Por lo tanto, para hacer la estimación de la señal de RR a partir de la señal de iPPG, la adquisición de los datos se llevó a cabo con el *flash* encendido y a la hora de procesar los videos se seleccionó el canal G. En la **Figura 3.2** se puede observar cómo funciona el formato RGB en los fotogramas y como solo se toma el canal de los verdes.

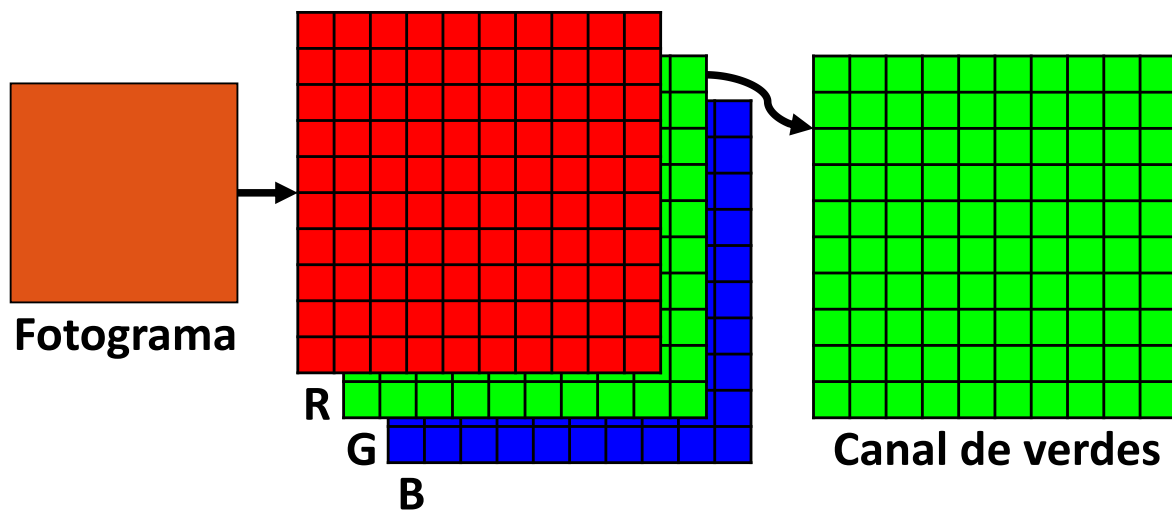


Figura 3.2. Ejemplificación de cómo se obtiene el canal verde. Los fotogramas en formato RGB (*red, green, blue*), están formados por tres canales de colores que pueden combinarse para formar un espectro de colores dependiendo de su intensidad.

3.2.2 Región de interés

La selección de la ROI no solo trata sobre su tamaño sino también sobre su ubicación espacial en el video. Esto se debe a que una ROI cercana al *flash* de la cámara, mejora la estimación de la señal iPPG [78]. Sin embargo, en el mercado de *smartphones* el *flash* se puede encontrar en distintas posiciones con respecto a la cámara. Por lo que, la ROI se estableció para los píxeles centrales.

Para establecer un tamaño adecuado de la ROI, debe considerarse que una ROI grande no asegura una mayor amplitud pico a pico en la señal [14]. Por lo que el grupo de trabajo hizo pruebas piloto con distintos porcentajes de la cantidad de píxeles utilizados para estimar la señal de iPPG. Como resultado, se estableció una ROI del 40% (128 x 96 píxeles), como se ilustra en la **Figura 3.3**, la cual reporto una buena amplitud pico a pico de 3.38×10^{-2} u.a. y un tiempo de procesamiento bajo de 14.25 segundos en comparación con otros porcentajes, por ejemplo, con una ROI del 100 % se había obtenido una amplitud pico a pico de 3.27×10^{-2} y un tiempo de procesamiento de 23.74 segundos, a diferencia de con una ROI del 40%, lo que

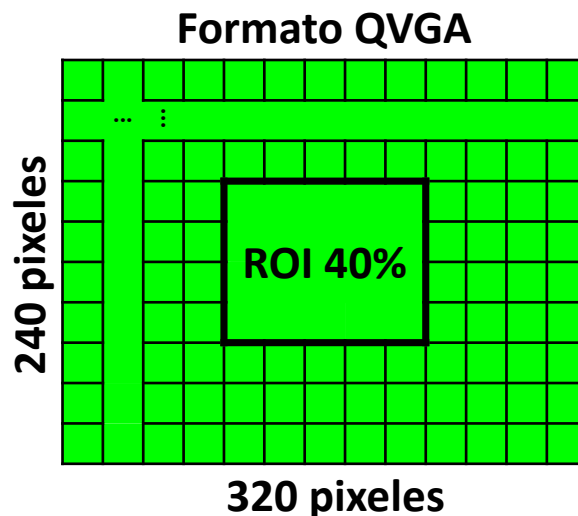


Figura 3.3 Elección de la ROI en un 40 % de los píxeles (128 x 96) y en el centro del canal de los verdes, el cual originalmente está en un formato QVGA de 320 x 240 píxeles.

concuenda con que el tamaño de la ROI no garantiza una mayor amplitud pico a pico y que aumenta el tiempo de procesamiento [13].

Al tener la ROI seleccionada, puede realizarse el promediado de canal G, para cambiar el formato bidimensional de la señal a uno unidimensional, i. e., que tome un solo valor de amplitud en el tiempo. Esto se hace siguiendo la ecuación

$$iPPG[n] = \frac{1}{|ROI|} \sum_{[i,j] \in ROI} g_{i,j}[n] \quad (3.2.1)$$

donde $iPPG$ corresponde a la señal extraída, $|ROI|$ es el numero de pixeles seleccionados del canal G, con pixeles $g_{i,j}$, y n denota el n – ésimo fotograma [73].

De esta forma se realizó la extracción de las señales de iPPG a partir de las grabaciones del dedo índice en contacto directo con la cámara del *smartphone*. Una vez obtenida la señal, esta pasó a la etapa de preprocesamiento.

3.3 Preprocesamiento de señales de iPPG

Antes de iniciar la obtención de la iRR, es imperativo que la señal de iPPG y la señal de referencia pasen por una etapa de acondicionamiento. El preprocesamiento se hizo en MATLAB.

3.3.1 Remuestreo

En primera instancia, la señal de iPPG fue remuestreada considerando que las señales de iPPG obtenidas en la etapa anterior tienen la misma frecuencia que los videos, i. e., 30 Hz. No obstante, en algunas ocasiones los videos no se grabaron a esta frecuencia constante, sino que algunos videos presentan variaciones entre 27.5 y 30 Hz debido a procesos internos del *smartphone* durante la grabación. Además, la referencia está a una frecuencia de 1000 Hz. Como la señal de iPPG y la referencia están a diferentes frecuencias, las señales de iPPG inicialmente se

remuestrearon a 30 Hz (frecuencia de la señal de iPPG) mediante *splines* cúbicos para que coincidieran con su correspondiente señal de referencia respiratoria.

3.3.2 Filtrado digital

Una vez que las señales se encontraban a la misma frecuencia de muestreo, siguió la etapa de filtrado necesaria para reducir el ruido en componentes de frecuencia que no son de interés y mantener las que sí lo son. Ambas señales fueron filtradas con un filtro IIR (infinite impulse response, respuesta al impulso infinita) pasabanda de 4° orden, con una banda de paso diferente para cada señal. En el caso de la iPPG, la banda de paso fue de 0.3 a 5 Hz y fue seleccionada tomando en cuenta la frecuencia cardiaca, la cual normalmente es de 0.65 a 4 Hz. Del mismo modo, para la señal respiratoria se eligió una banda de paso de 0.05 a 1.5 Hz, debido a que en condiciones normales los componentes de la respiración van de 0.1 a 1 Hz.

3.3.3 Normalización

Las señales de iPPG y respiración se pasaron por una etapa de normalización, con el objetivo de que todas ellas estén al mismo nivel de amplitud. En este caso, las señales quedaron en un rango de valores de 0 a 1, siguiendo la ecuación

$$S_{norm}(t) = \frac{S(t) - \min(S(t))}{\max(S(t)) - \min(S(t))} \quad (3.3.1)$$

donde, $S(t)$ denota a la señal que resultante de la etapa anterior de filtrado y $S_{norm}(t)$ a la señal normalizada.

3.3.4 Alineación de la señal de iPPG y la señal de respiración

Las señales de respiración e iPPG fueron adquiridas con diferentes sistemas. Así, a pesar de procurar que la adquisición en ambos empezara al mismo tiempo, se tenían casos donde uno de ellos comenzó la adquisición un momento después que el otro. Como resultado, al superponerlas en realidad no coincidían temporalmente,

i.e., una de ellas estaba temporalmente adelantada o retrasada en relación con la otra.

Para que ambas señales pudieran compararse adecuadamente se realizó una etapa de alineación, donde se utilizó un algoritmo desarrollado por el grupo de trabajo, el cual emplea la señal de ECG, adquirida simultáneamente con la respiración por el mismo equipo, y la señal de iPPG. En la mayoría de los casos, el retardo la señal de iPPG estaba retrasada con respecto a la respiración, debido al encendido y la estabilización de la cámara. Este retardo ocasiono que la señal de iPPG fuese más corta, pero no por más de tres segundos. Por consiguiente, las señales de iPPG fueron las que generalmente debían adelantarse con respecto a las señales de referencia, mientras que estas últimas eran recortadas en sus límites, logrando que todas las señales fueran del mismo tamaño.

La alineación automática se estableció utilizando la función de correlación cruzada de señales, la cual mide el grado de similitud entre ambas. Como las señales de ECG e iPPG son morfológicamente diferentes, se utilizaron las series de HR y PR, que se extraen de las primeras, respectivamente, puesto que estas series deben coincidir en amplitud y en temporalidad. La función de correlación lineal está definida por:

$$r_{x_1x_2}[n_1] = \frac{1}{L} \sum_{n=0}^{N-1} x_1[n]x_2[n - n_1], \quad 0 \leq n_1 < Q_1 \quad (3.3.2)$$

donde Q_1 denota el tamaño de la señal $x_1[n]$, la cual es la señal más corta, mientras que $x_2[n]$ denota la señal más larga. Cuando hay un pico predominante en $r_{x_1x_2}[n_1]$, en ese retardo de n_1 muestras, la similitud entre ambas señales es mayor [79]. Entonces, la segunda señal se desplaza n_1 muestras y en un vector se guardan los índices de desplazamiento de x_2 , los cuales se utilizan para obtener la versión desplazada de la señal, resultando en la alineación de ambas señales. Finalmente,

este vector de índices también se utilizó para desplazar y definir los límites en la señal de respiración de referencia.

3.3.5 Detección de picos de iPPG

Después, se hizo una búsqueda de los picos de la señal de iPPG, con el objetivo de hacer una interpolación y de ese modo obtener la serie PR, la cual presenta las oscilaciones provocadas por la respiración. Los picos de la señal de iPPG se encontraron con un algoritmo desarrollado por el grupo de trabajo [80], el cual toma en cuenta la morfología de la señal, para encontrar correctamente los máximos locales de la señal. A continuación, se explica paso a paso el algoritmo.

1. Primero, se establece una ventana de dos segundos que toma segmentos de toda la señal de iPPG de forma iterativa, con una superposición entre ventanas adyacentes del 50% del tamaño de esta ventana, i. e., de un segundo.
2. Cada uno de los segmentos de la señal ingresan a un filtro IIR pasabanda de 4° orden con banda de paso de 0.5 a 3 Hz. A posteriori, se hace un análisis espectral, basado en el periodograma de Welch y se calcula la frecuencia promedio que se utiliza para fijar el tamaño de dos sub-ventanas, la primera del tamaño de un periodo de la frecuencia promedio y la segunda con tamaño de dos periodos.
3. Las ventanas de 2 s originales se toman otra vez para ser filtradas de la misma forma que en el paso anterior, pero ahora con una banda de paso de 0.8 a 2 Hz. En seguida, los segmentos ingresan a la primera búsqueda de candidatos a máximos locales, los cuales cumplan con ser mayores que las muestras contiguas.
4. Se toma la primer sub-ventana del paso 2, tomando como muestra central los candidatos a máximo local y se aplica un umbral del 50 % de la ventana original. Si el umbral es superado, se seleccionará el máximo de mayor amplitud en cada sub-ventana y se establecerán como los segundos

candidatos a máximos locales. Si ningún punto supera el umbral, el candidato a máximo local se descarta.

5. Se utiliza la segunda sub-ventana para evaluar si los segundos candidatos inmediatos entre si se encuentran dentro de un rango de frecuencia cardiaca. Con esta sub-ventana se tienen al menos dos máximos locales, y si se establece un intervalo de medición para PR de 40 a 240 BPM (beats-per-minute, latidos-por-minuto), i. e., 0.67 a 4 Hz. En el momento que se encuentran dos puntos máximos que cumplen con esta condición, se toma el de mayor amplitud y el otro es descartado. Esto se hace primero en las ventanas originales y luego en la señal integrada de iPPG, donde se obtiene un vector que almacena los máximos globales finales de la señal de iPPG.

Con este algoritmo se logra encontrar los picos de la señal de iPPG, minimizando los efectos de las oscilaciones en la línea base, las reducciones en la amplitud pico-pico de los pulsos y los pulsos falsos, luego con los picos se hace una interpolación de *splines* cúbicos a 4 Hz, para obtener la serie PR. No obstante, cuando el dedo con el que se adquiere la señal se aleja de la cámara o hay algún otro tipo de perturbación inesperado, regularmente existe una reducción parcial o total de la amplitud pico-pico de los pulsos. Con la intención de enmendar este tipo de situaciones, después de encontrar los picos se realiza una corrección automática de estos, para encontrar picos faltantes. Así, cuando un máximo local no es detectado, hay presencia de cambios repentinos en la serie instantánea de PR que no pertenecen a variaciones en la respiración. Este algoritmo se lleva a cabo a través de los siguientes tres pasos:

1. Se inicia considerando ventanas de 15 s de duración con 5 s de traslape de la serie PR, para obtener su primera derivada. Entonces, se utiliza un umbral en la derivada y si este umbral es superado quiere decir que hay un pico faltante, por lo que comienza la búsqueda de ese máximo faltante.
2. La búsqueda del máximo local se hace tomando la distancia entre los máximos que se encuentran en el intervalo de 15 s de la señal de iPPG. Si

hay una distancia menor a 40 BPM se establece una sub-ventana, que inicia en la zona donde el umbral fue superado y termina donde la derivada vuelve a estar debajo del umbral.

3. Finalmente, se calcula la segunda derivada de la sub-ventana de la serie PR, en la cual se buscó un máximo que este en el rango de valores normales de frecuencia de pulso. En caso de que la amplitud de un pulso no exista, o mejor dicho, no hubiera un máximo, se establece como máximo local faltante la muestral central de la sub-ventana.

Al finalizar esta corrección automática, se espera que la serie de tiempo PR contenga los mínimos errores. Aun así, después de estos dos algoritmos pudiera requerirse una etapa de corrección manual de picos, donde se deban retirar picos de la serie, en caso de que los picos estén muy cercanos y realmente no pertenezcan a la frecuencia del pulso o sean picos repetidos, o también se pueden agregar picos.

3.4 Simulaciones

Se generaron señales sintéticas de respiración para simular diferentes maniobras respiratorias, usando una señal base y agregándoles diversos niveles de ruido. Cabe mencionar que las señales sintéticas buscan emular las señales de respiración adquiridas con los sensores de las bandas de respiración empleadas como referencia, así como las series de tiempo PR extraída de señales de iPPG adquiridas con la cámara del *smartphone*, i. e., el componente de la señal de iPPG que es modulado en frecuencia por la respiración mediante el mecanismo de la arritmia sinusal respiratoria.

3.4.1 Maniobras respiratorias simuladas

Se simularon señales de respiración correspondientes a cuatro tipos de maniobras: espontánea, metrónoma o constante, cambios abruptos y disminución gradual después del ejercicio. Con cada una de estas maniobras se buscó cubrir las

diferentes situaciones respiratorias de la base de datos disponible. Adicionalmente, se hizo la simulación de la TFR ideal para cada maniobra, i. e., se simuló tanto la forma de onda como el espectro variante en el tiempo de las señales de respiración. Las señales y TFR se simularon con una duración de 3 minutos, con exclusión de la maniobra con cambios abruptos que fue de 5 minutos, de acuerdo con la longitud de los datos de las señales adquiridas. A continuación, se describe la simulación de las maniobras respiratorias:

- **Espontanea:** Correspondió a una señal sinusoidal con una modulación en la frecuencia de tipo sinusoidal que varía entre los 0.13 y 0.15 Hz, i. e., entre (7.8 y 9 bpm, respectivamente. Se empleó la siguiente ecuación [81]:

$$s(t) = \sin\left(2\pi \frac{14}{100}t + 6\sin\left(2\pi \frac{1}{60}t\right)\right) \quad (3.4.1)$$

donde $s(t)$ denota la señal de respiración sinusoidal en el dominio del tiempo, con un período de 60 s, una frecuencia media de 0.14 Hz y una amplitud de 0.01 Hz, correspondientes a los valores entre los que oscila la señal de respiración.

- **Metronomas:** Se simularon señales sinusoidales con frecuencias fijas de 6,12,18 y 24 bpm. Por lo tanto, su TFR ideal corresponde a un único componente horizontal constante a la frecuencia dada, como se especifica en la siguiente ecuación:

$$s(t) = \sin(2\pi ft) \quad (3.4.2)$$

donde $s(t)$ denota la señal en el tiempo y f es el valor constante en frecuencia de la señal, el cual tomará valores para coincidir con las cuatro frecuencias respiratorias constantes mencionadas anteriormente.

- **Cambios abruptos:** En el dominio del tiempo, dado en amplitud, es la concatenación de señales sinusoidales de duración finita y con frecuencias

fijas, simuladas en el mismo orden y frecuencias que en las adquisiciones de la base de datos, i. e., 0.1, 0.3, 0.4, 0.2 y 0.1 Hz. En la **Ec. (3.4.3)**, se muestra cuando la señal $s(t)$ tiene como resultado la frecuencia, según el tiempo t .

$$s(t) = \begin{cases} \sin(2\pi \cdot 0.1t), & 0 \leq t < 60 \\ \sin(2\pi \cdot 0.3t), & 60 \leq t < 120 \\ \sin(2\pi \cdot 0.4t), & 120 \leq t < 180 \\ \sin(2\pi \cdot 0.2t), & 180 \leq t < 240 \\ \sin(2\pi \cdot 0.1t), & 240 \leq t < 300 \end{cases} \quad (3.4.3)$$

Esta es la única maniobra respiratoria en la que la adquisición tuvo una duración de cinco minutos (300 s), por lo cual su simulación tiene una duración que coincide.

- **Cambios graduales en la frecuencia:** Señal sinusoidal con una variación exponencial en la frecuencia instantánea con pendiente negativa la primera mitad de su duración y con modulación sinusoidal la segunda mitad, para simular la disminución y el retorno a la línea de base al terminar una actividad física. Al final la fórmula es:

$$s(t) = \begin{cases} \sin(2\pi(-23.805 + 24.129t^{0.999})), & 0 \leq t < 60 \\ \sin\left(2\pi \frac{11.5}{100}t + 6\sin\left(2\pi \frac{1}{60}t\right)\right), & 60 \leq t < 180 \end{cases} \quad (3.4.4)$$

donde $s(t)$ es la señal en el dominio del tiempo, cuya respuesta está en Hz. En la primera parte de la señal, la frecuencia máxima empieza en 0.3 Hz, con una caída exponencial hacia 0.12 Hz durante en 60 s [82]. La segunda parte de la señal es como las simulaciones espontaneas, pero con una frecuencia central de 0.115 Hz para que coincida con el final de la primera parte.

Cabe resaltar que también se simularon las TFR ideales para cada maniobra. Para ello, se generaron vectores de tiempo y frecuencia con una frecuencia de muestreo de 30 Hz y 4096 intervalos de frecuencia. Las imágenes TF se crearon a

partir de una matriz de ceros (0) a la que se les sumó un 1 en los valores de amplitud en las ubicaciones de frecuencia de acuerdo con los patrones de respiración de cada maniobra, i.e., basándose en las **Ecs. (3.4.1) a (3.4.4)**, previo a que se les sumara el ruido aditivo descrito a continuación.

3.4.2 Ruido aditivo

Las señales ideales no representan completamente a las adquisiciones reales, debido a que en ellas no se consideran las perturbaciones que pueden ocurrir en el ambiente o condiciones de adquisición, como algún cambio en la luz ambiental o un movimiento fortuito del voluntario, que se reflejen como ruido en las señales de iPPG y de respiración. Por esto, se planteó alterar las señales ideales en función de las características encontradas en la base de datos. Así, se calculó la relación señal-a-ruido (*signal-to-noise ratio*, SNR) de las señales, siguiendo la **Ec. (3.4.5)**

$$SNR_{dB} = 10 \log_{10} \frac{\sum_{n=1}^N |s_n|^2}{\sum_{n=1}^N |r_n|^2} \quad (3.4.5)$$

donde s_n denota la n – ésima muestra de la señal y r_n la n – ésima muestra del ruido aditivo, ambos con N muestras de longitud [83].

El valor promedio de SNR encontrado en la base de datos fue de 31 dB y se optó por usar ruido gaussiano aditivo en todas las simulaciones. Para propósitos de seleccionar el rango de valores de SNR que se emplearon en las simulaciones, se graficaron SPs de señales reales y de simulaciones con diferentes niveles de ruido. Al observarlos simultáneamente, se notó que algunos SPs reales tenían similitud con SPs de simulaciones para valores SNR menores al promedio, como se ve en la **Figura 3.4**, donde las TFR de la señal de respiración real tienen componentes de interferencia que aún se presentan en la imagen TF de la simulación obtenida con -10 dB de SNR y debido a esto se optó por tomar niveles de ruido menores. Además, para que las señales simuladas con ruido también tuvieran una TFR más cercada a lo ideal, también se tomó un valor SNR mayor al promedio. Así, los valores de SNR

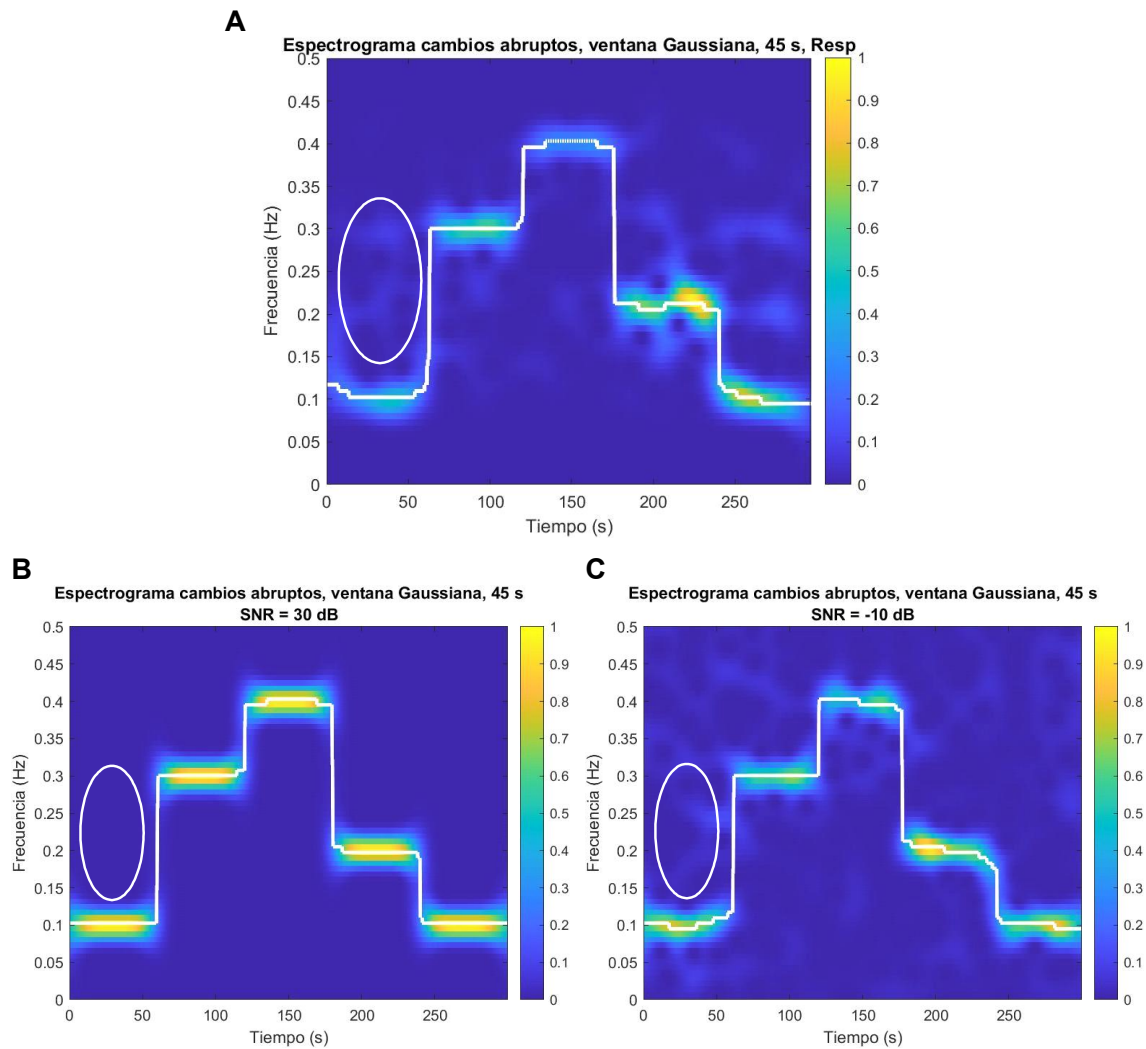


Figura 3.4 Espectrogramas con ventana Gaussiana de 45 s. A) Señal de respiración real. B) Simulación con SNR de 30 dB. C) Simulación con SNR de -10 dB. En los óvalos blancos en A) y C) es notable la presencia de términos de interferencia, mientras que en B) es menos notable este fenómeno.

empleados en las simulaciones se fijaron en 10, 20, 30 y 40 dB, como se muestra en la **Tabla 3.2**, dando un total de 28 simulaciones entre todas las maniobras respiratorias.

Tabla 3.2 Señales simuladas. A cada maniobra respiratoria se le adicionaron cuatro cantidades de relación señal-a-ruido (SNR).

Maniobra	SNR			
Espontanea	10	20	30	40
Cambios abruptos	10	20	30	40
Después del ejercicio	10	20	30	40
6 bpm	10	20	30	40
12 bpm	10	20	30	40
18 bpm	10	20	30	40
24 bpm	10	20	30	40

3.5 Selección de parámetros de la TFR

La selección de los parámetros implica que se debe hacer una comparación, preferentemente cuantitativa, entre las imágenes TF obtenidas con diferentes parámetros para cada técnica. Entonces, en cada una de las simulaciones de las maniobras respiratorias se calcularon las TFR con diferentes parámetros, los cuales se procuró que fueran variados y adecuados para entre ellos encontrar valores adecuados para finalmente usarlos en las señales reales de la base de datos.

Para el SP y la FSST se variaron el tipo y tamaño de ventana. La duración de ventana se eligió en un rango de 1 s al tamaño de la señal, i. e., 180 s o 300 s, dependiendo de la maniobra respiratoria. Entonces, se tomaron valores de 1, 4, 8, 10, 15, 20, 30, 45, 60, 120 y 300 s (180 s). Las ventanas que se usaron fueron la Blackman-Harris, la Gaussiana, la Hamming, la Hanning y la Bartlett (rectangular). Cada ventana se ajustó a diferentes tamaños, lo que dió un total de 55

combinaciones de ventanas para cada una de las dos técnicas, para cada una de las distintas simulaciones (4 tipos de maniobras con 4 valores de SNR cada una).

En el SC, los parámetros se variaron en función a sus propiedades y al área de Heisenberg, la cual comienza cuando $\gamma \geq 3$ y $\beta > \frac{1}{2}$, además el producto de ambos debe ser mayor a γ [60] y β no puede ser mayor a 40 [84], con el fin de no tener una resolución en el tiempo y la frecuencia distorsionada. Por estos motivos a γ se le asignaron valores de 3, 10, 20 y 30, y a β se le asignaron valores de 2, 10, 20, 30 y 40, lo que provocó que se pudieran ver los resultados entre las posibilidades disponibles para esta técnica. Cabe mencionar que hay estudios utilizando la *wavelet* Morlet, la cual se asemeja a las *wavelets* Morse cuando $\gamma = 3$ [14], [16]. Resultando en un total de 20 combinaciones para cada una de las simulaciones.

El HHS vía CEEMDAN fue la única técnica en la que se variaron tres parámetros. Los primeros dos están relacionados con la técnica CEEMDAN, donde se varió la desviación estándar (SD) del ruido, tomando valores de 0.1, 1 y 10, para observar el cambio por décadas. También se varió el número de realizaciones, del cual se usaron 1, 5, 10, 15 y 20, porque ese es un rango que va desde el valor mínimo de 1 a un valor de 20, a partir del cual se dejan de presentar mayores cambios en el resultado de la descomposición. Aunque Colominas et al. mencionan que el error aumenta con el tamaño de realizaciones, lo que le atribuyen al hecho de que diferentes copias ruidosas de la señal producen un número diferente de modos, lo que aumenta la probabilidad de tener un número diferente de modos, y también lo hace el error de reconstrucción, sin mencionar que el uso de más realizaciones tiene un costo computacional mayor [22]. El último parámetro corresponde al uso de un filtro gaussiano pasabajas con SD de 1.5 en las TFRs, después de entrar al algoritmo del HHS, para mejorar la legibilidad de esta TFR. Así, para esta técnica TF se usaron 30 combinaciones de parámetros para cada simulación.

Tabla 3.3 Parámetros utilizados en cada técnica tiempo-frecuencia.

Técnica TF	Parámetros		
Espectrograma	Ventana	Longitud de ventana	
	Blackman-Harris, Hamming, Hanning, Gaussiana y Bartlett	1, 4, 8, 10, 15, 20, 30, 45, 60, 120 y 300(180) s	
Transformada Synchrosqueezing basada en Fourier	Ventana	Longitud de ventana	
	Blackman-Harris, Hamming, Hanning, Gaussiana y Bartlett	1, 4, 8, 10, 15, 20, 30, 45, 60, 120 y 300(180) s	
Escalograma	Gamma	Beta	
	3, 10, 20 y 30	2, 10, 20, 30 y 40	
Espectrograma Hilbert-Huang	SD del ruido	Realizaciones	Filtro
	0.1, 1 y 10	1, 5, 10, 15 y 20	Si y no

En la **Tabla 3.3** se muestra cómo se relacionan los parámetros mencionados, donde cada tipo de parámetro se utilizó con cada uno de los demás tipos de parámetro, por ejemplo, en el escalograma una $\gamma = 3$, se utilizó con $\beta = 2, 10, 20, 30$ y 40 , lo que dio 5 TFR por simulación con una $\gamma = 3$. De la sumatoria de todas las combinaciones de parámetros utilizadas se obtiene que en cada simulación respiratoria se obtuvieron 160 TFR y como son 28 simulaciones al final se obtuvieron 4,480 TFR.

3.5.1 Obtención de la iRR a partir de las TFR

Para estimar la frecuencia respiratoria instantánea a partir de una imagen TF se utilizó un algoritmo que analiza la matriz TFR, donde las columnas representan a los instantes en el tiempo y las filas a los valores del vector de frecuencia (frequency *bins*). Las estimaciones de la iRR se obtuvieron de cada uno de los resultados de las simulaciones y también de las TFR ideales para comparar los resultados entre

ambas. A continuación, se describe el algoritmo empleado para estimar iRR a partir de una TFR, donde se restringen los saltos de frecuencia con una penalización que es proporcional a la distancia entre los *bins* de frecuencia.

- Primero, en una matriz A se transforma a su logaritmo negativo y también se define una nueva matriz B con J filas e I columnas que almacena nuevos valores, la primera columna es igual a la primera columna de la matriz TF.
- Después, se utiliza cada uno de los elementos de la matriz TF, comenzando con el primer elemento de la segunda columna, para evaluar las posibles soluciones vecinas al actualizar el valor del elemento seleccionado, en este caso con el valor mínimo resultante de la adición de la columna anterior (para la segunda columna sería la primera columna de la matriz), con la distancia entre cada *bin* al cuadrado multiplicada por la penalización. Para este trabajo de tesis se fijó una penalización con valor de 10. Esto se repite para cada uno de los siguientes elementos de la matriz con *i* como la fila e *j* como la columna, avanzando primero por las filas, de la siguiente forma:

$$B(i, j) = A(i, j) + \min(B(:, j) + penalizacion * distancia^2) \quad (3.5.1)$$

- Adicionalmente, se guardan los índices del número de donde proviene el mínimo, para conocer el elemento de donde proviene cada punto.
- Finalmente, se obtiene el valor mínimo de la última columna y se regresa al índice de la columna anterior de donde este proviene, resultando en un único valor por columna, proporcionando una estimación suavizada.

El funcionamiento del algoritmo se ilustra en la **Figura 3.5** con un ejemplo de cómo se estima la iRR a partir de una imagen TF de 3x3, donde se puede ver cómo funciona la **Ec. (3.5.1)**, utilizando una penalización de 10 para castigar en gran medida valores que no son cercanos a 1. El resultado obtenido fue el esperado, puesto que el valor mínimo de la columna final es 0, al que se le asigno el índice

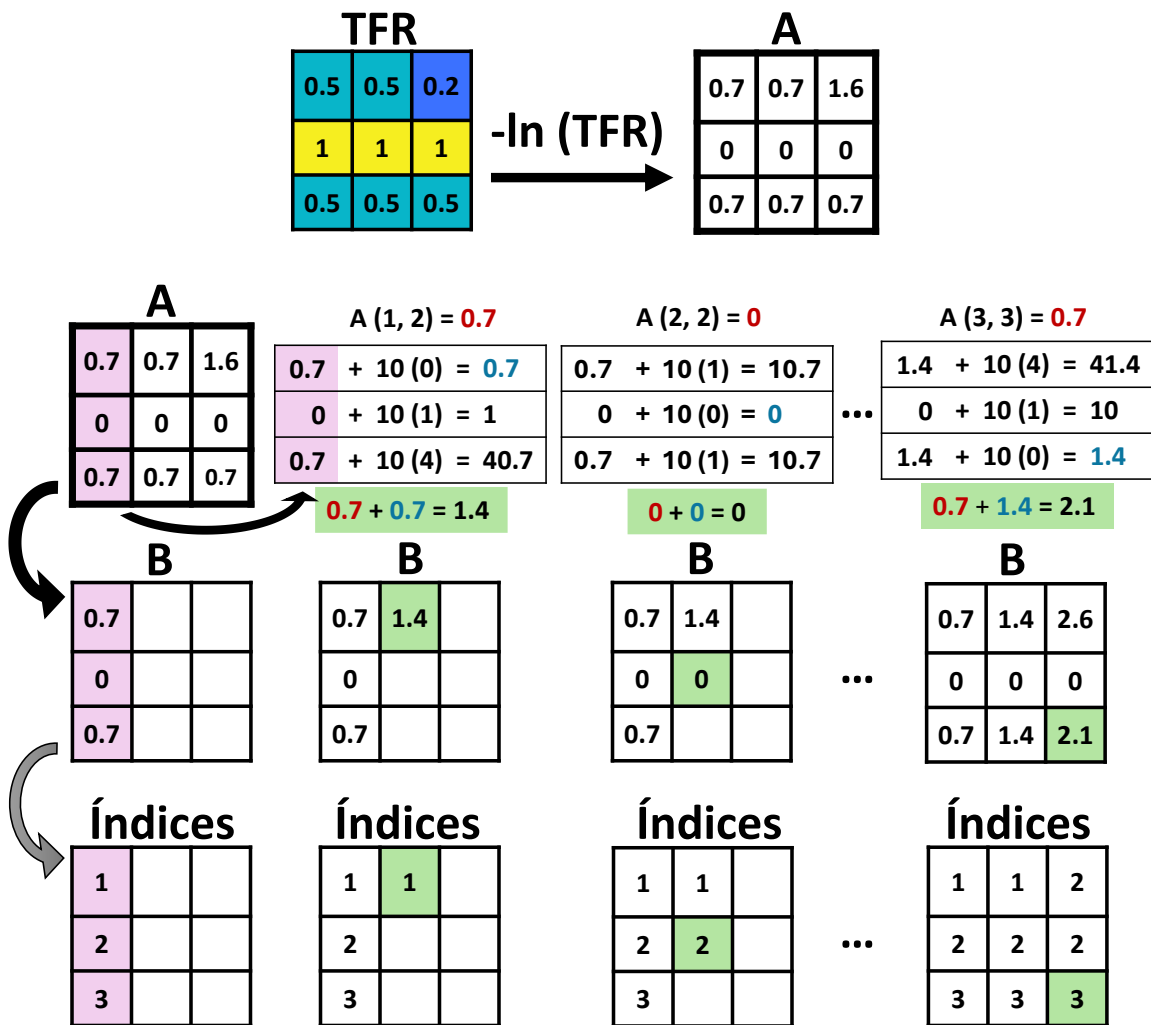


Figura 3.5 Funcionamiento del algoritmo greedy de búsqueda heurística hacia adelante y hacia atrás. Se presenta un ejemplo de una representación tiempo-frecuencia (TFR), a partir de la cual se obtiene la matriz A. Se define la matriz B, con su primera columna igual a la matriz A y la matriz donde se pondrán los índices. En seguida, se llenan los pixeles de la matriz B siguiendo la **Ec. (3.5.1)**, donde la penalización es 10 y sus resultados son la suma de los números en rojo y en azul, aparecen sombreados en verde. Al concluir, de la matriz B, se toma el mínimo de la última columna, en este caso 0, cuyo índice de origen es 2, por lo que se regresa al elemento 2 y así sucesivamente, donde aquí los índices de la marginal serían 2, 2, 2, lo que corresponde, a los elementos de mayor energía en la TFR.

2, lo que indica que debemos regresar a la fila 2 de la penúltima columna, donde también se asignó un valor de 2.

3.5.2 Índices de desempeño

Para seleccionar los mejores parámetros de cada técnica de análisis TF, se compararon las estimaciones de iRR obtenidas de cada TFR con la obtenida de su TF ideal, empleando tres índices de desempeño: 1) el índice de correlación cruzada de dos señales, 2) el error absoluto promedio (*mean absolute error*, MAE), y 3) la raíz del error cuadrático medio (*Root Mean Square Error*, RMSE), descritos en las **ecuaciones (3.5.2), (3.5.3) y (3.5.4)**, respectivamente, donde iRR_{TFR} denota la iRR obtenida a partir de una TFR, iRR_{ref} es la señal de respiración instantánea que se desea obtener, i. e., la simulación ideal y N es el tamaño de las señales iRR involucradas.

El primer índice empleado, el índice de correlación cruzada, nos revela la similitud lineal entre dos señales. Si el resultado es cero las señales son totalmente diferentes, contrario a cuando es uno, caso en el que las señales serían exactamente iguales. Este índice de desempeño es adimensional. El índice de correlación cruzada ρ_{IF} está definido por

$$\rho_{IF} = \frac{\sum_{n=1}^N iRR_{TFR}(n) iRR_{ref}(n)}{\sqrt{\sum_{n=1}^N iRR_{TFR}(n)^2 \sum_{n=1}^N iRR_{ref}(n)^2}} \quad (3.5.2)$$

Por su parte, el MAE es un error que mide las diferencias absolutas entre el valor calculado y el valor deseado, i. e., entre los resultados de las técnicas TF y el escenario ideal. En nuestro caso, las unidades del MAE serían bpm. Este error está dado por:

$$MAE = \frac{1}{N} \sum_{n=1}^N |iRR_{TFR}(n) - iRR_{ref}(n)| \quad (3.5.3)$$

Finalmente, el RMSE también es un error, pero este se considera más sensible a errores grandes que el MAE y al igual que este estará expresado en unidades de bpm. Se define por:

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (iRR_{TFR}(n) - iRR_{ref}(n))^2} \quad (3.5.4)$$

Estos índices de desempeño han sido usados por varios trabajos relacionados con este proyecto de tesis, donde se usaron técnicas TF para estimar la frecuencia respiratoria a partir de señales de PPG o iPPG [11], [12], [14], [17], [18], [20]. Lo anterior facilitará la comparación con otras investigaciones. Para cada maniobra, a partir de los resultados de los índices se hizo la selección de parámetros para cada técnica TF, con el objetivo final de obtener las estimaciones de la frecuencia respiratoria instantánea a partir de las señales de iPPG de la base de datos.

3.6 Estimación de RR instantánea en datos reales

En los datos reales se hizo la estimación de la iRR usando los parámetros seleccionados en la etapa anterior de simulación. Para verificar que las estimaciones fueran adecuadas, se compararon con la iRR obtenida de las señales respiratorias de referencia. Así, se realizó la estimación de la iRR en las TFR seleccionadas y en las señales de referencia.

3.6.1 Estimación de iRR de referencia

A partir de la señal de respiración adquirida, se calculó la iRR de referencia mediante la obtención de los periodos entre los picos sucesivos de la señal, sacando su recíproco e interpolando a 4 Hz. Esta señal de iRR servirá para compararla con las iRR obtenidas de las TFR con los parámetros seleccionados.

3.6.2 Estimación de iRR con las TFR seleccionadas

Esta etapa se aplicó el mismo algoritmo de estimación de iRR que para las señales simuladas descrito en la **sección 3.5.1** de este mismo capítulo.

3.7 Análisis de desempeño en señales reales

Se hizo un análisis de desempeño final con las estimaciones de iRR, para lo que se utilizaron los mismos índices de desempeño descritos en la **sección 3.5.2**. A partir de los resultados de esta etapa se determinó cuáles de los parámetros seleccionados de las TFR ofrecieron mejores resultados a los demás para cada una de las maniobras respiratorias y se analizó si alguna técnica TF se ajusta mejor a diferentes situaciones respiratorias en las que se encuentre una persona al intentar estimar su frecuencia respiratoria utilizando un dispositivo móvil.

Los resultados también fueron comparados con análisis estadístico, para ello primero se hicieron pruebas de normalidad en los datos. Como la base de datos es pequeña ($n = 9$), se utilizó la prueba Shapiro-Wilk para evaluar la normalidad de menos de 50 muestras [85]. A partir de los resultados se seleccionó el tipo de prueba que se le deben aplicar a los resultados. Debido a que los datos resultaron no normales y son de medidas repetidas, se utilizó la prueba Friedman, en la que el nivel de significancia se estableció en 5% [86].

Capítulo 4

Resultados y discusión

En este capítulo se presentan y discuten los resultados obtenidos después de seguir cada paso de la metodología propuesta en el capítulo anterior.

4.1 Preprocesamiento de señales de iPPG

A manera de ejemplo, en la **Figura 4.1** se muestran dos etapas del procesamiento de una señal de iPPG correspondiente a una respiración de 12 bpm de un voluntario. En la **Figura 4.1.A** se muestra como se ve la señal de iPPG después de extraerla de los videos, filtrarla y normalizarla y en la **Figura 4.1.B** se muestra la señal de PRV después de la interpolación de los picos a una frecuencia de 4 Hz utilizando *splines* cúbicos. Para contrastar como lucen las señales de respiración extraídas del iPPG junto a la señal respiratoria de referencia, la **Figura 4.2** muestra dos secciones con una longitud de 60 s, de dos maniobras respiratorias provenientes del mismo voluntario, en ellas se superponen las señales de PR (iPPG) y las señales respiratorias de referencia. En la **Figura 4.2.A** está la maniobra de cambios abruptos, donde se aprecia uno de los cambios en frecuencia, de 6 bpm a 18 bpm, propios de estas adquisiciones. En la **Figura 4.2.B** se observa la maniobra constante en 0.1 Hz (6 bpm) y como se espera no se ven cambios significativos en la frecuencia. En ambas adquisiciones es evidente que las amplitudes de las señales no son iguales, pero estas oscilan a la misma frecuencia. Es importante señalar que la estimación de la iRR se basa en analizar las variaciones en la frecuencia de la señal de PR derivada del iPPG.

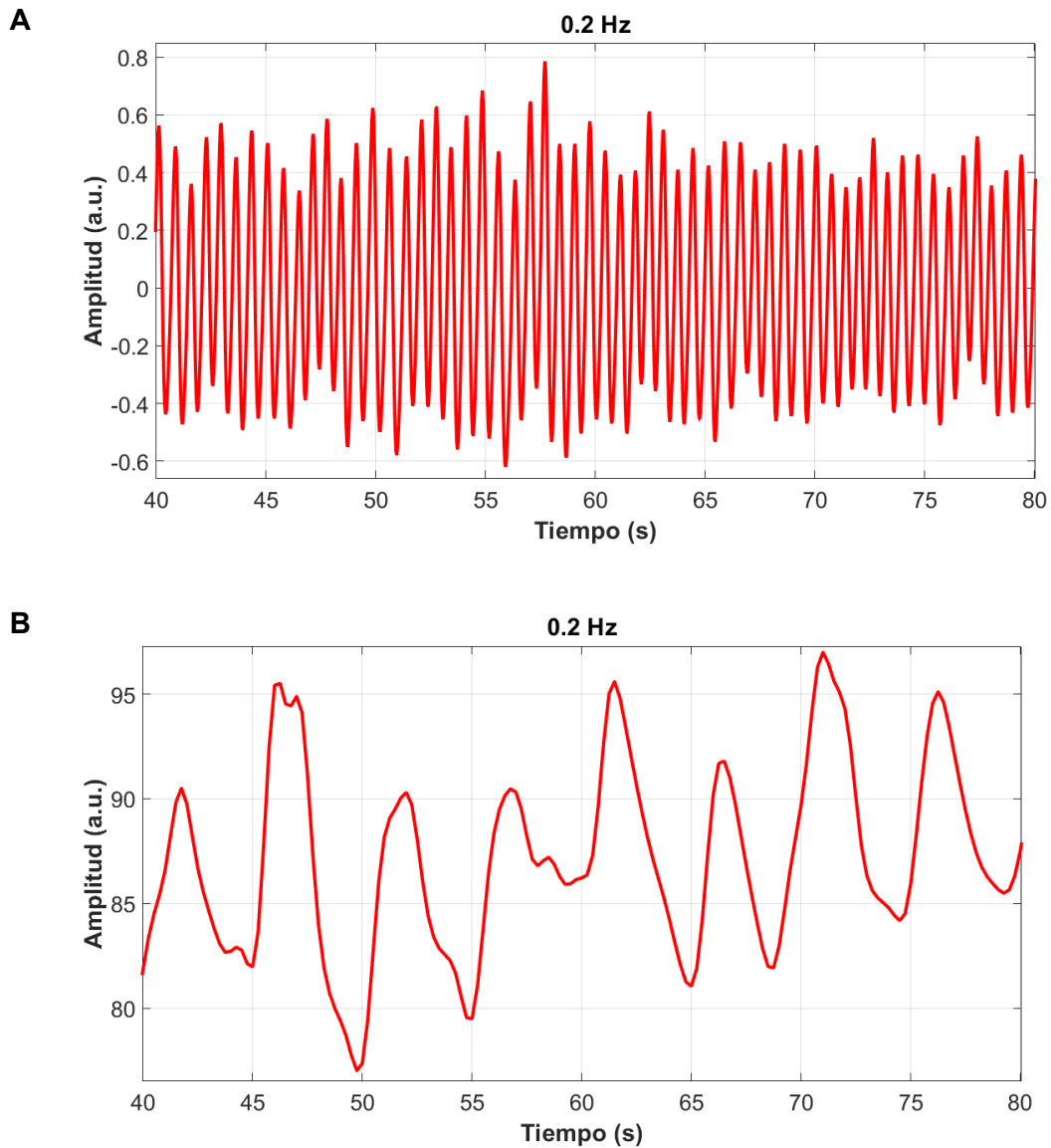


Figura 4.1 Preprocesamiento de señales. A) Señal iPPG filtrada y normalizada. B) Señal de frecuencia del pulso (extraída del iPPG).

4.2 Simulaciones

En esta sección se muestran las simulaciones hechas por medio de las ecuaciones presentadas en la sección de metodología. En primer lugar, se muestra como resultado la emulación de cada maniobra respiratoria en un escenario ideal y luego se ilustra

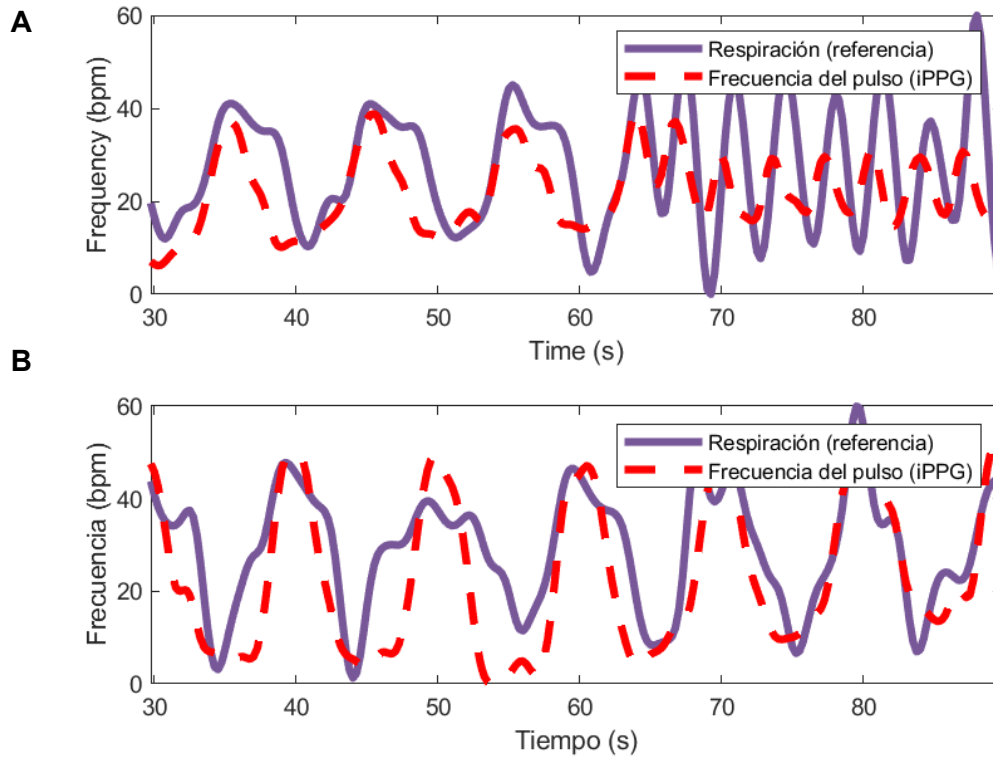


Figura 4.2 Ejemplo de la señal de respiración filtrada superpuesta en la serie de tiempo de la frecuencia de pulso obtenida a partir del iPPG. A) Segmento de la maniobra con un cambio abrupto de 6 a 18 bpm. B) Segmento de la maniobra constante a 6 bpm.

cómo las simulaciones se ven afectadas por el ruido aditivo, con el fin de que se parezcan más a datos reales.

4.2.1 Maniobras respiratorias

A continuación, se detallan los cuatro tipos de maniobras respiratorias que se simularon, estas están mostradas en la **Figura 4.3**, en la que se presentan las señales y las TFR ideales, las cuales tienen un mapa de colores donde en color amarillo esta la energía máxima (normalizada para tener un valor máximo de 1). Así, para las TFR ideales los valores máximos en un determinado tiempo corresponden a la frecuencia en ese instante de un componente sinusoidal. En la **Figura 4.3.A** se muestra la maniobra de cambios abruptos en la que se observa que cada 60

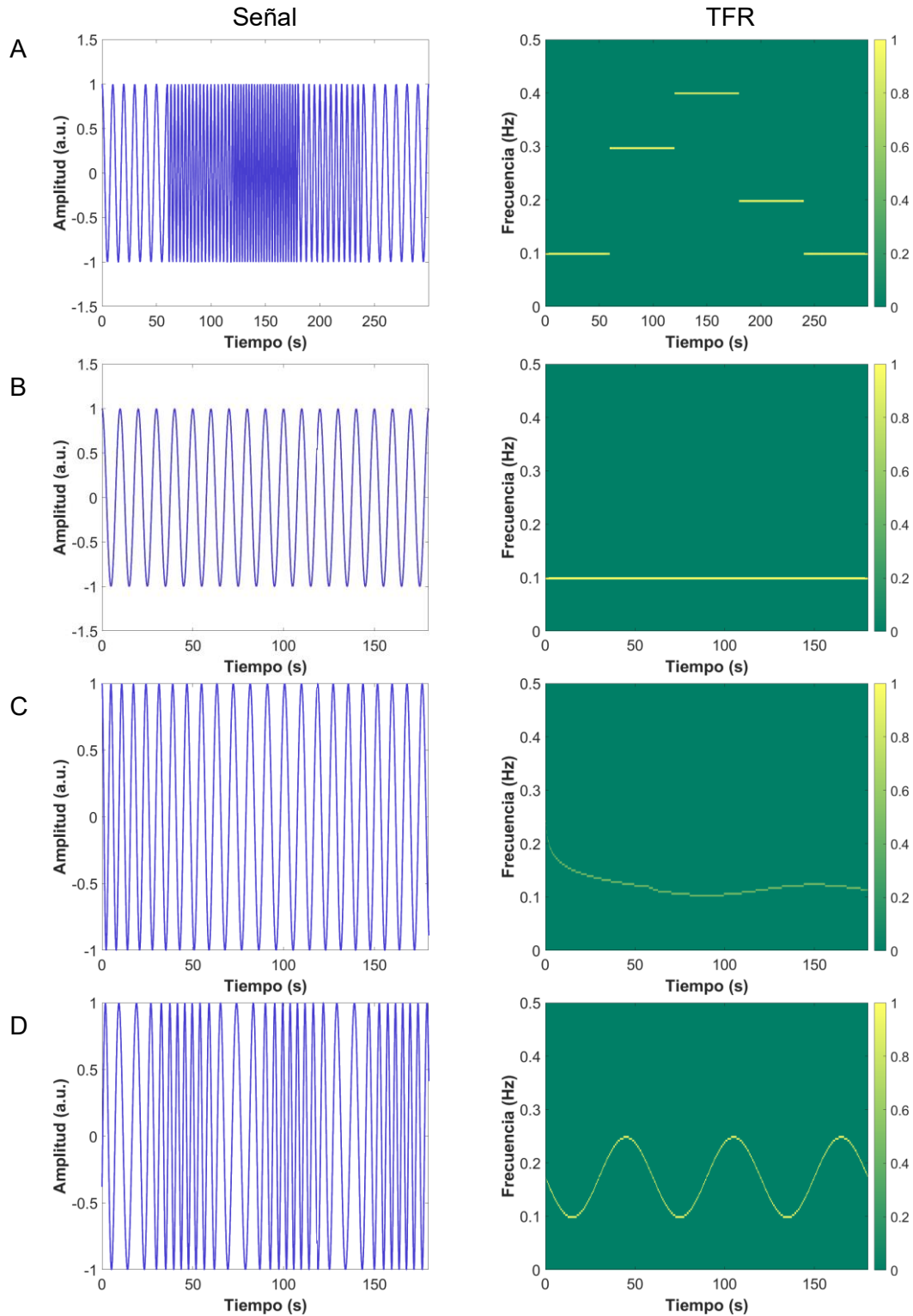


Figura 4.3 Señales y representaciones tiempo-frecuencia (TFR) ideales. A) Cambios abruptos. B) Constante en 0.3 Hz. C) Después del ejercicio. D) Espontánea.

segundos la frecuencia respiratoria se modifica abruptamente. En **Figura 4.3.B** solo se muestra una de las maniobras constantes, que es la de 0.3 Hz (18 bpm) donde solo se ve una línea horizontal amarilla en esa frecuencia. Para la maniobra después del ejercicio se observa un decaimiento en la frecuencia conforme avanza el tiempo, simulando el retorno a una respiración basal de 6 bpm (**Figura 4.3.C**). Finalmente, para la maniobra espontánea, en la **Figura 4.3.D**, se exhibe una modulación senoidal en la frecuencia. De cada TFR ideal, se extrajo la iRR para utilizarla como punto de comparación contra las extraídas al aplicar las técnicas TF a las señales de respiración simuladas adicionadas con ruido.

4.2.2 Ruido aditivo

Al sumarle el ruido de tipo Gaussiano a las señales de respiración simulaciones estas cambiaron en menor o mayor medida. En la **Figura 4.4** se presentan todas las simulaciones con ruido aditivo. En la **Figura 4.4.A** se muestra una parte de la simulación ideal con cambios abruptos, específicamente en el primer cambio de frecuencia que sufre la señal. En la **Figura 4.4.B** se ve una de las señales metrónomas, con 18 bpm, mientras que en la **Figura 4.4.C** están las simulaciones después del ejercicio y en la **Figura 4.4.D** están las simulaciones a una respiración espontáneas. En estas imágenes, cuando la SNR disminuye a 10 dB, la señal resulta más afectada, en cambio, cuando las simulaciones tienen 40 dB de SNR, casi no se altera a la señal, siendo estas las que más se parecen a las señales ideales.

4.3 Selección de parámetros de la TFR

Para hacer la selección de los parámetros se siguió la metodología presentada en la **sección 3.5**, donde el primer paso consistió en extraer la iRR de las TFR, el segundo paso en calcular los índices de desempeño (MAE, RMSE y coeficiente de correlación cruzada) tomando como punto de comparación las iRRs de TFR ideales y el último paso en seleccionar los mejores parámetros para cada maniobra.

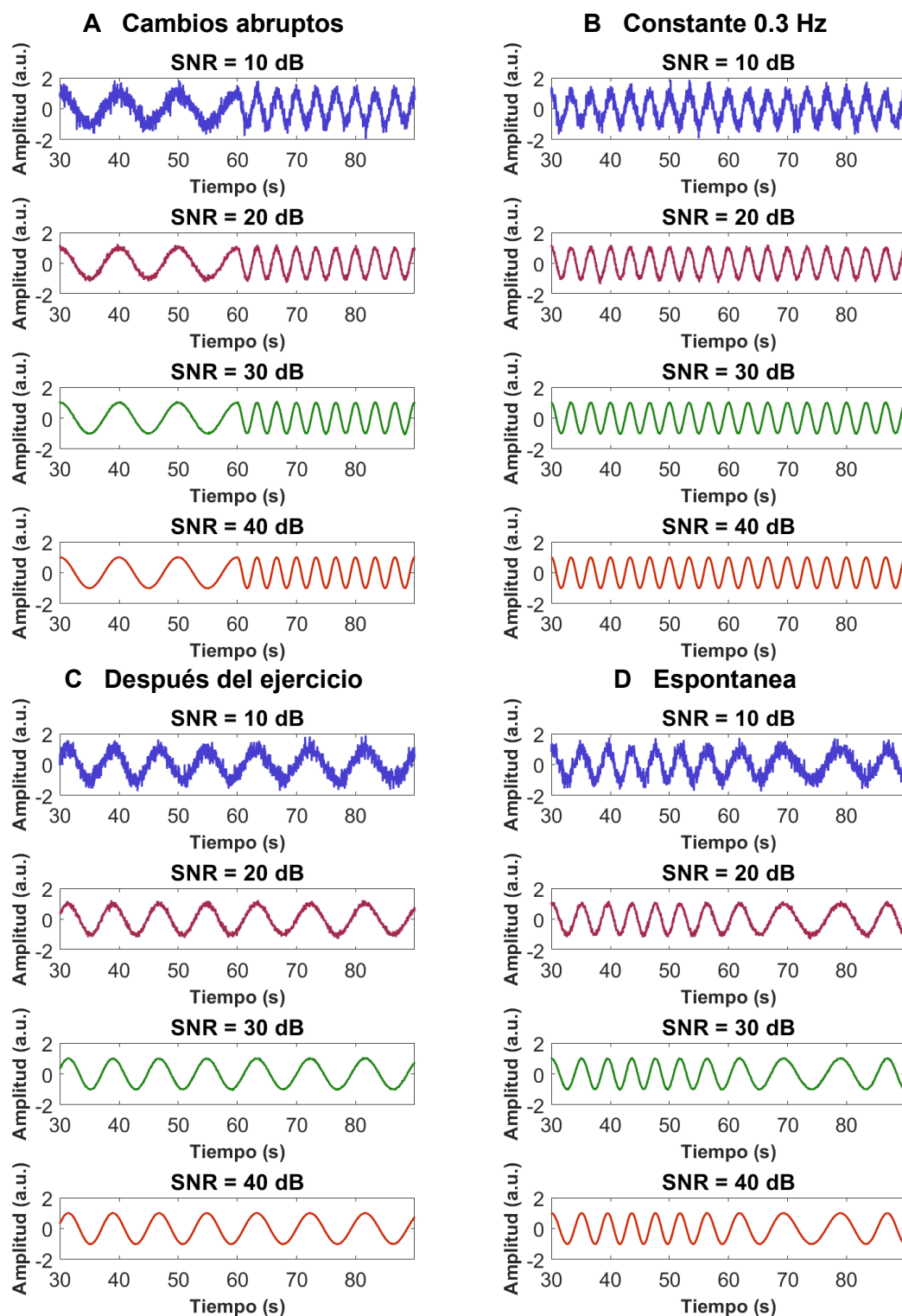


Figura 4.4 Simulaciones del mismo segmento temporal, pero con diferentes SNR. A) Cambios abruptos. B) Constante en 0.3 Hz. C) Después del ejercicio. D) Espontanea.

4.3.1 Marginal temporal de la TFR

Para poder seleccionar los parámetros más adecuados de cada técnica TF, se extrajo la marginal de cada TFR generada utilizando el algoritmo *greedy* de búsqueda heurística hacia adelante y hacia atrás, descrito en la **sección 3.5.1**. En la **Figura 4.5**, se ilustra el cálculo de algunas de las marginales del escalograma, donde se ve que las líneas no coinciden totalmente, esto se debe a la distinta SNR de la señal simulada, a la técnica TF empleada y a los parámetros que se utilizan. Es por esto que es fundamental buscar los mejores parámetros, para que las marginales, i.e., la iRR estimada, se aproximen lo más posible a la serie de tiempo de la frecuencia respiratoria real. En el **Referencias**

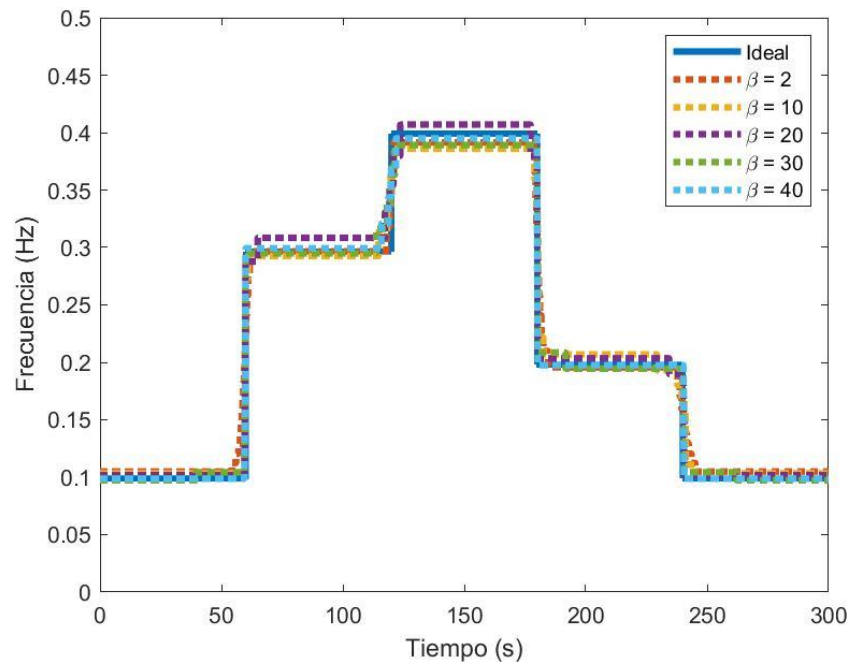


Figura 4.5 Marginales (iRR) de los escalogramas de la simulación de cambios abruptos con 20 dB de SNR. Se utilizó una $\gamma = 3$ y se observan todos los valores de β en las líneas punteadas. La señal ideal es la línea azul sólida.

- [1] “Las 10 principales causas de defunción.” Accessed: Aug. 14, 2024. [Online]. Available: <https://www.who.int/es/news-room/fact-sheets/detail/the-top-10-causes-of-death>
- [2] T. Ferkol and D. Schraufnagel, “The global burden of respiratory disease,” *Ann. Am. Thorac. Soc.*, vol. 11, no. 3, pp. 404–406, Mar. 2014, doi: 10.1513/AnnalsATS.201311-405PS.
- [3] A. Zaccaro *et al.*, “How Breath-Control Can Change Your Life: A Systematic Review on Psycho-Physiological Correlates of Slow Breathing,” *Front. Hum. Neurosci.*, vol. 12, p. 353, 2018, doi: 10.3389/fnhum.2018.00353.
- [4] F. S. Cardoso, C. J. Karvellas, N. M. Kneteman, G. Meeberg, P. Fidalgo, and S. M. Bagshaw, “Respiratory rate at intensive care unit discharge after liver transplant is an independent risk factor for intensive care unit readmission within the same hospital stay: A nested case-control study,” *J. Crit. Care*, vol. 29, no. 5, pp. 791–796, Oct. 2014, doi: 10.1016/j.jcrc.2014.03.038.
- [5] H. Liu, J. Allen, D. Zheng, and F. Chen, “Recent development of respiratory rate measurement technologies,” *Physiol. Meas.*, vol. 40, no. 7, p. 07TR01, Aug. 2019, doi: 10.1088/1361-6579/ab299e.
- [6] S. Stankoski, I. Kiprijanovska, I. Mavridou, C. Nduka, H. Gjoreski, and M. Gjoreski, “Breathing Rate Estimation from Head-Worn Photoplethysmography Sensor Data Using Machine Learning,” *Sensors*, vol. 22, no. 6, p. 2079, Jan. 2022, doi: 10.3390/s22062079.
- [7] E. Jonathan and M. Leahy, “Investigating a smartphone imaging unit for photoplethysmography,” *Physiol. Meas.*, vol. 31, pp. N79-83, Nov. 2010, doi: 10.1088/0967-3334/31/11/N01.
- [8] “Más del 75% de la población mundial tiene un teléfono celular y más del 65% usa el internet | Noticias ONU.” Accessed: Aug. 30, 2024. [Online]. Available: <https://news.un.org/es/story/2023/12/1526712>

- [9] D. O. Lemos and C. A. Siebra, "Methods for Breathing Rate Measurement through Mobile Platform: a Review," *Anais Principais do XIX Simpósio Brasileiro de Computação Aplicada à Saúde*, pp. 288–293, 2019.
- [10] H. Sharma, "Extraction of respiratory rate from PPG using ensemble empirical mode decomposition with Kalman filter," *Electron. Lett.*, vol. 56, no. 13, pp. 650–653, Jun. 2020, doi: 10.1049/el.2020.0566.
- [11] W. Karlen, A. Garde, D. Myers, C. Scheffer, M. Ansermino, and G. Dumont, "Estimation of Respiratory Rate From Photoplethysmographic Imaging Videos Compared to Pulse Oximetry," *IEEE J. Biomed. Health Inform.*, vol. 19, May 2015, doi: 10.1109/JBHI.2015.2429746.
- [12] D. Huang *et al.*, "Camera-Based Respiratory Imaging for Intelligent Rehabilitation Assessment of Thoracic Surgery Patients," *IEEE Internet Things J.*, vol. 11, no. 15, pp. 26253–26265, Aug. 2024, doi: 10.1109/JIOT.2024.3395364.
- [13] E. G. Hernández de la Cruz, "Estimación de las frecuencias cardiaca y respiratoria instantáneas empleando métodos de fotoplethysmografía de imagen en un smartphone," Maestría en Ingeniería Electrónica, Universidad Autónoma de San Luis Potosí, San Luis Potosí, 2021.
- [14] Y. Nam, J. Lee, and K. H. Chon, "Respiratory Rate Estimation from the Built-in Cameras of Smartphones and Tablets," *Ann. Biomed. Eng.*, vol. 42, no. 4, pp. 885–898, Apr. 2014, doi: 10.1007/s10439-013-0944-x.
- [15] M. Alafeef and M. Fraiwan, "Smartphone-Based Respiratory Rate Estimation Using Photoplethysmographic Imaging and Discrete Wavelet Transform," *J. Ambient Intell. Humaniz. Comput.*, vol. 11, Feb. 2020, doi: 10.1007/s12652-019-01339-6.
- [16] K. H. Chon, S. Dash, and K. Ju, "Estimation of Respiratory Rate From Photoplethysmogram Data Using Time–Frequency Spectral Estimation," *IEEE*

Trans. Biomed. Eng., vol. 56, no. 8, pp. 2054–2063, Aug. 2009, doi: 10.1109/TBME.2009.2019766.

- [17] J. Lázaro, Y. Nam, E. Gil, P. Laguna, and K. H. Chon, “Respiratory rate derived from smartphone-camera-acquired pulse photoplethysmographic signals,” *Physiol. Meas.*, vol. 36, no. 11, pp. 2317–2333, Nov. 2015, doi: 10.1088/0967-3334/36/11/2317.
- [18] P. Dehkordi, A. Garde, B. Molavi, J. M. Ansermino, and G. A. Dumont, “Extracting Instantaneous Respiratory Rate From Multiple Photoplethysmogram Respiratory-Induced Variations,” *Front. Physiol.*, vol. 9, 2018.
- [19] M. Pirhonen, M. Peltokangas, and A. Vehkaoja, “Acquiring Respiration Rate from Photoplethysmographic Signal by Recursive Bayesian Tracking of Intrinsic Modes in Time-Frequency Spectra,” *Sensors*, vol. 18, no. 6, p. 1693, Jun. 2018, doi: 10.3390/s18061693.
- [20] A. Garde, W. Karlen, P. Dehkordi, J. Ansermino, and G. Dumont, “Empirical mode decomposition for respiratory and heart rate estimation from the photoplethysmogram,” in *Computing in Cardiology 2013*, Sep. 2013, pp. 799–802.
- [21] Md. N. I. Shuzan *et al.*, “A Novel Non-Invasive Estimation of Respiration Rate From Motion Corrupted Photoplethysmograph Signal Using Machine Learning Model,” *IEEE Access*, vol. 9, pp. 96775–96790, 2021, doi: 10.1109/ACCESS.2021.3095380.
- [22] M. A. Colominas, G. Schlotthauer, and M. E. Torres, “Improved complete ensemble EMD: A suitable tool for biomedical signal processing,” *Biomed. Signal Process. Control*, vol. 14, pp. 19–29, Nov. 2014, doi: 10.1016/j.bspc.2014.06.009.

- [23] A. Nicolò, C. Massaroni, E. Schena, and M. Sacchetti, "The Importance of Respiratory Rate Monitoring: From Healthcare to Sport and Exercise," *Sensors*, vol. 20, no. 21, p. 6396, Jan. 2020, doi: 10.3390/s20216396.
- [24] E. Vanegas, R. Igual, and I. Plaza, "Sensing Systems for Respiration Monitoring: A Technical Systematic Review," *Sensors*, vol. 20, no. 18, p. 5446, Jan. 2020, doi: 10.3390/s20185446.
- [25] G. Yuan, N. A. Drost, and R. A. McIvor, "Respiratory Rate and Breathing Pattern".
- [26] M. Buist, S. Bernard, T. V. Nguyen, G. Moore, and J. Anderson, "Association between clinically abnormal observations and subsequent in-hospital mortality: a prospective study," *Resuscitation*, vol. 62, no. 2, pp. 137–141, Aug. 2004, doi: 10.1016/j.resuscitation.2004.03.005.
- [27] G. J. Tortora and B. Derrickson, *Principios de Anatomía y Fisiología*, 13th ed. Editorial médica panamericana, Fecha de creación/publicación de la obra. [Online]. Available: <http://archive.org/details/tortora-13-principios-de-anatomia-y-fisiologia>
- [28] D. R. Hess and R. M. Kacmarek, "Physiologic Effects of Mechanical Ventilation," in *Essentials of Mechanical Ventilation*, 4th ed., New York, NY: McGraw-Hill Education, 2019. Accessed: Aug. 15, 2024. [Online]. Available: accessanesthesiology.mhmedical.com/content.aspx?aid=1157418259
- [29] M. Elstad, E. L. O'Callaghan, A. J. Smith, A. Ben-Tal, and R. Ramchandra, "Cardiorespiratory interactions in humans and animals: rhythms for life," *Am. J. Physiol.-Heart Circ. Physiol.*, vol. 315, no. 1, pp. H6–H17, Jul. 2018, doi: 10.1152/ajpheart.00701.2017.
- [30] P. D. Larsen, Y. C. Tzeng, P. Y. W. Sin, and D. C. Galletly, "Respiratory sinus arrhythmia in conscious humans during spontaneous respiration," *Respir. Physiol. Neurobiol.*, vol. 174, no. 1, pp. 111–118, Nov. 2010, doi: 10.1016/j.resp.2010.04.021.

- [31] P. Y. W. Sin, D. C. Galletly, and Y. C. Tzeng, "Influence of breathing frequency on the pattern of respiratory sinus arrhythmia and blood pressure: old questions revisited," *Am. J. Physiol.-Heart Circ. Physiol.*, vol. 298, no. 5, pp. H1588–H1599, May 2010, doi: 10.1152/ajpheart.00036.2010.
- [32] I. Nederend, M. R. M. Jongbloed, E. J. C. De Geus, N. A. Blom, and A. D. J. Ten Harkel, "Postnatal Cardiac Autonomic Nervous Control in Pediatric Congenital Heart Disease," *J. Cardiovasc. Dev. Dis.*, vol. 3, no. 2, Art. no. 2, Jun. 2016, doi: 10.3390/jcdd3020016.
- [33] F. Yasuma and J. Hayano, "Respiratory Sinus Arrhythmia: Why Does the Heartbeat Synchronize With Respiratory Rhythm?," *CHEST*, vol. 125, no. 2, pp. 683–690, Feb. 2004, doi: 10.1378/chest.125.2.683.
- [34] J. A. Hirsch and B. Bishop, "Respiratory sinus arrhythmia in humans: how breathing pattern modulates heart rate," *Am. J. Physiol.*, vol. 241, no. 4, pp. H620–629, Oct. 1981, doi: 10.1152/ajpheart.1981.241.4.H620.
- [35] M. A. Russo, D. M. Santarelli, and D. O'Rourke, "The physiological effects of slow breathing in the healthy human," *Breathe*, vol. 13, no. 4, pp. 298–309, Dec. 2017, doi: 10.1183/20734735.009817.
- [36] M. Y. Balban *et al.*, "Brief structured respiration practices enhance mood and reduce physiological arousal," *Cell Rep. Med.*, vol. 4, no. 1, Jan. 2023, doi: 10.1016/j.xcrm.2022.100895.
- [37] S.-H. Li, B.-S. Lin, C.-H. Tsai, C.-T. Yang, and B.-S. Lin, "Design of Wearable Breathing Sound Monitoring System for Real-Time Wheeze Detection," *Sensors*, vol. 17, no. 1, p. 171, Jan. 2017, doi: 10.3390/s17010171.
- [38] K. Storck, M. Karlsson, P. Ask, and D. Loyd, "Heat transfer evaluation of the nasal thermistor technique," *IEEE Trans. Biomed. Eng.*, vol. 43, no. 12, pp. 1187–1191, Dec. 1996, doi: 10.1109/10.544342.

- [39] F. Q. Al-Khalidi, R. Saatchi, D. Burke, H. E. Elphick, and S. Tan, "Respiration rate monitoring methods: a review," *Pediatr. Pulmonol.*, vol. 46, no. 6, Jun. 2011, doi: 10.1002/ppul.21416.
- [40] C. Massaroni, A. Nicolò, D. Presti, M. Sacchetti, S. Silvestri, and E. Schena, "Contact-Based Methods for Measuring Respiratory Rate," *Sensors*, vol. 19, p. 908, Feb. 2019, doi: 10.3390/s19040908.
- [41] M. A. F. Pimentel *et al.*, "Toward a Robust Estimation of Respiratory Rate From Pulse Oximeters," *IEEE Trans Biomed Eng*, vol. 64, no. 8, pp. 1914–1923, Nov. 2016, doi: 10.1109/TBME.2016.2613124.
- [42] P. H. Charlton *et al.*, "Breathing Rate Estimation From the Electrocardiogram and Photoplethysmogram: A Review," *IEEE Rev. Biomed. Eng.*, vol. 11, pp. 2–20, 2018, doi: 10.1109/RBME.2017.2763681.
- [43] E. F. Greneker, "Radar sensing of heartbeat and respiration at a distance with applications of the technology," in *Radar 97 (Conf. Publ. No. 449)*, Oct. 1997, pp. 150–154. doi: 10.1049/cp:19971650.
- [44] Z. Zhu, J. Fei, and I. Pavlidis, "Tracking human breath in infrared imaging," in *Fifth IEEE Symposium on Bioinformatics and Bioengineering (BIBE'05)*, Oct. 2005, pp. 227–231. doi: 10.1109/BIBE.2005.55.
- [45] A. A. Kamshilin and N. B. Margaryants, "Origin of Photoplethysmographic Waveform at Green Light," *Phys. Procedia*, vol. 86, pp. 72–80, Jan. 2017, doi: 10.1016/j.phpro.2017.01.024.
- [46] R. Jolivot, Y. Benezeth, and F. Marzani, "Skin Parameter Map Retrieval from a Dedicated Multispectral Imaging System Applied to Dermatology/Cosmetology," *Int. J. Biomed. Imaging*, vol. 2013, p. 978289, 2013, doi: 10.1155/2013/978289.
- [47] R. R. Anderson and J. A. Parrish, "The Optics of Human Skin," *J. Invest. Dermatol.*, vol. 77, no. 1, pp. 13–19, Jul. 1981, doi: 10.1111/1523-1747.ep12479191.

- [48] Y. Sun and N. Thakor, "Photoplethysmography Revisited: From Contact to Noncontact, From Point to Imaging," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 463–477, Mar. 2016, doi: 10.1109/TBME.2015.2476337.
- [49] P. S. Addison and J. N. Watson, "Secondary transform decoupling of shifted nonstationary signal modulation components: application to photoplethysmography," *Int. J. Wavelets Multiresolution Inf. Process.*, vol. 02, no. 01, pp. 43–57, Mar. 2004, doi: 10.1142/S0219691304000329.
- [50] M. van Gastel, S. Stuijk, and G. de Haan, "Robust respiration detection from remote photoplethysmography," *Biomed. Opt. Express*, vol. 7, no. 12, pp. 4941–4957, 2016, doi: 10.1364/BOE.7.004941.
- [51] L. Stanković, I. Djurović, S. Stanković, M. Simeunović, S. Djukanović, and M. Daković, "Instantaneous frequency in time–frequency analysis: Enhanced concepts and performance of estimation algorithms," *Digit. Signal Process.*, vol. 35, pp. 1–13, Dec. 2014, doi: 10.1016/j.dsp.2014.09.008.
- [52] B. A. Reyes, N. Reljin, and K. H. Chon, "Tracheal Sounds Acquisition Using Smartphones," *Sensors*, vol. 14, no. 8, pp. 13830–13850, Aug. 2014, doi: 10.3390/s140813830.
- [53] T. Oberlin, S. Meignen, and V. Perrier, "Second-Order Synchrosqueezing Transform or Invertible Reassignment? Towards Ideal Time-Frequency Representations," *IEEE Trans. Signal Process.*, vol. 63, no. 5, pp. 1335–1344, Mar. 2015, doi: 10.1109/TSP.2015.2391077.
- [54] S. Meignen, D. H. Pham, and M. Colominas, "On the Use of Short-Time Fourier Transform and Synchrosqueezing-Based Demodulation for the Retrieval of the Modes of Multicomponent Signals," *Signal Process.*, vol. 178, p. 107760, Aug. 2020, doi: 10.1016/j.sigpro.2020.107760.
- [55] O. Poisson, P. Rioual, and M. Meunier, "New signal processing tools applied to power quality analysis," *IEEE Trans. Power Deliv.*, vol. 14, no. 2, pp. 561–566, Apr. 1999, doi: 10.1109/61.754104.

- [56] A. V. Oppenheim and R. W. Schaffer, *Discrete-time Signal Processing*. Prentice Hall, 1999.
- [57] M. Rhif, A. Abbes, I. Farah, B. Martinez, and Y. Sang, “Wavelet Transform Application for/in Non-Stationary Time-Series Analysis: A Review,” *Appl. Sci.*, vol. 9, p. 1347, Mar. 2019, doi: 10.3390/app9071345.
- [58] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*. Academic Press, 2008.
- [59] J. M. Lilly and S. C. Olhede, “Higher-Order Properties of Analytic Wavelets,” *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 146–160, Jan. 2009, doi: 10.1109/TSP.2008.2007607.
- [60] E. A. Martinez-Ríos, R. Bustamante-Bello, S. Navarro-Tuch, and H. Perez-Meana, “Applications of the Generalized Morse Wavelets: A Review,” *IEEE Access*, vol. 11, pp. 667–688, 2023, doi: 10.1109/ACCESS.2022.3232729.
- [61] S. C. Olhede and A. T. Walden, “Generalized Morse wavelets,” *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2661–2670, Nov. 2002, doi: 10.1109/TSP.2002.804066.
- [62] A. Silik, M. Noori, W. Altabey, R. Ghiasi, Z. Wu, and J. Dang, “Evaluation of analytic wavelet parameters effect for data analyses in civil structural health monitoring,” presented at the Conference: 10th International Conference on Structural Health Monitoring of Intelligent Infrastructure, Porto, Portugal, Jun. 2021.
- [63] S. Maes and I. Daubechies, “A Nonlinear Squeezing of the Continuous Wavelet Transform Based on Auditory Nerve Models,” in *Wavelets in Medicine and Biology*, Akram Aldroubi, Michael Unser., CRC Press, 1996, pp. 527–546.
- [64] F. Auger *et al.*, “Time-Frequency Reassignment and Synchrosqueezing: An overview,” *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 32–41, Nov. 2013, doi: 10.1109/MSP.2013.2265316.

- [65] H.-T. Wu, G. F. Lewis, M. I. Davila, I. Daubechies, and S. W. Porges, "Optimizing Estimates of Instantaneous Heart Rate from Pulse Wave Signals with the Synchrosqueezing Transform," *Methods Inf. Med.*, vol. 55, no. 5, pp. 463–472, Oct. 2016, doi: 10.3414/ME16-01-0026.
- [66] T. Oberlin, S. Meignen, and V. Perrier, "The Fourier-based Synchrosqueezing Transform," presented at the ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, May 2014, pp. 315–319. doi: 10.1109/ICASSP.2014.6853609.
- [67] S. Sandoval and P. L. De Leon, "Recasting the (Synchrosqueezed) Short-Time Fourier Transform as an Instantaneous Spectrum," *Entropy*, vol. 24, no. 4, p. 518, Apr. 2022, doi: 10.3390/e24040518.
- [68] D. C. Bowman and J. M. Lees, "The Hilbert-Huang Transform: A High Resolution Spectral Method for Nonlinear and Nonstationary Time Series," *Seismol. Res. Lett.*, vol. 84, no. 6, pp. 1074–1080, Nov. 2013, doi: 10.1785/0220130025.
- [69] N. E. Huang *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. Lond. Ser. Math. Phys. Eng. Sci.*, pp. 903–995, 1998.
- [70] M. A. Motin, C. Kumar Karmakar, D. K. Kumar, and M. Palaniswami, "PPG Derived Respiratory Rate Estimation in Daily living Conditions," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.*, vol. 2020, pp. 2736–2739, Jul. 2020, doi: 10.1109/EMBC44109.2020.9175682.
- [71] A. Waugh and A. Grant, *Ross & Willson Anatomy And Physiology in Health and Illness*, 12th ed. Elsevier, 2014. Accessed: May 08, 2024. [Online]. Available: <http://archive.org/details/RossWillsonAnatomyAndPhysiology>
- [72] J. Lázaro, Y. Nam, E. Gil, P. Laguna, and K. H. Chon, "Smartphone-camera-acquired pulse photoplethysmographic signal for deriving respiratory rate," in *2014 8th Conference of the European Study Group on Cardiovascular*

- Oscillations (ESGCO)*, Trento, Italy, May 2014, pp. 121–122. doi: 10.1109/ESGCO.2014.6847549.
- [73] L.-M. Po, X. Xu, L. Feng, Y. Li, K.-W. Cheung, and C.-H. Cheung, “Frame adaptive ROI for photoplethysmography signal extraction from fingertip video captured by smartphone,” in *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*, Lisbon, Portugal: IEEE, May 2015, pp. 1634–1637. doi: 10.1109/ISCAS.2015.7168963.
- [74] J. Lee, K. Matsumura, K. Yamakoshi, P. Rolfe, S. Tanaka, and T. Yamakoshi, “Comparison between red, green and blue light reflection photoplethysmography for heart rate monitoring during motion,” *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.*, vol. 2013, pp. 1724–1727, 2013, doi: 10.1109/EMBC.2013.6609852.
- [75] Y. Maeda, M. Sekine, and T. Tamura, “The advantages of wearable green reflected photoplethysmography,” *J. Med. Syst.*, vol. 35, no. 5, pp. 829–834, Oct. 2011, doi: 10.1007/s10916-010-9506-z.
- [76] D. Grimaldi, Y. Kurylyak, F. Lamonaca, and A. Nastro, “Photoplethysmography detection by smartphone’s videocamera,” in *Proceedings of the 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems*, Sep. 2011, pp. 488–491. doi: 10.1109/IDAACS.2011.6072801.
- [77] R. B. Lagido, J. Lobo, S. Leite, C. Sousa, L. Ferreira, and J. Silva-Cardoso, “Using the smartphone camera to monitor heart rate and rhythm in heart failure patients,” in *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, Jun. 2014, pp. 556–559. doi: 10.1109/BHI.2014.6864425.
- [78] A. Visvanathan, R. Banerjee, A. Misra, A. D. Choudhury, and A. Pal, “Effects of fingertip orientation and flash location in smartphone Photoplethysmography,” in *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Sep. 2014, pp. 468–471. doi: 10.1109/ICACCI.2014.6968536.

- [79] R. J. Schilling and S. L. Harris, *Fundamentals of digital signal processing using MATLAB*. Nelson, a division of Thomson Canada Limited, 2004. Accessed: Aug. 06, 2024. [Online]. Available: <http://archive.org/details/fundamentalsofdi0000schi>
- [80] J. Lee, B. A. Reyes, D. D. McManus, O. Maitas, and K. H. Chon, "Atrial Fibrillation Detection Using an iPhone 4S," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 1, pp. 203–206, Jan. 2013, doi: 10.1109/TBME.2012.2208112.
- [81] M. Lozano, J. A. Fiz, and R. Jané, "Performance evaluation of the Hilbert-Huang transform for respiratory sound analysis and its application to continuous adventitious sound characterization," *Signal Process.*, vol. 120, pp. 99–116, Mar. 2016, doi: 10.1016/j.sigpro.2015.09.005.
- [82] François Auger, Patrick Flandrin, Paulo Gonçalves, and Olivier Lemoine, "Time-Frequency Toolbox For Use with MATLAB Reference Guide." CNRS (Centre National de la Recherche Scientifique), Oct. 26, 2005.
- [83] K. Sozański, "Signal-to-noise ratio in power electronic digital control circuits," in *2016 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, Sep. 2016, pp. 162–171. doi: 10.1109/SPA.2016.7763606.
- [84] "Visualize scalogram of signals - MATLAB - MathWorks América Latina." Accessed: Aug. 13, 2024. [Online]. Available: <https://la.mathworks.com/help/wavelet/ref/wavelettimefrequencyanalyzer-app.html>
- [85] K. Rani Das, "A Brief Review of Tests for Normality," *Am. J. Theor. Appl. Stat.*, vol. 5, no. 1, p. 5, 2016, doi: 10.11648/j.ajtas.20160501.12.
- [86] R. Parsad, "Non-parametric methods in analysis of experimental data," Nov. 2014.
- [87] J. Lázaro, S. Kontaxis, R. Bailón, P. Laguna, and E. Gil, "Respiratory Rate Derived from Pulse Photoplethysmographic Signal by Pulse Decomposition Analysis," in *2018 40th Annual International Conference of the IEEE*

Engineering in Medicine and Biology Society (EMBC), Jul. 2018, pp. 5282–5285. doi: 10.1109/EMBC.2018.8513188.

Anexo A se muestran iRRs de todas las maniobras respiratorias simuladas.

4.3.2 Índices de desempeño

Al tener listas las iRRs obtenidas a partir del análisis TF de las simulaciones de las señales respiratorias, el siguiente paso consistió en hacer la comparación de la estimación de la iRR considerando la obtenida de la TFR ideal, con el fin de seleccionar los parámetros para cada una de las técnicas TF en cada maniobra respiratoria. Con el apoyo de los índices de desempeño, el MAE, el RMSE y el coeficiente de correlación cruzada, se hicieron diagramas de caja y bigotes con los distintos parámetros. La selección se hizo con ayuda de los diagramas de caja y los valores promedio de los índices. A continuación, se presenta como fue la selección para cada técnica.

4.3.2.1 Espectrograma

Para el SP, los parámetros que se variaron fueron el tipo de ventana y el tamaño de esta. Comenzando con las simulaciones de cambios abruptos, en la **Figura 4.6** se muestran los diagramas de caja y bigotes de los datos, donde por visualización solo se incluyen los tamaños con los RMSE más pequeños, i.e., se exceptúan las longitudes de 1, 4, 8, 10, y 300 s. Se encontró que las ventanas Blackman-Harris tienen una dispersión muy pequeña. Aunado a esto, los errores más pequeños se encuentran en ventanas de longitud de 45 y 60 s, donde en estos intervalos las ventanas con errores menores son la Gaussiana y la Hamming. Por lo que las mejores ventanas están entre la Blackman-Harris, la Gaussiana y la Hamming en tamaños de 45 y 60 s, pero para terminar de tomar una decisión se procedió a observar cómo se ven afectadas estas ventanas ante el nivel de ruido aditivo. La ventana Hamming resultó ser la más afectada por SNR pequeños. Además, la

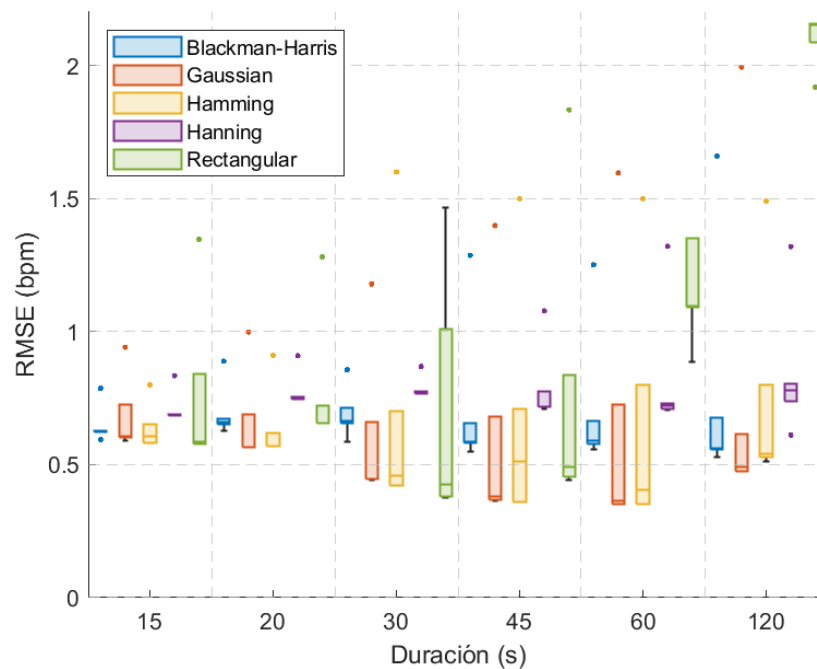


Figura 4.6 Diagramas de cajas y bigotes del RMSE (bpm) de las iRRs obtenidas a partir de las simulaciones de la respiración con cambios abruptos. En el eje x se muestra la duración de ventana (excepto 1, 4, 8, 10, y 300 s).

ventana Blackman-Harris obtuvo una mediana del valor MAE mayor a la Gaussiana. Entonces, se optó por utilizar la ventana Gaussiana, con tamaño de 45 s y no de 60 s, porque, aunque la mediana del RMSE en 60 s es de 0.366 bpm y en 45 s es de 0.379 bpm, la desviación estándar es de 0.412 y 0.498 bpm respectivamente, lo que indica que hay menor dispersión en un tamaño de 45 s.

En las simulaciones de la maniobra después del ejercicio, la ventana con el MAE promedio de los niveles de ruido más pequeño resultó ser la rectangular de 20 s con $0.177 \pm 5.349 \times 10^{-4}$ bpm (ver **Tabla 4.1**), de igual forma, su RMSE de $0.345 \pm 3.406 \times 10^{-4}$ bpm también fue el más pequeño y su índice de correlación cruzada de 9.990×10^{-1} fue la segunda mejor, después de la Rectangular de 15 s de 9.992×10^{-1} . Esta última ventana, también tiene valores MAE y RMSE muy pequeños, pero sus desviaciones estándares son mayores que la de 20 s y al hacer una prueba *tStudent* entre las dos ventanas se encontraron diferencias estadísticas significativas entre ellas, en el MAE se obtuvo una $p = 5.688 \times 10^{-3}$, en el RMSE una $p = 1.376 \times 10^{-4}$, y en el índice de correlación cruzada de $p = 6.342 \times 10^{-7}$. Entonces, la ventana rectangular de 20 s fue la ventana seleccionada.

En el caso de la maniobra espontanea, una ventana con duración de 30 s resultó en el mejor rendimiento para la estimación de la iRR, como se muestra en la

Tabla 4.1 Ventanas con errores más pequeños con el espectrograma en la maniobra simulada después del ejercicio.

Ventana	RMSE (bpm)	MAE (bpm)	Índice de correlación cruzada
Rectangular 15 s	$0.186 \pm 2.501 \times 10^{-3}$	$0.324 \pm 1.692 \times 10^{-3}$	$9.992 \times 10^{-1} \pm 9.960 \times 10^{-6}$
Rectangular 20 s	$0.177 \pm 5.349 \times 10^{-4}$	$0.345 \pm 3.406 \times 10^{-4}$	$9.990 \times 10^{-1} \pm 9.375 \times 10^{-6}$

Valor de los índices presentados como media $\pm$ SD

Figura 4.7. En esa longitud de ventana, el tipo de ventana con menor MAE fue la Hanning, la cual presenta poca dispersión, logrando un promedio de MAE de 0.114 bpm, RMSE de 0.168 bpm e índice de correlación cruzada de 0.999. Además, se encontró una diferencia estadística significativa con otros tipos de ventana ($p < 0.05$). Por lo tanto, se seleccionó la ventana de Hanning de 30 s. Cabe mencionar que la iRR estimado para las simulaciones espontáneas prácticamente no se vio afectado por el nivel de ruido, tal vez porque no presenta grandes cambios en frecuencia como ocurre en la maniobra de cambios bruscos.

Por último, en las simulaciones con frecuencia constante, todas las ventanas, sin importar su tipo o tamaño dieron el mismo resultado, lo que infiere que no importa que ventana se use el resultado sería el mismo. Lo anterior es debido a la naturaleza de estas señales respiratorias deterministas, puesto que, en cada ventaneo, la

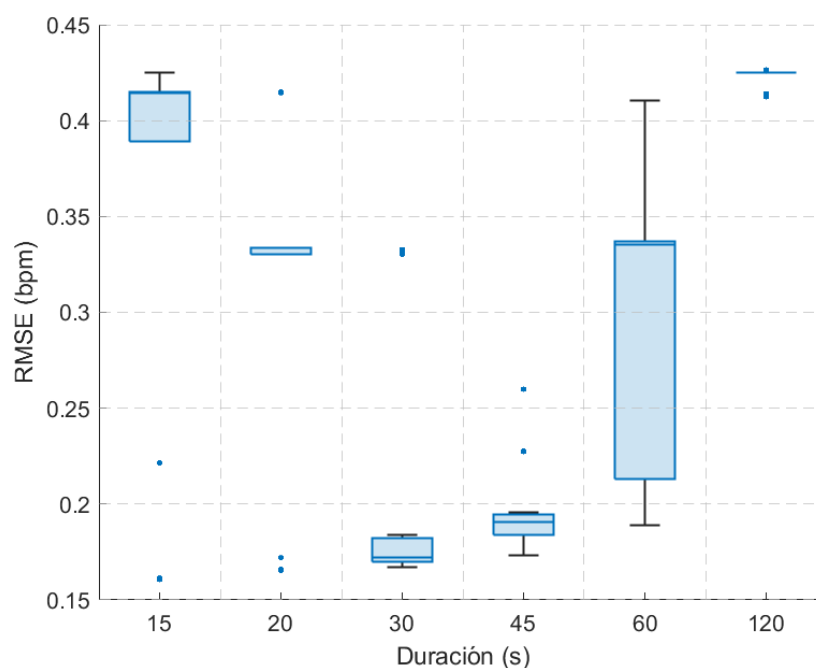


Figura 4.7 Diagrama de cajas y bigotes del RMSE (bpm) de las iRRs obtenidas a partir de las simulaciones de la respiración espontánea. En el eje x se muestra la duración de ventana (excepto 1, 4, 8, 10, y 180 s).

frecuencia respiratoria estimada sería exactamente la misma. Así, se usó una ventana Hamming de 18 s, como se hizo anteriormente por el grupo de trabajo [13].

En la **Tabla 4.2** se muestran todos los parámetros que se usaron en la base de datos para cada maniobra con la técnica del SP. En ninguna maniobra se eligieron los mismos parámetros, excepto para las frecuencias constantes.

4.3.2.2 Escalograma

En el caso del escalograma se variaron dos parámetros de la wavelet: el valor de γ y el valor de β . Para la maniobra de cambios abruptos, en cuestión de γ , el mejor RMSE se obtuvo para $\gamma = 3$, aunque, en el MAE el promedio más pequeño es para gamma igual a 20 ($MAE_{\gamma=20} = 0.325$ bpm, $MAE_{\gamma=3} = 0.362$ bpm), pero hay una gran desviación estándar a diferencia de la obtenida con un valor de $\gamma = 3$.

Tabla 4.2 Parámetros seleccionados en el espectrograma para todas las maniobras.

Maniobra respiratoria	Tipo de ventana	Duración (s)
Cambios abruptos	Gaussiana	45
Fija, 6 bpm	Hamming	18
Fija, 12 bpm	Hamming	18
Fija, 18 bpm	Hamming	18
Fija, 24 bpm	Hamming	18
Después del ejercicio	Rectangular	20
Espontánea	Hanning	30

Cuando se fijó el valor de $\beta = 30$ se obtuvo el valor RMSE más pequeño y el índice de correlación cruzada más grande, en cambio para el MAE su valor más chico ocurre cuando es igual a 40, pero por una diferencia mínima ($MAE_{\beta=30} = 0.233$ bpm, $MAE_{\beta=40} = 0.228$ bpm). Aun así, al ver la **Figura 4.8** con ambos parámetros la mejor configuración es para $\gamma = 3$ y $\beta = 40$, con un MAE de 0.143 bpm y poca desviación estándar a comparación de las demás cajas, además su RMSE es el más pequeño con valor de 0.461 bpm y su índice de correlación cruzada es de 0.999. También se puede observar que la *wavelet* con $\gamma = 20$ y $\beta = 30$ es muy cercana a la mencionada anteriormente, por lo cual se les hizo una prueba *t-Student*, con un resultado que indico que había una diferencia en todos los índices de desempeño. Por estos motivos se seleccionó la *wavelet* con $\gamma = 3$ y $\beta = 40$.

Con respecto a los resultados de todas las simulaciones constantes, no hay afectación alguna causada por el ruido, i. e., para cada *wavelet* los índices de

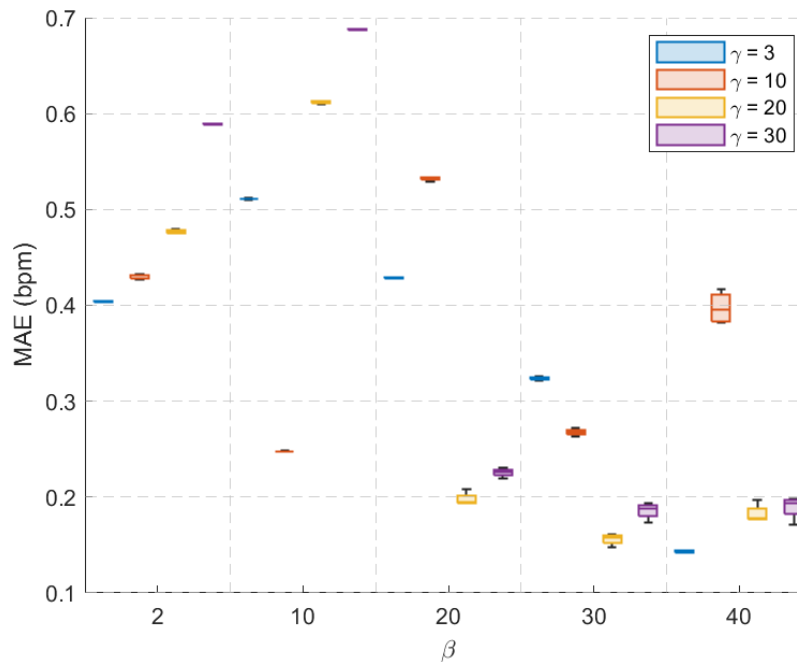


Figura 4.8 Diagrama de cajas y bigotes del MAE (bpm) de las iRRs obtenidas a partir de las simulaciones con cambios abruptos. En el eje x se encuentra el parámetro β y con el código de colores se identifica el parámetro γ .

desempeño fueron los mismos en las cuatro simulaciones. En la **Tabla 4.3** se presentan los índices de desempeño de los parámetros que se mencionan a la brevedad. La selección se hizo siguiendo el camino que se usó para la maniobra de cambios abruptos.

Pasando a las simulaciones constantes, la selección de parámetros fue la misma para las maniobras en 6 y 12 bpm, con $\gamma = 3$ y $\beta = 40$. Aunque, para 18 bpm hubo dos conjuntos de parámetros con el MAE y RMSE muy cercanos, con apenas 2 centésimas de diferencia, que son el de $\gamma = 3$, $\beta = 2$ y $\gamma = 20$ y $\beta = 30$, seleccionando el de $\gamma = 3$, $\beta = 2$, por tener los errores más pequeños. En el caso de la maniobra a 0.4 Hz (24 bpm) la selección fue de $\gamma = 30$ y $\beta = 2$, pese a que globalmente el RMSE más pequeño, sin importar que β se usara, fue con $\gamma = 3$.

Para la maniobra después del ejercicio, los RMSE y MAE más pequeños son para las combinaciones de $\gamma = 3$, $\beta = 40$, y de $\gamma = 10$, $\beta = 2$. En la **Tabla 4.4** se muestran los índices de desempeño de cada una de las *wavelets* mencionadas. Al realizar una prueba *t-Student* en las dos configuraciones, se encontró una diferencia

Tabla 4.3 Índices de desempeño de algunos resultados de las simulaciones constantes con el escalograma.

Maniobra constante	γ	β	RMSE (bpm)	MAE (bpm)	Índice de correlación cruzada
6 bpm	3	40	$8.413 \times 10^{-3} \pm 0$	$8.413 \times 10^{-3} \pm 0$	1 ± 0
12 bpm	3	40	$1.682 \times 10^{-2} \pm 0$	$1.682 \times 10^{-2} \pm 0$	1 ± 0
18 bpm	3	2	$2.043 \times 10^{-3} \pm 0$	$2.043 \times 10^{-3} \pm 0$	1 ± 0
18 bpm	20	30	$4.272 \times 10^{-3} \pm 0$	$.272 \times 10^{-3} \pm 0$	1 ± 0
24 bpm	30	2	$7.932 \times 10^{-2} \pm 0$	$7.932 \times 10^{-2} \pm 0$	1 ± 0

Valor de los índices presentados como media $\pm$ SD

Tabla 4.4 Índices de desempeño de dos wavelets de las simulaciones de la maniobra después del ejercicio.

γ	β	RMSE (bpm)	MAE (bpm)	Índice de correlación cruzada
3	40	$0.304 \pm 5 \times 10^{-4}$	$0.197 \pm 5 \times 10^{-4}$	$0.999 \pm 1.016 \times 10^{-6}$
10	2	$0.294 \pm 2.8 \times 10^{-3}$	$0.205 \pm 2.9 \times 10^{-3}$	$0.999 \pm 6.028 \times 10^{-6}$

Valor de los índices presentados como media $\pm$ SD

significativa para los tres índices de desempeño, en el MAE hubo un $p = 9.429 \times 10^{-3}$, en el RMSE de $p = 4.409 \times 10^{-3}$, y en el índice de correlación cruzada de $p = 2.625 \times 10^{-4}$. Se seleccionó la combinación de $\gamma = 3$, $\beta = 40$, porque a pesar de tener un RMSE más grande y una correlación más pequeña, la diferencia realmente es poca y su desviación estándar en estos índices si es más pequeña.

En la maniobra espontánea la decisión fue más sencilla, puesto que para ambos errores la configuración que produjo los mejores promedios de los índices fue la de $\gamma = 3$, $\beta = 20$, en la **Figura 4.9** se puede observar esto para el RMSE, donde el $RMSE_{\gamma=3,\beta=20} = 0.309$ bpm y con un $MAE_{\gamma=3,\beta=20} = 0.24$ bpm. En el caso del índice de correlación cruzada, el promedio más grande fue de 9.9941×10^{-1} para $\gamma = 20$ y $\beta = 2$, pero su MAE y su RMSE, no están entre los mejores. El segundo más grande fue de 9.9939×10^{-1} para $\gamma = 3$, $\beta = 20$ y se debe señalar que la diferencia es de diezmilésimas contra el primer lugar, por este motivo y dado que sus errores son los más pequeños fue que se seleccionaron estos últimos parámetros. Nótese, que para cuatro de las siete maniobras se seleccionaron exactamente los mismos parámetros ($\gamma = 3$, $\beta = 40$). Además, la $\gamma = 3$ fue la más

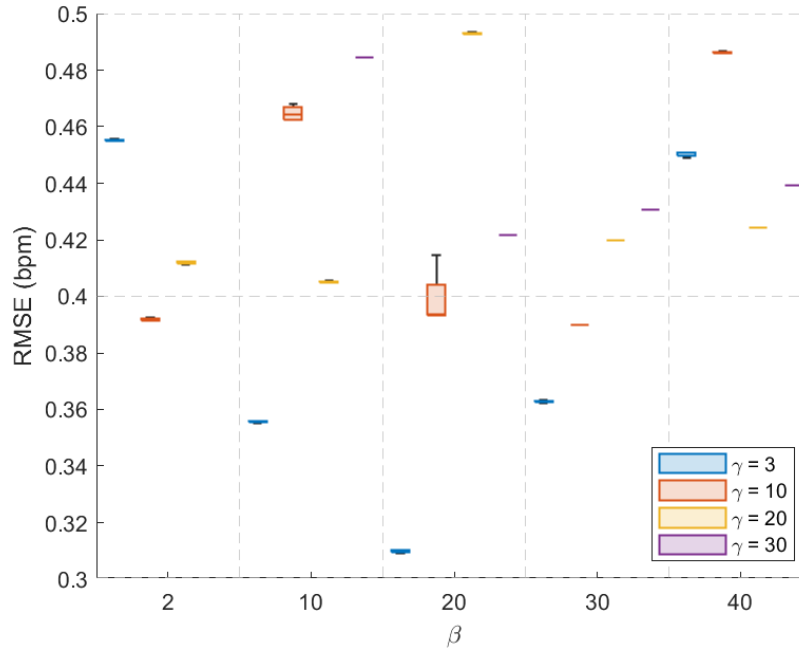


Figura 4.9 Diagrama de cajas y bigotes del RMSE (bpm) de los escalogramas de las simulaciones espontaneas.

escogida, excepto para la respiración constante de 24 bpm y para una de las dos *wavelets* elegidas para la respiración constante de 18 bpm, ver **Tabla 4.6**.

4.3.2.3 FSST

Comenzando con la maniobra de cambios abruptos, el análisis inicio con el tamaño de las ventanas, donde en promedio el mejor MAE es de 0.175 bpm para ventanase de 30 s y la segunda mejor es la de 60 s con un MAE de 0.184 bpm. En la **Tabla 4.5** se muestran las ventanas con los errores más pequeños, entre esas ventanas, la que tiene el promedio más pequeño de MAE y de RMSE, además de tener el índice de correlación cruzada más grande, es la ventana Gaussiana de 60 s, con 0.123 bpm, 0.254 bpm y un índice de 9.998×10^{-1} , respectivamente. Por estos motivos es la ventana seleccionada.

En la maniobra constante a 6 bpm, las ventanas con el mejor desempeño son la rectangular de 180 s y de 120 s, además de la gaussiana de 120 s, siendo la de 180 s la que es un poco mejor, en comparación a las otras dos, como se muestra

Tabla 4.6 Parámetros seleccionados en el escalograma para cada una de las maniobras.

Maniobra respiratoria	γ	β
Cambios abruptos	3	40
Fija, 6 bpm	3	40
Fija, 12 bpm,	3	40
Fija, 18 bpm	3	2
Fija, 24 bpm	30	2
Después del ejercicio	3	40
Espontánea	3	20

Tabla 4.5. Índices de desempeño de algunas ventanas utilizadas en el FSST con las simulaciones de cambios abruptos.

Ventana	RMSE (bpm) SD RMSE (bpm)	MAE (bpm) SD MAE (bpm)	Índice de correlación cruzada
Gaussiana 30 s	$0.327 \pm 7.863 \times 10^{-3}$	$0.153 \pm 2.831 \times 10^{-3}$	$9.997 \times 10^{-1} \pm 1.266 \times 10^{-5}$
Hamming 30 s	$0.327 \pm 8.787 \times 10^{-3}$	$0.154 \pm 2.745 \times 10^{-3}$	$9.997 \times 10^{-1} \pm 1.219 \times 10^{-5}$
Gaussiana 60 s	$0.254 \pm 3.244 \times 10^{-2}$	$0.122 \pm 2.597 \times 10^{-3}$	$9.998 \times 10^{-1} \pm 8.310 \times 10^{-4}$
Hamming 60 s	$0.260 \pm 4.215 \times 10^{-3}$	$0.123 \pm 3.07 \times 10^{-3}$	$9.998 \times 10^{-1} \pm 8.64126 \times 10^{-4}$

Valor de los índices presentados como media $\pm$ SD.

en la **Tabla 4.7** donde aparecen los índices de desempeño de estas ventanas. Al hacer una prueba *t-Student*, para saber cuáles de las ventanas se usarían, se encontró que hay diferencia entre los errores para las ventanas rectangulares de 180 s y de 120 s, inclinando la balanza a favor de la de 180 s. Adicionalmente, entre esta última y la Gaussiana de 120 s también hay diferencia, por lo que la mejor es la rectangular de 180 s. Debido a que una ventana del tamaño de nuestra señal sería poco o nada útil en aplicaciones de la vida diaria se propuso otra ventana de tamaño inferior. En ventanas más pequeñas, hay errores pequeños para las ventanas de 30, 20 y 10 s, siendo la de 30 s, con los errores más pequeños, en la cual todas las ventanas, excepto la Blackman-Harris, tienen un MAE y RMSE de 0.074 bpm y un índice de correlación cruzada de 1.

En la maniobra a 12 bpm, la ventana que tiene el MAE y RMSE más pequeño es la Blackman-Harris de 180 s con un MAE = 0.124 bpm y RMSE = 0.127 bpm. La segunda ventana con errores más pequeños es la Hanning de 180 s con MAE = 0.128 bpm y RMSE = 0.134 bpm, la cual no tiene diferencia con la Blackman-Harris de 180 s. Adicionalmente, conforme las ventanas disminuyen de tamaño, los errores poco a poco aumentan como se muestra en la **Figura 4.10**, pero al igual que en la

Tabla 4.7 Ventanas utilizadas en la transformada *synchrosqueezing* basada en Fourier con las simulaciones a 6 bpm.

Ventana	RMSE (bpm)	MAE (bpm)	Índice de correlación cruzada
Gaussiana 120 s	$6.905 \times 10^{-2} \pm 0$	$6.905 \times 10^{-2} \pm 0$	1 ± 0
Rectangular 120 s	$6.905 \times 10^{-2} \pm 0$	$6.905 \times 10^{-2} \pm 0$	1 ± 0
Rectangular 180 s	$6.804 \times 10^{-2} \pm 0$	$6.804 \times 10^{-2} \pm 0$	1 ± 0
Rectangular 30 s	$7.405 \times 10^{-2} \pm 0$	$7.405 \times 10^{-2} \pm 0$	1 ± 0

Valor de los índices presentados como media $\pm$ SD.

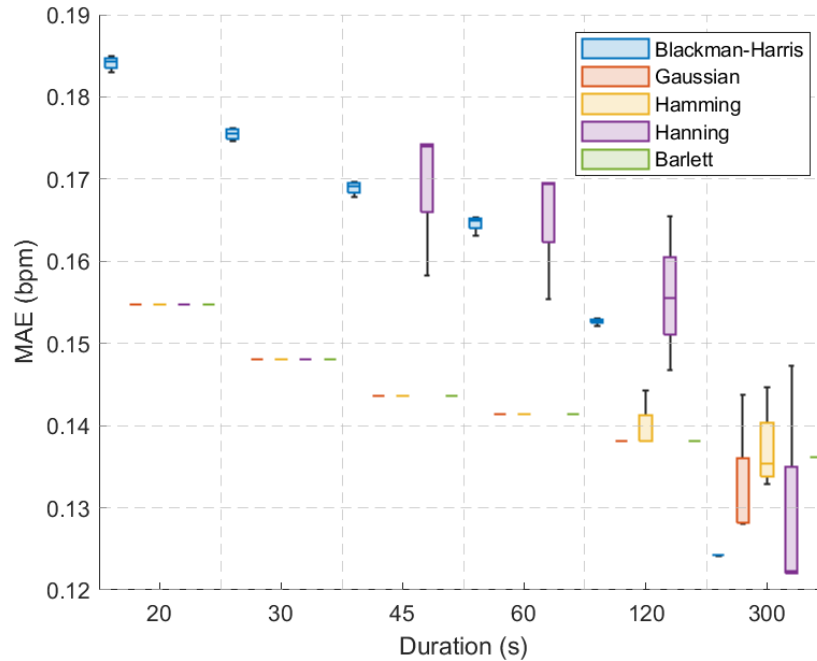


Figura 4.10 Diagrama de cajas y bigotes del MAE para las simulaciones constantes en 12 bpm de la transformada *synchrosqueezing* basada en Fourier.

maniobra constante a 6 bpm y por los mismos motivos también se usará una ventana de 30 s las cuales resultan en un MAE y RMSE de 0.148 bpm y un índice de correlación cruzada de 1, a excepción de la Blackman-Harris.

Para la maniobra de 18 bpm la mejor ventana es la misma que en 6 bpm, i.e., la rectangular de 180 s (RMSE y MAE = 0.172 bpm), no muy lejos las ventanas de 30 s tienen un RMSE y MAE = 0.222 bpm. Algo similar ocurre en los 24 bpm, puesto que la ventana con los errores más pequeños resulto la Blackman-Harris de 180 s (RMSE y MAE = 0.052), la misma que para 0.2 Hz, pero el problema del tamaño de la ventana prevalece. En esta última maniobra, tres de las ventanas de 30 s tienen un RMSE y MAE de 0.076 bpm. En estos casos las ventanas largas no tienen errores muy alejados de las ventanas de 30 s, por lo que, en todas las constantes se decidió usar una ventana rectangular de 30 s, debido a que es una ventana más corta con buenos resultados y es del mismo tipo que las de 180 s para las maniobras respiratorias de 6 y 18 bpm.

En las simulaciones después del ejercicio las mejores ventanas están en 120 s para todos los índices de desempeño, como se ilustra en la **Figura 4.11** para el índice RMSE. Particularmente, la que tiene los menores errores y la correlación más grande es la ventana Hamming (MAE = 0.255 bpm, RMSE = 0.487 bpm, índice de correlación cruzada = 9.98×10^{-1}). Las siguientes mejores ventanas son la Gaussiana y la Blackman-Harris de 120 s. Al comparar la Hamming con la Gaussiana, utilizando una prueba *t-Student*, no hay diferencia ($p_{MAE} = 0.263, p_{RMSE} = 0.435, p_x = 0.435$), pero al compararla con la Blackman-Harris si la hay. Pese a no haber diferencia con la Gaussiana, la ventana Hamming de 120 s fue la seleccionada, por tener los errores más chicos y el índice de correlación cruzada más grande.

Por último, en las simulaciones espontaneas, las mejores ventanas se encuentran en la longitud de 60 s, de las cuales se muestran sus índices de desempeño en la

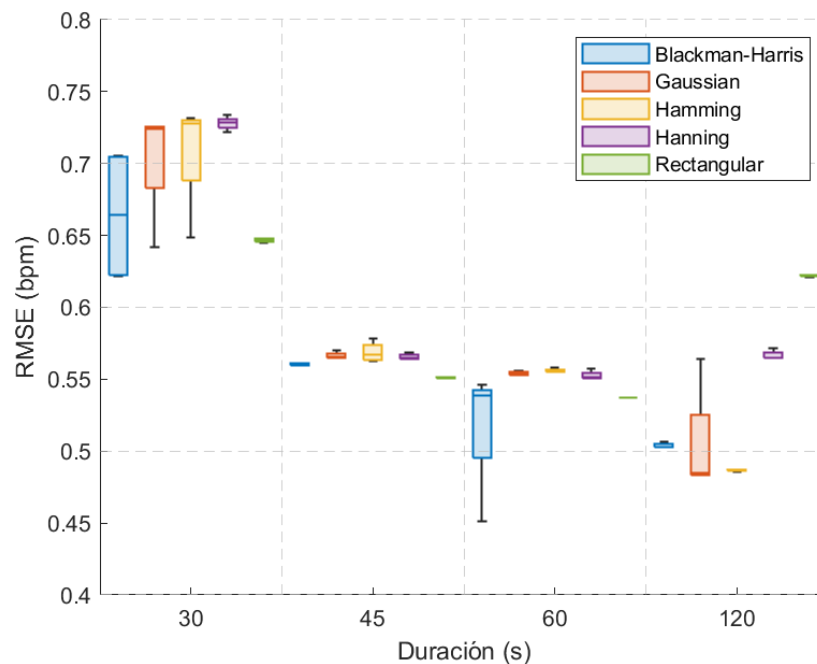


Figura 4.11 Diagrama de cajas y bigotes del RMSE (bpm) de las iRRs obtenidas a partir de las simulaciones después del ejercicio y la FSST. En el eje x se encuentra la duración de la ventana (solo de 30 a 120 s) y en el código de colores se encuentra el tipo de ventana.

Tabla 4.8. En ese tamaño, la ventana Hamming es la mejor, luego le siguen la ventana Gaussiana y la rectangular. Al comparar la Hamming con esas dos ventanas con un *t-Student*, no se encontró diferencia con la Gaussiana ($p_{MAE} = 0.057, p_{RMSE} = 0.058, p_x = 0.058$), sin embargo, con la rectangular si hubo diferencia. Al igual que en la maniobra después del ejercicio, se tuvo que seleccionar una sola ventana, la cual fue la Hamming. En la **Tabla 4.9** se muestran los parámetros seleccionados en esta técnica para cada maniobra, donde se puede observar que en prácticamente todas las maniobras respiratorias constantes se seleccionó la ventana rectangular y la segunda más seleccionada fue la Gaussiana. También se destaca que los tamaños de ventana son mínimo de 30 s.

4.3.2.4 HHS vía CEEMDAN

En lo que se refiere al HHS vía CEEMDAN, se utilizaron tres diferentes parámetros, los cuales son el número de realizaciones, la desviación estándar (DS) del ruido y el uso o no de un filtro Gaussiano con el fin de mejorar la visualización de la TFR en los resultados del CEEMDAN. A continuación, se presenta la seleccionaron los parámetros para cada maniobra.

En la maniobra de cambios abruptos, los errores más pequeños estaban en las simulaciones con una desviación estándar de ruido igual a 1, sin importar si se

Tabla 4.8 Índices de desempeño de las iRRs obtenidas de la transformada *synchrosqueezing* basada en Fourier de las simulaciones espontaneas utilizando un tamaño de 60 s en las ventanas.

Ventana (60 s)	RMSE (bpm)	MAE (bpm)	Índice de correlación cruzada
Gaussiana	$0.319 \pm 1.979 \times 10^{-2}$	$0.273 \pm 9.783 \times 10^{-3}$	$9.994 \times 10^{-1} \pm 1.087 \times 10^{-4}$
Hamming	$0.289 \pm 3.667 \times 10^{-4}$	$0.259 \pm 2.121 \times 10^{-4}$	$9.996 \times 10^{-1} \pm 1.035 \times 10^{-6}$
Rectangular	$0.297 \pm 1.190 \times 10^{-3}$	$0.263 \pm 7.095 \times 10^{-4}$	$9.994 \times 10^{-1} \pm 7.369 \times 10^{-6}$

Valor de los índices presentados como media $\pm$ SD.

Tabla 4.9 Parámetros seleccionados en la transformada *synchrosqueezing* basada en Fourier.

Maniobra respiratoria	Tipo de ventana	Duración de la ventana (s)
Cambios abruptos	Gaussiana	60
6 bpm	Rectangular	30
12 bpm	Rectangular	30
18 bpm	Rectangular	30
24 bpm	Rectangular	30
Después del ejercicio	Hamming	120
Espontanea	Hamming	60

usó o no el filtro, como se ve en la **Figura 4.12**. Aun así, al comparar los datos con una SD de ruido igual a 1, los valores más bajos estaban en las simulaciones donde si se aplicó filtro, como se puede observar en la **Figura 4.12.A** en las cajas de color rojo, pese a que las diferencias son mínimas con las de la **Figura 4.12.B**. En especial, los MAE y RMSE más pequeños estaban en 10, 15 y 20 realizaciones. A pesar de esto, se decidió solo utilizar la de 15 realizaciones, debido a que su MAE de 0.360 bpm y RMSE de 0.628 bpm son más pequeños que los demás e incluso su índice de correlación cruzada = 9.991×10^{-1} es la mayor.

En la maniobra respiratoria constante a 6 bpm, el MAE y RMSE más pequeños ocurren para una SD de ruido de 1, tanto con filtro, como sin filtro y los valores con 10 y 15 realizaciones se muestran en la **Tabla 4.10**, donde se observa que los errores menores están en el grupo donde se utilizó filtro. Entre estos parámetros, los mejores estaban en 15 realizaciones. Entonces, se determinó utilizar el filtro, con SD de ruido = 1 y 15 realizaciones. Lo mismo ocurrió en la maniobra respiratoria a 12 bpm.

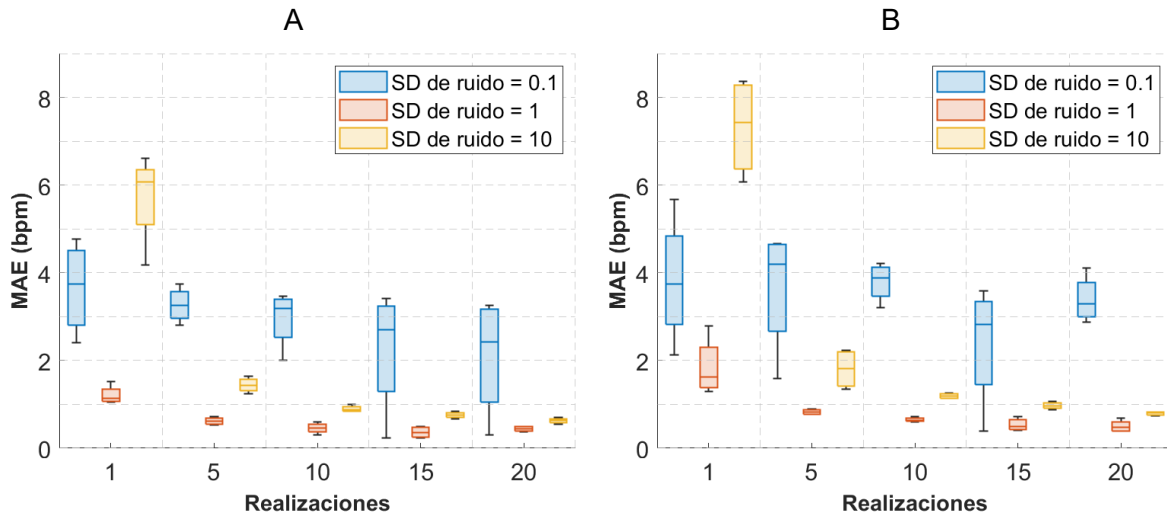


Figura 4.12 Diagramas de cajas y bigotes de las simulaciones con cambios abruptos de los resultados del espectro Hilbert-Huang. A) Con filtro Gaussiano y B) sin filtro.

Tabla 4.10 Índices de desempeño de cuatro espectros Hilbert-Huang de la maniobra constante a 6 bpm. La SD de ruido esta fija en 1.

Realiza- ciones	Sin filtro			Con filtro		
	RMSE (bpm)	MAE (bpm)	Índice de correlación cruzada	RMSE (bpm)	MAE (bpm)	Índice de correlación cruzada
10	0.403 $\pm 1.602 \times 10^{-2}$	0.312 $\pm$ 9.369 $\times 10^{-3}$	9.983×10^{-1} $\pm 2.033 \times 10^{-4}$	0.330 $\pm$ 7.366 $\times 10^{-2}$	0.247 $\pm$ 7.182 $\times 10^{-2}$	9.987×10^{-1} $\pm 3.513 \times 10^{-4}$
15	0.391 $\pm$ 1.155 $\times 10^{-2}$	0.308 $\pm$ 1.169 $\times 10^{-2}$	9.985×10^{-1} $\pm 9.033 \times 10^{-5}$	0.308 $\pm$ 6.72 $\times 10^{-2}$	0.222 $\pm$ 7.276 $\times 10^{-2}$	9.988×10^{-1} $\pm 2.44.2 \times 10^{-4}$

Valor de los índices presentados como media $\pm$ SD.

Para las simulaciones en 18 bpm y 24 bpm, los mejores índices de desempeño se encuentran utilizando el filtro, con SD del ruido de 15 y 20 realizaciones, ver **Figura 4.13** para los 24 bpm y la **Tabla 4.11** para ambos tipos de

simulaciones. Aunque, no se encontró diferencia entre usar 15 o 20 realizaciones,

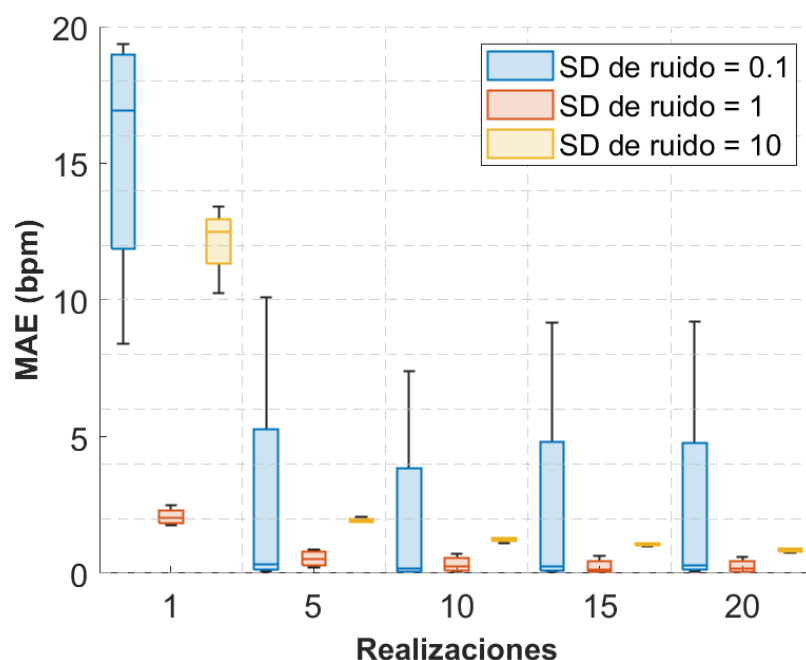


Figura 4.13 Diagrama de cajas y bigotes del MAE (bpm) de las iRRs obtenidas a partir de las simulaciones en 24 bpm. En el eje x se encuentra el número de realizaciones y con el código de colores se denota la desviación estándar del ruido.

Tabla 4.11 Índices de desempeño de las maniobras constantes a 18 bpm y a 24 bpm. La SD de ruido esta fija en 1.

Realizaciones	18 bpm			24 bpm		
	RMSE (bpm)	MAE (bpm)	Índice de correlación cruzada	RMSE (bpm)	MAE (bpm)	Índice de correlación cruzada
15	0.294 ± 0.131	0.272 ± 0.099	9.999 x 10 ⁻¹ ± 1.488 x 10 ⁻⁴	0.341 ± 0.333	0.251 ± 0.271	9.998 x 10 ⁻¹ ± 2.429 x 10 ⁻⁴
20	0.276 ± 0.069	0.262 ± 0.057	9.999 x 10 ⁻¹ ± 8.061 x 10 ⁻⁵	0.313 ± 0.349	0.243 ± 0.26	9.998 x 10 ⁻¹ ± 2.407 x 10 ⁻⁴

Valor de los índices presentados como media ± SD.

excepto para 24 bpm en el índice de correlación cruzada, se determinó usar 20 realizaciones para ambas maniobras por tener los valores menores en los errores.

Los mejores resultados para la maniobra después del ejercicio, nuevamente, se encuentran en una SD del ruido de 1, con el uso del filtro. El problema es que en las realizaciones hay muchas con buenos resultados, como los mostrados en la **Tabla 4.12**, en los cuales no se encontró diferencia al hacer la *t-Student*, puesto que en todos los índices se obtuvo una $p < 0.5$. De todas formas, se seleccionó la de 20 realizaciones por tener dos de los mejores índices (RMSE e índice de correlación cruzada).

En el caso de la maniobra espontanea ocurre lo mismo que en la de ejercicio, puesto que hay muy buenos grupos de parámetros, en los cuales no se encuentran diferencias y de nuevo en todos esos grupos se utilizó el filtro y una SD de ruido de 1, en la **Tabla 4.13** se muestran estos grupos. Finalmente, para elegir las realizaciones también se tomó en cuenta la SD, por la mínima diferencia entre los valores de los índices, entonces se seleccionó utilizar 20 realizaciones por tener la mínima SD.

Los parámetros de SD de ruido y el uso del filtro en esta técnica fueron los mismos en todas las maniobras, i. e., SD de ruido igual a 1 y si usar el filtro

Tabla 4.12 Índices de desempeño de las simulaciones de después del ejercicio utilizando el espectro Hilbert-Huang. Solo se muestran datos donde se usó el filtro. El SD de ruido es 1.

Realizaciones	RMSE (bpm)	MAE (bpm)	Índice de correlación cruzada
10	0.499 ± 0.107	0.285 ± 0.058	$0.998 \times 10^{-1} \pm 9.971 \times 10^{-4}$
15	0.547 ± 0.133	0.271 ± 0.034	$0.997 \times 10^{-1} \pm 1.368 \times 10^{-3}$
20	0.479 ± 0.060	0.273 ± 0.040	$0.998 \times 10^{-1} \pm 6.103 \times 10^{-4}$

Valor de los índices presentados como media $\pm$ SD.

solamente hubo cambios en la selección de las realizaciones, como se muestra en la **Tabla 4.14**, donde en tres casos se seleccionaron 15 realizaciones y en cuatro se seleccionaron 20. Estos resultados nos indican que al menos para las señales simuladas podría haber un rango óptimo de parámetros en esta técnica.

Tabla 4.13 Índices de desempeño de las simulaciones espontaneas utilizando el espectro Hilbert-Huang. Se muestran los datos con uso del filtro y SD de ruido unitario.

Realizaciones	RMSE (bpm)	MAE (bpm)	Índice de correlación cruzada
10	$0.410 \pm 6.312 \times 10^{-1}$	$0.330 \pm 4.464 \times 10^{-1}$	$9.992 \times 10^{-1} \pm 4.382 \times 10^{-4}$
15	$0.376 \pm 8.616 \times 10^{-1}$	$0.300 \pm 4.494 \times 10^{-1}$	$9.992 \times 10^{-1} \pm 6.053 \times 10^{-4}$
20	$0.377 \pm 2.969 \times 10^{-1}$	$0.302 \pm 2.271 \times 10^{-1}$	$9.994 \times 10^{-1} \pm 1.441 \times 10^{-4}$

Valor de los índices presentados como media $\pm$ SD.

Tabla 4.14 Realizaciones seleccionadas en el HSS vía CEEMDAN.

Maniobra respiratoria	Realizaciones
Cambios abruptos	15
6 bpm	15
12 bpm	15
18 bpm	20
24 bpm	20
Después del ejercicio	20
Espontanea	20

4.4 Estimación de RR instantánea en datos reales

Al tener los parámetros seleccionados para cada técnica TF, estos se aplicaron en los datos de la base de datos. Al igual que se hizo en las simulaciones, se extrajo la iRR de las TFR para hacer un análisis de desempeño de los resultados, pero esta vez se utilizó la señal de respiración de la base de datos como referencia, como se ilustra en la **Figura 4.14**, donde aparecen un ejemplo de TFR de cada una de las técnicas TF aplicada a una maniobra respiratoria. En la **Figura 4.14.A** se muestra el HHS vía CEEMDAN y como se observa, la iRR estimada oscila demasiado provocando que no coincida totalmente con la referencia. En la **Figura 4.14.B** se encuentra la maniobra respiratoria constante a 6 bpm con la FSST, donde la iRR estimada está prácticamente empalmada con la referencia. En la **Figura 4.14.C** aparece la maniobra después del ejercicio con el escalograma, donde se ve que la iRR estimada pudo seguir casi todos los cambios en la frecuencia de la referencia, excepto en el extremo final de la señal. En la **Figura 4.14.D** se muestra el SP de la maniobra espontánea, donde también se observa que la iRR estimada coincide muy bien con la referencia respiratoria.

4.4.1 Estimación de iRR con las TFR seleccionadas

La estimación de la iRR se hizo de la misma forma que para las simulaciones, cuya metodología se detalla en la **sección 3.5**. Algunos de los resultados no correspondían con la maniobra respiratoria o alguno de los archivos estaban dañados, por lo que se tuvieron que descartar. Aun así, en algunas de las señales de iPPG, de los voluntarios cuyas señales que no concordaban con la maniobradespiratoria, se aplicaron las técnicas seleccionadas, e.g., en la **Figura 4.15**, la señal de referencia la mayoría del tiempo está por encima de los 6 bpm, a pesar de que esa fue la frecuencia de respiración que llevó a lo largo de la adquisición y curiosamente, a partir de la señal PR extraída del iPPG, si se obtuvo una frecuencia respiratoria alrededor de los 6 bpm. A pesar de estas situaciones las señales se descartaron porque no había una referencia adecuada para comparar

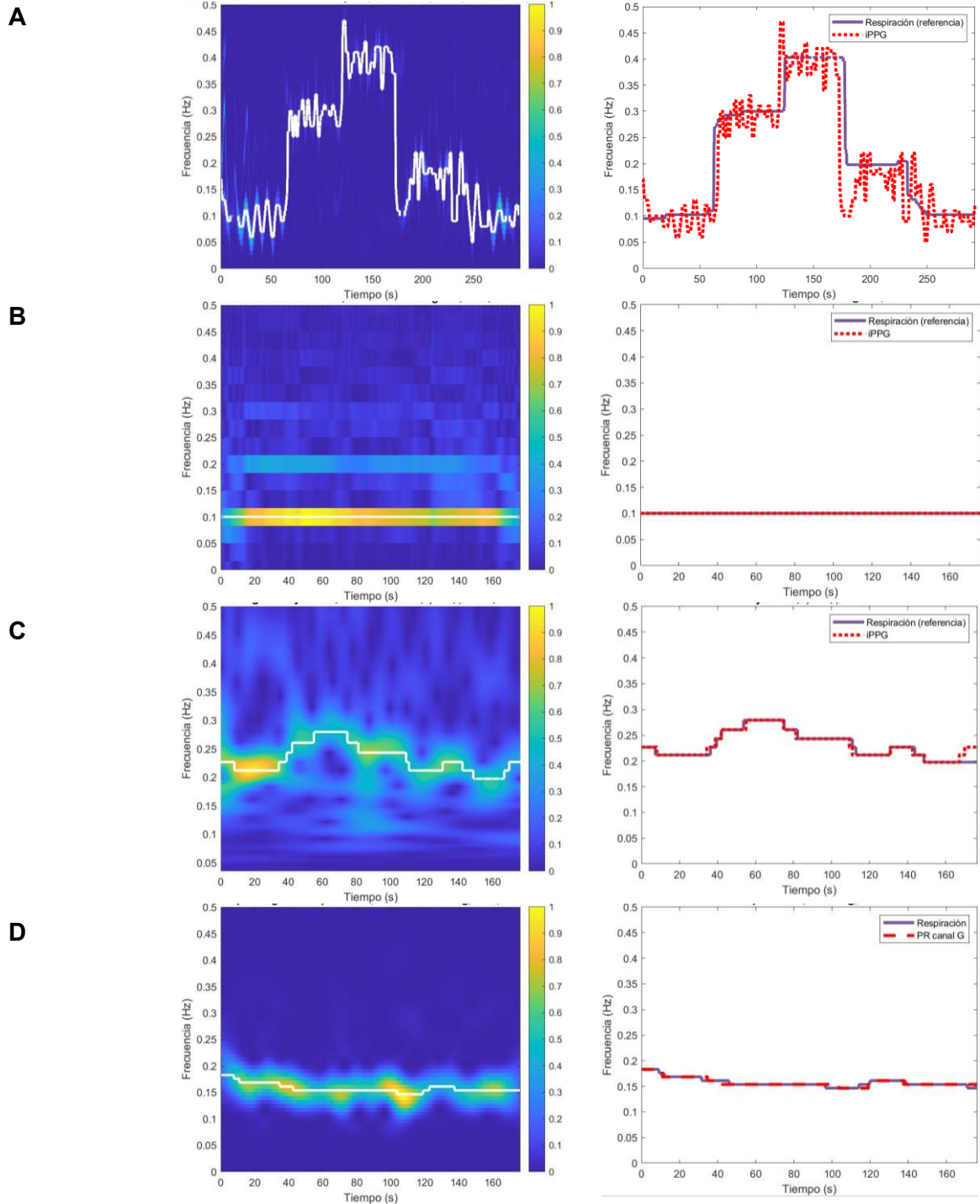


Figura 4.14 Ejemplos de TFR e iRR de la base de datos. A) Espectro Hilbert-Huang en la maniobra de cambios abruptos. B) Transformada *synchrosqueezing* basada en Fourier de la maniobra a 6 bpm. C) Escalograma de la maniobra después del ejercicio. D) Espectrograma de la maniobra espontanea.

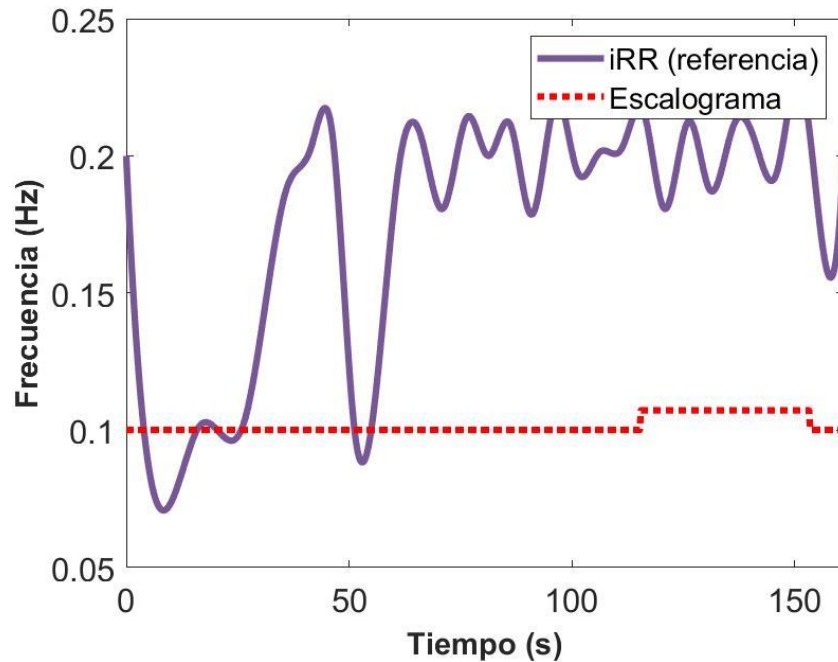


Figura 4.15 iRR de la señal respiratoria de referencia (en morado) y iRR estimada mediante el escalograma ($\gamma = 30$, $\beta = 2$) de la señal PR (en línea roja discontinua) del mismo voluntario al respirar a 6 bpm (0.1 Hz).

los resultados del estimador de la iRR basado en técnicas TF.

4.4.2 Análisis de desempeño

En este caso, el análisis de desempeño se hizo para cada maniobra. Se compararon los resultados de la estimación con la ventana utilizada en esfuerzos anteriores con las ventanas seleccionadas a partir de los resultados de las simulaciones realizadas en esta tesis. Al realizar las pruebas de normalidad en los datos, ninguno de ellos siguió una distribución normal, además los datos son de medidas repetidas, por lo que la prueba estadística utilizada para encontrar si existía una diferencia entre los grupos de parámetros fue la Friedman.

En la comparativa, también se incluyeron las estimaciones de la iRR siguiendo la metodología anteriormente propuesta por el grupo de trabajo, donde solo se hizo uso del SP con una ventana Hamming con tamaño del 10 % de la señal.

Además, en esa metodología, la marginal se obtuvo con un algoritmo que seleccionaba el valor máximo de energía por cada instante de tiempo, sin tomar en cuenta la posición de los instantes de tiempo vecinos, como se hace en el algoritmo *greedy* de búsqueda heurística penalizado hacia adelante y hacia atrás que se usó en esta tesis.

4.4.2.1 Cambios abruptos en la respiración

Para la maniobra de cambios abruptos se utilizaron siete combinaciones diferentes de parámetros entre las cuatro técnicas TF, cuyos resultados están en la **Tabla 4.15**, en la cual se muestra que los mejores índices de desempeño ocurren para el SP usando una ventana Gaussiana de 45 s, seguido por los del escalograma. Se realizó

Tabla 4.15 Índices de desempeño de la iRR estimada a partir de la señal de iPPG de la maniobra de cambios abruptos.

Técnica tiempo-frecuencia	RMSE (bpm)	MAE (bpm)	Índice de correlación cruzada
Espectrograma (Gaussiana 45 s)	2.951 ± 1.152 3.256(1.279, 4.062)	1.518 ± 0.856 1.424 (0.508, 2.597)	9.808 x 10 ⁻¹ ± 0.012 0.980 (0.968, 0.996)
Espectrograma (Hamming 30 s)	4.625 ± 2.801 4.215(1.432,10.174)	3.068 ± 2.368 2.510 (0.6, 7.795)	9.634 ± 0.028 0.968 (0.909, 0.996)
FSST (Gaussiana 60 s)	3.503 ± 1.647 3.692 (1.359, 5.50)	1.754 ± 1.159 1.658 (0.352, 3.451)	9.704 x 10 ⁻¹ ± 0.024 0.973 (0.934, 0.996)
Escalograma ($\gamma = 3$, $\beta = 40$)	3.250 ± 1.463 3.423(1.315, 5.306)	1.799 ± 1.005 1.819 (0.499, 3.267)	9.770 x 10 ⁻¹ ± 0.016 0.978 (0.951, 0.996)
HSS (DS = 1, realizaciones = 15)	4.439 ± 2.643 3.348(1.588, 9.204)	3.117 ± 2.039 2.223 (1.109, 7.133)	9.489 x 10 ⁻¹ ± 0.052 0.974 (0.846, 0.994)

Valores de los índices presentados como media +/- SD en el primer renglón y mediana (mínimo, máximo) en el segundo renglón.

una prueba Friedman con los datos y no se encontró que hubiera una diferencia entre al menos uno de los grupos de parámetros, con un valor $p = 0.157$ para el RMSE, $p = 0.073$ para el MAE y $p = 0.570$ para el índice de correlación cruzada. En la **Figura 4.16** se muestra la dispersión del RMSE (bpm) de los grupos de parámetros para la maniobra de cambios abruptos, donde cuatro de las cinco técnicas tienen un RMSE medio por debajo de los 4 bpm. Por otro lado, la técnica con los resultados menos deseables fue el HSS vía CEEMDAN, donde, aunque para 15 realizaciones se obtuvo una mediana de 3.348 bpm, su promedio es de 4.439 bpm y presenta la dispersión más grande. Otro de los peores resultados fue el SP con la ventana Hamming de 30 s con mediana de 4.215 bpm.

4.4.2.2 Respiración a una frecuencia constante

En la **Tabla 4.16** están los resultados de la prueba de Friedman en las maniobras constantes, donde de 12 a 24 bpm hubo una $p < 0.05$, por lo que al menos una media de los grupos es diferente al resto, en todos los casos la HHS con 20

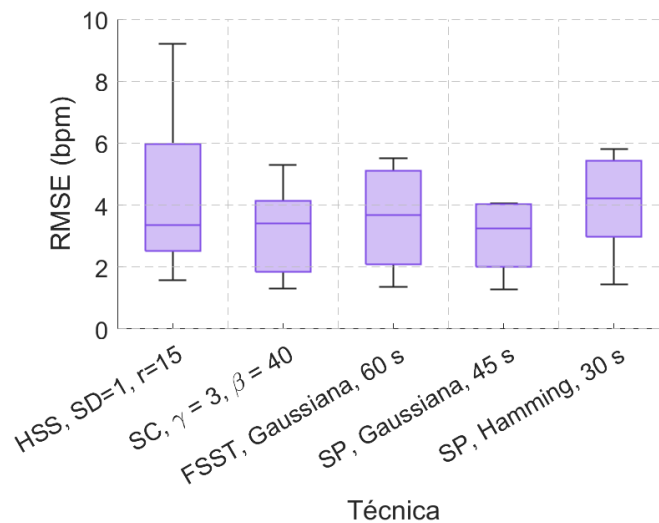


Figura 4.16 RMSE (bpm) de las técnicas TF y parámetros utilizados en la maniobra de cambios abruptos de la base de datos. Las abreviaciones son por: Espectro Hilbert-Huang (HSS), Escalograma (SC), transformada *synchrosqueezing* basada en Fourier (FSST), Espectrograma (SP).

realizaciones fue diferente a todos los demás grupos y para la respiración a 18 bpm también hubo diferencia entre el SP con ventana Hamming de 18 s y la FSST con ventana rectangular de 30 s. En la **Figura 4.17** se observan los resultados de la comparación para la correlación cruzada de la maniobra a 12 bpm, donde el mejor

Tabla 4.16 p valores de las maniobras metrónomas para cada índice de desempeño.

Técnica	RMSE	MAE	Índice de correlación cruzada
6 bpm	7.515×10^{-2}	0.164	5.101×10^{-2}
12 bpm	1.914×10^{-2}	4.59×10^{-5}	8.694×10^{-5}
18 bpm	7.569×10^{-5}	2.224×10^{-5}	2.224×10^{-5}
24 bpm	3.19×10^{-3}	1.213×10^{-3}	3.19×10^{-3}

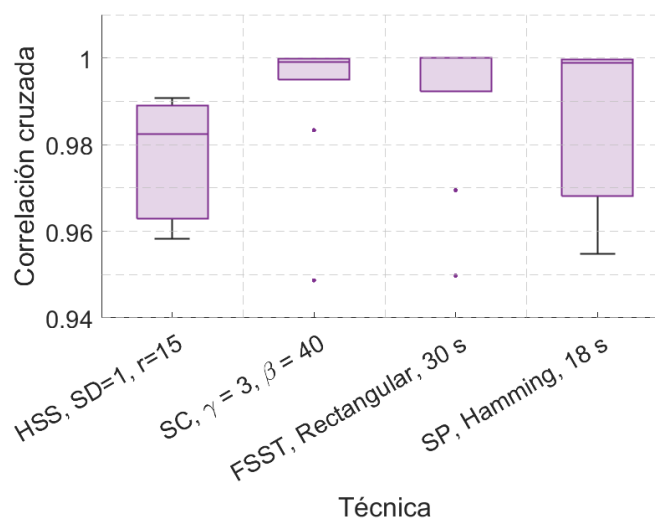


Figura 4.17 Índice de correlación cruzada de las técnicas y parámetros utilizados en la maniobra constante en 12 bpm de la base de datos. Las abreviaciones son por: Espectro Hilbert-Huang (HSS), Escalograma (SC), transformada *synchrosqueezing* basada en Fourier (FSST), Espectrograma (SP).

resultado le atañe al escalograma, seguido por la FSST, aunque no haya diferencia estadística entre estas dos técnicas. En la **Tabla 4.17** se presentan los resultados en dos de las maniobras metrónomas, 6 y 18 bpm, donde la FSST tiene errores menores a los del escalograma, pero el índice de correlación cruzada es mayor en este último para 6 bpm e igual para 18 bpm. En la maniobra respiratoria a 18 bpm, cabe destacar que en el SC y en la FSST hay medianas y mínimos de cero, lo que

Tabla 4.17 Índices de desempeño de la iRR estimada a partir de las señales de iPPG de 6 y 18 bpm.

Técnica tiempo-frecuencia	RMSE (bpm)	MAE (bpm)	Índice de correlación cruzada
6 bpm			
Espectrograma (Hamming 18 s)	4.629 ± 2.909 4.602 (1.166, 8.145)	3.773 ± 2.803 3.549 (0.632, 7.363)	0.956 ± 0.020 0.951 (0.939, 0.982)
FSST (Rectangular 30 s)	1.952 ± 2.322 3.692 (1.359, 5.50)	1.297 ± 1.634 1.658 (0.352, 3.451)	0.969 ± 0.037 0.975 (0.926, 1)
Escalograma ($\gamma = 3$, $\beta = 40$)	2.109 ± 2.419 1.609 (0, 4.589)	1.503 ± 1.99 0.896 (0, 3.394)	0.975 ± 0.030 0.981 (0.938, 0.999)
HSS (DS = 1, realizaciones = 15)	3.106 ± 1.162 3.468 (1.417, 4.069)	2.218 ± 0.807 2.510 (1.033, 2.819)	0.917 ± 0.036 0.901 (0.894, 0.971)
18 bpm			
Espectrograma (Hamming 18 s)	1.280 ± 0.889 0.847 (0.373, 2.803)	0.697 ± 0.383 0.633 (0.275, 1.440)	0.996 ± 0.004 0.998 (0.987, 0.999)
FSST (Rectangular 30 s)	0.289 ± 0.511 0 (0, 1.429)	0.087 ± 0.157 0 (0, 0.379)	0.999 ± 0.001 1 (0.997, 1)
Escalograma ($\gamma = 3$, $\beta = 2$)	0.663 ± 0.674 0.672 (0, 1.791)	0.287 ± 0.378 0.238 (0, 1.145)	0.999 ± 0.001 0.999 (0.997, 1)
HSS (DS = 1, realizaciones = 20)	3.036 ± 1.836 2.581 (1.576, 7.452)	2.344 ± 1.373 1.908 (1.214, 5.629)	0.967 ± 0.029 0.975 (0.900, 0.991)

Valores de los índices presentados como media +/- SD en el primer renglón y mediana (mínimo, máximo) en el segundo renglón

habla muy bien de los parámetros seleccionados, pero también de la naturaleza de la frecuencia constante de las señales. Los resultados de la **Figura 4.17** y de la **Tabla 4.17** son muy parecidos para las demás maniobras respiratorias constantes.

Para la frecuencia a 0.4 Hz, el RMSE medio más bajo fue de 2.099 bpm, que le corresponde a la FSST, siendo ese el RMSE más grande entre los mínimos. Esta pérdida en la resolución al aumentar la frecuencia se ha registrado en otros trabajos, con diversas técnicas como el escalograma, la VCFDM o al descomponer la señal en sus tres modulaciones [14], [17], [87]. Esto puede deberse a que la modulación inducida por la respiración pueda tener un efecto menos intenso a frecuencias respiratorias más altas [87]. Por otra parte, los errores en las estimaciones en 0.1 Hz también tendieron a ser más grandes que en 12 y 18 bpm, esto debido a irregularidades que se encuentran en las señales a bajas frecuencias, posiblemente debido a las variaciones en la profundidad y la velocidad de la respiración, el sensor puede causar irregularidades en la señal que no reflejan con precisión la ventilación del voluntario.

4.4.2.3 Respiración después del Ejercicio

En la maniobra después del ejercicio también se usaron siete grupos de parámetros, de los cuales sus índices de desempeño se encuentran en la **Tabla 4.19**. En la prueba Friedman se encontró que, no hay una diferencia entre los grupos para el índice RMSE con $p = 0.060$ y para el MAE con $p = 5.004 \times 10^{-2}$, en cambio para el índice de correlación cruzada la $p = 2.562 \times 10^{-2}$. Los mejores resultados están entre la ventana Hamming de 120 s de la FSST y el escalograma, pero se tiene que tomar en cuenta que la primera es una ventana muy larga y la técnica tiene mayor costo computacional, e. g., que el SP. El otro buen resultado, que fue el del escalograma, donde se obtuvo una mediana en el MAE de 0.362 bpm, como se observa en la **Figura 4.19**, aunque su dispersión es un poco más grande, debido a que para tres voluntarios se obtuvieron índices MAE más grandes que el resto, pero menores a 6.5 bpm, Además, la mediana de su RMSE = 0.684 bpm es menor a la

Tabla 4.19 Índices de desempeño de la iRR estimada a partir de la señal de iPPG después del ejercicio.

Técnica	RMSE (bpm)	MAE (bpm)	Índice de correlación cruzada
Espectrograma (Hamming 18 s)	3.170 ± 1.682 2.865 (0.491, 5.278)	1.775 ± 1.213 1.172 (0.311, 3.666)	9.739 x 10 ⁻¹ ± 0.015 0.971 (0.953, 0.999)
Espectrograma (Rectangular 20 s)	2.968 ± 2.426 2.093 (0.331, 6.293)	1.773 ± 1.632 0.852 (0.206, 4.235)	9.749 x 10 ⁻¹ ± 0.022 0.980 (0.942, 0.999)
FSST (Hamming 120 s)	2.394 ± 2.861 1.107 (0.266, 7.316)	1.743 ± 2.372 0.592 (0.051, 6.058)	9.828 x 10 ⁻¹ ± 0.026 0.996 (0.941, 0.999)
Escalograma ($\gamma = 3$, $\beta = 40$)	2.783 ± 2.781 0.684 (0.432, 6.467)	1.785 ± 1.984 0.362 (0.139, 4.458)	9.771 x 10 ⁻¹ ± 0.027 0.997 (0.939, 0.999)
HSS (DS = 1, realizaciones = 20)	4.672 ± 1.884 4.757 (2.270, 7.049)	3.393 ± 1.610 3.303 (1.400, 5.640)	9.298 x 10 ⁻¹ ± 0.046 0.952 (0.857, 0.982)

Valores de los índices presentados como media +/- SD en el primer renglón y mediana (mínimo, máximo) en el segundo renglón.

de la FSST con un RMSE = 1.107 bpm. Por otra parte, nuevamente los peores índices de desempeño se encuentran en el HSS vía CEEMDAN.

4.4.2.4 Respiración espontánea

Al igual que en los demás casos, en esta maniobra también se encontró una diferencia entre los grupos de parámetros, donde en este caso los p valores fueron muy pequeños, donde para el RMSE fue de 6.778×10^{-5} , para el MAE de 2.854×10^{-5} y para el índice de correlación cruzada fue de 4.492×10^{-5} . Los grupos donde hubo diferencia son entre el HHS y todos los demás, excepto por la Hamming de 18 s del SP; también hubo una diferencia ente la ventana Hanning de 30 s y la Hamming de 18 s del SP. En la **Figura 4.18** se muestra el índice de correlación cruzada, donde se observa que la mediana más grande ocurre para la ventana

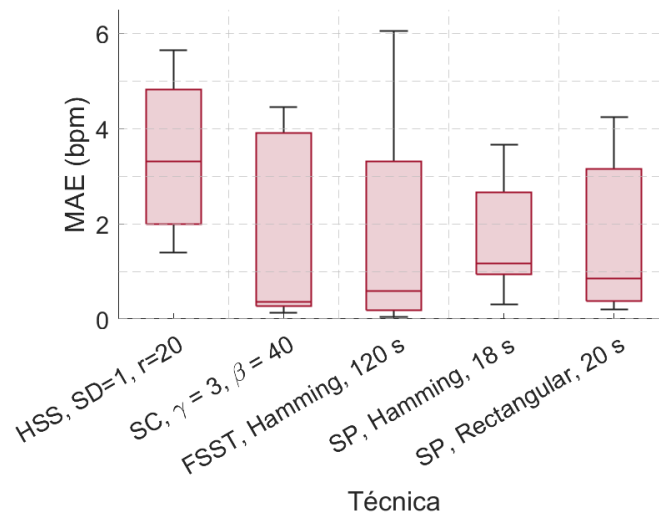


Figura 4.19 Diagrama de cajas y bigotes de los MAE (bpm) de la maniobra respiratoria después del ejercicio. Las abreviaciones son por: Espectro Hilbert-Huang (HSS), Escalograma (SC), transformada *synchrosqueezing* basada en Fourier (FSST), Espectrograma (SP).

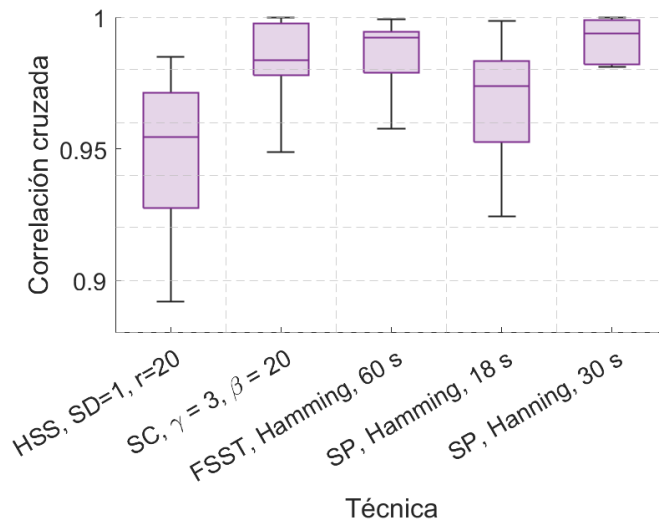


Figura 4.18 Diagrama de cajas y bigotes del coeficiente de correlación cruzada de la maniobra respiratoria espontanea. Las abreviaciones son por: Espectro Hilbert-Huang (HSS), Escalograma (SC), transformada *synchrosqueezing* basada en Fourier (FSST), Espectrograma (SP).

Hanning del SP. De igual manera, en la **Tabla 4.20** se presentan los promedios de

Tabla 4.20 Índices de desempeño de la maniobra espontanea de la base de datos.

Técnica	RMSE (bpm)	MAE (bpm)	Índice de correlación cruzada
Espectrograma (Hamming 18 s)	3.191 ± 1.798 2.459 (0.509, 6.488)	1.935 ± 1.237 1.357 (0.304, 4.036)	9.691 x 10 ⁻¹ ± 0.023 0.974 (0.924, 0.998)
Espectrograma (Hanning 30 s)	1.621 ± 1.120 1.851 (0.183, 3.273)	1.052 ± 0.799 0.974 (0.076, 2.174)	9.914 x 10 ⁻¹ ± 8.115 x 10 ⁻³ 0.994 (0.981, 0.999)
FSST (Hamming 60 s)	2.241 ± 1.487 2.103 (0.370, 4.948)	1.351 ± 0.940 1.319 (0.137, 2.696)	9.853 x 10 ⁻¹ ± 0.014 0.992 (0.957, 0.999)
Escalograma (γ = 3, β = 20)	2.449 ± 1.964 2.220 (0.229, 6.545)	1.459 ± 1.268 1.060 (0.078, 3.794)	9.836 x 10 ⁻¹ ± 0.016 0.983 (0.949, 0.999)
HSS (DS = 1, realizaciones = 20)	4.022 ± 1.982 3.591 (1.623, 6.811)	3.078 ± 1.640 2.699 (1.191, 5.479)	9.465 x 10 ⁻¹ ± 0.030 0.954 (0.892, 0.985)

Valores de los índices presentados como media +/- SD en el primer renglón y mediana (mínimo, máximo) en el segundo renglón.

los índices de desempeño para esta maniobra, se observa que la misma ventana tiene los mejores índices de desempeño y es seguida por la ventana Hamming de 60 segundos de la FSST. Además, se observa que la ventana Hamming de 18 s del SP y el HHS tienen los MAE y RMSE más grandes y las correlaciones cruzadas más pequeñas, lo que concuerda con las diferencias que se encontraron entre los grupos.

A modo de resumen, se destaca que en general la selección de los parámetros si ayudo a obtener estimaciones de la iRR más precisas en comparación con intentos anteriores por el grupo de trabajo, donde se usó el SP con una ventana Hamming, e. g., en la **Figura 4.20.A** están los resultados de un solo voluntario de la maniobra con cambios abruptos con cada una de las técnicas TF que se seleccionaron con ayuda de las simulaciones, donde para ese caso en particular el

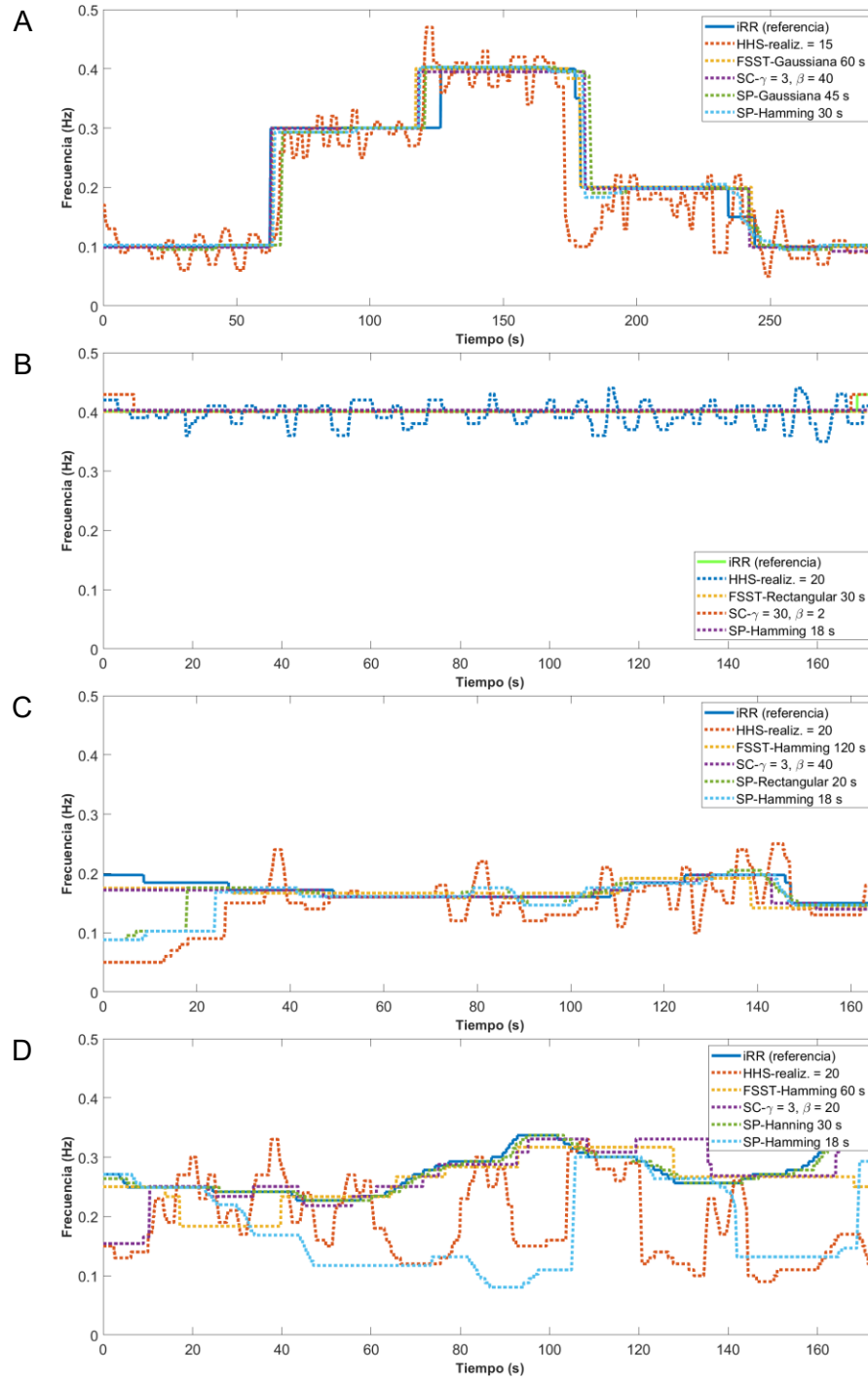


Figura 4.20 Ejemplo de estimación de la iRR con las técnicas tiempo-frecuencia aplicadas a los datos de un voluntario para A) la maniobra de cambios abruptos, B) la maniobra a 24 bpm, C) la maniobra después del ejercicio y D) la maniobra espontánea. Las abreviaciones son por: Espectro Hilbert-Huang (HSS), Escalograma (SC), transformada *synchrosqueezing* basada en Fourier (FSST) y Espectrograma (SP).

SC y la FSST pudieron seguir la mayor parte del tiempo a la referencia y más de cerca que la ventana Hamming de 30 segundos del SP, aun así, hay cambios que ninguna técnica pudo seguir como sucede en la **Figura 4.20.A** cerca de los 120 s.

En la **Figura 4.20.B** se presenta la maniobra respiratoria constante a 24 bpm, la cual es la frecuencia más alta utilizada. Se puede ver que casi todos los resultados están empalmados e incluso aunque al final la referencia cambia un poco de frecuencia, la FSST y la SP continúan sobre la frecuencia respiratoria de 24 bpm (0.4 Hz).

En la **Figura 4.20.C** se presentan los resultados de la maniobra después del ejercicio de un sujeto, donde se puede observar que al principio de la señal los dos SPs se acercaron a la referencia prácticamente en todo momento.

En la **Figura 4.20.D** se muestran los resultados de un voluntario de la maniobra espontánea, en donde se observa que la señal de referencia pudo ser seguida por la mayoría de las técnicas, aunque para este voluntario la ventana Hamming de 18 s del SP no pudo seguir todos los cambios de la señal de referencia. Aquí, se debe mencionar al HSS vía CEEMDAN, el cual en todos los casos obtuvo un menor desempeño que la técnica mencionada anteriormente.

4.5 Comparación con estudios reportados en la literatura

En la literatura sobre la estimación de la iRR a partir de la señal de iPPG, se ha reportado el uso de las técnicas VFCDM y CWT tomando ventaja de la modulación en la frecuencia de estas señales con la respiración. En específico, Nam et al. usaron maniobras metrónicas de una base de datos propia, con los datos obtenidos con diferentes dispositivos móviles, de 10 voluntarios. Ellos utilizaron tanto el canal de rojos, como el de verdes para estimar la señal de iPPG, con una ROI de 12x12 píxeles y el canal de verdes les resultó en mejores aproximaciones en todos los dispositivos [14]. Sus resultados están dados en términos de las medianas del RMSE. Para una respiración metrónica a 12 bpm, el RMSE más pequeño fue de

2.1 $\pm$ 1.05 bpm, para la maniobra a 18 bpm se obtuvo 2.08 $\pm$ 1.14 bpm y para la maniobra a 24 bpm fue de 2.38 $\pm$ 1.29 bpm. En nuestro caso, utilizando la ventana rectangular de 30 s en la FSST, para 12 bpm se obtuvo un RMSE de 1.009 $\pm$ 2.032, para 18 bpm se obtuvo un RMSE promedio de 0.289 $\pm$ 0.511 y para 24 bpm fue de 2.099 $\pm$ 3.369.

Karlen et al. reportaron el uso de algoritmos de fusión inteligente, al promediar el espectro de potencia de las tres modulaciones en las señales de iPPG de la respiración(en frecuencia, línea de base y amplitud), tomando las señales de iPPG del canal de rojos con una ROI central de 40x40 pixeles de 19 voluntarios mientras respiraban espontáneamente, para obtener una estimación de la RR de dos diferentes modelos de *smartphone*, donde con un dispositivo se obtuvo un RMSE de 2.27 bpm y con el otro de 2.29 bpm [11]. En contraste, en esta tesis, con dos de las selecciones hechas para esta maniobra se redujo el error; con la ventana Hamming de 60 s del FSST se obtuvo un RMSE = 2.241 bpm y aun mejor con el SP y la ventana Hanning de 30 s donde se obtuvo un RMSE = 1.621 bpm.

Gastel et al. adquirieron la iPPG mediante luz visible e infrarroja a distancia de 12 voluntarios, donde para extraer la señal iPPG, se utilizó un algoritmo en los pixeles que elegia cuales tendrían el menor SNR en la señal PR en todos los canales de color, Luego, para adquirir la iRR, se usó el SP con una ventana Hanning de 8 s, donde para una respiración espontanea los autores reportan MAEs de 4.72 bpm y 7.40 bpm, para luz visible e infrarroja, respectivamente[50]. Lo anterior resalta la adecuada selección de parámetros de las técnicas TF, donde incluso con el HHS se obtuvo un MAE = 3.078 bpm. Además, con respiraciones guiadas, que se asemejan a la de cambios abruptos utilizada aquí, se obtuvo un MAE de 1.74 bpm con luz visible y de 2.27 bpm con luz infrarroja [50]. En esta tesis, el primero de esos valores fue superado por el SP con una ventana Gaussiana de 45 s con un MAE = 1.518 bpm, y el segundo también fue superado por la FSST (MAE = 1.754 bpm) y el escalograma (MAE = 1.799 bpm). Cabe mencionar que los resultados reportados

en [50] emplean iPPG en modalidad remota, mientras que en esta tesis se consideró la modalidad de contacto.

Dehkordi et al utilizaron la transformada *synchrosqueezing* con una *wavelet* Morlet en señales de PPG de dos bases de datos publicas donde los voluntarios estaban anestesiados o dormidos (*CapnoBase* y *Sleep Data Set*). La transformada se usó primero en las señales de PPG y a partir de los resultados aplicaron nuevamente la transformada para obtener las tres modulaciones de la señal de iPPG. Finalmente se usó un algoritmo de fusión para combinar las estimaciones de iRR simultáneas [18]. Se hizo un análisis de los resultados partiendo de cada una de las modulaciones y de la fusión. Para una base de datos el menor RMSE fue para la modulación basada en la variación en la frecuencia con un valor de 3.7 bpm y para la otra base de datos fue un RMSE = 4.1 bpm para la modulación basada en la línea de base. Estos resultados contrastan con los dos RMSE más grandes obtenidos en esta tesis, donde en la maniobra espontanea se obtuvo un RMSE = 4.022 bpm para la HSS y un RMSE = 3.191 bpm para la ventana Hamming de 18 s en el SP. Los autores también usaron un algoritmo *peak-conditioned fusión*, que utiliza las tres modulaciones donde obtuvieron un RMSE = 1.8 bpm y de 2.3 bpm para cada base de datos, los cuales son más grandes a los obtenidos aquí por el SP con la ventana Hanning de 30 s (RMSE = 1.621) y a la Hamming de 60 s de la FSST (RMSE = 2.241).

Además, en casos donde se ha usado la PPG Pirhonen et al. llevaron a cabo una metodología, en la base de datos “*VORTAL dataset*” que contiene datos de 39 voluntarios sanos mientras respiraban de forma espontánea. Se utilizó la modulación en amplitud de la PPG, luego se usó la FSST con una ventana Gaussiana, y después se aplicó un filtrado de partículas en las TFRs, donde se obtuvo una RMSE = 4.28 ± 3.42 bpm el cual es mayor a todos los obtenidos aquí para la misma maniobra, y un MAE = 2.88 ± 0.85 bpm, el cual solo es más pequeño que el obtenido en esta tesis por la HHS, i.e., MAE = 3.078 ± 1.640 bpm [19].

Alafeef et al. estimaron la iRR de las señales de iPPG de 19 voluntarios durante la maniobra respiratoria espontánea, con el escalograma aplicándolo a señales basadas en el canal de intensidad rojo para extraer la señal de iPPG, donde se usó la WT discreta con la *wavelet* de Daubechies de orden 2 y obtuvieron una precisión promedio de 97.8% [15]. Al respecto, en esta tesis se encontraron índices de correlación cruzada = 0.991 para el epectrograma con la ventana Hanning de 30 s, para la FSST de 0.985 y el escalograma de 0.984.

En lo que respecta al HHS, Karlen et al., además de usar la fusión inteligente, previamente mencionada, usaron la técnica EMD para extraer cuatro modos de oscilación intrínseca (IMF) de señales de iPPG para una maniobra de respiración espontanea, obtenidas de una cámara conectada a un oxímetro, donde se obtuvo un RMSE = 7.17 bpm [11], el cual es mucho mayor que los RMSE obtenidos aquí con la selección de parámetros.

Capítulo 5

Conclusiones

El objetivo de este trabajo de tesis se centró en estimar la serie de tiempo de la frecuencia respiratoria instantánea (iRR) con diversas técnicas tiempo-frecuencia (TF) con parámetros preseleccionados a partir del análisis de simulaciones de diversas maniobras respiratorias de interés. Las técnicas TF exploradas en esta tesis fueron el clásico espectrograma (SP), la transformada *synchrosquezing* basada en Fourier (FSST), el escalograma (SC) y el espectro de Hilbert-Huang (HHS).

En la primera etapa, donde se seleccionaron los parámetros, fue evidente que hubo parámetros que se repitieron más que otros, e. g., en el SC una $\gamma = 3$ se eligió seis de ocho veces o en el HHS la DS de ruido fue de 1 para todas las selecciones. En el caso del SP se seleccionaron ventanas de 20 a 60 s y la ventana Gaussiana se seleccionó dos veces. Para la FSST, la ventana más seleccionada fue la rectangular, debido a que en las maniobras constantes tendió a ser de las mejores, mientras que en las longitudes hubo un rango de 30 a 120 s.

En las maniobras constantes no fue de sorprender que la selección de ventanas perteneciese a las de longitudes largas, puesto que solo poseen un único componente en frecuencia, y sería equivalente a obtener su espectro tradicional para la señal completa. Por su parte, en las maniobras con cambios en la frecuencia la selección resultó en ventanas con tamaños intermedios, lo que concuerda con lo reportado en la literatura [41]. Otro aspecto explorado fue el empleo de ruido aditivo,

el cual ayudó a seleccionar algunos de los parámetros, puesto que algunos de estos resultaban en estimaciones de la iRR que se veían muy afectadas incluso a bajos niveles de ruido.

La segunda etapa consistió en usar los parámetros seleccionados a partir de las simulaciones con los datos reales de la base de datos del grupo de trabajo. En general, la técnica con los resultados más prometedores fue la FSST, aunque con ventanas de gran longitud, e.g., en la maniobra después del ejercicio, donde la ventana Hamming de 120 s tuvo de los mejores índices de desempeño. El SP y el SC le siguieron en buenos resultados a la FSST, incluso el primero de ellos fue el mejor para la maniobra espontánea. La técnica con los resultados más pobres fue el HHS, donde hasta los parámetros usados en el esfuerzo anterior por el grupo de trabajo la superaron, los cuales fueron seleccionados de forma heurística.

La búsqueda de los parámetros de las técnicas de análisis tiempo-frecuencia con ayuda de las simulaciones con ruido aditivo si pudo mejorar la estimación de la frecuencia respiratoria instantánea en la gran mayoría de los casos, excluyendo al HHS, en comparación con el empleo de parámetros seleccionados únicamente de forma heurística. La frecuencia respiratoria es un signo vital muy relevante para identificar eventos adversos graves en el sistema cardiorrespiratorio. En tal sentido, la población en general podría verse beneficiada de una herramienta automática que brinde información acerca de los cambios temporales de la frecuencia respiratoria en situaciones de la vida diaria y que no dependa de instrumentación biomédica o de capacitación especializada. Consideramos que el trabajo desarrollado en esta tesis permite acercarse más a una estimación respiratoria instantánea de ese tipo a partir del adecuado análisis tiempo-frecuencia de señales de fotopletimografía de imagen adquiridas directamente con un dispositivo inteligente presente en la gran mayoría de los hogares.

5.1 Limitaciones

La principal limitación de este trabajo fue el empleo de una base de datos adquirida con pocos voluntarios, de los cuales además se tuvieron que descartar algunos registros porque los videos, a partir de los cuales se estima la iRR a partir del iPPG, o las señales respiratorias de referencia no eran adecuados. Además, sería beneficioso adquirir datos con sensores diferentes para la señal de respiración, como sensores de flujo diferenciales o turbina, o utilizando anemómetros de hilo caliente, como es aconsejado para la estimación de parámetros respiratorios [40]. Sin embargo, no se dispone de estos instrumentos biomédicos en el grupo de investigación actual.

5.2 Trabajo a futuro

Como trabajo a futuro, se plantera incorporar los parámetros seleccionados en esta tesis para calcular el SP en la *app* móvil desarrollada previamente por el grupo de investigación, para mejorar la estimación actual de la iRR. Actualmente, esta aplicación permite extraer las señales de iPPG a partir de la información capturada con la cámara principal, estimar la serie de tiempo PR y desplegar su SP, para observar las variaciones de RR a través del tiempo. La aplicación le pide al usuario establecer el tiempo de adquisición de iPPG. Luego, al presionar el botón de inicio, la cámara y el *flash* se encienden, y se comienzan a extraer las señales de iPPG, las cuales son graficadas en la pantalla y almacenadas. Hay dos opciones que permiten visualizar la información de la RR, una es por su frecuencia dominante con su PSD y la otra a partir de su SP, para observar la iRR [13]. Así, resultará importante incorporar el ajuste de parámetros al SP de acuerdo con lo encontrado en este trabajo de tesis.

También, se plantea explorar algoritmos que detecten automáticamente cual es la mejor ROI en cualquier dispositivo sin importar la posición del *flash* con respecto a la cámara. Adicionalmente, se propone adquirir datos durante maniobras respiratorias diferentes como frecuencias respiratorias por encima de los 24 bpm, y

analizar si parámetros parecidos a los aquí seleccionados también funcionan en otras condiciones. También, se podría explorar hacer ajustes en la técnica del HHS al momento de producir el espectro de Hilbert, puesto que se utilizaron todos los IMFs obtenidos para hacer la TFR y se plantea que podría ser beneficioso hacer una selección de ciertos IMFs para mejorar la estimación de iRR.

Además, cabe resaltar que la metodología planteada en esta tesis podría aplicarse a otras técnicas TF como la transformada *synchrosqueezing* basada en *wavelets*.

Finalmente, con el auge de la inteligencia artificial en los últimos años, la posibilidad de diseñar una red neuronal que pueda detectar automáticamente la frecuencia respiratoria espontánea a partir de la señal de iPPG es otro camino que se recomienda explorar.

Referencias

- [1] “Las 10 principales causas de defunción.” Accessed: Aug. 14, 2024. [Online]. Available: <https://www.who.int/es/news-room/fact-sheets/detail/the-top-10-causes-of-death>
- [2] T. Ferkol and D. Schraufnagel, “The global burden of respiratory disease,” *Ann. Am. Thorac. Soc.*, vol. 11, no. 3, pp. 404–406, Mar. 2014, doi: 10.1513/AnnalsATS.201311-405PS.
- [3] A. Zaccaro *et al.*, “How Breath-Control Can Change Your Life: A Systematic Review on Psycho-Physiological Correlates of Slow Breathing,” *Front. Hum. Neurosci.*, vol. 12, p. 353, 2018, doi: 10.3389/fnhum.2018.00353.
- [4] F. S. Cardoso, C. J. Karvellas, N. M. Kneteman, G. Meeberg, P. Fidalgo, and S. M. Bagshaw, “Respiratory rate at intensive care unit discharge after liver transplant is an independent risk factor for intensive care unit readmission within the same hospital stay: A nested case-control study,” *J. Crit. Care*, vol. 29, no. 5, pp. 791–796, Oct. 2014, doi: 10.1016/j.jcrc.2014.03.038.
- [5] H. Liu, J. Allen, D. Zheng, and F. Chen, “Recent development of respiratory rate measurement technologies,” *Physiol. Meas.*, vol. 40, no. 7, p. 07TR01, Aug. 2019, doi: 10.1088/1361-6579/ab299e.
- [6] S. Stankoski, I. Kiprijanovska, I. Mavridou, C. Nduka, H. Gjoreski, and M. Gjoreski, “Breathing Rate Estimation from Head-Worn Photoplethysmography Sensor Data Using Machine Learning,” *Sensors*, vol. 22, no. 6, p. 2079, Jan. 2022, doi: 10.3390/s22062079.
- [7] E. Jonathan and M. Leahy, “Investigating a smartphone imaging unit for photoplethysmography,” *Physiol. Meas.*, vol. 31, pp. N79-83, Nov. 2010, doi: 10.1088/0967-3334/31/11/N01.

- [8] “Más del 75% de la población mundial tiene un teléfono celular y más del 65% usa el internet | Noticias ONU.” Accessed: Aug. 30, 2024. [Online]. Available: <https://news.un.org/es/story/2023/12/1526712>
- [9] D. O. Lemos and C. A. Siebra, “Methods for Breathing Rate Measurement through Mobile Platform: a Review,” *Anais Principais do XIX Simpósio Brasileiro de Computação Aplicada à Saúde*, pp. 288–293, 2019.
- [10] H. Sharma, “Extraction of respiratory rate from PPG using ensemble empirical mode decomposition with Kalman filter,” *Electron. Lett.*, vol. 56, no. 13, pp. 650–653, Jun. 2020, doi: 10.1049/el.2020.0566.
- [11] W. Karlen, A. Garde, D. Myers, C. Scheffer, M. Ansermino, and G. Dumont, “Estimation of Respiratory Rate From Photoplethysmographic Imaging Videos Compared to Pulse Oximetry,” *IEEE J. Biomed. Health Inform.*, vol. 19, May 2015, doi: 10.1109/JBHI.2015.2429746.
- [12] D. Huang *et al.*, “Camera-Based Respiratory Imaging for Intelligent Rehabilitation Assessment of Thoracic Surgery Patients,” *IEEE Internet Things J.*, vol. 11, no. 15, pp. 26253–26265, Aug. 2024, doi: 10.1109/JIOT.2024.3395364.
- [13] E. G. Hernández de la Cruz, “Estimación de las frecuencias cardiaca y respiratoria instantáneas empleando métodos de fotopletismografía de imagen en un smartphone,” Maestría en Ingeniería Electrónica, Universidad Autónoma de San Luis Potosí, San Luis Potosí, 2021.
- [14] Y. Nam, J. Lee, and K. H. Chon, “Respiratory Rate Estimation from the Built-in Cameras of Smartphones and Tablets,” *Ann. Biomed. Eng.*, vol. 42, no. 4, pp. 885–898, Apr. 2014, doi: 10.1007/s10439-013-0944-x.
- [15] M. Alafeef and M. Fraiwan, “Smartphone-Based Respiratory Rate Estimation Using Photoplethysmographic Imaging and Discrete Wavelet Transform,” *J. Ambient Intell. Humaniz. Comput.*, vol. 11, Feb. 2020, doi: 10.1007/s12652-019-01339-6.

- [16] K. H. Chon, S. Dash, and K. Ju, "Estimation of Respiratory Rate From Photoplethysmogram Data Using Time–Frequency Spectral Estimation," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 8, pp. 2054–2063, Aug. 2009, doi: 10.1109/TBME.2009.2019766.
- [17] J. Lázaro, Y. Nam, E. Gil, P. Laguna, and K. H. Chon, "Respiratory rate derived from smartphone-camera-acquired pulse photoplethysmographic signals," *Physiol. Meas.*, vol. 36, no. 11, pp. 2317–2333, Nov. 2015, doi: 10.1088/0967-3334/36/11/2317.
- [18] P. Dehkordi, A. Garde, B. Molavi, J. M. Ansermino, and G. A. Dumont, "Extracting Instantaneous Respiratory Rate From Multiple Photoplethysmogram Respiratory-Induced Variations," *Front. Physiol.*, vol. 9, 2018.
- [19] M. Pirhonen, M. Peltokangas, and A. Vehkaoja, "Acquiring Respiration Rate from Photoplethysmographic Signal by Recursive Bayesian Tracking of Intrinsic Modes in Time-Frequency Spectra," *Sensors*, vol. 18, no. 6, p. 1693, Jun. 2018, doi: 10.3390/s18061693.
- [20] A. Garde, W. Karlen, P. Dehkordi, J. Ansermino, and G. Dumont, "Empirical mode decomposition for respiratory and heart rate estimation from the photoplethysmogram," in *Computing in Cardiology 2013*, Sep. 2013, pp. 799–802.
- [21] Md. N. I. Shuzan *et al.*, "A Novel Non-Invasive Estimation of Respiration Rate From Motion Corrupted Photoplethysmograph Signal Using Machine Learning Model," *IEEE Access*, vol. 9, pp. 96775–96790, 2021, doi: 10.1109/ACCESS.2021.3095380.
- [22] M. A. Colominas, G. Schlotthauer, and M. E. Torres, "Improved complete ensemble EMD: A suitable tool for biomedical signal processing," *Biomed. Signal Process. Control*, vol. 14, pp. 19–29, Nov. 2014, doi: 10.1016/j.bspc.2014.06.009.

- [23] A. Nicolò, C. Massaroni, E. Schena, and M. Sacchetti, "The Importance of Respiratory Rate Monitoring: From Healthcare to Sport and Exercise," *Sensors*, vol. 20, no. 21, p. 6396, Jan. 2020, doi: 10.3390/s20216396.
- [24] E. Vanegas, R. Igual, and I. Plaza, "Sensing Systems for Respiration Monitoring: A Technical Systematic Review," *Sensors*, vol. 20, no. 18, p. 5446, Jan. 2020, doi: 10.3390/s20185446.
- [25] G. Yuan, N. A. Drost, and R. A. McIvor, "Respiratory Rate and Breathing Pattern".
- [26] M. Buist, S. Bernard, T. V. Nguyen, G. Moore, and J. Anderson, "Association between clinically abnormal observations and subsequent in-hospital mortality: a prospective study," *Resuscitation*, vol. 62, no. 2, pp. 137–141, Aug. 2004, doi: 10.1016/j.resuscitation.2004.03.005.
- [27] G. J. Tortora and B. Derrickson, *Principios de Anatomía y Fisiología*, 13th ed. Editorial médica panamericana, Fecha de creación/publicación de la obra. [Online]. Available: <http://archive.org/details/tortora-13-principios-de-anatomia-y-fisiologia>
- [28] D. R. Hess and R. M. Kacmarek, "Physiologic Effects of Mechanical Ventilation," in *Essentials of Mechanical Ventilation*, 4th ed., New York, NY: McGraw-Hill Education, 2019. Accessed: Aug. 15, 2024. [Online]. Available: accessanesthesiology.mhmedical.com/content.aspx?aid=1157418259
- [29] M. Elstad, E. L. O'Callaghan, A. J. Smith, A. Ben-Tal, and R. Ramchandra, "Cardiorespiratory interactions in humans and animals: rhythms for life," *Am. J. Physiol.-Heart Circ. Physiol.*, vol. 315, no. 1, pp. H6–H17, Jul. 2018, doi: 10.1152/ajpheart.00701.2017.
- [30] P. D. Larsen, Y. C. Tzeng, P. Y. W. Sin, and D. C. Galletly, "Respiratory sinus arrhythmia in conscious humans during spontaneous respiration," *Respir. Physiol. Neurobiol.*, vol. 174, no. 1, pp. 111–118, Nov. 2010, doi: 10.1016/j.resp.2010.04.021.

- [31] P. Y. W. Sin, D. C. Galletly, and Y. C. Tzeng, "Influence of breathing frequency on the pattern of respiratory sinus arrhythmia and blood pressure: old questions revisited," *Am. J. Physiol.-Heart Circ. Physiol.*, vol. 298, no. 5, pp. H1588–H1599, May 2010, doi: 10.1152/ajpheart.00036.2010.
- [32] I. Nederend, M. R. M. Jongbloed, E. J. C. De Geus, N. A. Blom, and A. D. J. Ten Harkel, "Postnatal Cardiac Autonomic Nervous Control in Pediatric Congenital Heart Disease," *J. Cardiovasc. Dev. Dis.*, vol. 3, no. 2, Art. no. 2, Jun. 2016, doi: 10.3390/jcdd3020016.
- [33] F. Yasuma and J. Hayano, "Respiratory Sinus Arrhythmia: Why Does the Heartbeat Synchronize With Respiratory Rhythm?," *CHEST*, vol. 125, no. 2, pp. 683–690, Feb. 2004, doi: 10.1378/chest.125.2.683.
- [34] J. A. Hirsch and B. Bishop, "Respiratory sinus arrhythmia in humans: how breathing pattern modulates heart rate," *Am. J. Physiol.*, vol. 241, no. 4, pp. H620–629, Oct. 1981, doi: 10.1152/ajpheart.1981.241.4.H620.
- [35] M. A. Russo, D. M. Santarelli, and D. O'Rourke, "The physiological effects of slow breathing in the healthy human," *Breathe*, vol. 13, no. 4, pp. 298–309, Dec. 2017, doi: 10.1183/20734735.009817.
- [36] M. Y. Balban *et al.*, "Brief structured respiration practices enhance mood and reduce physiological arousal," *Cell Rep. Med.*, vol. 4, no. 1, Jan. 2023, doi: 10.1016/j.xcrm.2022.100895.
- [37] S.-H. Li, B.-S. Lin, C.-H. Tsai, C.-T. Yang, and B.-S. Lin, "Design of Wearable Breathing Sound Monitoring System for Real-Time Wheeze Detection," *Sensors*, vol. 17, no. 1, p. 171, Jan. 2017, doi: 10.3390/s17010171.
- [38] K. Storck, M. Karlsson, P. Ask, and D. Loyd, "Heat transfer evaluation of the nasal thermistor technique," *IEEE Trans. Biomed. Eng.*, vol. 43, no. 12, pp. 1187–1191, Dec. 1996, doi: 10.1109/10.544342.

- [39] F. Q. Al-Khalidi, R. Saatchi, D. Burke, H. E. Elphick, and S. Tan, "Respiration rate monitoring methods: a review," *Pediatr. Pulmonol.*, vol. 46, no. 6, Jun. 2011, doi: 10.1002/ppul.21416.
- [40] C. Massaroni, A. Nicolò, D. Presti, M. Sacchetti, S. Silvestri, and E. Schena, "Contact-Based Methods for Measuring Respiratory Rate," *Sensors*, vol. 19, p. 908, Feb. 2019, doi: 10.3390/s19040908.
- [41] M. A. F. Pimentel *et al.*, "Toward a Robust Estimation of Respiratory Rate From Pulse Oximeters," *IEEE Trans Biomed Eng*, vol. 64, no. 8, pp. 1914–1923, Nov. 2016, doi: 10.1109/TBME.2016.2613124.
- [42] P. H. Charlton *et al.*, "Breathing Rate Estimation From the Electrocardiogram and Photoplethysmogram: A Review," *IEEE Rev. Biomed. Eng.*, vol. 11, pp. 2–20, 2018, doi: 10.1109/RBME.2017.2763681.
- [43] E. F. Greneker, "Radar sensing of heartbeat and respiration at a distance with applications of the technology," in *Radar 97 (Conf. Publ. No. 449)*, Oct. 1997, pp. 150–154. doi: 10.1049/cp:19971650.
- [44] Z. Zhu, J. Fei, and I. Pavlidis, "Tracking human breath in infrared imaging," in *Fifth IEEE Symposium on Bioinformatics and Bioengineering (BIBE'05)*, Oct. 2005, pp. 227–231. doi: 10.1109/BIBE.2005.55.
- [45] A. A. Kamshilin and N. B. Margaryants, "Origin of Photoplethysmographic Waveform at Green Light," *Phys. Procedia*, vol. 86, pp. 72–80, Jan. 2017, doi: 10.1016/j.phpro.2017.01.024.
- [46] R. Jolivot, Y. Benezeth, and F. Marzani, "Skin Parameter Map Retrieval from a Dedicated Multispectral Imaging System Applied to Dermatology/Cosmetology," *Int. J. Biomed. Imaging*, vol. 2013, p. 978289, 2013, doi: 10.1155/2013/978289.
- [47] R. R. Anderson and J. A. Parrish, "The Optics of Human Skin," *J. Invest. Dermatol.*, vol. 77, no. 1, pp. 13–19, Jul. 1981, doi: 10.1111/1523-1747.ep12479191.

- [48] Y. Sun and N. Thakor, "Photoplethysmography Revisited: From Contact to Noncontact, From Point to Imaging," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 463–477, Mar. 2016, doi: 10.1109/TBME.2015.2476337.
- [49] P. S. Addison and J. N. Watson, "Secondary transform decoupling of shifted nonstationary signal modulation components: application to photoplethysmography," *Int. J. Wavelets Multiresolution Inf. Process.*, vol. 02, no. 01, pp. 43–57, Mar. 2004, doi: 10.1142/S0219691304000329.
- [50] M. van Gastel, S. Stuijk, and G. de Haan, "Robust respiration detection from remote photoplethysmography," *Biomed. Opt. Express*, vol. 7, no. 12, pp. 4941–4957, 2016, doi: 10.1364/BOE.7.004941.
- [51] L. Stanković, I. Djurović, S. Stanković, M. Simeunović, S. Djukanović, and M. Daković, "Instantaneous frequency in time–frequency analysis: Enhanced concepts and performance of estimation algorithms," *Digit. Signal Process.*, vol. 35, pp. 1–13, Dec. 2014, doi: 10.1016/j.dsp.2014.09.008.
- [52] B. A. Reyes, N. Reljin, and K. H. Chon, "Tracheal Sounds Acquisition Using Smartphones," *Sensors*, vol. 14, no. 8, pp. 13830–13850, Aug. 2014, doi: 10.3390/s140813830.
- [53] T. Oberlin, S. Meignen, and V. Perrier, "Second-Order Synchrosqueezing Transform or Invertible Reassignment? Towards Ideal Time-Frequency Representations," *IEEE Trans. Signal Process.*, vol. 63, no. 5, pp. 1335–1344, Mar. 2015, doi: 10.1109/TSP.2015.2391077.
- [54] S. Meignen, D. H. Pham, and M. Colominas, "On the Use of Short-Time Fourier Transform and Synchrosqueezing-Based Demodulation for the Retrieval of the Modes of Multicomponent Signals," *Signal Process.*, vol. 178, p. 107760, Aug. 2020, doi: 10.1016/j.sigpro.2020.107760.
- [55] O. Poisson, P. Rioual, and M. Meunier, "New signal processing tools applied to power quality analysis," *IEEE Trans. Power Deliv.*, vol. 14, no. 2, pp. 561–566, Apr. 1999, doi: 10.1109/61.754104.

- [56] A. V. Oppenheim and R. W. Schaffer, *Discrete-time Signal Processing*. Prentice Hall, 1999.
- [57] M. Rhif, A. Abbes, I. Farah, B. Martinez, and Y. Sang, “Wavelet Transform Application for/in Non-Stationary Time-Series Analysis: A Review,” *Appl. Sci.*, vol. 9, p. 1347, Mar. 2019, doi: 10.3390/app9071345.
- [58] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*. Academic Press, 2008.
- [59] J. M. Lilly and S. C. Olhede, “Higher-Order Properties of Analytic Wavelets,” *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 146–160, Jan. 2009, doi: 10.1109/TSP.2008.2007607.
- [60] E. A. Martinez-Ríos, R. Bustamante-Bello, S. Navarro-Tuch, and H. Perez-Meana, “Applications of the Generalized Morse Wavelets: A Review,” *IEEE Access*, vol. 11, pp. 667–688, 2023, doi: 10.1109/ACCESS.2022.3232729.
- [61] S. C. Olhede and A. T. Walden, “Generalized Morse wavelets,” *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2661–2670, Nov. 2002, doi: 10.1109/TSP.2002.804066.
- [62] A. Silik, M. Noori, W. Altabey, R. Ghiasi, Z. Wu, and J. Dang, “Evaluation of analytic wavelet parameters effect for data analyses in civil structural health monitoring,” presented at the Conference: 10th International Conference on Structural Health Monitoring of Intelligent Infrastructure, Porto, Portugal, Jun. 2021.
- [63] S. Maes and I. Daubechies, “A Nonlinear Squeezing of the Continuous Wavelet Transform Based on Auditory Nerve Models,” in *Wavelets in Medicine and Biology*, Akram Aldroubi, Michael Unser., CRC Press, 1996, pp. 527–546.
- [64] F. Auger *et al.*, “Time-Frequency Reassignment and Synchrosqueezing: An overview,” *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 32–41, Nov. 2013, doi: 10.1109/MSP.2013.2265316.

- [65] H.-T. Wu, G. F. Lewis, M. I. Davila, I. Daubechies, and S. W. Porges, "Optimizing Estimates of Instantaneous Heart Rate from Pulse Wave Signals with the Synchrosqueezing Transform," *Methods Inf. Med.*, vol. 55, no. 5, pp. 463–472, Oct. 2016, doi: 10.3414/ME16-01-0026.
- [66] T. Oberlin, S. Meignen, and V. Perrier, "The Fourier-based Synchrosqueezing Transform," presented at the ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, May 2014, pp. 315–319. doi: 10.1109/ICASSP.2014.6853609.
- [67] S. Sandoval and P. L. De Leon, "Recasting the (Synchrosqueezed) Short-Time Fourier Transform as an Instantaneous Spectrum," *Entropy*, vol. 24, no. 4, p. 518, Apr. 2022, doi: 10.3390/e24040518.
- [68] D. C. Bowman and J. M. Lees, "The Hilbert-Huang Transform: A High Resolution Spectral Method for Nonlinear and Nonstationary Time Series," *Seismol. Res. Lett.*, vol. 84, no. 6, pp. 1074–1080, Nov. 2013, doi: 10.1785/0220130025.
- [69] N. E. Huang *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. Lond. Ser. Math. Phys. Eng. Sci.*, pp. 903–995, 1998.
- [70] M. A. Motin, C. Kumar Karmakar, D. K. Kumar, and M. Palaniswami, "PPG Derived Respiratory Rate Estimation in Daily living Conditions," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.*, vol. 2020, pp. 2736–2739, Jul. 2020, doi: 10.1109/EMBC44109.2020.9175682.
- [71] A. Waugh and A. Grant, *Ross & Willson Anatomy And Physiology in Health and Illness*, 12th ed. Elsevier, 2014. Accessed: May 08, 2024. [Online]. Available: <http://archive.org/details/RossWillsonAnatomyAndPhysiology>
- [72] J. Lázaro, Y. Nam, E. Gil, P. Laguna, and K. H. Chon, "Smartphone-camera-acquired pulse photoplethysmographic signal for deriving respiratory rate," in *2014 8th Conference of the European Study Group on Cardiovascular*

Oscillations (ESGCO), Trento, Italy, May 2014, pp. 121–122. doi: 10.1109/ESGCO.2014.6847549.

- [73] L.-M. Po, X. Xu, L. Feng, Y. Li, K.-W. Cheung, and C.-H. Cheung, “Frame adaptive ROI for photoplethysmography signal extraction from fingertip video captured by smartphone,” in *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*, Lisbon, Portugal: IEEE, May 2015, pp. 1634–1637. doi: 10.1109/ISCAS.2015.7168963.
- [74] J. Lee, K. Matsumura, K. Yamakoshi, P. Rolfe, S. Tanaka, and T. Yamakoshi, “Comparison between red, green and blue light reflection photoplethysmography for heart rate monitoring during motion,” *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.*, vol. 2013, pp. 1724–1727, 2013, doi: 10.1109/EMBC.2013.6609852.
- [75] Y. Maeda, M. Sekine, and T. Tamura, “The advantages of wearable green reflected photoplethysmography,” *J. Med. Syst.*, vol. 35, no. 5, pp. 829–834, Oct. 2011, doi: 10.1007/s10916-010-9506-z.
- [76] D. Grimaldi, Y. Kurylyak, F. Lamonaca, and A. Nastro, “Photoplethysmography detection by smartphone’s videocamera,” in *Proceedings of the 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems*, Sep. 2011, pp. 488–491. doi: 10.1109/IDAACS.2011.6072801.
- [77] R. B. Lagido, J. Lobo, S. Leite, C. Sousa, L. Ferreira, and J. Silva-Cardoso, “Using the smartphone camera to monitor heart rate and rhythm in heart failure patients,” in *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, Jun. 2014, pp. 556–559. doi: 10.1109/BHI.2014.6864425.
- [78] A. Visvanathan, R. Banerjee, A. Misra, A. D. Choudhury, and A. Pal, “Effects of fingertip orientation and flash location in smartphone Photoplethysmography,” in *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Sep. 2014, pp. 468–471. doi: 10.1109/ICACCI.2014.6968536.

- [79] R. J. Schilling and S. L. Harris, *Fundamentals of digital signal processing using MATLAB*. Nelson, a division of Thomson Canada Limited, 2004. Accessed: Aug. 06, 2024. [Online]. Available: <http://archive.org/details/fundamentalsofdi0000schi>
- [80] J. Lee, B. A. Reyes, D. D. McManus, O. Maitas, and K. H. Chon, "Atrial Fibrillation Detection Using an iPhone 4S," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 1, pp. 203–206, Jan. 2013, doi: 10.1109/TBME.2012.2208112.
- [81] M. Lozano, J. A. Fiz, and R. Jané, "Performance evaluation of the Hilbert-Huang transform for respiratory sound analysis and its application to continuous adventitious sound characterization," *Signal Process.*, vol. 120, pp. 99–116, Mar. 2016, doi: 10.1016/j.sigpro.2015.09.005.
- [82] François Auger, Patrick Flandrin, Paulo Gonçalves, and Olivier Lemoine, "Time-Frequency Toolbox For Use with MATLAB Reference Guide." CNRS (Centre National de la Recherche Scientifique), Oct. 26, 2005.
- [83] K. Sozański, "Signal-to-noise ratio in power electronic digital control circuits," in *2016 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, Sep. 2016, pp. 162–171. doi: 10.1109/SPA.2016.7763606.
- [84] "Visualize scalogram of signals - MATLAB - MathWorks América Latina." Accessed: Aug. 13, 2024. [Online]. Available: <https://la.mathworks.com/help/wavelet/ref/wavelettimefrequencyanalyzer-app.html>
- [85] K. Rani Das, "A Brief Review of Tests for Normality," *Am. J. Theor. Appl. Stat.*, vol. 5, no. 1, p. 5, 2016, doi: 10.11648/j.ajtas.20160501.12.
- [86] R. Parsad, "Non-parametric methods in analysis of experimental data," Nov. 2014.
- [87] J. Lázaro, S. Kontaxis, R. Bailón, P. Laguna, and E. Gil, "Respiratory Rate Derived from Pulse Photoplethysmographic Signal by Pulse Decomposition Analysis," in *2018 40th Annual International Conference of the IEEE*

Engineering in Medicine and Biology Society (EMBC), Jul. 2018, pp. 5282–5285. doi: 10.1109/EMBC.2018.8513188.

Anexo A

En este anexo se muestran algunas de las iRR estimadas de las simulaciones con diferentes técnicas tiempo-frecuencia. En la **Figura A.1** se muestran las iRR estimadas de las simulaciones de la maniobra de cambios abruptos, con un SNR de 10 dB, donde se fijo uno de los parametros y otro parametro de varió.

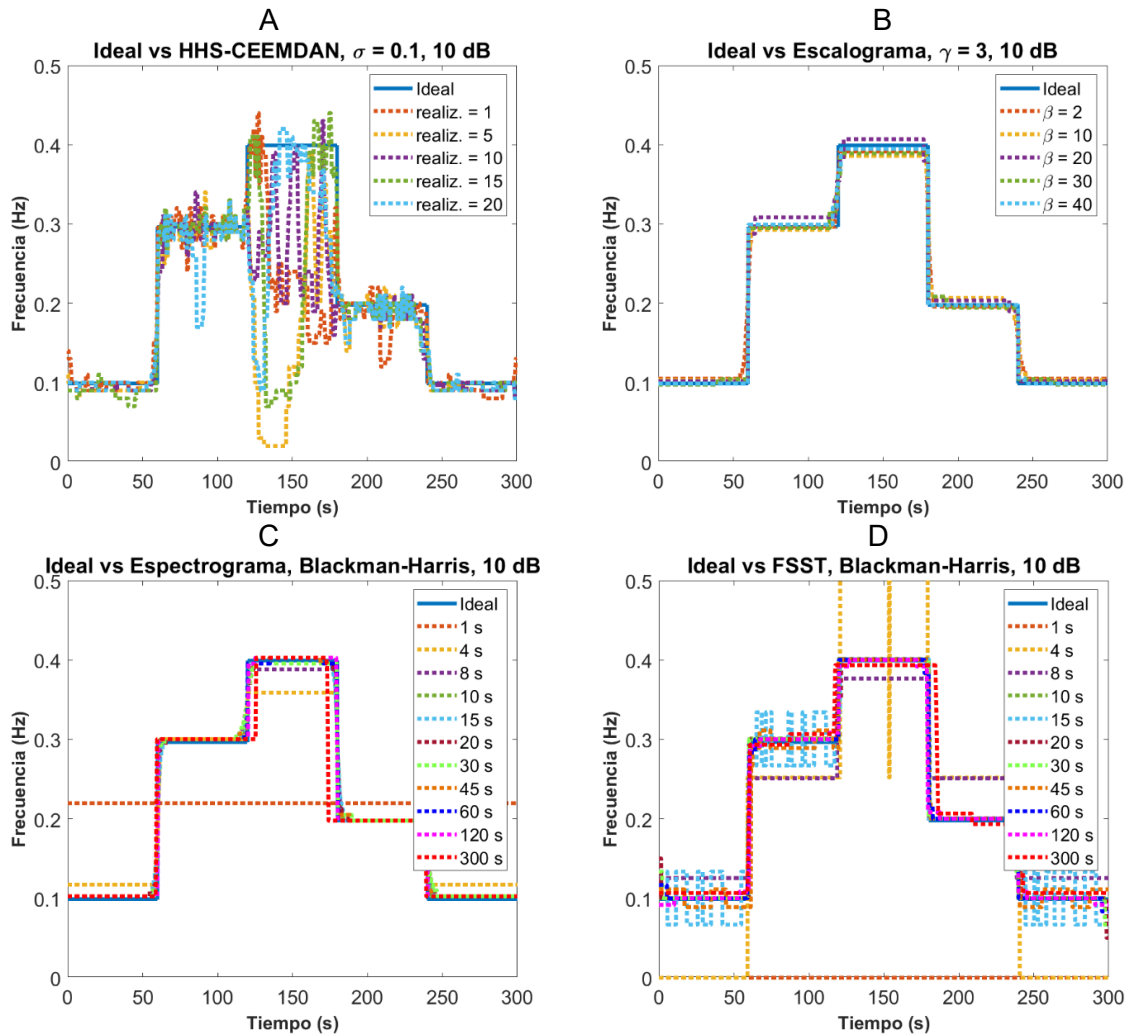


Figura A.1 iRRs de las simulaciones con cambios abruptos con 10 dB. A) Espectro Hilbert-Huang (HHS) vía CEEMDAN. B) Escalograma. C) Espectrograma. D) Transformada *synchrosqueezing* basada en Fourier (FSST).

Lo mismo se hizo en las demas figuras del anexo, donde se varia la maniobra simulada y el SNR. En la **Figura A.2** esta la maniobra constante a 12 bpm con SNR de 20 dB, en la **Figura A.3** se muestra la maniobra despues del ejercicio a con SNR de 30 dB y en la **Figura A.4** esta la maniobra espontanea con SNR de 40 dB.

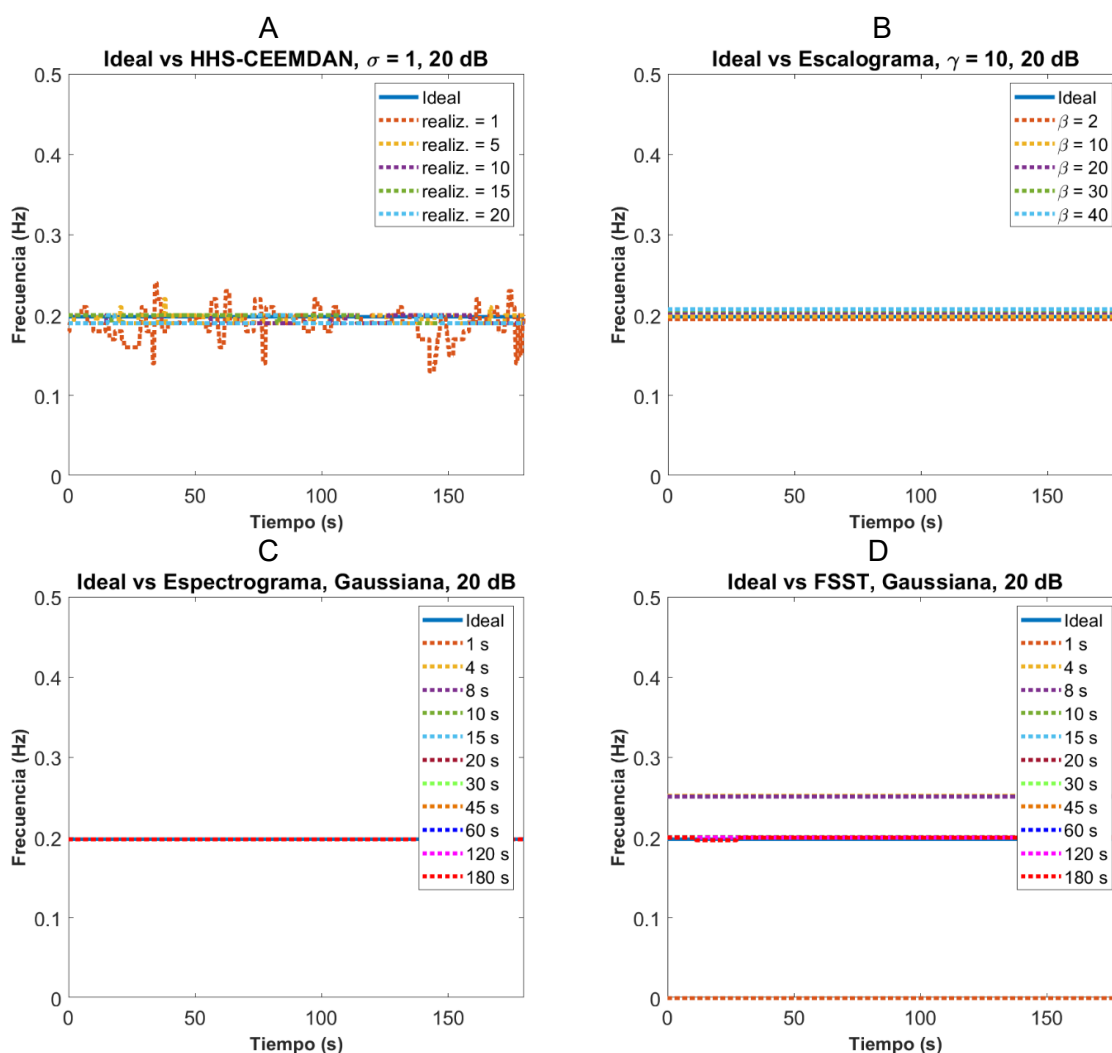


Figura A.2 iRRs de las simulaciones con Constante 12 bpm con 20 dB. A) Espectro Hilbert-Huang (HHS) vía CEEMDAN. B) Escalograma. C) Espectrograma. D) Transformada *synchrosqueezing* basada en Fourier (FSST).

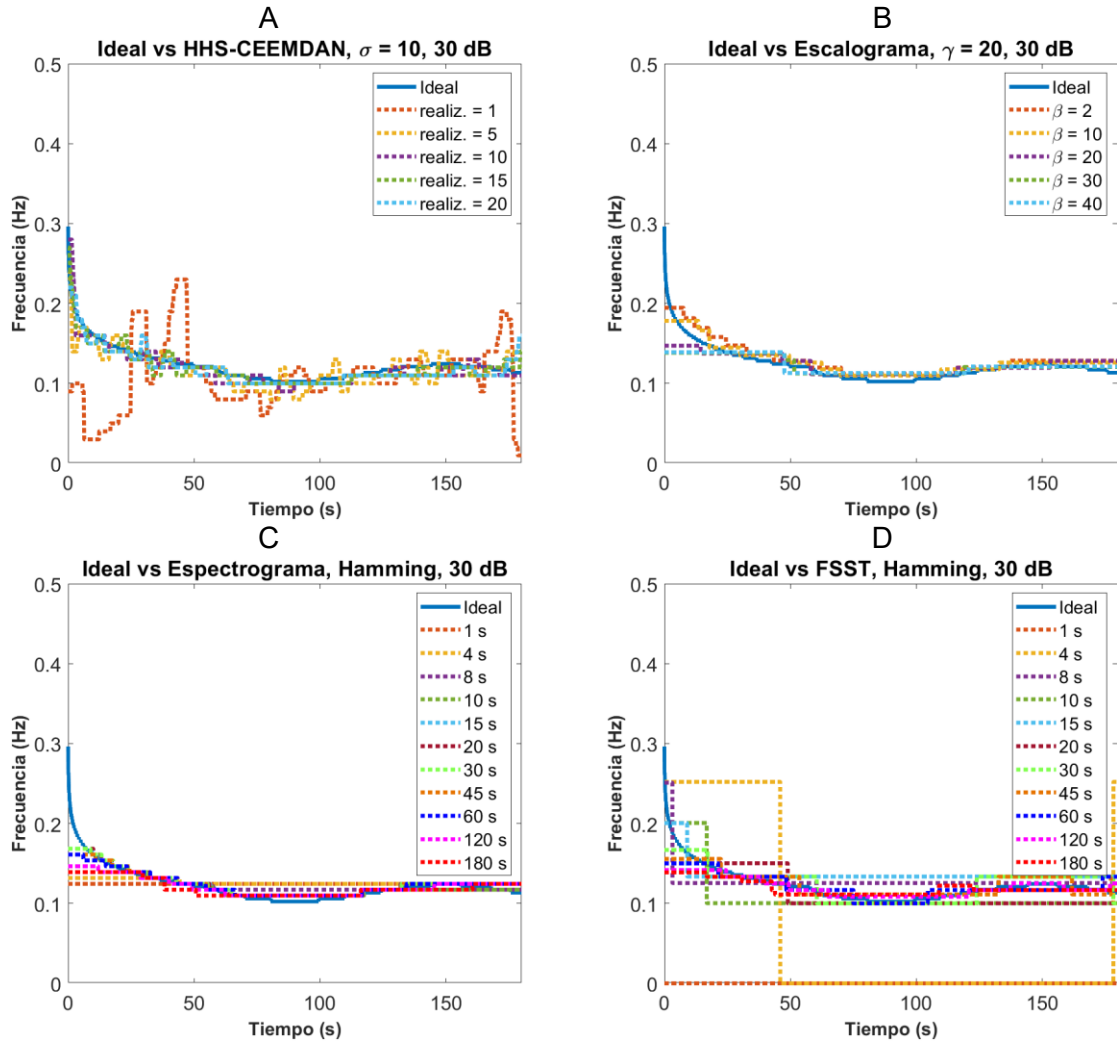


Figura A.3 iRRs de las simulaciones con después del ejercicio con 30 dB. A) Espectro Hilbert-Huang (HHS) vía CEEMDAN. B) Escalograma. C) Espectrograma. D) Transformada *synchrosqueezing* basada en Fourier (FSST).

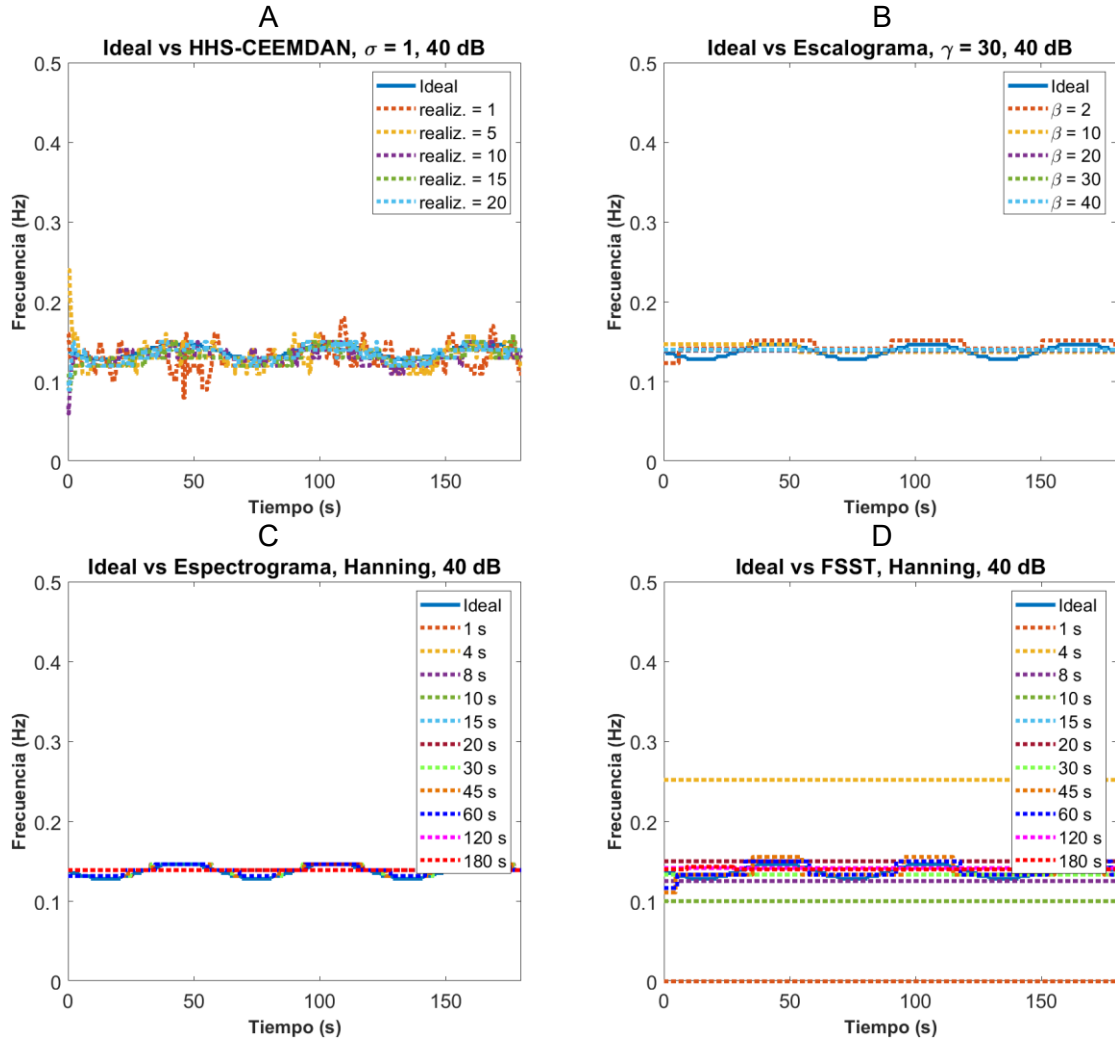


Figura A.4 iRRs de las simulaciones con espontanea con 40 dB. A) Espectro Hilbert-Huang (HHS) vía CEEMDAN. B) Escalograma. C) Espectrograma. D) Transformada *synchrosqueezing* basada en Fourier (FSST).

Anexo B

En el *Congreso Internacional sobre Investigación en STEM y Ciencias Sociales: Desafíos Transdisciplinarios Actuales* se participó con la presentación de un poster científico titulado *Improving the Estimation of Instantaneous Respiratory Rate Obtained via the Spectrogram of Smartphone-based Image Photoplethysmography*. En él se presentaron los resultados del espectrograma de las maniobras con cambios abruptos y la espontánea. A continuación, se anexa la miniatura del póster presentado.

Improving the Estimation of Instantaneous Respiratory Rate Obtained via the Spectrogram of Smartphone-based Image Photoplethysmography



FACULTAD DE
CIENCIAS
UASLP

I.B. Susana Morales Lugo
a275529@alumnos.uaslp.mx

Dr. Bersain A. Reyes
bersain.reyes@uaslp.mx



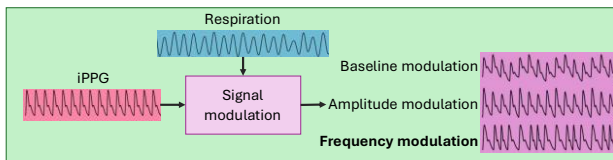
International Congress on STEM research and
social sciences: current transdisciplinary challenges

ABSTRACT

- Respiratory rate is a physiological parameter that reflects health status, but in daily life its measurement has not been permeated to the general population.
- To achieve a ubiquitous, simple and reliable measurement, the time-frequency analysis of smartphone-based image photoplethysmography has been proposed.
- In this study, the spectrogram is tested using synthetic respiratory signals to find which parameters achieve better performance when applied to real acquired data.

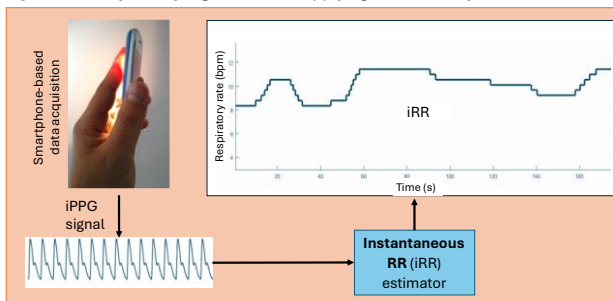
INTRODUCTION

- Abnormal breathing** is indicative of respiratory and cardiorespiratory diseases [1], where **respiratory rate** (RR) has been found helpful.
- RR can be estimated from **image photoplethysmography** (iPPG) by taking advantage of **iPPG signal modulation by the respiration** [2].
- iPPG signals can be conveniently acquired by anyone using **smartphones**.
- Time-frequency techniques, e.g., the **spectrogram** (SP), **allow** to extract the frequency modulation, and to **estimate instantaneous RR (iRR)** [3].



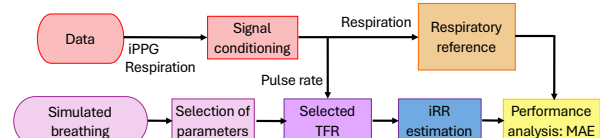
GENERAL AIM

To improve the **estimation of RR time series** from iPPG signals, acquired using a smartphone, by **quantitative selecting the parameters** of the SP via **synthetic respiratory signals** before applying it to **real acquired data**.



METHODOLOGY

- Real data** of respiratory and iPPG signals from $N = 9$ volunteers [4].
- Two synthetic respiratory signals were simulated:** 1) abrupt changes in breathing rate, and 2) spontaneous breathing.
- Gaussian noise was added** to the signals using signal-to-noise ratio (SNR) values of 10, 20, 30 and 40 dB, to emulate real conditions.

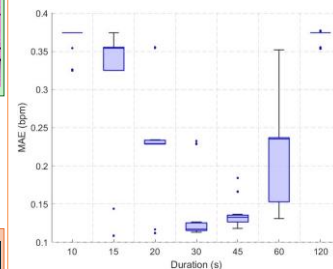


- Simulated data was **processed via the SP**, where **two parameter were varied:** 1) type, and 2) duration of the window.

Window's duration: 1, 4, 8, 10, 15, 20, 30, 45, 60, 120, and 300 (180) s. **Window's types:** Gaussian, Blackman-Harris, Hamming, Hanning, and Bartlett.

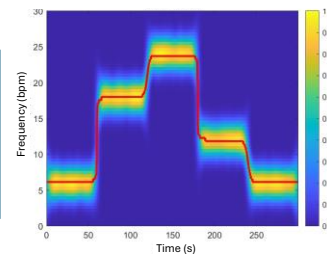
- Best parameters of SP** were selected with a penalized forward-backward greedy algorithm was used.
- The **iPPG-based iRR estimation** were compared to the one from **respiratory reference**, using the **mean absolute error (MAE)**.

RESULTS



Example of results of the SP window selection for the spontaneous breathing simulation. MAE was computed from iRR estimates for all types of windows at different lengths

Example of estimated iRR from SP of smartphone-based iPPG for the abrupt changes maneuver simulation with a SNR of 10 dB, obtained with a 30 s Blackman-Harris window.



iRR estimation results of real acquired data (N=9) volunteers.

Respiratory Maneuver	Window used	Study	iRR estimates	
			MAE (bpm)	SD (bpm)
Abrupt changes	Hamming, 30 s	Previous [4]	3.07	2.37
	Gaussian, 60 s	This one	2.59	1.92
Spontaneous	Hamming, 18 s	Previous [4]	1.93	1.24
	Hanning, 30 s	This one	1.05	0.79

CONCLUSIONS

- Quantitative selection of SP parameters results in better RR estimates.**
- It is worth exploring more advanced time-frequency techniques.**

ACKNOWLEDGMENTS

Study supported by a CONAHCYT graduate scholarship (No. 1240307).



REFERENCES

- [1] T. Ferkol and D. Schraufnager, "The global burden of respiratory disease," *Ann. Am. Thorac. Soc.*, vol. 11, no. 3, pp. 404-406, Mar. 2014.
- [2] H. Sharma, "Extraction of respiratory rate from PPG using ensemble empirical mode decomposition with Kalman filter," *Electron. Lett.*, vol. 56, no. 13, pp. 650-653, Jun. 2020.
- [3] W. Katten, A. Gerd, D. Myers, C. Schaffer, M. Ansermino, and G. Dumont, "Estimation of Respiratory Rate From Photoplethysmographic Imaging Videos Compared to Pulse Oximetry," *IEEE J. Biomed. Health Inform.*, vol. 19, May 2015, doi: 10.1109/JBHI.2015.2429746.
- [4] E. G. Hernández de la Cruz, "Estimation of instantaneous heart and respiratory rates using imaging photoplethysmography methods on a smartphone", Master thesis in Electronics Engineering, UASLP, 2021.