

UNIVERSIDAD PANAMERICANA  
FACULTAD DE INGENIERÍA  
Con estudios incorporados a la  
Secretaría de Educación Pública

**“MONITOREO, SIN CONTACTO, DE LA FRECUENCIA  
RESPIRATORIA EN NEONATOS EN UNIDAD DE CUIDADO  
INTENSIVO MEDIANTE VISIÓN POR COMPUTADORA”**

**TESIS**

QUE PARA OBTENER EL GRADO DE  
**MAESTRÍA EN INGENIERÍA**

P R E S E N T A

**ORLANDO YAEL RIVAS SCOTT**

**ASESOR:**

**DR. JORGE EDUARDO BRIEVA RICO**

**COASESOR:**

**DR. ERNESTO MOYA ALBOR**

# Dedicatoria

Con dilección perenne a María Estela Scott Becerril y a José Javier Rivas Salas, mis dotadores de vida.

A María Estela Becerril Peregrina, mi amada abuelita, quien siempre ha procurado por mí.

A Alejandra Poleth Rivas Scott, mi querida hermana.

A Mario Fernando Rodríguez Palacios, mi excelso amigo.

Y a quien maniere cavilando para coadyuvar a desavezar la saudade.

# Agradecimientos

Al Dr. Jorge Eduardo Brieva Rico, el tutor de esta tesis, quien me ha inspirado y me ha guiado en este camino maravilloso de la investigación. Le agradezco profundamente: su apoyo, sus enseñanzas, sus comentarios y recomendaciones, pero sobre todo el tiempo que ha dedicado en mí para forjar un mejor estudiante, investigador y persona.

Al Dr. Ernesto Moya Albor, el cotutor de esta tesis, por sus comentarios, correcciones y su disponibilidad para resolver mis preguntas.

A la Dra. Sandra L. Gómez Coronel, por el tiempo que dedicó en leer esta tesis, por las aportaciones que me sugirió para corregir y mejorar este documento.

A la Dra. María Victoria Carreras Cruz, por su paciencia, por escuchar mis inquietudes y orientarme.

A la Facultad de Ingeniería por confiar en mí y apoyarme con una beca.

A mi alma máter, la Universidad Panamericana, por permitirme ser parte de su historia, por su compromiso con: el conocimiento, la integridad de las personas y su desarrollo humano. *Ubi spiritus libertas.*

Al Laboratorio LTSI y a la Universidad de Rennes por facilitar la base datos.

A la Mtra. Анна Евгеньевна Бобровник (Anna Evgenevna Bobrovnik) quien siempre estuvo muy atenta al avance y mostró interés durante toda la travesía de la tesis, lo cual fue un aliciente.

A mi madre, María Estela Scott Becerril, por todo su amor sempiterno.

A mi padre, José Javier Rivas Salas, por su apoyo incondicional.

A Dios.

# Resumen

La frecuencia respiratoria en recién nacidos permite conocer el estado de salud de los pacientes. Los monitores clásicos permiten monitorear este signo vital, sin embargo, al tener contacto con el bebé pueden causar lesiones como irritaciones en la piel, por lo cual en este trabajo se propone una metodología para estimar la frecuencia respiratoria sin contacto en condiciones reales de adquisición en neonatos en una unidad de terapia intensiva. El método elaborado se basa en la magnificación de movimiento utilizando la transformada de Hermite y se compara con la estrategia de magnificación utilizando la estrategia de Laplace. Nuestra propuesta, calcula la región de interés de manera automática y permite cambiar parámetros para los diferentes tiempos calmos de manera adaptativa. Se obtuvieron resultados de coincidencia con respecto a la referencia de  $\approx 92\%$  en ambas estrategias.

# Índice general

<b>Dedicatoria</b>	<b>i</b>
<b>Agradecimientos</b>	<b>ii</b>
<b>Resumen</b>	<b>iii</b>
<b>Introducción</b>	<b>1</b>
<b>1 Antecedentes y estado de la cuestión</b>	<b>4</b>
<b>2 Materiales y métodos</b>	<b>11</b>
2.1 Descripción de la base de datos .....	11
2.2 Descripción general del método .....	16
2.2.1 Procesamiento de la señal respiratoria de referencia.....	20
2.2.1.1 Filtrado de la señal de referencia .....	20
2.2.1.2 Extracción de la señal de frecuencia respiratoria de referencia.....	21
2.2.2 Procesamiento del video .....	22
2.2.2.1 Magnificación del video.....	22
2.2.2.2 Extracción de la región de interés.....	26
2.2.2.3 Extracción de la señal respiratoria en la región de interés	35
2.2.2.4 Extracción de la señal de frecuencia respiratoria .....	42
2.3 Experimentos .....	44
2.3.1 Configuración de parámetros.....	44

2.3.2 Métricas de evaluación del método .....	44
<b>3 Resultados</b>	<b>46</b>
<b>4 Discusión</b>	<b>60</b>
<b>5 Conclusiones</b>	<b>64</b>
<b>6 Sugerencias para trabajos futuros</b>	<b>65</b>

# Índice de figuras

2.1	Cuadro del video del sujeto 010087.....	12
2.2	Señal de referencia (Tiempo calmo 4 del sujeto 010087, primeros 60 segundos, 3750 muestras).....	16
2.3	Metodología propuesta para la estimación respiratoria.....	17
2.4	Flujo de datos de la metodología propuesta para la estimación respiratoria.....	18
2.5	Flujo de datos para la extracción de la región de interés .....	19
2.6	Señal de referencia filtrada del sujeto 010087, tiempo calmo 4 (3750 muestras que representan 60 segundos, con $F_{Ck} = 43$ respiraciones por minuto y $w = 5$ ).....	21
2.7	Área de trabajo del sujeto 010087.....	26
2.8	Regiones de interés con magnificación Laplaciana con $\alpha = 15$ y $\nu = 25$ para el sujeto 010087.....	29
2.9	Regiones de interés con magnificación Laplaciana con $\alpha = 20$ y $\nu = 25$ para el sujeto 010087.....	30
2.10	Regiones de interés con magnificación Laplaciana con $\alpha = 30$ y $\nu = 25$ para el sujeto 010087.....	31
2.11	Regiones de interés con magnificación Hermitiana con $\alpha = 15$ y $\nu = 25$ para el sujeto 010087.....	32
2.12	Regiones de interés con magnificación Hermitiana con $\alpha = 20$ y $\nu = 25$ para el sujeto 010087.....	33

2.13 Regiones de interés con magnificación Hermitiana con $\alpha = 30$ y $\nu = 25$ para el sujeto 010087.....	34
2.14 Señal extraída de la región de interés del video magnificado con Hermite del tiempo calmo 4, sujeto 010087 ( $\alpha = 30$ ) .....	35
2.15 Señal extraída del video sin magnificar de la región de interés de encontrada en el video magnificado con Hermite del tiempo calmo 4, sujeto 010087 .....	36
2.16 Señal extraída de la región de interés del video magnificado con Laplace ( $\alpha = 30$ ) del tiempo calmo 4, sujeto 010087 .....	36
2.17 Señal extraída del video sin magnificar de la región de interés de encontrada en el video magnificado con Laplace ( $\alpha = 30$ ) del tiempo calmo 4, sujeto 010087 .....	37
2.18 Minuto 9 de la señal extraída de la región de interés del video magnificado con Hermite ( $\alpha = 30$ ) del tiempo calmo 4, sujeto 010087 .....	37
2.19 Minuto 9 de la señal extraída del video sin magnificar de la región de interés de encontrada en el video magnificado con Hermite ( $\alpha = 30$ ) del tiempo calmo 4, sujeto 010087 .....	38
2.20 Minuto 9 de la señal extraída de la región de interés del video magnificado con Laplace ( $\alpha = 30$ ) del tiempo calmo 4, sujeto 010087 .....	38
2.21 Minuto 9 de la señal extraída del video sin magnificar de la región de interés de encontrada en el video magnificado con Laplace ( $\alpha = 30$ ) del tiempo calmo 4, sujeto 010087 .....	39
2.22 Señal extraída y filtrada de Hermite en el primer minuto .....	40
2.23 Señal extraída y filtrada de Laplace en el primer minuto.....	40
2.24 Señal extraída y filtrada de Hermite en el minuto 10.....	41
2.25 Señal extraída y filtrada de Laplace en el minuto 10.....	41
2.26 Señal de frecuencia respiratoria ( $\tilde{R}$ ) y señal de frecuencia respiratoria de referencia ( $RR$ ) del sujeto 010076 del tiempo calmo 1 .....	43
2.27 Señal de frecuencia respiratoria ( $\tilde{R}$ ) y señal de frecuencia respiratoria de referencia ( $RR$ ) del sujeto 010077 del tiempo calmo 13 .....	43

3.1	Figuras del análisis de Bland-Altman $\alpha = 15$ .....	54
3.2	Figuras del análisis de Bland-Altman $\alpha = 20$ .....	55
3.3	Figuras del análisis de Bland-Altman $\alpha = 30$ .....	56
3.4	Análisis de Bland-Altman para todos los sujetos utilizando la estrategia de Laplace y la estrategia de Hermite .....	59

# Introducción

## Identificación del problema de estudio

La medición de la frecuencia respiratoria es un signo vital de suma importancia debido a que permite hacer un seguimiento de la salud del paciente (Liu et al., 2019), en los neonatos y en recién nacidos prematuros es un signo vital importante para el seguimiento de su salud durante los primeros días.

Los monitores de signos vitales por contacto son intrusivos, por lo que pueden causar lesiones, irritaciones y molestar el estado del bebé en períodos continuos de monitoreo. También los neonatos en ocasiones, durante largos períodos de adquisición de la frecuencia respiratoria llegan a desconectar los electrodos, lo que inhabilita el monitoreo o se enredan con los cables, lo cual puede ser peligroso para el recién nacido.

## Justificación

Los métodos sin contacto toman importancia al ser una alternativa para el monitoreo de signos vitales que evita: molestar al sujeto, tener problemas con los cables y conectar los electrodos. Además, la adquisición de la señal respiratoria puede ser prolongada, evitando así la interrupción del monitoreo.

En este trabajo proponemos un algoritmo de estimación de la frecuencia respiratoria a partir de videos adquiridos en condiciones reales en unidades de cuidados intensivos. Nos basamos en la técnica de magnificación de movimiento de los videos en una región de interés, la cual es seleccionada de manera automática a partir de la optimización de una métrica de error con respecto a una medida de referencia de frecuencia respiratoria por contacto

## **Hipótesis**

La frecuencia respiratoria en neonatos que se encuentran en unidades de cuidado intensivo puede ser monitoreada mediante técnicas de procesamiento de video y su estimación es comparable a las medidas suministradas por los sensores clásicos de contacto.

## **Objetivos generales y específicos**

### **Objetivos generales**

Estimar mediante visión por computadora la frecuencia respiratoria durante periodos de sueños calmos en neonatos en unidades de cuidado intensivo, para monitorear el signo vital de manera no intrusiva.

### **Objetivos específicos**

- Segmentar los tiempos de sueño calmo en la totalidad del video para establecer una correspondencia temporal entre la señal de referencia respiratoria y el video, dado que las frecuencias de muestreo de las dos fuentes de información son diferentes.
- Implementar una estrategia de solución algorítmica mediante procesamiento de imágenes para extraer la señal respiratoria a partir del video.
- Diseñar una estrategia de selección de la región de interés en el video para extraer la señal respiratoria.
- Evaluar el método de estimación con respecto a las medidas de referencia obtenidas con sensores de contacto para verificar su confiabilidad.

## **Resumen de las secciones de este documento**

En el capítulo 1 se muestra el estado del arte y se exponen diferentes soluciones al monitoreo de signos vitales en neonatos sin contacto.

En el capítulo 2 se presenta el método propuesto en este trabajo incluyendo la descripción de la base de datos, la metodología algorítmica y las métricas que se utilizan para la validación del método.

El capítulo 3 describe los resultados obtenidos a partir de la metodología presentada.

En el capítulo 4 se detallan algunos datos importantes de las experimentaciones realizadas, se analizan los resultados, las parametrizaciones elegidas y se comparan las metodologías de diversos autores. Finalmente, se abordan las conclusiones y el trabajo futuro.

# Capítulo 1

## Antecedentes y estado de la cuestión

A través los años se han propuesto diversas soluciones para el monitoreo de signos vitales en neonatos de forma no intrusiva. Gracias a estos desarrollos ha sido posible estimar la frecuencia respiratoria (Alinovi et al., 2018), la oxigenación y la frecuencia cardíaca. También se ha permitido identificar apneas (Geertsema et al., 2020), epilepsias (Ntonfo, 2014), convulsiones (Cattani et al., 2017) y estados del sueño (Aoki et al., 2005). Algunos de estos métodos comparten procesamientos similares por lo cual, es de interés conocer las estrategias existentes en esta diversidad de aplicaciones.

La frecuencia respiratoria es un parámetro fisiológico de gran importancia, puesto que si presenta anomalías es un buen indicador de enfermedades graves (Liu et al., 2019). Por ejemplo, mediante el monitoreo de la frecuencia respiratoria es posible detectar trastornos del sistema corporal que provoquen hipoxemia o hipercapnia (Liu et al., 2019).

Los dispositivos que monitorean la frecuencia respiratoria convencionales contienen artefactos (electrodos, cables) que pueden ser movidos por los recién nacidos o por circunstancias ajenas al monitor lo que puede causar falsas alertas. Por otro lado, aunque las mediciones sin contacto son estimadas a partir de un video, el cual registra los movimientos, es posible identificar los movimientos propios del signo vital de aquellos que son ajenos al estudio de interés (Tarassenko et al., 2014).

La medición de los signos vitales en recién nacidos, en las unidades de cuidados intensivos se dificulta si el paciente tiene una condición patológica, por la cual es

imposible el contacto con el sistema de medición. Puede ser que el monitor de signos vitales le cause estrés, dolor, daño a causa de que la piel del recién nacido sea frágil (Villarreal et al., 2014). Por ello, es importante desarrollar sistemas que no tengan contacto con el bebé.

A continuación, describiremos algunos trabajos de monitoreo a distancia para diferentes aplicaciones.

- **Apneas**

“Un episodio de apnea se define generalmente como un cese de la respiración durante 20 o más segundos de pausa o una pausa más corta acompañada de bradicardia (<100 latidos por minuto)” cianosis o palidez (Eichenwald, 2016).

Para la detección de apneas existen varias soluciones, una de ellas es mediante el procesamiento de la señal de video, la detección de los eventos y la clasificación. A través de la evaluación del flujo óptico, las velocidades en grupo y el espectro de frecuencia de tiempo de Gabor (Geertsema et al., 2020) es una propuesta que nos permite la detección de apneas, debido a que realiza una modulación multiescalar sigmoidea, estandariza las características y clasifica utilizando una máquina de soporte vectorial.

“La apnea prematura es uno de los diagnósticos más frecuentes en las unidades de cuidados intensivos de recién nacidos.” Para su identificación es necesario el monitoreo de la frecuencia respiratoria, la oxigenación y la frecuencia cardíaca (Eichenwald, 2016), es por ello que toma importancia la supervisión de los signos vitales, en el caso de este trabajo el de la frecuencia respiratoria específicamente.

- **Convulsiones**

Para la detección de convulsiones, el procesamiento de la imagen mediante la erosión y la selección de una máscara cuadrada permite visualizar mejor los bordes y elimina el ruido. Aunque este procedimiento tiene pérdidas en la señal, ésta es despreciable. Se obtiene el promedio de luminancia de la señal y el promedio de

movimiento de señal la cual se compara con los resultados del electroencefalograma (EEG) (Ntonfo et al., 2012).

- **Frecuencia Cardíaca**

Una estrategia es calcular la intensidad promedio de una región de interés, utilizando el Análisis de Componentes Independientes (ACI) de los tres canales de color para obtener para la forma de onda de una señal de fotopletimografía de reflectancia (PPG) para realizar un modelo de coeficientes autorregresivos. Se realiza un filtrado en el dominio de la frecuencia para delimitar el rango de interés (Villarreal et al., 2014).

- **Frecuencia Respiratoria**

Cuando los bebés duermen presentan una frecuencia respiratoria menor comparada con la frecuencia en estado de vigilia, a su vez la frecuencia respiratoria es más compleja de estimar en este estado debido a que los bebés se mueven y pueden causar ruido en los sistemas de medición (Morley et al., 1990).

Tarassenko et al. presentan una estrategia para la eliminación de ruido en la grabación debido a la iluminación artificial. Los autores proponen un método para cancelar las frecuencias que se generan por la iluminación, usando un modelo de autorregresivo y de la cancelación de polos (Tarassenko et al., 2014).

Lukac et al. proponen una solución mediante la implementación del método de flujo óptico de Lucas-Kanade para obtener las velocidades de los movimientos del tórax. Además, calculan una relación de la señal a ruido de los píxeles para un seguimiento en bloques y las velocidades de varias áreas son consideradas y promediadas (Lukac et al., 2014).

Se ha implementado la transformada de Fourier para delimitar las frecuencias que se quieren detectar del movimiento torácico. Por otro lado, se ha utilizado cámaras que son capaces de detectar la luz cerca de los valores de infrarrojo con el fin de mejorar la adquisición de la señal y clasificar mediante después de elegir una máscara, filtrar, y segmentar los periodos que corresponden a la frecuencia respiratoria y a los que son propios de movimientos (Bartula et al., 2013).

Mediante el desarrollo de una interfaz gráfica de usuario y una cámara de video de alta velocidad (webcam) (Ghassemlooy et al., 2010) proponen una solución para estimar la velocidad mediante la diferencia entre imágenes consecutivas, e interpretando el movimiento de la magnitud en cierto tiempo, con ello logran obtener representaciones de la inhalación y la exhalación. Para reducir el ruido utilizan un filtro de suavizado con  $m = 2$ . Y para obtener la frecuencia respiratoria consideran el cambio en el gradiente de la señal obtenida (Ghassemlooy et al., 2010).

- **Magnificación**

Se ha utilizado la magnificación Euleriana en el video para amplificar los movimientos propios de la respiración, con lo cual es más fácil identificar los pequeños cambios de color en la piel, y evaluar los píxeles RGB (rojos, verdes y azules, por las siglas en inglés de Red, Green y Blue) en un área de interés en cada imagen del video, de esta manera se pueden obtener los picos de la señal y la densidad espectral (Antognoli et al., 2018). (Cattani et al., 2017) han propuesto la magnificación Euleriana aunada a la conversión de los píxeles a escala de grises, aplicando un filtro de respuesta finita al pulso y restando imágenes consecutivas para evaluar la presencia de un componente sinusoidal que les permita estimar la frecuencia.

La propuesta de Alinovi et al. analiza regiones de interés específicas y aplica la magnificación Euleriana en las zonas para resaltar los pequeños cambios que son invisibles al ojo humano, este método consta en una descomposición de cuadros multiescala espacial y un procesamiento temporal. Se crea una pirámide Laplaciana, “que divide cada cuadro del video en imágenes escaladas que representan diferentes bandas de frecuencias espaciales. Los niveles de la pirámide Laplaciana se ordenan desde un nivel más bajo a un nivel más alto” (Alinovi et al., 2015), se aplica un filtro temporal Butterworth y una amplificación de los movimientos

Koolen et al. se utilizan el flujo óptico, el ACI (definido en el apartado de la Frecuencia Cardíaca) o el Análisis de Componente Principal (ACP) para la detección y se comparó la correlación con la señal de referencia correspondiente (Koolen et al., 2015).

Otro enfoque de magnificación es el basado en las fases de movimiento, el cual magnifica solo las áreas de interés y filtros temporales. Para mantener baja la complejidad computacional se utiliza el filtro paso banda de Butterwoth de orden bajo. De esta manera es posible la señal respiratoria en función del tiempo (Alinovi et al., 2018).

En la Tabla 1.1 se realizó una síntesis de los trabajos presentados anteriormente. La tabla contiene los autores, la aplicación, la solución propuesta, la referencia utilizada, la cantidad de sujetos de estudio, el tipo de paciente adulto o neonato y la forma en se evaluó el desempeño de las propuestas.

Si bien los autores presentan resultados con base en los sujetos utilizados en sus experimentos, no existe una referencia común o videos universales con los que se pueda trabajar para comparar de una manera más objetiva los resultados presentados por los investigadores. Aunado a ello, los tiempos calmos son seleccionados por los autores, lo cual genera que el parámetro sea un tanto subjetivo y debido a ello, es posible que los resultados obtenidos por las diferentes estrategias tengan variaciones significativas en los resultados si se cambia de sujeto o adquisición.

Los algoritmos aún presentan complicaciones en el reconocimiento de inhalaciones y exhalaciones cuando reconstruyen la respiración, por lo que se han utilizado heurísticas, pero esto conlleva a fuertes requerimientos de procesamiento (Bartula et al., 2013).

En este trabajo se propone un algoritmo de magnificación utilizando la transformada de Hermite y de Laplace. Además, se presenta una estrategia de optimización para la selección automática de la región de interés. La estimación de la frecuencia respiratoria se realiza mediante el conteo de picos de la señal y se compara con una señal de referencia por contacto. El estudio se hace en tres bebés considerando únicamente los tiempos calmos de sueño. La evaluación del método se realiza utilizando métricas para comparar la estimación de la frecuencia respiratoria con respecto a la señal de referencia.

Tabla 1.1: Comparación de soluciones

Publicación	Aplicación	No. datos	Tipo de paciente	Evaluación	Referencia
(Geertsema et al., 2020)	Apnea	10	-	Sensibilidad y precisión	Datos verdaderos fundamentales
(Cabon et al., 2019a)	Clasificación de los estados del sueño	10	Neonatal	Precisión media	Análisis manual del video
(Sun et al., 2019)	Identificación de momentos incómodos	11	Infantes prematuros	RSME, auto-correlación	-
(Antognoli et al., 2018)	Frecuencia respiratoria	7	Neonatal	Correlación	Monitor estándar
(Alinovi et al., 2018)	Frecuencia respiratoria	6	Adultos e infantes	RSME	Acelerómetro y neumógrafo
(Cattani et al., 2017)	Apneas y convulsiones	-	Neonatal	Criterio de máxima verosimilitud, sensibilidad, especificidad	Neumógrafo
(Koolen et al., 2015)	Frecuencia respiratoria	7	Neonatal	Correlación	Señal de control
(Villaruel et al., 2014)	Clasificación de estados del sueño	30	Neonatal	MAPE	ECG
(Pisani et al., 2014)	Convulsiones	12	Neonatal	Sensibilidad, AUC y especificidad	EEG
(Ntonfo, 2014)	Epilepsia	-	Neonatal		EEG
(Lukac et al., 2014)	Frecuencia respiratoria	-	-	Velocidad de movimiento respiratorio	Inspección visual media
(Bartula et al., 2013)	Frecuencia respiratoria	5	Neonatal	Correlación	Pletismografía
(Ntonfo et al., 2012)	Convulsiones	-	Neonatal	Autocorrelación-Yin y sensibilidad y especificidad	EEG
(Alinovi et al., 2015)	Frecuencia respiratoria	-	Adultos y neonatal	Límites de tolerancia	Neumógrafo y cinta torácica

AUC = Área bajo la curva ROC, EEG = Electroencefalograma, ECG = Electrocardiograma, MAPE = Error absoluto porcentual medio, RSME = Raíz del error cuadrático medio

Tabla 1.1: Comparación de soluciones (continuación).

<b>Publicación</b>	<b>Aplicación</b>	<b>No. da- tos</b>	<b>Tipo de paciente</b>	<b>Evaluación</b>	<b>Referencia</b>
(Kouamou et al., 2011)	Convulsiones	-	Neonatal	Luminosidad diferencial media	EEG
(Ghassemlooy et al., 2010)	Frecuencia respiratoria	1	Adulto	-	Sistemas de monitoreo de respiración por contacto
(Tarassenko et al., 2014)	Frecuencia cardíaca, respiratoria y saturación de oxígeno	-	Adultos	Desviación estándar, correlación	Sensor de expansión torácica
(Zhao et al., 2013)	Frecuencia respiratoria y pulso cardíaco	15	Adultos	Desviación estándar, confianza, RSME, correlación	ECG, sensor de difluoruro de polivinilideno
(Janssen et al., 2016)	Frecuencia respiratoria	4,2	Adultos y neonatal	Correlación	ECG
(Aoki et al., 2005)	Estatus respiratorio	-	-	Correlación	Espirómetro

AUC = Área bajo la curva ROC, EEG = Electroencefalograma, ECG = Electrocardiograma,  
MAPE = Error absoluto porcentual medio, RSME = Raíz del error cuadrático medio

# Capítulo 2

## Materiales y métodos

En esta sección se describe la base de datos utilizada, el método propuesto y la estrategia de evaluación.

### 2.1 Descripción de la base de datos

En este trabajo, se utilizó una parte de la base de datos presentada en (Cabon et al., 2019b) y facilitada por el Laboratorio LTSI de la Universidad de Rennes. Los videos fueron adquiridos durante un proyecto realizado en el Hospital Universitario de Rennes, aprobado por el Comité de Protección de Individuos CPP Ouest 6-598 y cumpliendo con los estándares establecidos por la Declaración de Helsinki. En este estudio se incluyeron diez recién nacidos y se obtuvo la firma de un consentimiento informado por parte de los padres de cada uno de ellos. Durante los experimentos, se instaló una cámara en la habitación de los bebés para registrar la escena. Se instaló cerca de la cama para poder observar la mayor parte del cuerpo. Las grabaciones se realizaron en una oscuridad moderada. La cámara tenía una resolución de 640x480 píxeles y registró 25 fotogramas por segundo. Se pueden notar diferencias en las condiciones de luminosidad y distancias de cámara. Un experto en el Programa de Evaluación y Cuidado Individualizado del Desarrollo del Recién Nacido (NIDCAP, por sus siglas en inglés) llevó a cabo una puntuación de las etapas del sueño, basada en una observación conductual directa (Cabon et al., 2019b), sincrónicamente durante la grabación, con-

siderando cinco etapas: Sueño Calmo, Sueño activo, Somnolencia, Alerta Silenciosa y Alerta Activa. Además de los videos de los bebés (AV), la base de datos contiene las señales respiratorias de referencia (AR) dada por un monitor clásico de signos vitales adquirida con una frecuencia de muestreo de 62.5 muestras por segundo. Tanto los videos como las señales respiratorias se adquirieron de manera síncrona.

En este trabajo se utilizaron 3 sujetos (el 010076, el 010077 y el 010087) con un total de 21 horas 59 minutos y 16 segundos de grabación. Los sujetos contienen 479 minutos con 44 segundos, 329 minutos con 49 segundos y 509 minutos con 43 segundos de grabación respectivamente. Los sujetos fueron seleccionados ya que, contenían el mayor número de videos con datos útiles, es decir donde: la mayoría del tiempo el bebé permanecía en la incubadora y no era removido o movido por alguna persona y donde existían las señales análogas de referencia y de video. Para este trabajo, solamente se tuvieron en cuenta los períodos de sueño calmo de los tres bebés.

En la Figura 2.1 se puede apreciar un cuadro del video del sujeto 010087 del tiempo calmo 4 adquirido por la cámara en la unidad de cuidados intensivos. En la parte inferior izquierda se ven los cables del monitor de signos vitales clásico que llegan hasta el neonato, quien se encuentra en el centro de la imagen acostado con cobijas en una incubadora.



Figura 2.1: Cuadro del video del sujeto 010087

### Preprocesamiento de la base de datos

Dadas las anotaciones de las bases de datos, se seleccionan de los Sueños Calmos las partes de los videos en dónde el recién nacido permanece quieto, sin movimientos abruptos (por lo general este comportamiento es recurrente cuando los bebés duermen (Morley et al., 1990)). Realizamos esta selección de los datos debido a que, en ocasiones, el bebé se mueve demasiado lo que causa mucho ruido o por el simple hecho de que el recién nacido es removido de la incubadora.

Una vez seleccionadas las partes donde el bebé está tranquilo se tomaron sólo aquellos donde la duración del video fuese igual o mayor a 60 segundos, esto con el fin de tener datos suficientes para poder procesarlos y analizarlos, a los cuales les llamamos tiempos calmos, cada sujeto tiene una cantidad  $N$  de ellos.

Creamos la variable  $V$  la cual almacena los videos  $V_k$  que son las partes seleccionadas (tiempos calmos) de los videos de la base de datos ( $AV$ ) (definida en la Sección 2.1), cada una de ellas tienen una duración de interés ( $t_k$ ) para el tiempo calmo número  $k$ . También extrajimos de la señal  $AR$  (definida en la Sección 2.1) los datos análogos a cada  $V_k$  y los llamamos  $R_k$  con el fin de tener parejas de señales y poder comparar los resultados venideramente.

En la Tabla 2.1 mostramos el número de Sueños Calmos y de tiempos calmos ( $N$ ) de los sujetos utilizados y en las Tablas 2.2, 2.3 y 2.4 con la duración de cada tiempo calmo para el sujeto 010076, 010077 y 010087 respectivamente.

Tabla 2.1: Número de sueños calmos en la base de datos y sueños calmos seleccionados (tiempos calmos) de los sujetos

Sujeto	Sueños Calmos en la base de datos	Sueños Calmos seleccionados
010076	45	18
010077	52	11
010087	8	5

Tabla 2.2: Duración de tiempos calmos en el Sujeto 010076

<b>Nombre del tiempo calmo</b>	<b>Duración en segundos</b>
TSC1	64
TSC3	98
TSC4	85
TSC5	89
TSC6	187
TSC8	103
TSC10	67
TSC12	71
TSC14	501
TSC15	125
TSC16	128
TSC19	113
TSC20	60
TSC21	212
TSC25	173
TSC27	129
TSC28	71
TSC29	74
<b>Total</b>	<b>2350</b>

Tabla 2.3: Duración de tiempos calmos en el Sujeto 010077

<b>Nombre del tiempo calmo</b>	<b>Duración en segundos</b>
TSC2	171
TSC3	86
TSC5	220
TSC6	64
TSC7	295
TSC8	437
TSC9	164
TSC10	426
TSC11	293
TSC12	119
TSC13	191
<b>Total</b>	<b>2455</b>

Tabla 2.4: Duración de tiempos calmos en el Sujeto 010087

<b>Nombre del tiempo calmo</b>	<b>Duración en segundos</b>
TSC1	122
TSC2	66
TSC3	373
TSC4	813
TSC5	173
<b>Total</b>	1547

En la Figura 2.2 se puede apreciar la señal de referencia del tiempo calmo número 4 del sujeto 010087. En ella se observa la oscilación de la señal donde cada pico superior corresponde a una inhalación y cada pico inferior a una exhalación. En el eje de las abscisas se encuentran las primeras 3750 muestras que representan 60 segundos a una frecuencia de muestreo de 62.5 muestras por segundo. Por otro lado, el eje de las ordenadas representa la señal en bruto después de la adquisición, en niveles de cuantificación digitalizados.

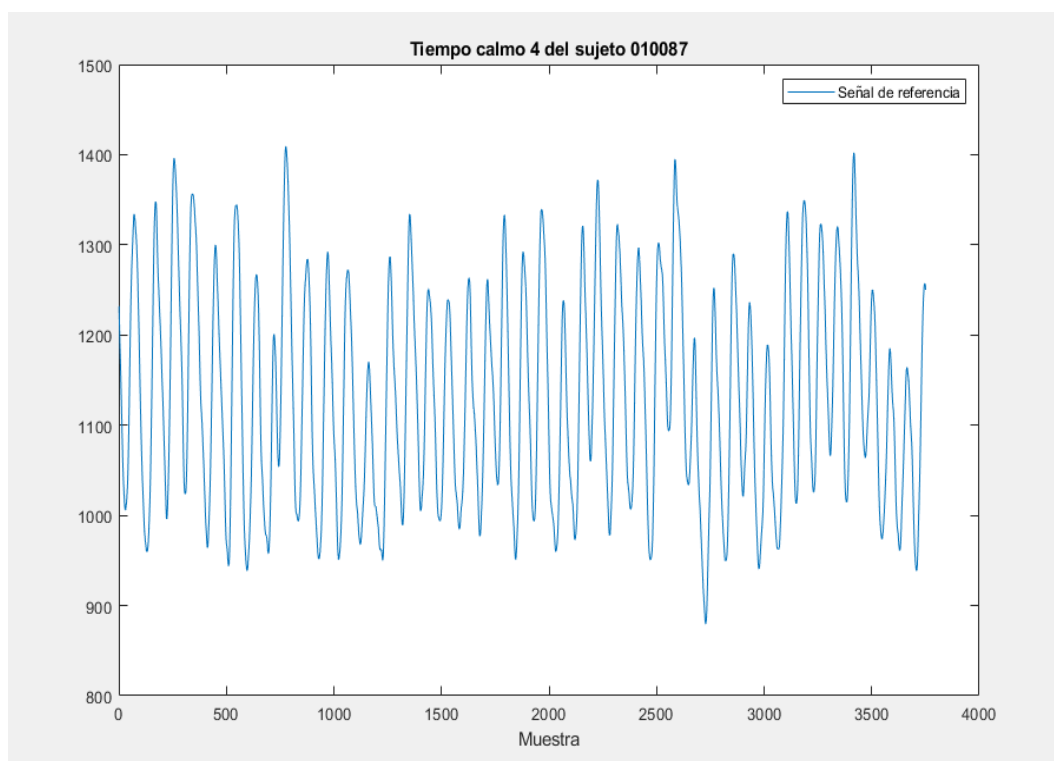


Figura 2.2: Señal de referencia (Tiempo calmo 4 del sujeto 010087, primeros 60 segundos, 3750 muestras)

## 2.2 Descripción general del método

En esta sección se describe la técnica utilizada de procesamiento de imágenes para detectar y estimar la frecuencia respiratoria. Además, evaluarla con respecto a la señal respiratoria de referencia. En la Figura 2.3 se observan los diferentes pasos de la metodología utilizada en este trabajo.

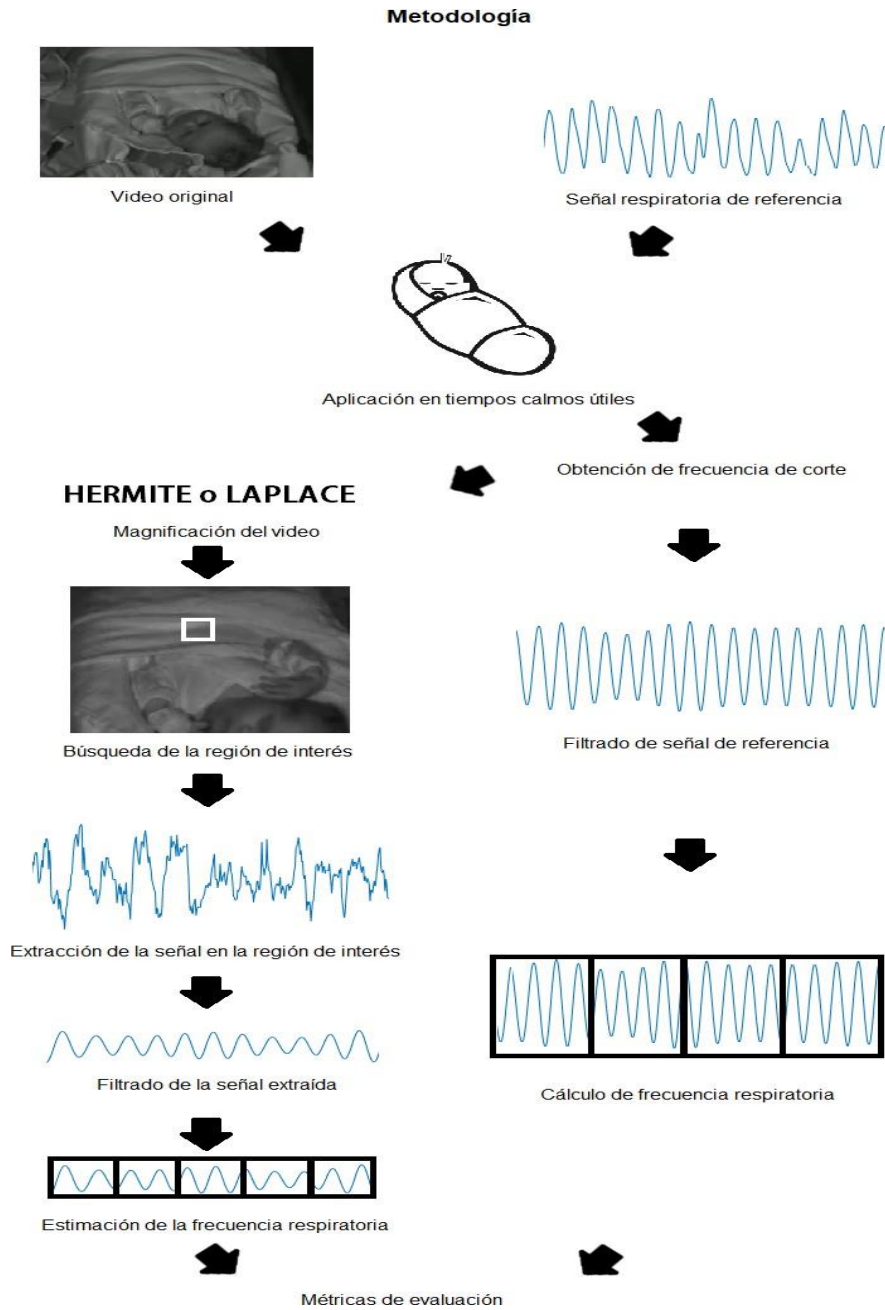


Figura 2.3: Metodología propuesta para la estimación respiratoria

En la Figura 2.4 se muestra el flujo de datos de la metodología. Las conexiones dentro del diagrama de flujo fueron representadas mediante un círculo y un identificador (C1 y C2).

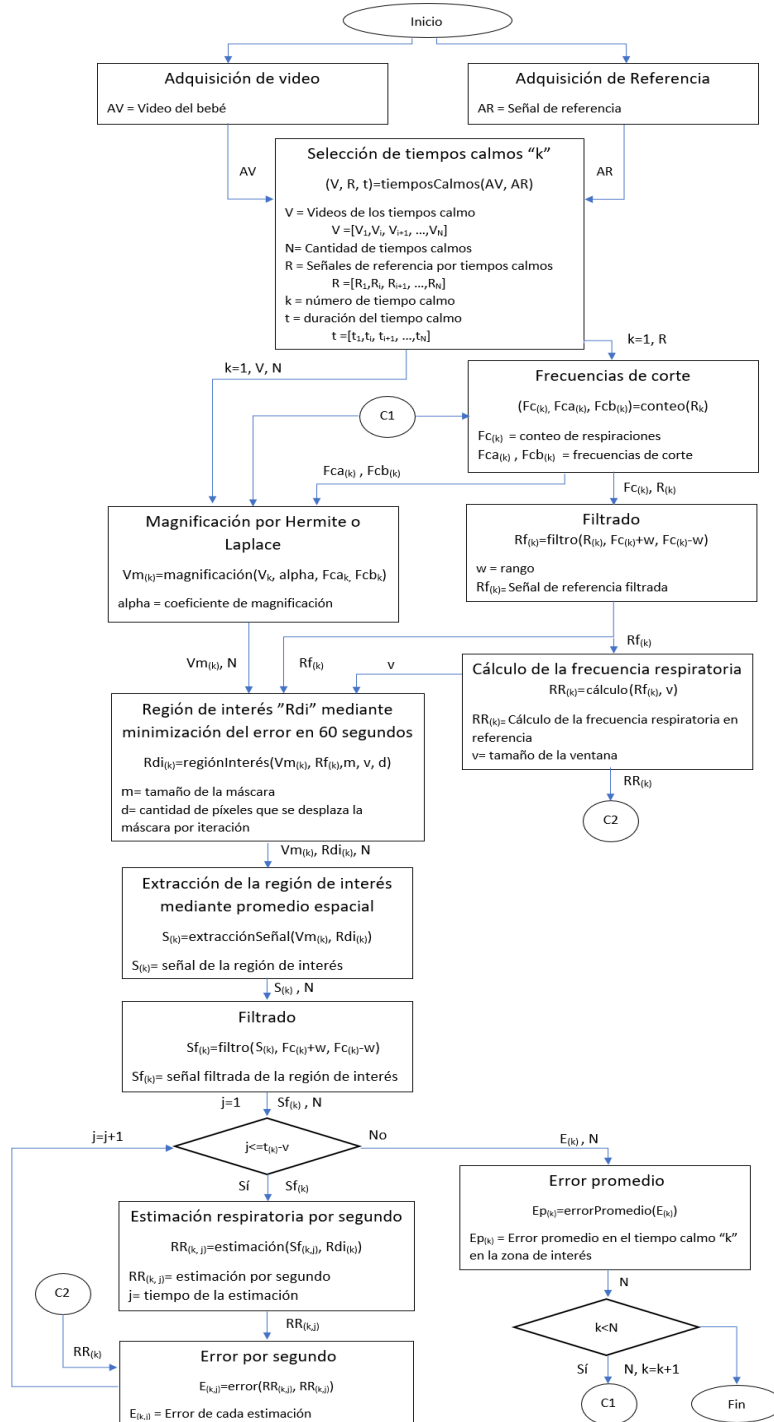


Figura 2.4: Flujo de datos de la metodología propuesta para la estimación respiratoria

En la figura 2.5 se muestra el flujo de datos en la extracción de la región de interés.

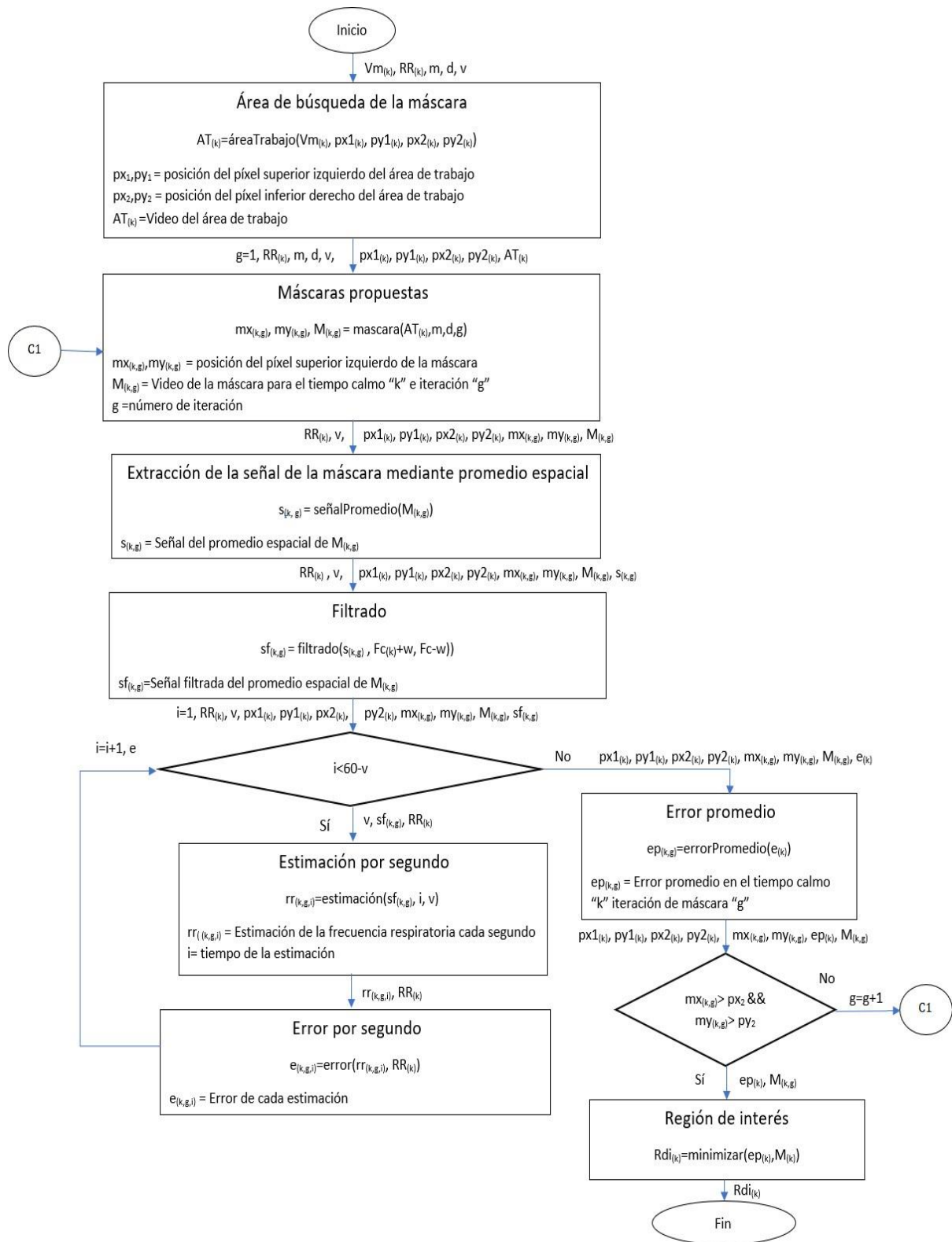


Figura 2.5: Flujo de datos para la extracción de la región de interés

## 2.2.1 Procesamiento de la señal respiratoria de referencia

En esta sección se describe el preprocesamiento realizado a la señal respiratoria para calcular la frecuencia respiratoria que servirá como estándar de oro para evaluar el método. Este procesamiento consta de una etapa de filtrado y de una etapa de estimación de la frecuencia mediante el conteo de picos en una ventana predeterminada.

### 2.2.1.1 Filtrado de la señal de referencia

Antes de realizar el filtrado calculamos las frecuencias de corte que utilizaremos, para ello analizamos los primeros 60 segundos de cada una de las señales de referencia ( $R_k$ ) (definida en la Sección 2.1) y contamos manualmente las respiraciones que pudimos identificar, para obtener la frecuencia central ( $F_{Ck}$ ) es decir, los picos que se presentan en la referencia, ya que sugieren una respiración: cada oscilación superior de la señal; una inhalación, y cada inferior; una exhalación. En ocasiones el conteo de los picos no es trivial, puesto que la señal presenta ruido, es por ello, que se considera un rango para obtener las frecuencias de corte  $F_{Cak}$  y  $F_{Cb_k}$  usadas en la magnificación. Donde  $F_{Cak}$  y  $F_{Cb_k}$  son  $F_{Ck} + a$  y  $F_{Cb_k} - b$  respectivamente, siendo  $a$  y  $b$  variables que definen el rango en cada tiempo calmo.

El filtrado de la referencia se realiza para eliminar ruido de las señales. Se utiliza un filtro paso bandas con las frecuencias de corte que son obtenidas dado el conteo de respiraciones  $F_{Ck} + w$  y  $F_{Ck} - w$  con las cuales obtuvimos la señal  $Rf_k$ . Ocupamos un valor de  $w = 5$ , ya que las señales de referencia presentan un comportamiento bastante claro, es decir, que analizando a simple vista los valores se pueden diferenciar las inhalaciones y las exhalaciones del recién nacido, sin embargo, necesitábamos deshacernos de las señales ajenas al estudio, puesto que en ocasiones se presentan ruidos minúsculos e identificables que fuimos capaces de limpiar con un valor experimental de  $w = 5$ . En la Figura 2.6 se muestran la señal  $Rf_{k=4}$  del sujeto 010087.

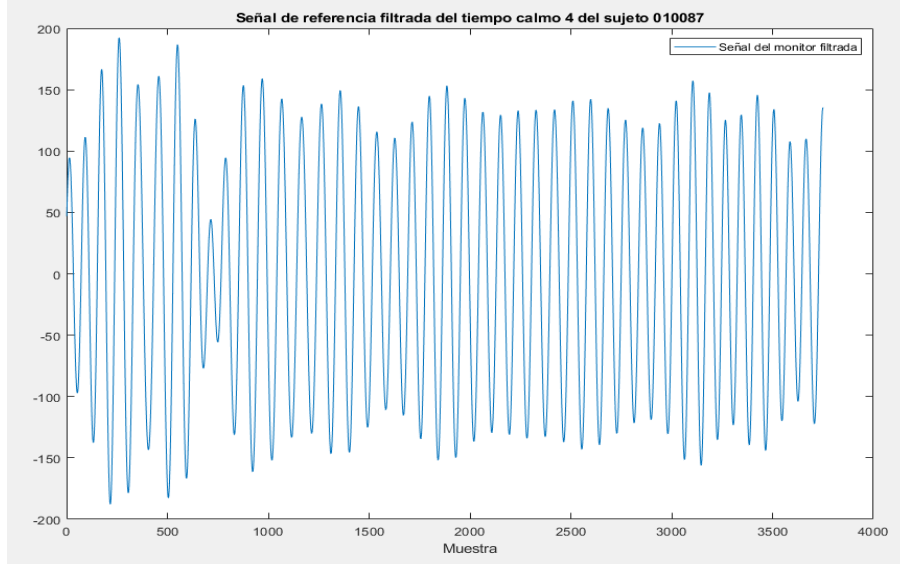


Figura 2.6: Señal de referencia filtrada del sujeto 010087, tiempo calmo 4 (3750 muestras que representan 60 segundos, con  $F_{Ck} = 43$  respiraciones por minuto y  $w = 5$ )

### 2.2.1.2 Extracción de la señal de frecuencia respiratoria de referencia

Utilizamos la señal filtrada de referencia  $Rf_k$  y optamos por tamaños de ventanas ( $v$  que se mencionan en el apartado 2.3.1) las cuales indican la duración de la muestra que se tomará para poder realizar el cálculo de las respiraciones por minuto en cada segundo.

La suputación de la frecuencia respiratoria se obtiene mediante el conteo de ( $\#muestras_k$ ), después lo multiplicamos por el período del muestreo de la referencia ( $\Delta tr = \frac{1}{62.5}$ ) y obtenemos  $Tr_{(k,z)}$ , donde  $z$  representa el número de la estimación y  $k$  el número de tiempo calmo. Posteriormente, obtenemos  $\tilde{T}r_{(k,z)}$  dividiendo  $Tr_{(k,z)}$  entre el número de picos ( $P$ ) en la ventana. Finalmente obtenemos el cálculo de la respiración por minuto en cada segundo ( $RR_{(k,z)}$ ) dividiendo 60 segundos entre ( $\tilde{T}r_{(k,z)}$ ). La señal ( $RR_k$ ) representa las respiraciones por minuto y contiene:  $t_k - v + 1$  datos.

$$Tr_{(k,z)} = \#muestras_{(k,z)} \times \Delta t \quad (2.1)$$

$$\tilde{T}r_{(k,z)} = \frac{Tr_{(k,z)}}{P} \quad (2.2)$$

$$RR(k, z) = \frac{60}{\overline{Tr}(k, z)} \quad (2.3)$$

## 2.2.2 Procesamiento del video

### 2.2.2.1 Magnificación del video

Primeramente, los videos  $V_k$  son convertidos de RGB a escala de grises ( $Vbn_k$ ), con el objetivo de simplificar los datos y reducir el tiempo de procesamiento. Utilizamos la magnificación debido a que es una técnica que sirve para enfatizar los movimientos de ciertas frecuencias, en nuestro caso magnificamos  $Vbn_k$  y obtuvimos  $Vm_k$ . Utilizamos las frecuencias de corte  $Fca_k$  y  $Fcb_k$  y en cada experimento nos valimos de una de las dos magnificaciones: la Laplaciana o la Hermitiana.

La magnificación es un método que se utiliza para amplificar los cambios en un video, los cuales son imperceptibles para el ojo humano. Sea  $L(X, t)$  una secuencia de imágenes, donde  $X = (x, y)^T$  representa la posición del píxel y  $W(t) = (\delta_x(t), \delta_y(t))^T$  el desplazamiento correspondiente dentro del dominio de la imagen con  $t$  como el tiempo asociado a cada cuadro del video. Las intensidades del píxel  $X$  en la secuencia de imágenes puede ser expresado como  $L(X, t) = f(X + W(t))$  con  $L(X, 0)$  como el primer cuadro del video. La propuesta de la magnificación de movimientos consiste en amplificar la función de desplazamiento  $W(t)$  con un factor  $\alpha$  para obtener una secuencia de imágenes amplificada  $\mathring{L}(X, t)$  (Brieva et al., 2021). Expandiendo  $L(X, t) = f(X + W(t))$  por una aproximación de Taylor de primer orden obtenemos:

$$L(X, t) \approx f(X) + W^T (\nabla f(X)), \quad (2.4)$$

donde  $\nabla f(X) := \left( \frac{\partial L(X, t)}{\partial x}, \frac{\partial L(X, t)}{\partial y} \right)^T$ .

Aplicando un filtro de banda temporal de banda ancha a (2.4) donde el vector de desplazamiento es  $W(t)$  obtenemos:

$$B(X, t) = W^T (\nabla f(X)). \quad (2.5)$$

Donde  $B(X, T)$  representa los cambios de la componente de intensidad de la secuencia de imágenes  $L(X, t)$ . Amplificando  $B(X, T)$  por un factor  $\alpha$  y sumando  $L(X, t)$  obtenemos la secuencia de imágenes amplificada:

$$\tilde{L}(X, t) = L(X, t) + \alpha B(X, t). \quad (2.6)$$

Resumiendo, el procedimiento de la magnificación Euleriana propuesto por (Wu et al., 2012) tenemos los siguientes pasos:

- Realizar una descomposición espacial de la secuencia de imágenes. Lo cual permite descomponer la secuencia de imágenes en diferentes bandas de frecuencia espaciales que están relacionadas con diferentes cambios de intensidades (2.4).
- Realizar un filtrado temporal de la descomposición espacial para retener los componentes de cambios de intensidad (2.5). Las frecuencias de corte del filtro se escogen para retener los componentes de cambio de intensidad.
- Amplificación de las bandas de frecuencia por el factor  $\alpha$ .
- Reconstrucción de la secuencia de imágenes amplificadas mediante el procedimiento inverso de la descomposición espacial.
- Agregar la secuencia de imágenes original por medio de (2.6).

Para obtener la descomposición de la pirámide usada en (Wu et al., 2012) primero una pirámide de paso baja Gaussiana de  $N$  niveles ( $L_0, \dots, L_{N-1}$ ), es necesaria:

$$L_j(x, y) = \sum_{y_0=-a}^a h(y_0) \sum_{x_0=-b}^b h(x_0) L_{j-1}(2x - x_0, 2y - y_0), \quad (2.7)$$

donde  $L_0(x, y)$  es la imagen original y  $L_j$ , en la  $2^j$  resolución, es la aproximación de la imagen original obtenida recursivamente de  $L_{j-1}$  por el filtro separable paso bajas con  $h(x, y)$  de tamaño  $k \times l$  seguido de un submuestreo de factor 2,  $a = (k - 1)/2$  y  $b = (l - 1)/2$  (Baaziz and Labit, 1990).

La pirámide Laplaciana de nivel  $N$ ,  $(\mathbf{L}_0, \dots, \mathbf{L}_{N-1})$  está dada por:

$$\begin{aligned}\mathbf{L}_{N-1} &= L_{N-1} \\ \mathbf{L}_j(x, y) &= L_j(x, y) - \bar{L}_{j+1}(x, y),\end{aligned}\quad (2.8)$$

donde  $\bar{L}_j$  es la interpolación de la imagen de  $L_j$  por un factor 2 en ambas direcciones definido por:

$$\bar{L}_{j+1}(x, y) = 4 \left[ \sum_{y_0 = -a}^a h(y_0) \sum_{x_0 = -b}^b h(x_0) L_{j-1} \left( \frac{x+x_0}{2}, \frac{y+y_0}{2} \right) \right]. \quad (2.9)$$

Para reconstruir la imagen original se utilizan las ecuaciones:

$$\begin{aligned}\mathbf{L}_{N-1} &= L_{N-1} \\ \mathbf{L}_j(x, y) &= L_j(x, y) + \bar{L}_{j+1}(x, y).\end{aligned}\quad (2.10)$$

En el caso de la transformada de Hermite (Martens, 1990), la descomposición de Hermite es obtenida mediante la convolución de la imagen  $L(x, y)$  con el banco de filtros de análisis  $D_{m,n-m}(x, y) = D_m(x)D_{n-m}(y)$  seguido por el submuestreo ( $T$ ) para obtener una transformada multiresolución (Brieva et al., 2021):

$$\begin{aligned}L_{m,n-m}(x, y) &= \sum_{y_0 = -a}^a D_{n-m}(y - y_0) \sum_{x_0 = -b}^b D_m(x - x_0) L(x_0, y_0) \\ n &= 0, \dots, N \\ m &= 0, \dots, n,\end{aligned}\quad (2.11)$$

donde  $L_{m,n-m}(x, y)$  los coeficientes cartesianos de Hermite son para el orden de análisis  $m$  y  $(n - m)$  en  $x$  y  $y$  respectivamente,  $N$  es el orden máximo de la expansión, y  $k \times l$  es el tamaño del filtro separable del análisis de Hermite  $D_{m,n-m}$  con  $a = (k-1)/2$  y  $b = (l-1)/2$ .

El análisis de los filtros de Hermite  $D_{m,n-m}$  son definidos por los polinomios de

1D de Hermite  $H_n$  (Martens, 1990):

$$D_k(x) = \frac{(-1)^k}{\sqrt{2^k k!}} \frac{1}{\sigma\sqrt{\pi}} H_k\left(\frac{x}{\sigma}\right) \exp\left(-\frac{x^2}{\sigma^2}\right)$$

$$k = 0, 1, \dots, N, \quad (2.12)$$

donde los polinomios de Hermite generalizados son:

$$H_n\left(\frac{x}{\sigma}\right) = (-1)^n \exp\left(-\frac{x^2}{\sigma^2}\right) \frac{d^n}{dx^n} \exp\left(-\frac{x^2}{\sigma^2}\right) \quad (2.13)$$

Para la reconstrucción de la imagen original  $L(x, y)$  se realiza un proceso similar usando los coeficientes de Hermite  $L_{m,n-m}(x, y)$ :

$$L(x, y) = \sum_n^{\infty} \sum_{m=0}^n \sum_{(x_0, y_0) \in S} L_{m,n-m}(x_0, y_0) P_{m,n-m}(x - x_0, y - y_0) \quad (2.14)$$

donde los filtros de Hermite son definidos como:

$$P_{m,n-m} = \frac{D_{m,n-m}(x, y)}{V(x, y)}, \quad (2.15)$$

con la función de peso:

$$V(x, y) = \sum_{(x_0, y_0) \in S} v^2(x - x_0, y - y_0) \neq 0 \quad (2.16)$$

### 2.2.2.2 Extracción de la región de interés

#### Área de trabajo

La región de interés es la zona en la cual nos vamos a enfocar en analizar para estimar la frecuencia respiratoria. Para ello, se elige un área de trabajo ( $AT$ ), la cual es una parte rectangular del video magnificado ( $Vm_k$ ) seleccionada manualmente que contiene la zona del pecho y abdomen del recién nacido, también pueden incluir zonas adyacentes que sugieran que es posible el cálculo de la frecuencia respiratoria a través de ellas y se utiliza para limitar el espacio donde el píxel superior izquierdo de las máscaras puede ser colocado. El  $AT$  se delimita por los píxeles: superior izquierdo ( $px_1, py_1$ ) y el inferior derecho ( $px_2, py_2$ ). En la figura 2.7 se aprecia mediante un cuadrado blanco el área de trabajo del sujeto 010087.



Figura 2.7: Área de trabajo del sujeto 010087

### Máscaras propuestas

Las máscaras son zonas cuadradas de tamaño  $m \times m$ , donde  $m$  es el número de píxeles por columnas y por filas, cuyos píxeles superiores izquierdos no salen del área de trabajo. La importancia de ellas radica en que la posición de la máscara en donde el error promedio sea menor en los primeros 60 segundos, será la región de interés que se utilizará para estimar la frecuencia respiratoria.

Debido a que tomar todas las máscaras posibles es altamente costoso computacionalmente, se optó por realizar  $d$  número de saltos de píxeles en cada iteración ( $g$ ) obteniendo de esta manera, una posición diferente para cada máscara  $M_{k,g}$  donde  $k$  indica el número de tiempo calmo.

En la primera iteración ( $g = 1$ ) se posiciona el píxel superior izquierdo ( $mx_{k,g}, my_{k,g}$ ) para la máscara ( $M_{k,g}$ ) en la misma posición del píxel superior izquierdo del AT es decir, ( $px_1, py_1$ ).

En las siguientes iteraciones ( $g = g + 1$ ) se desplaza  $d$  cantidad de píxeles a la derecha la máscara anterior ( $M_{k,g-1}$ ) hasta llegar al extremo derecho del AT, una vez que eso sucede, el píxel superior izquierdo ( $mx_{k,g}, my_{k,g}$ ) se coloca en el extremo izquierdo y se coloca  $d = 15$  píxeles por debajo de la posición anterior ( $mx_{k,g-1}, my_{k,g-1}$ ) y se repite el proceso hasta haber recorrido toda el área de trabajo. Ya que las máscaras son de tamaño  $m \times m$  donde  $m = 20$ , se optó por tener un desplazamiento que permitiese tener 25 % de los datos de la máscara anterior con el fin de no descartar todos los datos y mejorar la búsqueda en las zonas. Es legítimo sugerir un desplazamiento con mayor porcentaje de la máscara anterior para tener una búsqueda más detallada, sin embargo, el tiempo de procesamiento aumentaría demasiado y los resultados obtenidos en las experimentaciones que realizamos para decidir utilizar  $d = 15$ , es decir 25 % de superposición de las máscaras no mostraron cambios significativos o valorables para el tiempo demorado.

Debido a que resulta inviable procesar individualmente cada píxel a lo largo del tiempo del video ( $M_{k,g}$ ), calculamos la señal ( $s_{k,g}$ ), la cual es el resultado de promediar el valor de los píxeles en cada cuadro de ( $M_{k,g}$ ).

Posteriormente filtramos para eliminar todas aquellas partes de la señal que no son características de la frecuencia respiratoria, regularmente son pequeños movimientos en el infante, así como sombras que pueden generarse por la iluminación o por el enfoque de la cámara. Para obtener la señal filtrada ( $sf_{(k,g)}$ ) se emplean un filtro paso banda con frecuencias de corte ( $F_{c_k} + w$ ) y ( $F_{c_k} - w$ ) donde  $w = 5$ .

### Selección de la máscara óptima

Se calcula la señal de la frecuencia respiratoria ( $\widetilde{rr}_{(k,g,i)}$ ) de la señal ( $sf_{(k,g)}$ ) a partir del método descrito en la subsección 2.2.2.4 para cada máscara en la iteración ( $g$ ), con un  $ne = 60+1 - v$ , ya que, sólo se evalúan los primeros 60 segundos de las señales.

Para diferenciar las variables de este apartado de las descritas en la subsección 2.2.2.4 y 2.3.2 utilizamos letras minúsculas e  $i$  en vez de  $j$ .

Una vez que obtenemos ( $\widetilde{rr}_{(k,g)}$ ) se calcula  $e$ , el error relativo (descrito en el apartado 2.3.2), como se puede apreciar en la Ecuación 2.17.

$$e_{(k,g,i)} = \frac{|\widetilde{rr}_{(k,g,i)} - RR_{(k,g,i)}|}{RR_{(k,g,i)}} \quad (2.17)$$

Después calculamos el error absoluto porcentual medio ( $ep_{(k,g)}$ ) (definido en el apartado 2.3.2) de los errores por segundo ( $e_{(k,g)}$ ) para tener información más concreta y representativa sobre el desempeño de cada iteración de las máscaras como se muestra en la ecuación 2.18.

$$ep_{(k,g)} = \frac{\sum_{i=1}^{i=(60+1-v)} e_{(k,g,i)}}{60 + 1 - v} \quad (2.18)$$

Una vez calculado el error absoluto porcentual medio se continua con la iteración subsecuente, es decir  $g = g + 1$ , para proponer la máscara siguiente.

Una vez que se tienen almacenados en  $ep_{(k,g)}$  todos los errores absolutos porcentuales medios para cada iteración  $g$  en el tiempo calmo  $k$  y las máscaras almacenadas en  $M_{(k,g)}$ . Buscamos el error mínimo de  $ep_{(k,g)}$  y su correspondiente máscara  $M_{k,g}$  a la cual la definimos como la región de interés ( $Rdi_k$ ).

Se pueden apreciar en un cuadro pequeño blanco las regiones de interés encon-

tradas en los videos magnificados por Laplace en las Figuras 2.8, 2.9 y 2.10 y por Hermite en las Figuras 2.11, 2.12 y 2.13 con  $\alpha$  igual a: 15, 20 y 30 (conforme los valores mencionados en la sección 2.3.1) respectivamente en los 5 tiempos calmos del sujeto 010087. El rectángulo blanco muestra el área de trabajo. En el tiempo calmo 3 se puede ver que la región de interés sobresale del área de trabajo y esto es válido debido a que el área de trabajo sólo limita la posición del píxel superior izquierdo de la máscara que se propone.

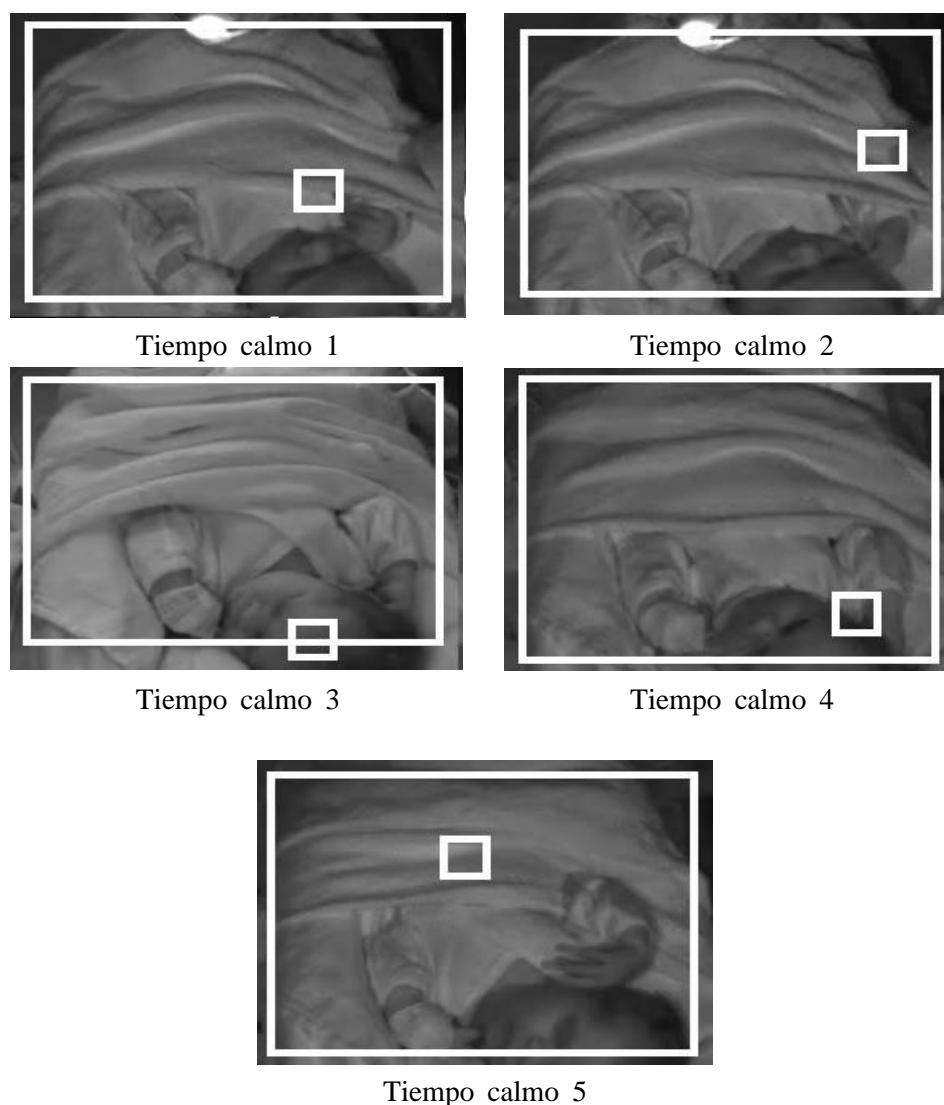


Figura 2.8: Regiones de interés con magnificación Laplaciana con  $\alpha = 15$  y  $\nu = 25$  para el sujeto 010087.

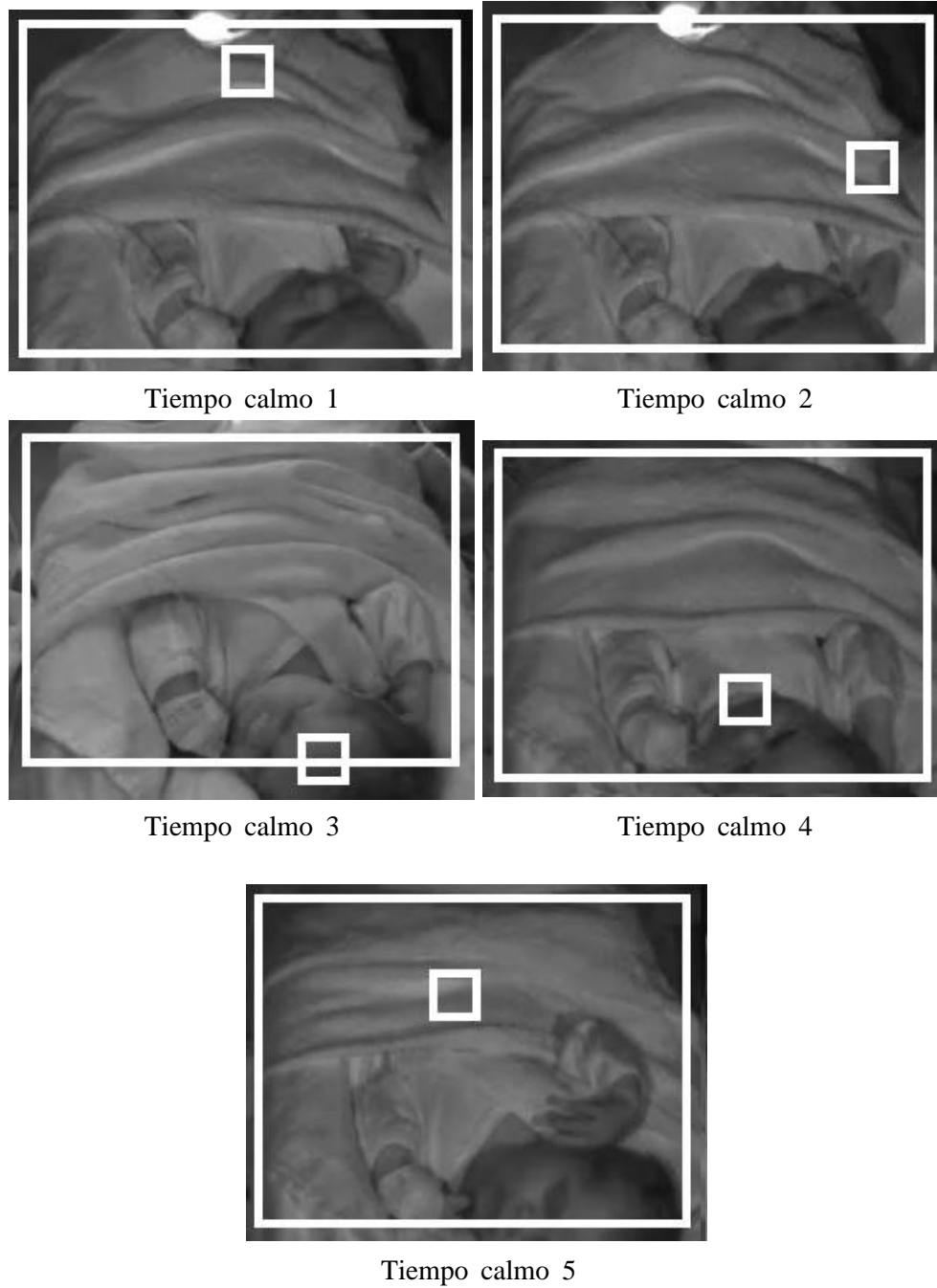


Figura 2.9: Regiones de interés con magnificación Laplaciana con  $\alpha = 20$  y  $\nu = 25$  para el sujeto 010087

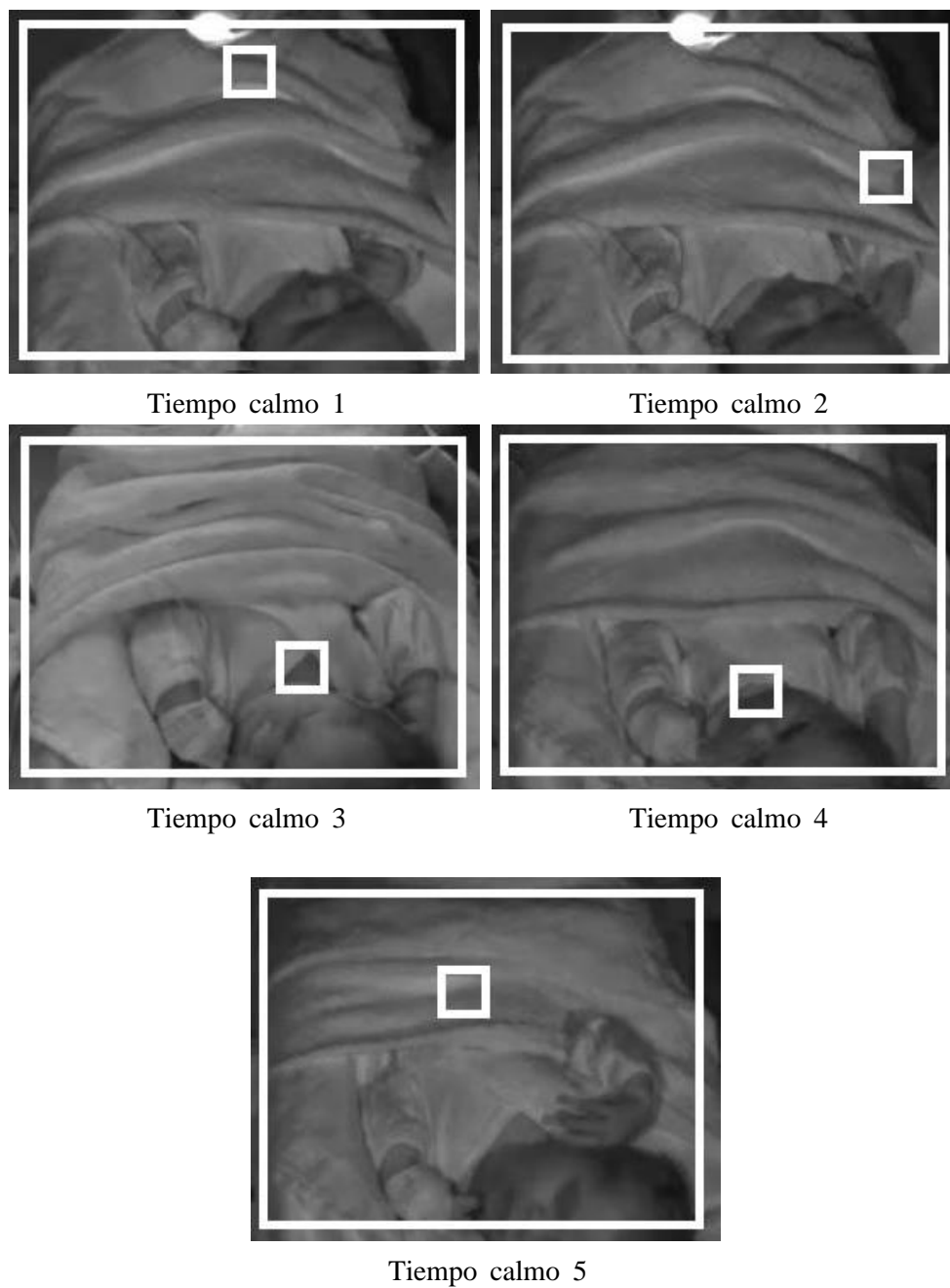


Figura 2.10: Regiones de interés con magnificación Laplaciana con  $\alpha = 30$  y  $\nu = 25$  para el sujeto 010087.

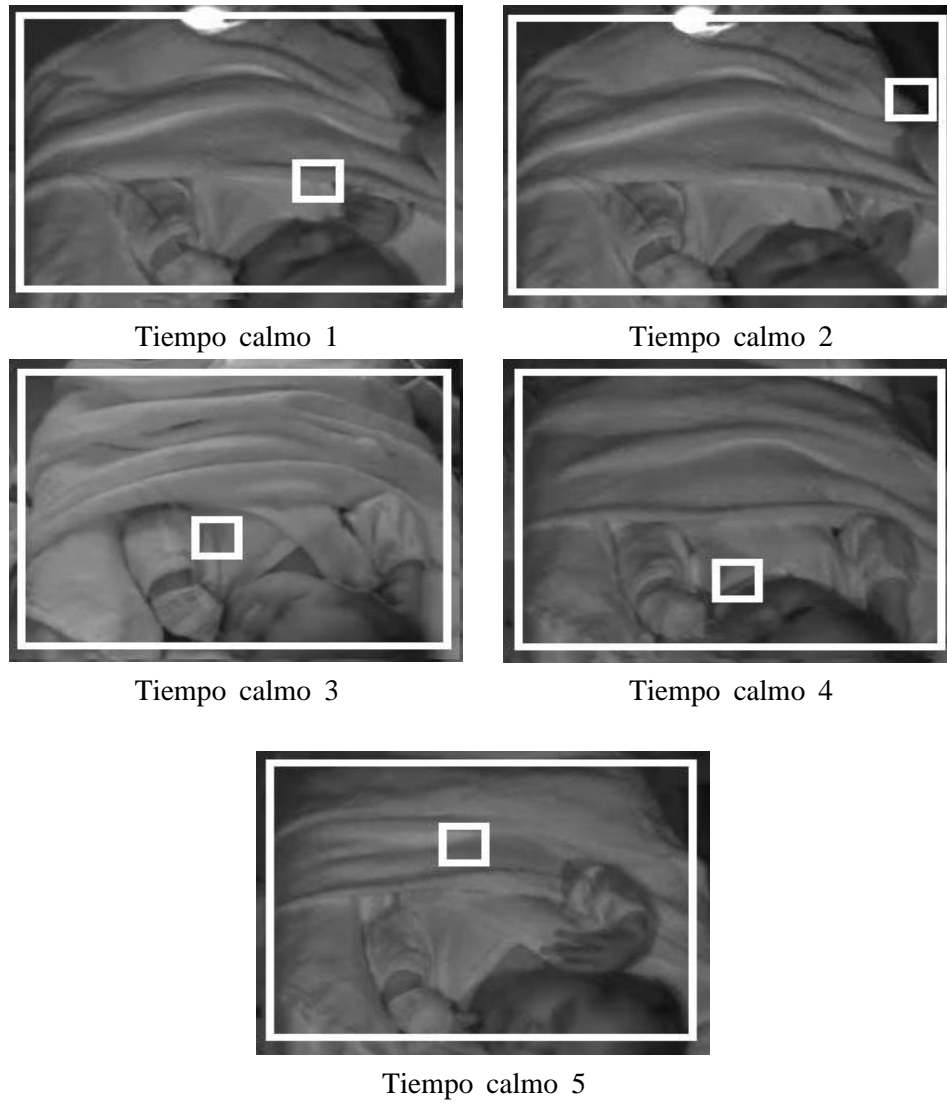


Figura 2.11: Regiones de interés con magnificación Hermitiana con  $\alpha = 15$  y  $\nu = 25$  para el sujeto 010087.

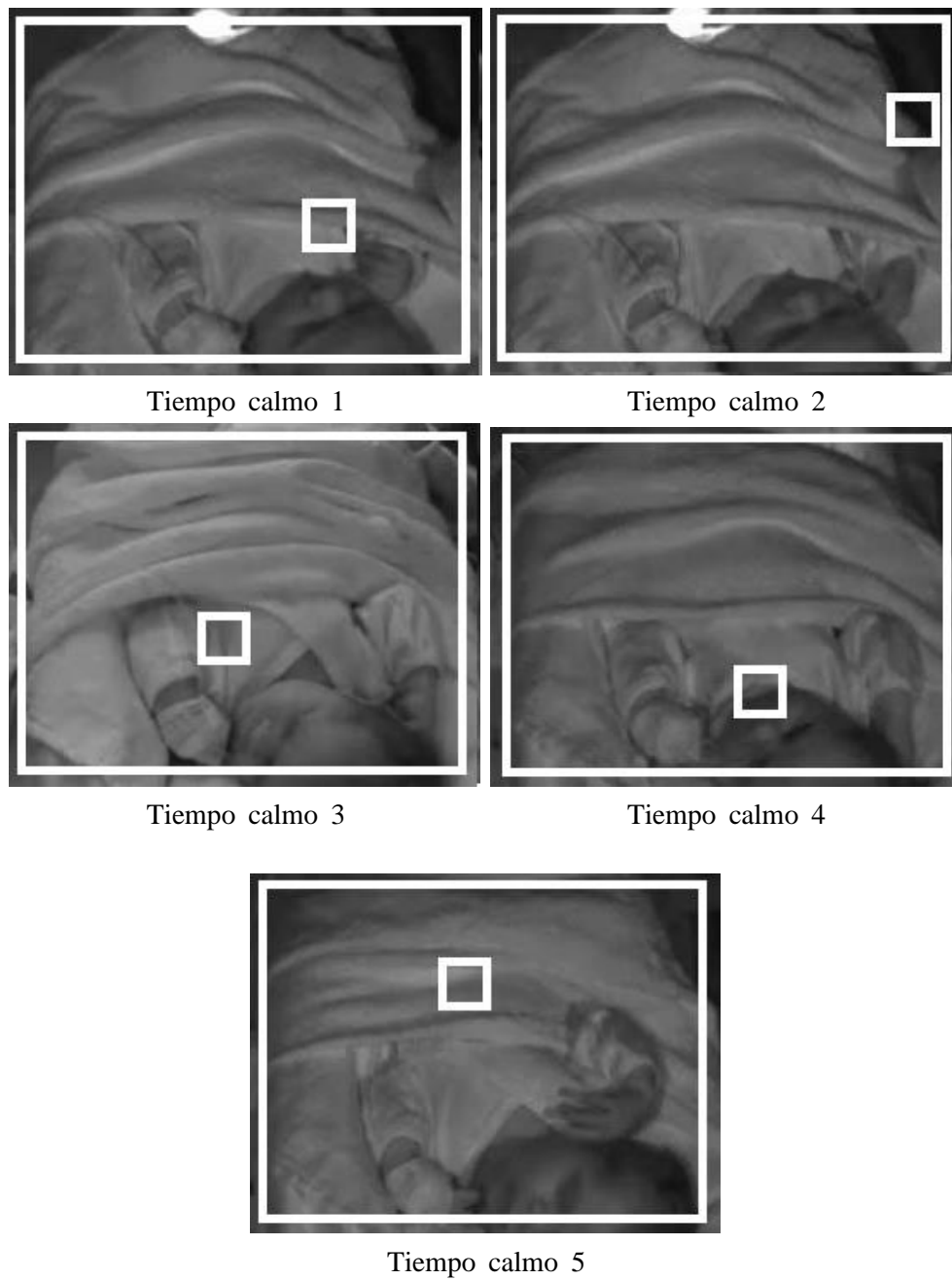


Figura 2.12: Regiones de interés con magnificación Hermitiana con  $\alpha = 20$  y  $\nu = 25$  para el sujeto 010087.

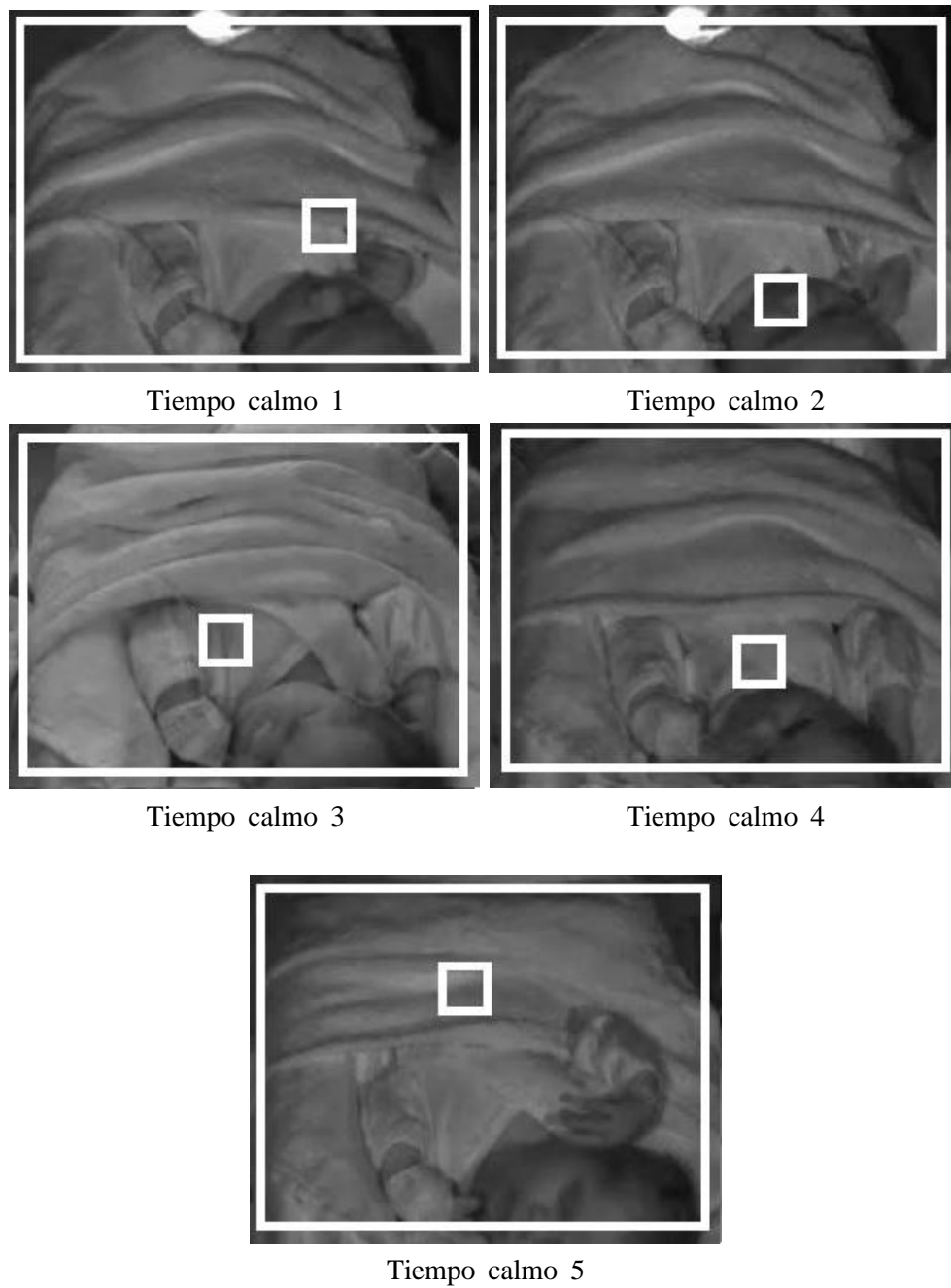


Figura 2.13: Regiones de interés con magnificación Hermitiana con  $\alpha = 30$  y  $\nu = 25$  para el sujeto 010087.

### 2.2.2.3 Extracción de la señal respiratoria en la región de interés

Calculamos un promedio espacial en cada cuadro de la región de interés ( $Rd_{ik}$ ) donde  $k$  es el tiempo calmo, obteniendo así la señal ( $S_k$ ). Es importante recordar que cada cuadro contiene  $m \times m$  píxeles con una escala de grises con valores del 0 al 255.

En la Figura 2.14 se pueden ver los primeros 60 segundos de la señal extraída de la región de interés del video magnificado con Hermite del tiempo calmo 4 del sujeto 010087. El eje de las abscisas representa el número de las muestras y el eje de las ordenadas representa la señal en bruto después de la adquisición, en niveles de cuantificación digitalizados.

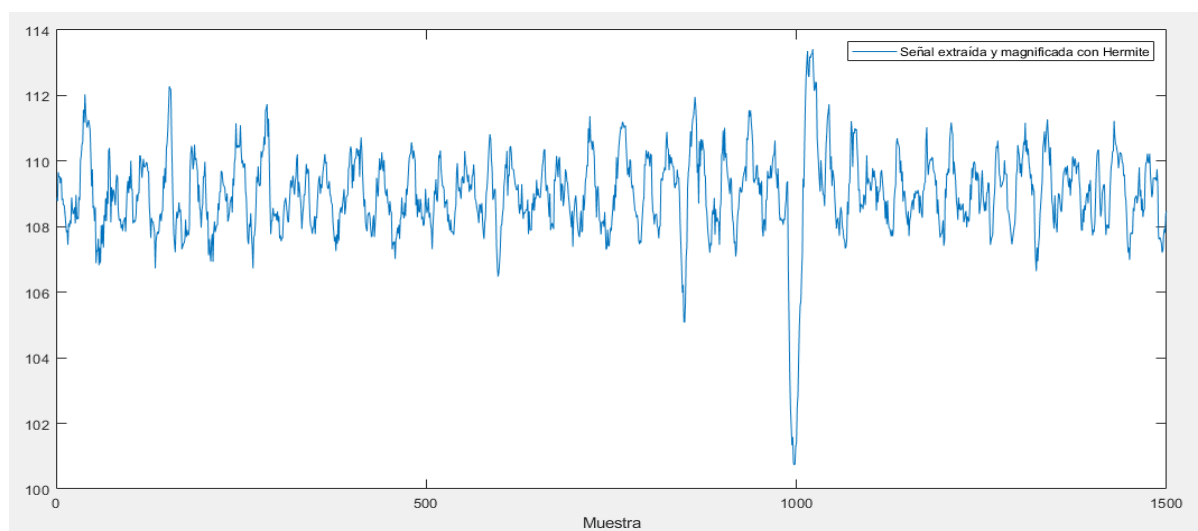


Figura 2.14: Señal extraída de la región de interés del video magnificado con Hermite del tiempo calmo 4, sujeto 010087 ( $\alpha = 30$ )

En la Figura 2.15 se pueden ver los primeros 60 segundos de la señal extraída del video sin magnificar en la región de interés encontrada mediante el video magnificado con Hermite ( $\alpha = 30$  y  $\nu = 25$ ) del tiempo calmo 4 del sujeto 010087.

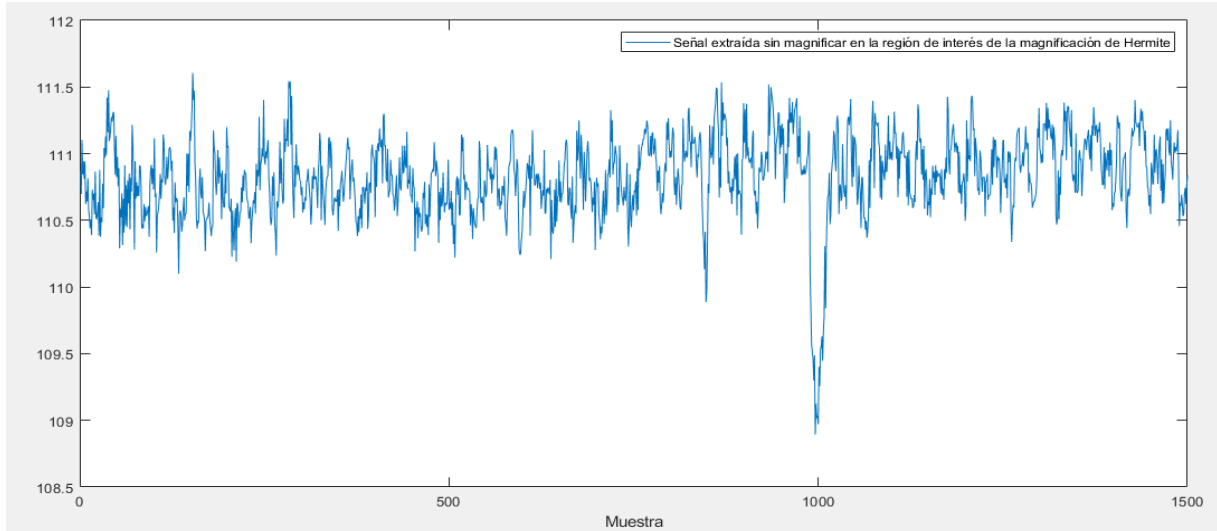


Figura 2.15: Señal extraída del video sin magnificar de la región de interés de encontrada en el video magnificado con Hermite del tiempo calmo 4, sujeto 010087

En la Figura 2.16 se pueden ver los primeros 60 segundos en la señal extraída de la región de interés del video magnificado con Laplace del tiempo calmo 4 del sujeto 010087.

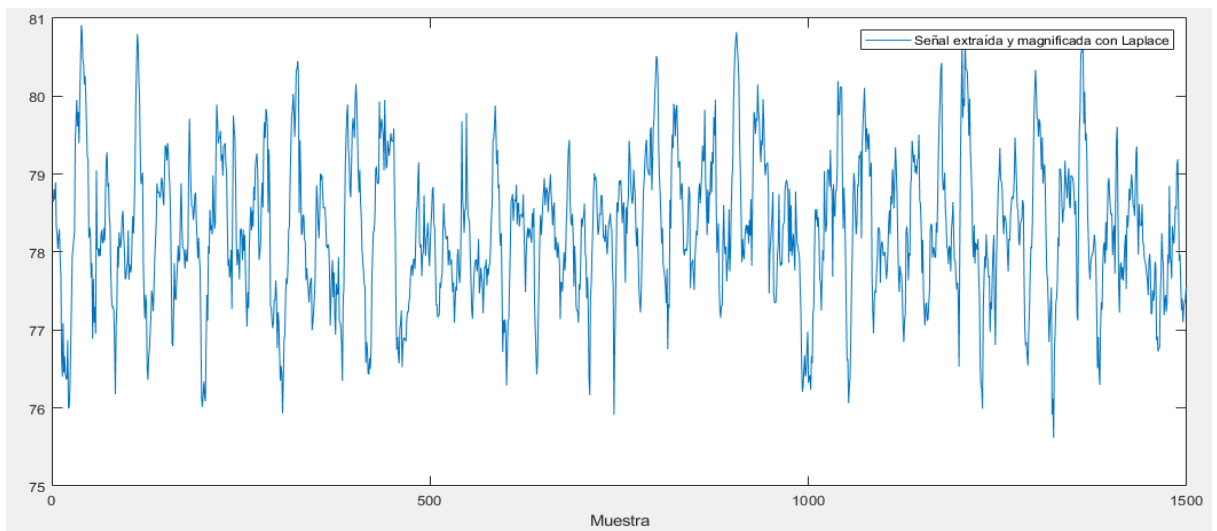


Figura 2.16: Señal extraída de la región de interés del video magnificado con Laplace ( $\alpha = 30$ ) del tiempo calmo 4, sujeto 010087

En la Figura 2.17 se pueden ver los primeros 60 segundos de la señal extraída del video sin magnificar en la región de interés encontrada mediante el video magnificado con Laplace del tiempo calmo 4 del sujeto 010087.

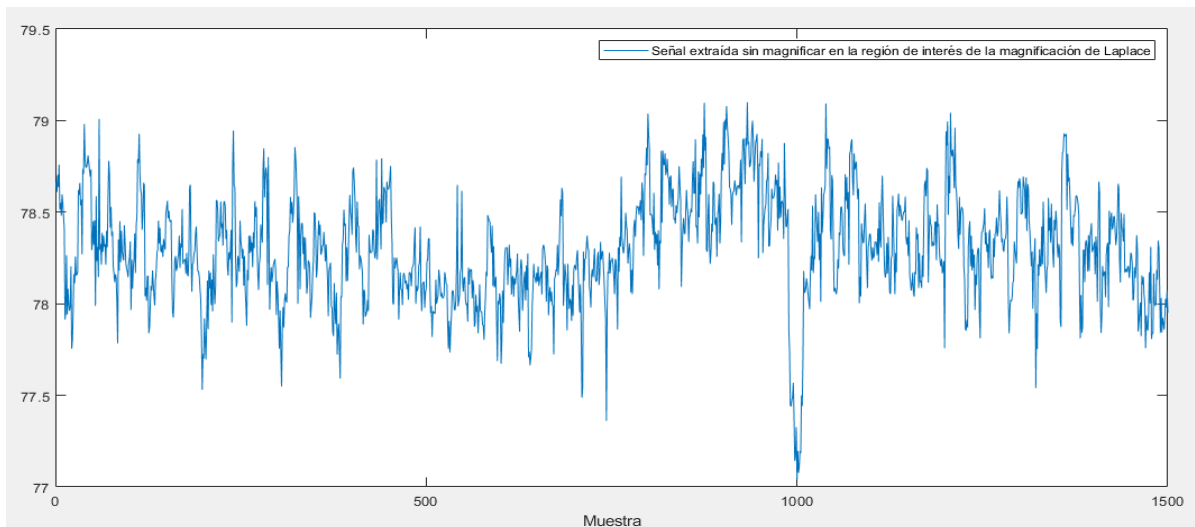


Figura 2.17: Señal extraída del video sin magnificar de la región de interés de encontrada en el video magnificado con Laplace ( $\alpha = 30$ ) del tiempo calmo 4, sujeto 010087

En la Figura 2.18 se puede ver el minuto 9 de la señal extraída de la región de interés del video magnificado con Hermite del tiempo calmo 4 del sujeto 010087.

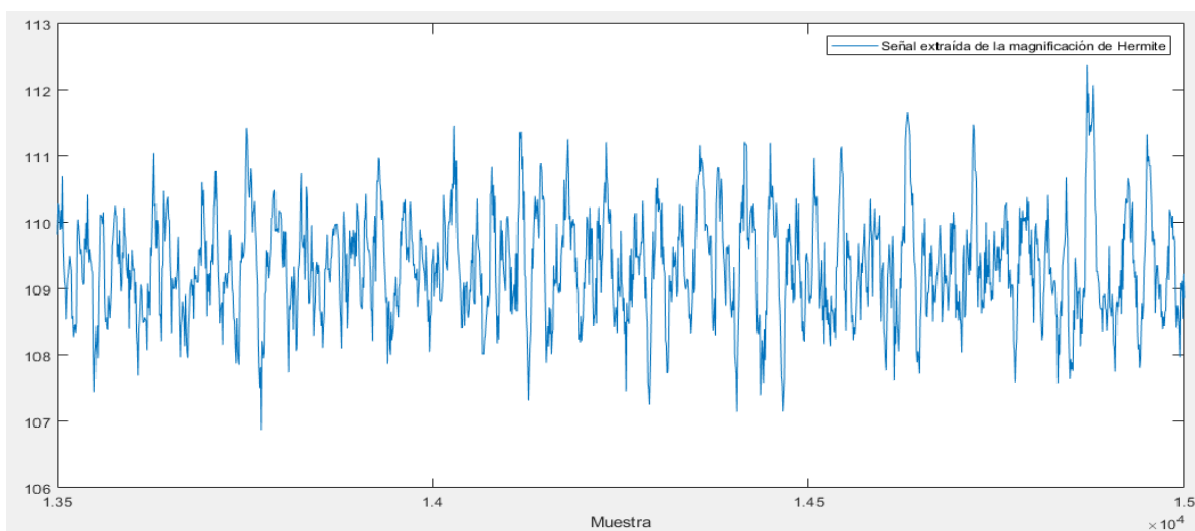


Figura 2.18: Minuto 9 de la señal extraída de la región de interés del video magnificado con Hermite ( $\alpha = 30$ ) del tiempo calmo 4, sujeto 010087

En la Figura 2.19 se puede ver se puede ver el minuto 9 de la señal extraída del

video sin magnificar en la región de interés encontrada mediante el video magnificado con Hermite del tiempo calmo 4 del sujeto 010087.

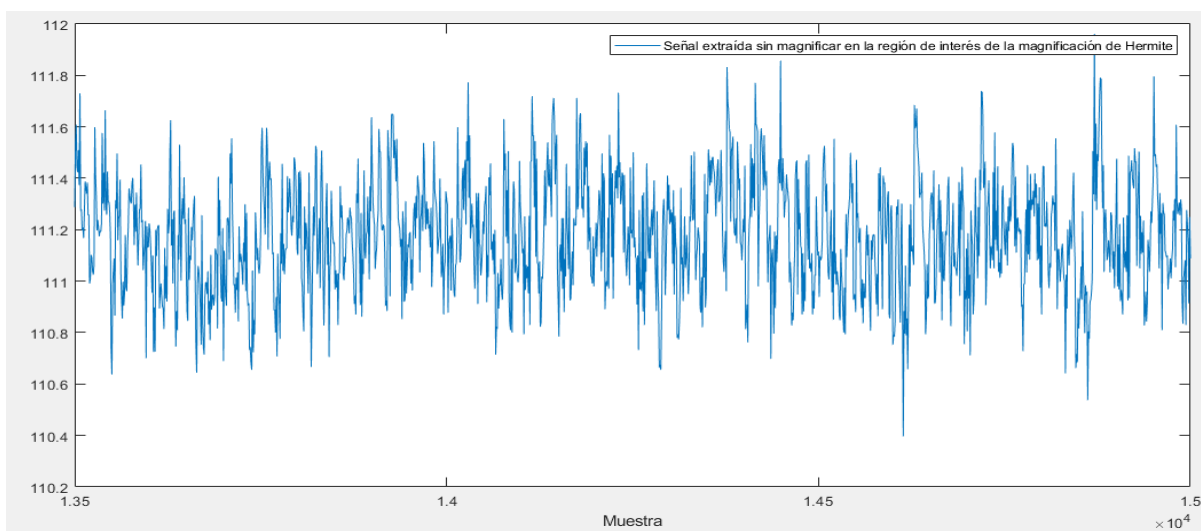


Figura 2.19: Minuto 9 de la señal extraída del video sin magnificar de la región de interés de encontrada en el video magnificado con Hermite ( $\alpha = 30$ ) del tiempo calmo 4, sujeto 010087

En la Figura 2.20 se pueden ver el minuto 9 de la señal extraída de la región de interés del video magnificado con Laplace del tiempo calmo 4 del sujeto 010087.

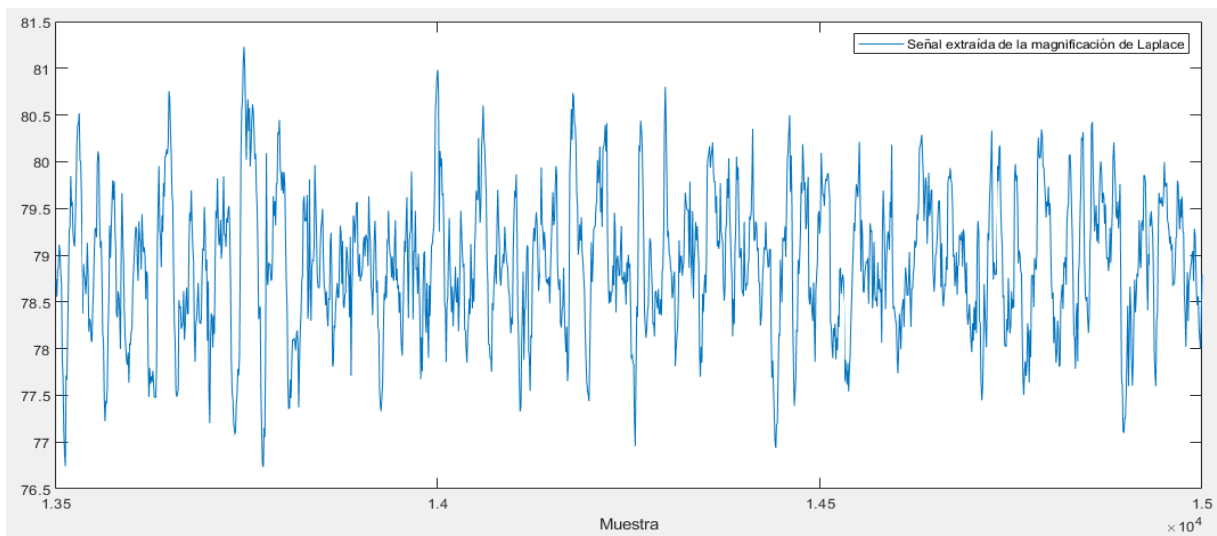


Figura 2.20: Minuto 9 de la señal extraída de la región de interés del video magnificado con Laplace ( $\alpha = 30$ ) del tiempo calmo 4, sujeto 010087

En la Figura 2.21 se pueden ver los primeros 60 segundo de la señal extraída del

video sin magnificar en la región de interés encontrada mediante el video magnificado con Laplace del tiempo calmo 4 del sujeto 010087. Es importante recordar que el eje de las ordenadas representa la señal en bruto después de la adquisición.

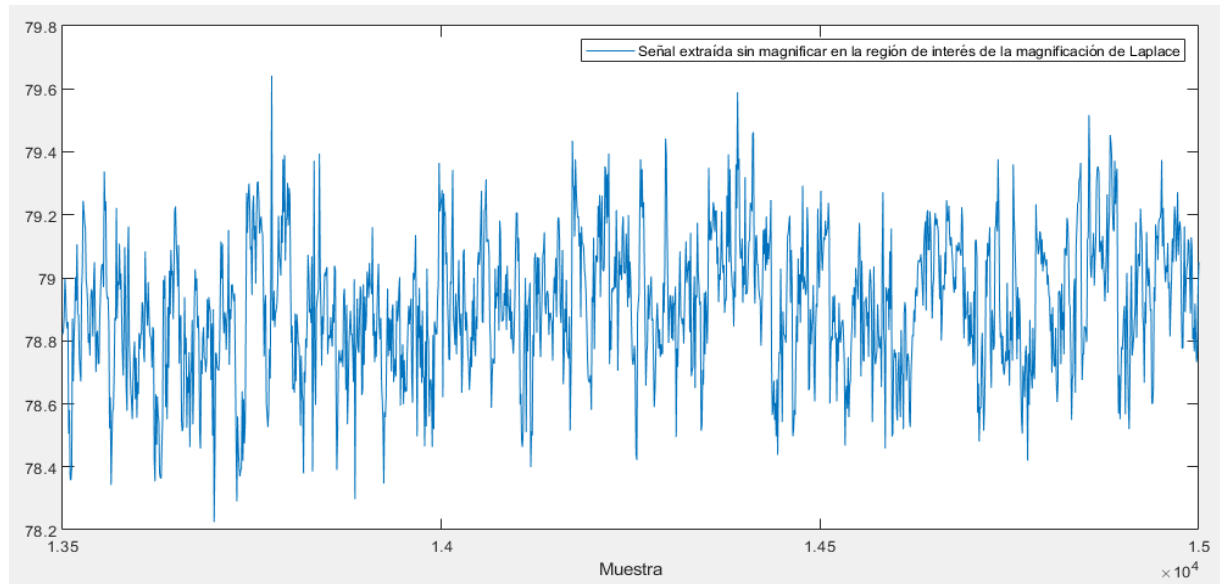


Figura 2.21: Minuto 9 de la señal extraída del video sin magnificar de la región de interés de encontrada en el video magnificado con Laplace ( $\alpha = 30$ ) del tiempo calmo 4, sujeto 010087

La señal  $S_k$  se filtra con las frecuencias de corte  $F_{ck} + w$  y  $F_{ck} - w$  utilizando un filtro pasobanda, el cual nos permitirá eliminar todas aquellas frecuencias que son ajenas a nuestro estudio. Por medio de esta operación obtenemos la señal filtrada ( $Sf_k$ ).

En la Figura 2.22 se pueden observar 1500 muestras correspondientes a los primeros 60 segundos de la señal magnificada con Hermite y filtrada de la región de interés del sujeto 010087 del tiempo calmo 4.

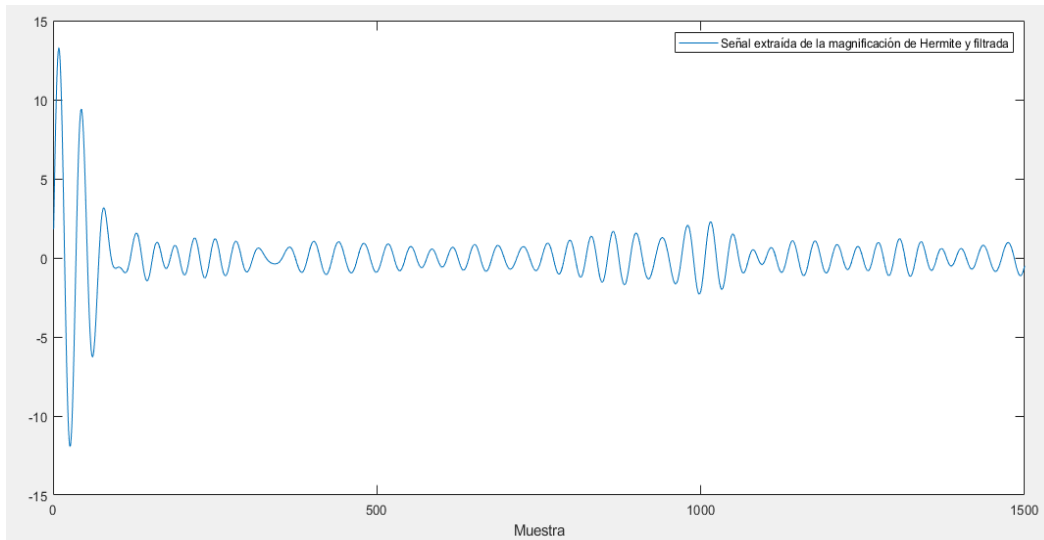


Figura 2.22: Señal extraída y filtrada de Hermite en el primer minuto

En la Figura 2.23 se pueden observar 1500 muestras correspondientes a los primeros 60 segundos de la señal magnificada con Laplace y filtrada de la región de interés del sujeto 010087 del tiempo calmo 4.

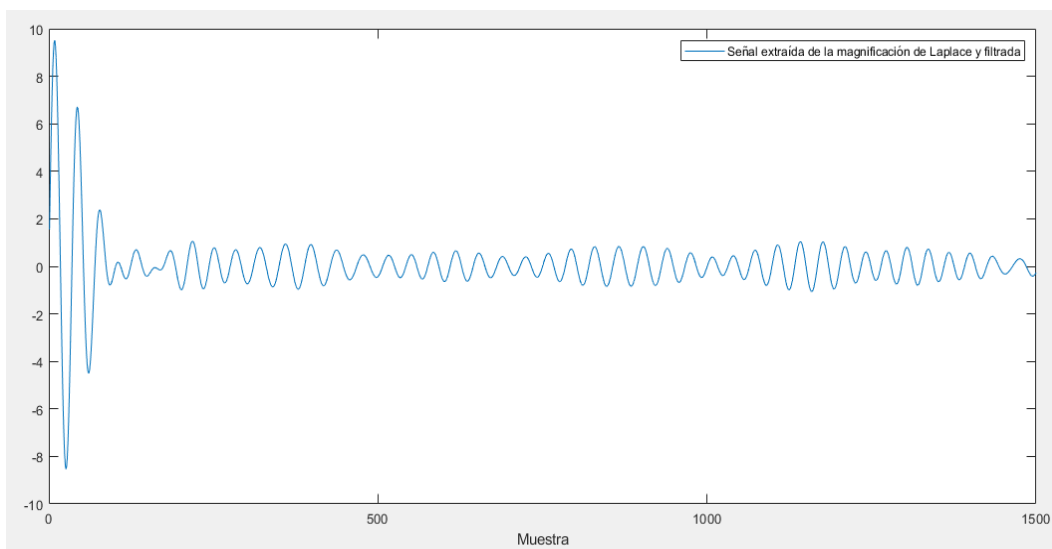


Figura 2.23: Señal extraída y filtrada de Laplace en el primer minuto

En la Figura 2.24 se pueden observar 1500 muestras correspondientes al minuto 10 de la señal magnificada con Hermite y filtrada de la región de interés del sujeto 010087 del tiempo calmo 4.

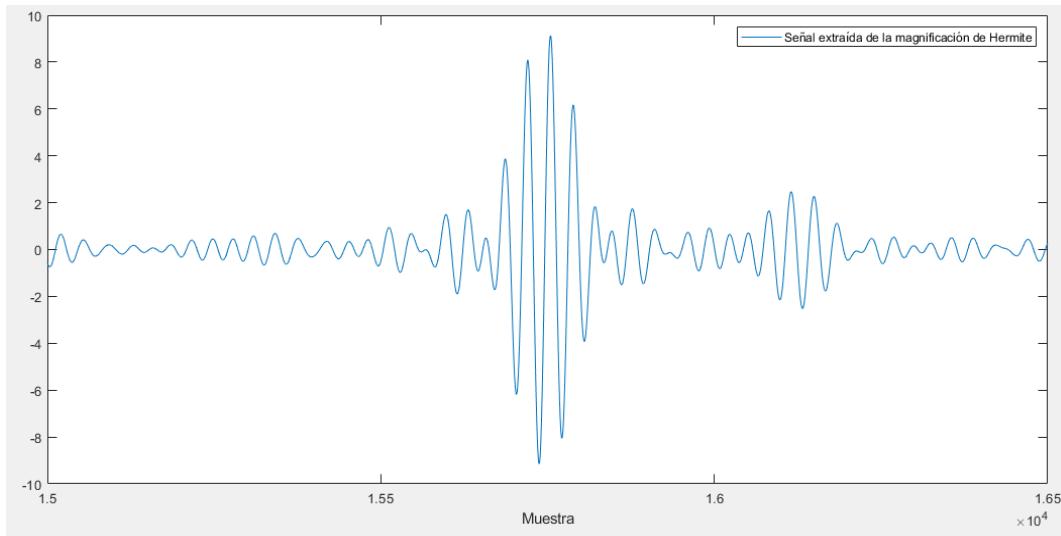


Figura 2.24: Señal extraída y filtrada de Hermite en el minuto 10

En la Figura 2.25 se pueden observar 1500 muestras correspondientes al minuto 10 de la señal magnificada con Hermite y filtrada de la región de interés del sujeto 010087 del tiempo calmo 4.

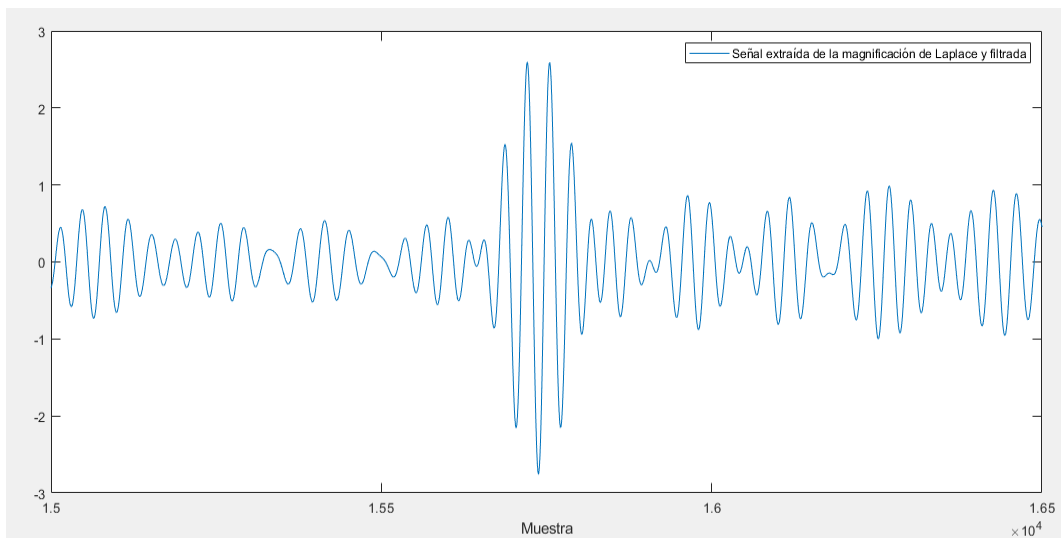


Figura 2.25: Señal extraída y filtrada de Laplace en el minuto 10

### 2.2.2.4 Extracción de la señal de frecuencia respiratoria

La extracción de la señal de frecuencia respiratoria se calcula a partir del tamaño de la ventana ( $v$ ), ésta indica el tiempo (en segundos) que se toman para poder estimar las respiraciones por minuto, así es posible saber la cantidad de muestras a analizar ( $cma$ ) de  $Sf_k$  por cada estimación.

Para conocer el valor de  $cma$  se divide la cantidad de segundos de la ventana ( $v$ ) entre el período del muestreo ( $\Delta t$ ) de la señal a tratar, en este caso  $Sf_k$  tiene una

$$\Delta t = \frac{1}{25}$$

$$cma = \frac{v}{\Delta t} \quad (2.19)$$

Para la primera estimación ( $j = 1$ ) se toman las primeras  $cma$  muestras de la señal  $Sf_k$  y se cuenta el número de muestras ( $\#muestras_{(k,j)}$ ) que hay entre los picos de esta ventana, después lo multiplicamos por el período del muestreo ( $\Delta t$ ), dando como resultado  $T_{(k,j)}$ . Después, obtenemos  $\tilde{T}_{(k,j)}$  dividiendo  $T_{(k,j)}$  entre el número de picos ( $N$ ) en la ventana. Obtenemos la estimación de la respiración por minuto ( $\widetilde{RR}_{(k,j)}$ ) dividiendo 60 segundos entre ( $\tilde{T}_{(k,j)}$ ). Se suma en uno el valor de  $j$  y se repite el procedimiento hasta que sea  $j = t_k + 1 - v$  ya que es el límite de estimaciones que se puede realizar, puesto que tras cada iteración se recorre la ventana un segundo es decir  $\Delta t^{-1}$  muestras.

$$T_{(k,j)} = \#muestras_{(k,j)} \times \Delta t \quad (2.20)$$

$$\tilde{T}_{(k,j)} = \frac{T_{(k,j)}}{N} \quad (2.21)$$

$$\widetilde{RR}_{(k,j)} = \frac{60}{\tilde{T}_{(k,j)}} \quad (2.22)$$

En la Figura 2.26 podemos observar  $\widetilde{RR}$  (en azul) y  $RR$  (en naranja) del sujeto 010076 del tiempo calmo 1, magnificado con la estrategia de Hermite con  $\alpha = 30$  y una ventana  $\nu = 25$ .

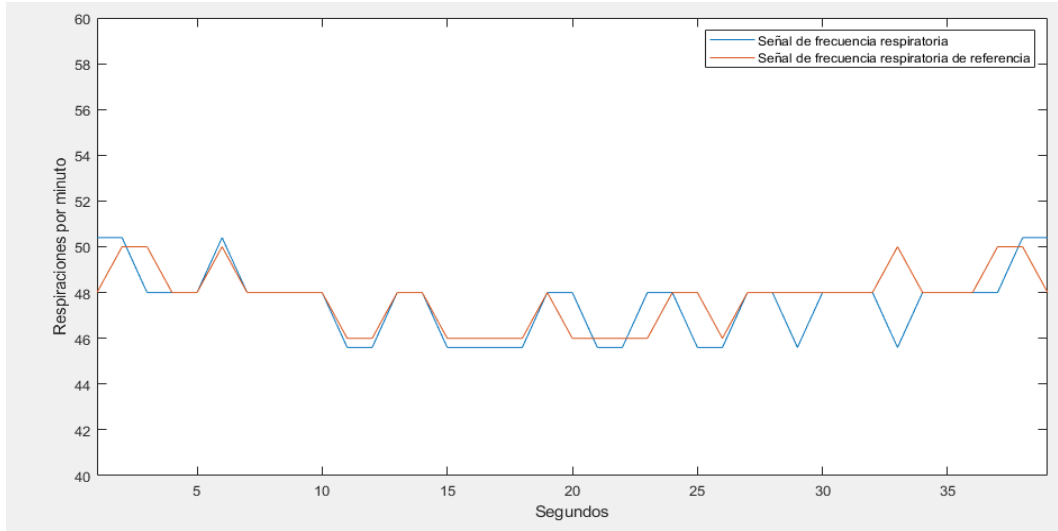


Figura 2.26: Señal de frecuencia respiratoria ( $\widetilde{RR}$ ) y señal de frecuencia respiratoria de referencia ( $RR$ ) del sujeto 010076 del tiempo calmo 1

En la Figura 2.27 podemos observar  $\widetilde{RR}$  (en azul) y  $RR$  (en naranja) del sujeto 010077 del tiempo calmo 13, magnificado con la estrategia de Laplace con  $\alpha = 20$  y una ventana  $\nu = 25$ .

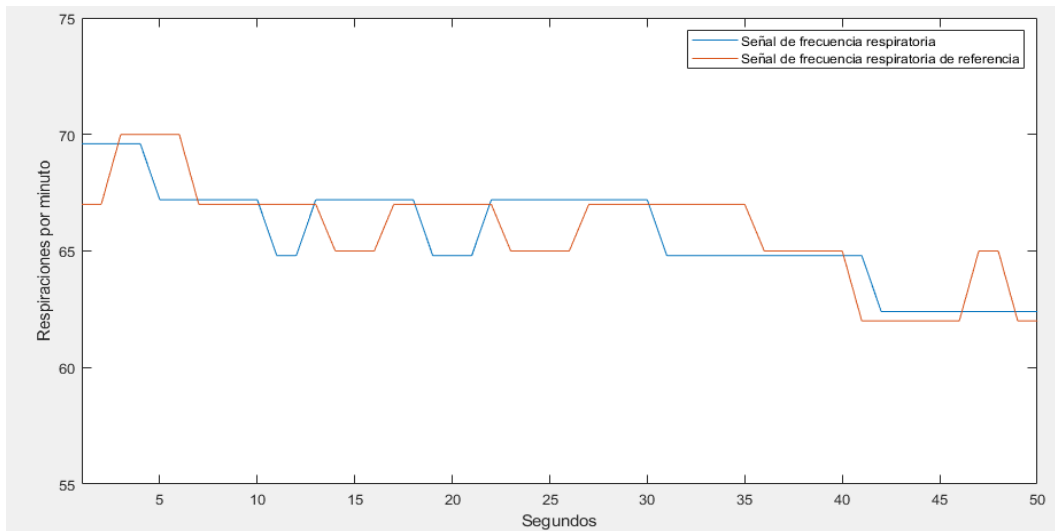


Figura 2.27: Señal de frecuencia respiratoria ( $\widetilde{RR}$ ) y señal de frecuencia respiratoria de referencia ( $RR$ ) del sujeto 010077 del tiempo calmo 13

## 2.3 Experimentos

### 2.3.1 Configuración de parámetros

Para las experimentaciones se utilizaron ventanas ( $v$ ) con valores de: 5, 15, 20 y 25 segundos. Las máscaras tienen un tamaño  $m \times m$  con  $m = 20$  píxeles y la cantidad de píxeles saltados por iteración con  $d = 15$

Para cada una de las magnificaciones (Hermitianas y Laplacianas) se realizaron experimentos con 3 valores:  $\alpha = 15$ ,  $\alpha = 20$  y  $\alpha = 30$ , las cuales se obtuvieron experimentalmente. En el trabajo de Koolen se utiliza un coeficiente de magnificación  $\alpha = 15$ . (Koolen et al., 2015), por lo cual también fue considerado en nuestro trabajo. Cabe mencionar que alfas con valores más altos distorsionan los datos y si son menores a 15 no alcanzan a magnificar la señal de la forma deseada.

### 2.3.2 Métricas de evaluación del método

#### Error absoluto porcentual medio (MAPE)

Para calcular el error absoluto porcentual medio primeramente se calcula el error relativo ( $E_{(k,j)}$ ), que está dado por la división del absoluto de la resta entre la señal de frecuencia respiratoria del video  $\widetilde{RR}_{(k,j)}$  y la de la referencia  $RR_{(k,j)}$  entre  $RR_{(k,j)}$  como se muestra en la Ecuación 2.23.

$$E_{(k,j)} = \frac{|\widetilde{RR}_{(k,j)} - RR_{(k,j)}|}{RR_{(k,j)}} \quad (2.23)$$

Donde  $k$  representa el tiempo calmo y  $j$  el segundo a analizar

El error promedio absoluto porcentual medio  $EP_{(k)}$  del tiempo calmo ( $k$ ) está dado por la sumatoria de los errores de  $E_{(k,j)}$  dividida entre la cantidad de elementos que contiene ( $ne$ ) la señal y multiplicada por 100, como se puede observar en la ecuación 2.24.

$$EP_{(k)} = \frac{\sum_{j=1}^{j=(ne)} E_{(k,j)}}{ne} \times 100 \quad (2.24)$$

**Error absoluto medio (MAE)**

El error absoluto medio (MAE) se obtiene mediante la sumatoria de los valores absolutos de la resta de cada uno de los valores de la señal de frecuencia respiratoria obtenida del video  $\widetilde{RR}_{(k,j)}$  menos la señal de frecuencia respiratoria de la referencia  $RR_{(k,j)}$  y finalmente dividiendo el resultado entre el número de elementos de las señales ( $ne$ ), como se muestra en 2.25.

$$MAE_{(k)} = \frac{\sum_{j=1}^{j=(ne)} |\widetilde{RR}_{(k,j)} - RR_{(k,j)}|}{ne} \quad (2.25)$$

**Raíz del error cuadrático medio (RMSE)**

Se calcula la raíz del error cuadrático medio (RMSE) dada la Ecuación 2.26. En donde: la frecuencia respiratoria de la referencia es  $(RR(k, j))$ , la frecuencia respiratoria estimada a partir del video es  $(\widetilde{RR}(k, j))$ ,  $k$  la iteración del tiempo calmo,  $j$  el número de elemento evaluado de  $(RR(k, j))$  y  $(\widetilde{RR}(k, j))$ ,  $ne$  cantidad de muestras en las señales y  $v$  el tamaño de la ventana.

$$RMSE_{(k)} = \sqrt{\sum_{j=1}^{j=(t_k+1-v)} \frac{(\widetilde{RR}(k, j) - RR(k, j))^2}{ne}} \quad (2.26)$$

Además, se realizó un análisis de Bland-Altman para evaluar el nivel de coincidencia entre la estimación de la frecuencia respiratoria a partir de esta propuesta y los valores de frecuencia respiratoria obtenidos a partir de la señal respiratoria de referencia. El 95 % de los límites de coincidencia fueron definidos como  $MOD \pm 1.96\sigma$ , donde  $MOD$  es la media de las diferencias entre el valor estimado y la referencia y  $\sigma$  es la desviación estándar de la  $MOD$ . Este análisis se hizo por sujeto y de manera global tanto para la estrategia que utiliza la transformada de Hermite como para la estrategia Laplaciana. Se calcularon, además, los coeficientes de correlación de Spearman (SCC) y de Pearson's (PCC) entre los valores estimados y la referencia. Estos coeficientes se calcularon por sujeto y de manera global tanto para la estrategia que utiliza la transformada de Hermite como para la estrategia Laplaciana.

# Capítulo 3

## Resultados

En las Tablas 3.1 y 3.2 se muestran el error MAPE (descrito en la Sección 2.3.2) de los tiempos calmos obtenidos en el sujeto 010076, utilizando la magnificación de Hermite y de Laplace respectivamente para cada tabla, con cuatro tipos de ventanas: 25, 20, 15 y 5 segundos; y con  $\alpha$  de: 30, 20 y 15 (conforme los valores mencionados en la Sección 2.3.1). Para la columna MAPE en 60 segundos, se utilizaron el primer minuto de la señal, por lo que  $ne = 60 + 1 - v$ , la cual representa la parte de la señal que fue evaluada para encontrar la región de interés y para MAPE en el tiempo calmo completo se utilizó toda la señal, siendo así  $ne = t_k + 1 - v$ .

Tabla 3.1: Errores promedio en los tiempos calmos del sujeto 010076 utilizando la estrategia de Hermite

<b>Ventana</b>	$\alpha$	<b>MAPE</b>
25	30	5.391
25	20	5.208
25	15	5.263
20	30	5.957
20	20	5.251
20	15	5.269
15	30	6.525
15	20	6.485
15	15	6.249
5	30	10.253
5	20	10.036
5	15	10.365

Tabla 3.2: Errores promedio en los tiempos calmos del sujeto 010076 utilizando la estrategia de Laplace

<b>Ventana</b>	$\alpha$	<b>MAPE</b>
25	30	5.239
25	20	5.149
25	15	5.326
20	30	5.888
20	20	5.479
20	15	5.533
15	30	6.046
15	20	5.723
15	15	5.947
5	30	9.998
5	20	9.794
5	15	10.458

En las Tablas 3.3 y 3.4 se muestran los promedios de los errores de los tiempos calmos obtenidos en el sujeto 010077, utilizando la magnificación de Hermite y de Laplace respectivamente para cada tabla, con cuatro tipos de ventanas: 25, 20, 15 y 5 segundos; y con  $\alpha$  de: 30, 20 y 15.

Tabla 3.3: Errores promedio en los tiempos calmos del sujeto 010077 utilizando la estrategia de Hermite

<b>Ventana</b>	$\alpha$	<b>MAPE</b>
25	30	4.838
25	20	5.218
25	15	5.070
20	30	5.099
20	20	5.460
20	15	5.022
15	30	5.455
15	20	5.734
15	15	5.365
5	30	8.848
5	20	8.961
5	15	8.911

Tabla 3.4: Errores promedio en los tiempos calmos del sujeto 010077 utilizando la estrategia de Laplace

<b>Ventana</b>	$\alpha$	<b>MAPE</b>
25	30	4.975
25	20	4.888
25	15	5.017
20	30	8.084
20	20	5.272
20	15	5.274
15	30	6.031
15	20	5.836
15	15	5.590
5	30	8.904
5	20	8.885
5	15	9.568

En las Tablas 3.5 y 3.6 se muestran los promedios de los errores de los tiempos calmos obtenidos en el sujeto 010087, utilizando la magnificación de Hermite y de Laplace respectivamente para cada tabla, con cuatro tipos de ventanas: 25, 20, 15 y 5 segundos; y con  $\alpha$  de: 30, 20 y 15.

Tabla 3.5: Errores promedio en los tiempos calmos del sujeto 010087 utilizando la estrategia de Hermite

<b>Ventana</b>	$\alpha$	<b>MAPE</b>
25	30	4.020
25	20	4.992
25	15	4.777
20	30	4.893
20	20	5.064
20	15	4.730
15	30	5.156
15	20	5.119
15	15	4.916
5	30	12.126
5	20	10.995
5	15	11.117

Tabla 3.6: Errores promedio en los tiempos calmos del sujeto 010087 utilizando la estrategia de Laplace

Ventana	$\alpha$	MAPE
25	30	4.704
25	20	4.616
25	15	4.588
20	30	4.505
20	20	4.728
20	15	4.468
15	30	4.712
15	20	4.852
15	15	4.579
5	30	11.754
5	20	11.876
5	15	11.803

En las Tablas 3.7 y 3.8 se muestran los errores desglosados en los tiempos calmos del sujeto 010087 que se obtuvieron en la región de interés en 60 segundos y los errores en todo el tiempo calmo utilizando la estrategia de Hermite y de Laplace respectivamente para cada tabla.

Tabla 3.7: Errores desglosados de los tiempos calmos del sujeto 010087 utilizando la estrategia de Hermite ( $\alpha = 30$  y  $\nu = 25$ )

Tiempo calmo k	Duración (s)	MAPE en 60 s	MAPE en tiempo calmo completo
1	122	<b>1.106</b>	2.560
2	66	1.747	<b>1.850</b>
3	373	1.686	4.929
4	813	1.310	3.737
5	173	1.992	4.947

Tabla 3.8: Errores desglosados de los tiempos calmos del sujeto 010087 utilizando la estrategia de Laplace ( $\alpha = 30$  y  $\nu = 25$ )

Tiempo calmo k	Duración (s)	MAPE en 60 s	MAPE en tiempo calmo completo
1	122	0.907	4.661
2	66	1.622	<b>2.492</b>
3	373	1.394	4.177
4	813	1.177	5.041
5	173	<b>0.852</b>	4.796

En las Tablas 3.9 y 3.10 se muestran los errores desglosados en los tiempos calmos del sujeto 010077 que se obtuvieron en la región de interés en 60 segundos y los errores en todo el tiempo calmo utilizando la estrategia de Hermite y de Laplace respectivamente para cada tabla.

Tabla 3.9: Errores desglosados de los tiempos calmos del sujeto 010077 utilizando la estrategia de Hermite ( $\alpha = 30$  y  $\nu = 25$ )

Tiempo calmo k	Duración	MAPE en 60 s	MAPE en tiempo calmo completo
2	171	1.214	4.964
3	86	1.164	2.946
5	220	1.886	4.934
6	64	1.726	<b>1.950</b>
7	295	1.610	4.693
8	437	1.567	6.450
9	164	1.416	2.203
10	426	2.029	7.304
11	293	<b>0.971</b>	2.798
12	119	1.649	3.136
13	191	2.047	2.741

Tabla 3.10: Errores desglosados de los tiempos calmos del sujeto 010077 utilizando la estrategia de Laplace ( $\alpha = 30$  y  $\nu = 25$ )

Tiempo calmo k	Duración	MAPE en 60 s	MAPE en tiempo calmo completo
2	171	1.275	4.313
3	86	1.156	2.888
5	220	1.862	5.408
6	64	1.684	<b>1.511</b>
7	295	1.584	3.994
8	437	1.384	7.257
9	164	1.285	2.277
10	426	1.305	7.411
11	293	<b>1.001</b>	3.790
12	119	1.308	3.034
13	191	2.503	1.962

En las Tablas 3.11 y 3.12 se muestran los errores desglosados en los tiempos calmos del sujeto 010076 que se obtuvieron en la región de interés en 60 segundos y los errores en todo el tiempo calmo utilizando la estrategia de Hermite y de Laplace respectivamente para cada tabla.

Tabla 3.11: Errores desglosados de los tiempos calmos del sujeto 010076 utilizando la estrategia de Hermite ( $\alpha = 30$  y  $\nu = 25$ )

Tiempo calmo k	Duración	MAPE en 60 s	MAPE en tiempo calmo completo
1	64	1.312	<b>1.429</b>
3	98	2.186	4.196
4	85	4.514	8.516
5	89	3.194	4.075
6	187	1.572	9.554
8	103	2.501	5.819
10	67	2.008	3.334
12	71	3.856	7.603
14	501	1.482	5.822
15	125	<b>1.106</b>	5.816
16	128	2.280	2.797
19	113	4.328	5.690
20	60	2.897	2.750
21	212	3.790	4.785
25	173	1.653	3.786
27	129	3.547	5.850
28	71	1.951	2.622
29	74	2.327	3.991

Tabla 3.12: Errores desglosados de los tiempos calmos del sujeto 010076 utilizando la estrategia de Laplace ( $\alpha = 30$  y  $\nu = 25$ )

Tiempo calmo k	Duración	MAPE en 60 s	MAPE en tiempo calmo completo
1	64	1.312	<b>1.429</b>
3	98	1.882	2.883
4	85	4.424	3.759
5	89	2.849	4.464
6	187	1.709	8.094
8	103	2.790	7.369
10	67	2.107	4.014
12	71	4.393	5.943
14	501	1.651	6.233
15	125	<b>1.106</b>	5.542
16	128	2.077	3.266
19	113	3.627	5.318
20	60	3.273	3.723
21	212	3.625	5.073
25	173	1.392	3.782
27	129	3.508	5.891
28	71	2.333	3.457
29	74	2.547	2.305

En la Figura 3.1 se observan las gráficas de Bland-Altman para  $\alpha = 15$  para los tres sujetos. Se muestran los resultados para el sujeto 010076 utilizando Hermite en a) y Laplace en b); para el sujeto 010077 utilizando Hermite en c) y Laplace en d); para el sujeto 010087 utilizando Hermite en e) y Laplace en f). En la Tabla 3.13 utilizando la estrategia de Hermite y en la Tabla 3.16 utilizando la estrategia de Laplace se observan los valores de MAE, MPE, las medias de las diferencias entre la estimación y la referencia (MOD), los límites de coincidencia (LOA+ y LOA-) del análisis de Bland-Altman y las correlaciones de Spearman (SCC) y Pearson (PCC) entre la estimación y la referencia.

En la Figura 3.2 se observan las gráficas de Bland-Altman para  $\alpha = 20$  para los tres sujetos. Se muestran los resultados para el sujeto 010076 utilizando Hermite en a) y Laplace en b); para el sujeto 010077 utilizando Hermite en c) y Laplace en d); para el sujeto 010087 utilizando Hermite en e) y Laplace en f). En la Tabla 3.14 utilizando la estrategia de Hermite y en la Tabla 3.17 utilizando la estrategia de

Laplace se observan los valores de MAE, MAPE, las medias de las diferencias entre la estimación y la referencia (MOD), los límites de coincidencia (LOA+ y LOA-) del análisis de Bland-Altman y las correlaciones de Spearman (SCC) y Pearson (PCC) entre la estimación y la referencia.

En la Figura 3.3 se observan las gráficas de Bland-Altman para  $\alpha = 30$  para los tres sujetos. Se muestran los resultados para el sujeto sujeto 010076 utilizando Hermite en a) y Laplace en b); para el sujeto 010077 utilizando Hermite en c) y Laplace en d); para el sujeto 010087 utilizando Hermite en e) y Laplace en f). En la Tabla 3.15 utilizando la estrategia de Hermite y en la Tabla 3.18 utilizando la estrategia de Laplace se observan los valores de MAE, MPE, las medias de las diferencias entre la estimación y la referencia (MOD), los límites de coincidencia (LOA+ y LOA-) del análisis de Bland-Altman y las correlaciones de Spearman (SCC) y Pearson (PCC) entre la estimación y la referencia.

Tabla 3.13: Resultados utilizando Hermite  $\alpha = 15$

Sujeto	MAE	MAPE	RMSE	$MOD \pm 1.96\sigma$	LoA+	LoA-	PCC	SCC
010076	3.235	5.263	4.347	$0.936 \pm 8.322$	9.258	-7.387	0.935	0.926
010077	3.183	5.070	4.186	$0.060 \pm 8.206$	8.266	-8.146	0.863	0.866
010087	2.193	4.777	2.969	$-0.989 \pm 5.490$	4.501	-6.479	0.733	0.669

Tabla 3.14: Resultados utilizando Hermite  $\alpha = 20$

Sujeto	MAE	MAPE	RMSE	$MOD \pm 1.96\sigma$	LOA+	LOA-	PCC	SCC
010076	3.228	5.207	4.163	$0.861 \pm 7.985$	8.846	-7.124	0.944	0.935
010077	3.267	5.218	4.271	$0.099 \pm 8.371$	8.470	-8.272	0.862	0.873
010087	2.280	4.992	3.010	$-0.611 \pm 5.778$	5.167	-6.389	0.708	0.648

En la Figura 3.4 se observan las gráficas de Bland-Altman para los tres sujetos de forma global utilizando la estrategia de Hermite con  $\alpha = 15$  en a), con  $\alpha = 20$  en c) y con  $\alpha = 30$  en e) y utilizando la estrategia de Laplace con  $\alpha = 15$  en b), con  $\alpha = 20$  en d) y con  $\alpha = 30$  en f). En la Tabla 3.19 se observan los valores de MAE, MAPE, las medias de las diferencias entre la estimación y la referencia (MOD), los límites de coincidencia (LOA+ y LOA-) del análisis de Bland-Altman y las correlaciones de

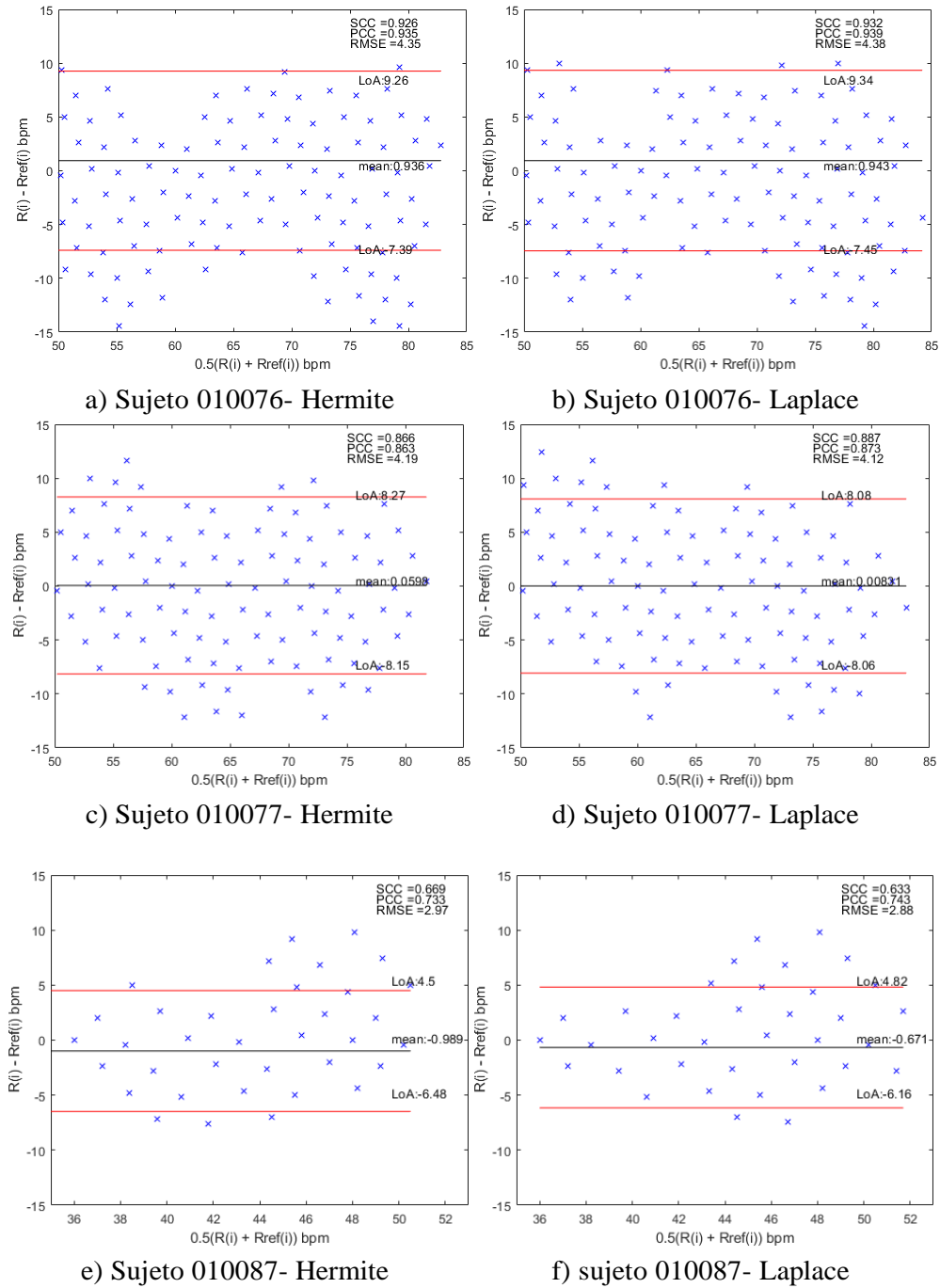


Figura 3.1: Figuras del análisis de Bland-Altman  $\alpha = 15$

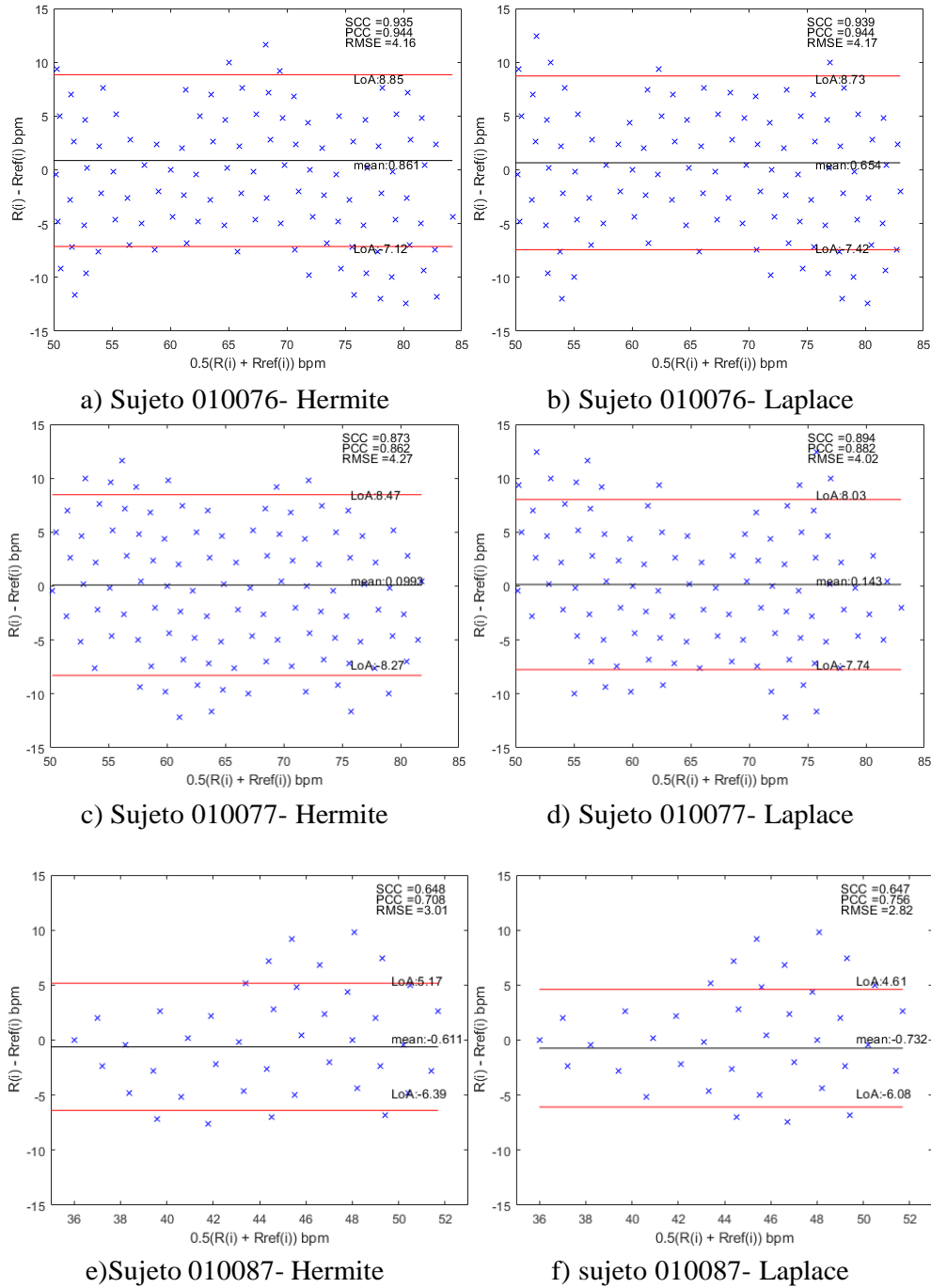


Figura 3.2: Figuras del análisis de Bland-Altman  $\alpha = 20$ .

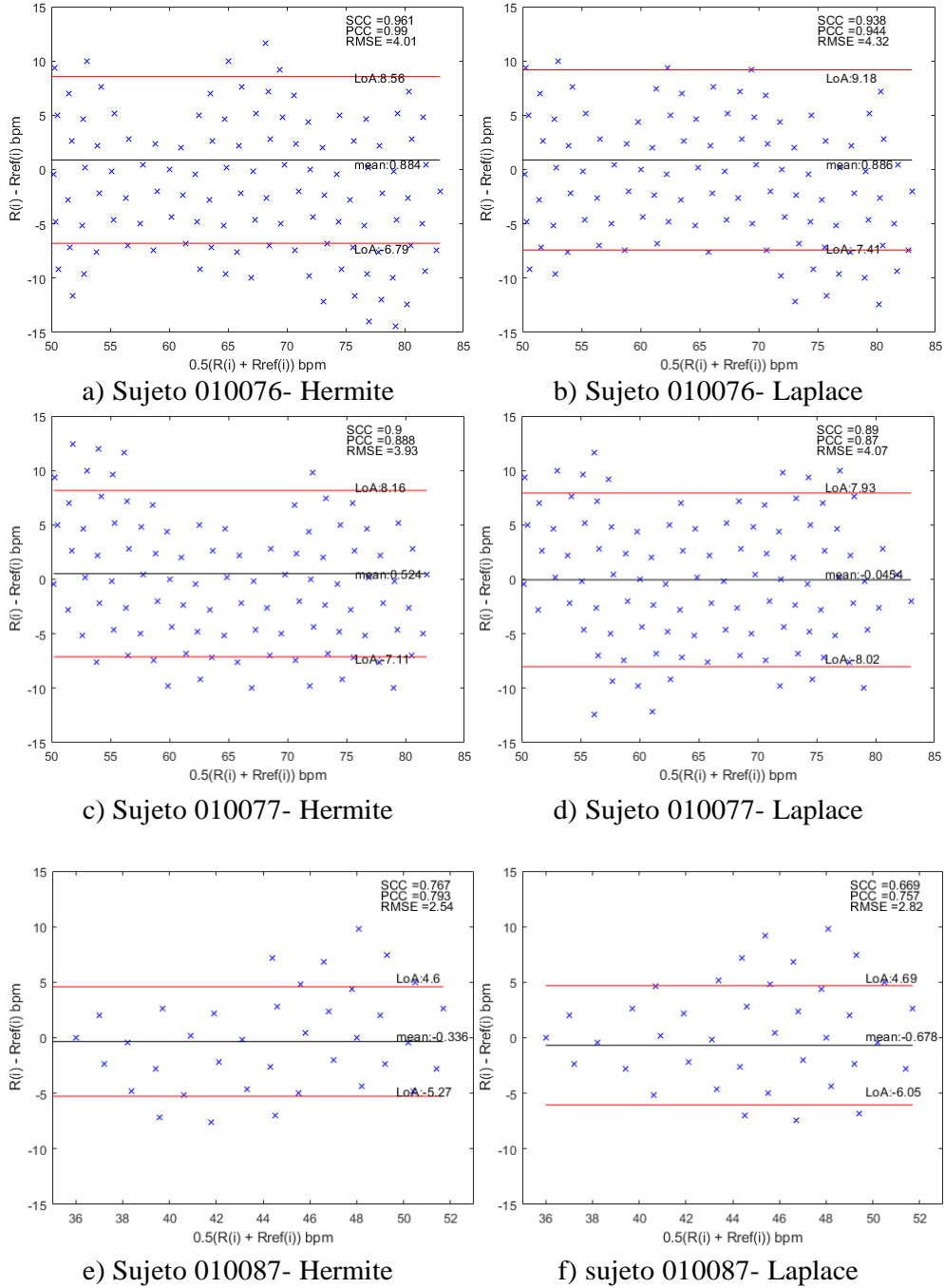


Figura 3.3: Figuras del análisis de Bland-Altman  $\alpha = 30$ .

Tabla 3.15: Resultados utilizando Hermite  $\alpha = 30$ 

Sujeto	MAE	MAPE	RMSE	$MOD \pm 1.96\sigma$	LOA+	LOA-	PCC	SCC
010076	3.351	5.391	4.414	$1.069 \pm 8.396$	9.466	-7.327	0.940	0.931
010077	3.011	4.838	3.932	$0.524 \pm 7.639$	8.163	-7.115	0.888	0.900
010087	1.818	4.020	2.537	$-0.336 \pm 4.931$	4.595	-5.266	0.793	0.767

Tabla 3.16: Resultados utilizando Laplace  $\alpha = 15$ 

Sujeto	MAE	MAPE	RMSE	$MOD \pm 1.96\sigma$	LOA+	LOA-	PCC	SCC
010076	3.228	5.207	4.163	$0.861 \pm 7.985$	8.846	-7.124	0.944	0.935
010077	3.267	5.218	4.271	$0.099 \pm 8.371$	8.470	-8.272	0.862	0.873
010087	2.280	4.992	3.010	$-0.611 \pm 5.778$	5.167	-6.389	0.708	0.648

Tabla 3.17: Resultados utilizando Laplace  $\alpha = 20$ 

Sujeto	MAE	MAPE	RMSE	$MOD \pm 1.96\sigma$	LOA+	LOA-	PCC	SCC
010076	3.194	5.149	4.170	$0.654 \pm 8.075$	8.729	-7.421	0.944	0.939
010077	3.044	4.888	4.023	$0.143 \pm 7.882$	8.025	-7.739	0.882	0.894
010087	2.137	4.616	2.823	$-0.732 \pm 5.347$	4.614	-6.079	0.756	0.647

Tabla 3.18: Resultados utilizando Laplace  $\alpha = 30$ 

Sujeto	MAE	MAPE	RMSE	$MOD \pm 1.96\sigma$	LOA+	LOA-	PCC	SCC
010076	3.306	5.239	4.323	$0.886 \pm 8.296$	9.182	-7.410	0.944	0.938
010077	3.102	4.975	4.066	$-0.045 \pm 7.971$	7.926	-8.017	0.870	0.890
010087	2.166	4.704	2.821	$-0.678 \pm 5.370$	4.691	-6.048	0.757	0.669

Spearman (SCC) y Pearson (PCC) entre la estimación y la referencia para los tres valores de magnificación utilizando la estrategia de Hermite.

En la Tabla 3.20 se observan los valores de MAE, MAPE, las medias de las diferencias entre la estimación y la referencia (MOD), los límites de coincidencia (LOA+ y LOA-) del análisis de Bland-Altman y las correlaciones de Spearman (SCC) y Pearson (PCC) entre la estimación y la referencia para los tres valores de magnificación utilizando la estrategia de Laplace.

Los mejores resultados según las métricas de MAE, MAPE y RMSE se obtuvieron en la estrategia que usa la transformada de Hermite para la magnificación para un valor de  $\alpha = 30$ . Se obtuvieron resultados comparables utilizando la estrategia de

Laplace para un valor de  $\alpha = 20$ . Esto se puede explicar dada la mayor robustez de la transformada de Hermite a los artefactos de reconstrucción debido a su inspiración en el sistema visual humano.

Según el análisis de Bland-Altman, se puede observar que se obtuvo un valor de coincidencia con respecto a la referencia de  $\approx 92\%$  en ambas estrategias.

La correlación entre la estimación y la referencia es fuerte y mayor del 90% para ambas estrategias.

Tabla 3.19: Resultados sobre todos los sujetos utilizando Hermite para los tres factores de magnificación

alfa	MAE	MAPE	RMSE	$MOD \pm 1.96\sigma$	LOA+	LOA-	PCC	SCC
15	2.946	5.061	3.969	<b>0.091 <math>\pm</math> 7.777</b>	<b>7.868</b>	-7.686	0.948	0.949
20	2.999	5.156	3.944	0.179 $\pm$ 7.724	7.902	-7.545	0.950	0.949
30	<b>2.820</b>	<b>4.818</b>	<b>3.810</b>	0.490 $\pm$ 7.407	7.897	<b>-6.917</b>	<b>0.955</b>	<b>0.956</b>

Tabla 3.20: Resultados sobre todos los sujetos utilizando Laplace para los tres factores de magnificación

alfa	MAE	MAPE	RMSE	$MOD \pm 1.96\sigma$	LOA+	LOA-	PCC	SCC
15	2.938	5.013	3.937	0.155 $\pm$ 7.711	7.866	-7.556	0.951	0.950
20	<b>2.862</b>	<b>4.908</b>	<b>3.805</b>	<b>0.093 <math>\pm</math> 7.458</b>	<b>7.551</b>	<b>-7.364</b>	<b>0.954</b>	0.952
30	2.931	4.996	3.882	0.113 $\pm$ 7.605	7.718	-7.493	0.952	<b>0.953</b>

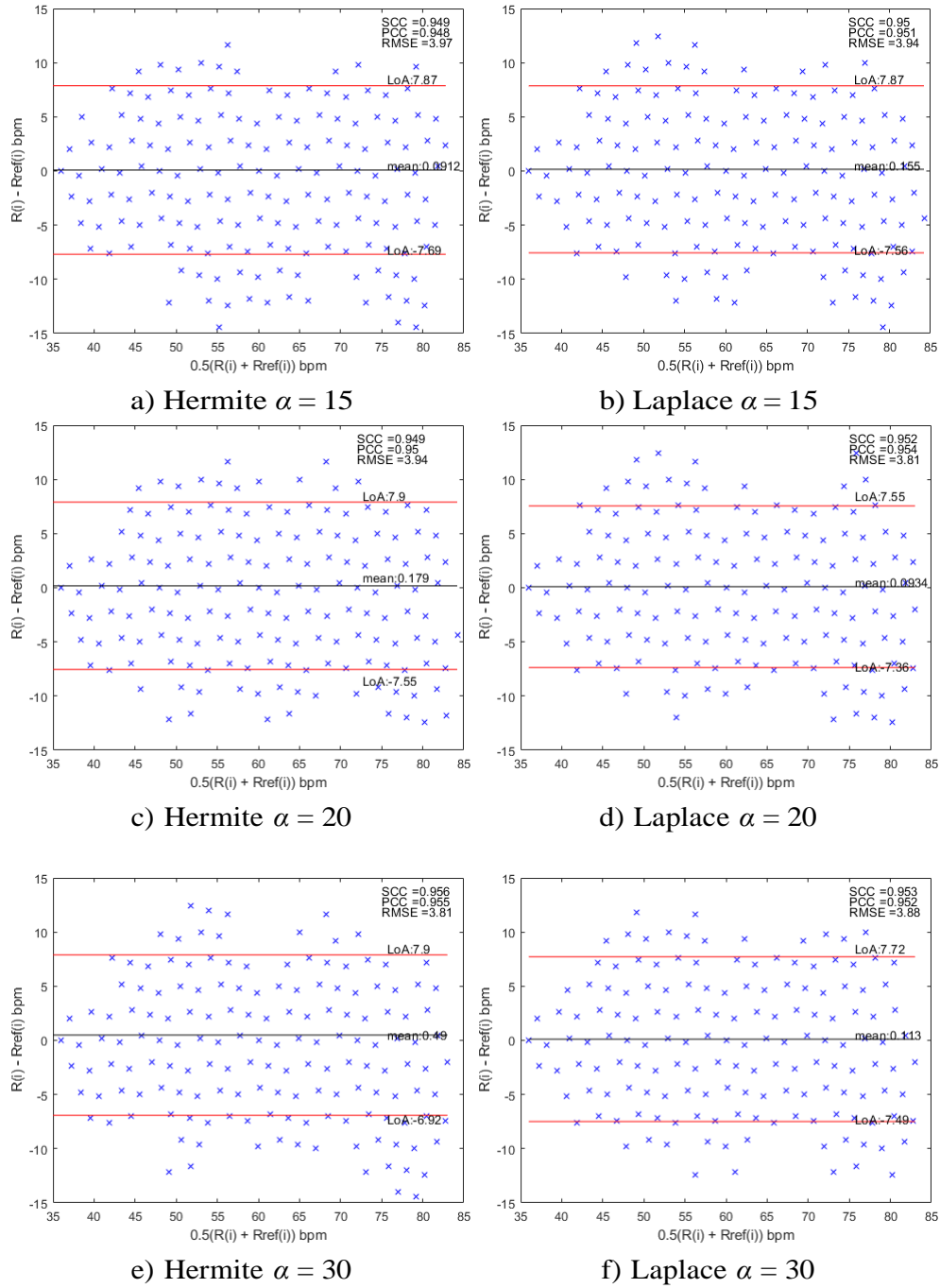


Figura 3.4: Análisis de Bland-Altman para todos los sujetos utilizando la estrategia de Laplace y la estrategia de Hermite

# Capítulo 4

## Discusión

En la base de datos utilizada, los sujetos son posicionados en diferentes espacios, la iluminación y los elementos grabados (incubadora, equipo clásico de monitor de signos vitales, electrodos, cables, etc.) son diferentes. Por otro lado, la cámara se colocó en distintos lugares así que la distancia focal, el ángulo de ataque y su enfoque se modifican en cada sujeto. Los resultados obtenidos a partir de la metodología propuesta muestran robustez a los cambios de iluminación y diferentes distancias de la cámara.

A diferencia de otros autores, utilizamos tres sujetos recién nacidos en sesiones de adquisición bastante largas y no se ocupó una vestimenta especial en los sujetos para mejorar la adquisición del video. Además, las condiciones ambientales no fueron controladas sino en condiciones reales.

Otros autores no mencionan como seleccionan los datos ni como excluyen grabaciones de los sujetos con los que trabajan. Nosotros seleccionamos aquellos tiempos de las grabaciones donde los bebés permanecen calmados, independientemente de la calidad y los cambios de iluminación en las imágenes adquiridas por la cámara.

Mientras que otros autores, seleccionan las frecuencias de corte, nosotros damos un rango de frecuencias mediante la información que se obtiene de la evaluación de los primeros 60 segundos de la referencia lo que hace nuestra metodología adaptativa a las variaciones frecuenciales de la señal respiratoria. Por otro lado, en esta metodología a diferencia de las propuestas anteriormente publicadas, se utilizan filtros después de

la magnificación para mejorar la calidad de la señal.

En la Tabla 4.1 se visualizan las diferentes parametrizaciones que utilizan las metodologías de otros autores y la de nuestra propuesta. Podemos notar similitudes en la duración de las ventanas, sin embargo, variaciones en los coeficientes de magnificación ( $\alpha$ ) los cuales difieren debido a la naturaleza de los datos en los videos y a la metodología elegida. En nuestro trabajo optamos por elegir frecuencias de corte dinámicas conforme al análisis del video, por lo cual nos permite adaptarnos mejor a las necesidades de cada video.

Tabla 4.1: Comparación de parametrizaciones de magnificaciones

Publicación	Ventana	Alpha	Frecuencia de corte Hz	Cuadros por segundo
(Alinovi et al., 2015)	20	1, 3.75, 6 9.25 y 12	0.25 a 1.05	25
(Koolen et al., 2015)	5, 10 y 20	15	0.5 a 2	-
(Antognoli et al., 2018)	-	-	0.5 a 4	30
(Cattani et al., 2017)	10	-	-	25
Nuestra propuesta	5, 15, 20, 25	15, 20, 30	$*F_{Ck} - *w$ a $*F_{Ck} + *w$	25

$F_{Ck}$  = Frecuencia de corte en el tiempo calmo  $k$ , donde el mayor  $F_{Ck}$  es 1.316Hz y el menor es 0.633Hz

$w = 0.083$  Hz

\* En Hertz.

En la Tabla 4.2 se comparan las metodologías propuestas por otros autores que utilizan la técnica de magnificación, se puede apreciar la duración de análisis, es decir la duración máxima de las partes analizadas de los videos individualmente, la superposición de las ventanas y la manera de seleccionar la región de interés.

Tabla 4.2: Comparación de metodologías

Autores	Duración del análisis	Superposición de ventanas	Región de interés
(Koolen et al., 2015)	-	-	Manual
(Antognoli et al., 2018)	Hasta 10''	-	Manual
(Alinovi et al., 2018)	-	50 %	Automática
Nuestra propuesta	Hasta 13'33''	80 %, 93 %, 95 % y 96 %	Automática

En la Tabla 4.3 se comparan los resultados obtenidos en diferentes metodologías mediante diferentes tipos de métricas utilizadas por los autores. Se puede observar

que los mejores resultados para la raíz del error cuadrático medio se obtuvieron en la metodología de (Alinovi et al., 2018). Empero, debe considerarse: la duración de los videos proporcionada en la Tabla 4.2 y los resultados de las otras métricas para poder tener un análisis más detallado del desempeño de las metodologías.

Tabla 4.3: Comparación de resultados de las metodologías

<b>Autores</b>	<b>Método</b>	<b>MAE</b>	<b>MAPE</b>	<b>Corr, PCC, SCC</b>	<b>RSME</b>	<b>MOD <math>\pm 1.96\sigma</math></b>	<b>LOA+</b>	<b>LOA-</b>
(Koolen et al., 2015)	Sumatoria de flujo óptico	-	-	0.941	-	-	-	-
	Análisis de componente independiente	-	-	0.928	-	-	-	-
	Análisis de componente principal	-	-	0.939	-	-	-	-
(Antognoli et al., 2018)	Extracción de señal dominante en el tórax	-	-	-	7.6	-	-	-
(Alinovi et al., 2018)	Magnificación de movimiento local	-	-	-	0.05	-	-	-
Nuestra propuesta	Magnificación Hermitiana y promedio espacial	2.820	4.818	$\bar{-}$ , 0.955, 0.956	3.810	0.490 +7.407	7.897	-6.917
	Magnificación Laplaciana y promedio espacial	2.862	4.908	$\bar{-}$ , 0.954, 0.952	3.805	0.093 +7.458	7.551	-7.364

Nuestra propuesta es validada con videos más largos y se calculan más métricas de evaluación. Las condiciones de adquisición no son controladas y tanto las frecuencias de corte como la región de interés son adaptativas a cada tiempo calmo de cada bebé. Nuestra metodología está pensada para adaptarse a las posibles variaciones de la señal respiratoria en diferentes tiempos calmos.

Computacionalmente el método es costoso, sobre todo requiere mucho tiempo

calcular la magnificación y la búsqueda de la región de interés, sin embargo, es posible mejorar el desempeño si se eficientiza el algoritmo de búsqueda, se magnifica solamente esa parte del video o se opta por un lenguaje de programación más veloz.

El algoritmo fue programado en MATLAB R2019a y ejecutado en un ordenador con: un procesador Intel(R) Core(TM) i7-10700 CPU @2.90GHz de 8 núcleos con 16 procesadores lógicos y una RAM de 64.0 GB.

# Capítulo 5

## Conclusiones

En este trabajo se propone una metodología para estimar la frecuencia respiratoria en condiciones reales de adquisición en neonatos en una unidad de terapia intensiva. El método se basa en la magnificación de movimiento utilizando la transformada de Hermite ya que, éste está basado en el modelo del sistema visual humano y nos permite compararlo con la estrategia de magnificación clásica utilizando la estrategia de Laplace. Nuestra propuesta calcula la región de interés de manera automática y permite cambiar parámetros para los diferentes tiempos calmos de manera adaptativa.

Observamos que la estrategia con la transformada de Hermite tiene un buen desempeño gracias a su inspiración en el sistema visual humano y es una alternativa a la Laplace. Los mejores resultados del algoritmo propuesto utilizando máscaras cuadradas de 20 x 20 píxeles se obtuvieron con una parametrización del coeficiente de magnificación  $\alpha = 30$  y una ventana de 25 segundos ( $v = 25$ ) para la estrategia utilizando Hermite y con un coeficiente de magnificación de  $\alpha = 20$  y una ventana de 25 segundos ( $v = 25$ ) para la estrategia de magnificación utilizando Laplace. Según el análisis de Bland-Altman, se puede observar que se obtuvo un valor de coincidencia con respecto a la referencia de  $\approx 92\%$  en ambas estrategias.

# Capítulo 6

## Sugerencias para trabajos futuros

Es posible implementar la metodología estudiada en este trabajo para obtener las etiquetas y las regiones de interés para entrenar a una red neuronal, la cual sea capaz de identificar la región de interés aunado a la estimación de la frecuencia respiratoria.

Experimentar con diferentes tipos y tamaños de máscaras podría mejorar el desempeño del algoritmo, así como entender si es posible que existan movimientos relacionados, que se encuentren en máscara no vecinas que permitan un mejor desempeño en la estimación.

También es posible investigar sobre qué tipo de algoritmo de búsqueda tiene mejor desempeño al momento de encontrar la región de interés, y dada esa información realizar un clasificador de regiones de interés para que en futuros trabajos se conozca con mayor facilidad y se comprenda las características que comparten las zonas donde es posible estimar la frecuencia respiratoria con errores mínimos.

Considerar diferentes parametrizaciones y agregar heurísticas a las magnificaciones podría develar el movimiento característico de la frecuencia respiratoria en los videos de mejor manera.

Considerar un aumento del espectro de color capturado por la cámara, pasar de 8 bits a 12 bits o 14 bits, para evaluar y estudiar si la región de interés mejora. Así como, eliminar áreas y frecuencias que no son de interés de una manera más rigurosa.

# Bibliografía

- Alinovi, D., Cattani, L., Ferrari, G., Pisani, F., and Raheli, R. (2015). Spatio-temporal video processing for respiratory rate estimation. In *2015 IEEE International Symposium on Medical Measurements and Applications, MeMeA 2015, Torino, Italy, May 7-9, 2015*, pages 12–17. IEEE.
- Alinovi, D., Ferrari, G., Pisani, F., and Raheli, R. (2018). Respiratory rate monitoring by video processing using local motion magnification. In *26th European Signal Processing Conference, EUSIPCO 2018, Roma, Italy, September 3-7, 2018*, pages 1780–1784. IEEE.
- Antognoli, L., Marchionni, P., Nobile, S., Carnielli, V. P., and Scalise, L. (2018). Assessment of cardio-respiratory rates by non-invasive measurement methods in hospitalized preterm neonates. In *2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pages 1–5. IEEE.
- Aoki, H., Koshiji, K., Nakamura, H., Takemura, Y., and Nakajima, M. (2005). Study on respiration monitoring method using near-infrared multiple slit-lights projection. pages 291–296. IEEE.
- Baaziz, N. and Labit, C. (1990). Laplacian pyramid versus wavelet decomposition for image sequence coding. In *International Conference on Acoustics, Speech, and Signal Processing*, pages 1965–1968 vol.4.
- Bartula, M., Tigges, T., and Muehlsteff, J. (2013). Camera-based system for contactless monitoring of respiration. In *35th Annual International Conference of the*

- IEEE Engineering in Medicine and Biology Society, EMBC 2013, Osaka, Japan, July 3-7, 2013*, pages 2672–2675. IEEE.
- Brieva, J., Moya-Albor, E., and Ponce, H. (2021). A non-contact spo2 estimation using a video magnification technique. In *17th International Symposium on Medical Information Processing and Analysis*, volume 12088, pages 10–18. SPIE.
- Cabon, S., Porée, F., Simon, A., Rosec, O., Pladys, P., and Carrault, G. (2019a). Video and audio processing in paediatrics: a review. *Physiological measurement*, 40(2):02TR02.
- Cabon, S., Porée, F., Simon, A., Met-Montot, B., Pladys, P., Rosec, O., Nardi, N., and Carrault, G. (2019b). Audio- and video-based estimation of the sleep stages of newborns in neonatal intensive care unit. *Biomedical Signal Processing and Control*, 52:362–370.
- Cattani, L., Alinovi, D., Ferrari, G., Raheli, R., Pavlidis, E., Spagnoli, C., and Pisani, F. (2017). Monitoring infants by automatic video processing: A unified approach to motion analysis. *Comp. in Bio. and Med.*, 80:158–165.
- Eichenwald, E. C. (2016). Apnea of prematurity. *Pediatrics*, 137.
- Geertsema, E. E., Visser, G. H., Sander, J. W., and Kalitzin, S. N. (2020). Automated non-contact detection of central apneas using video. *Biomed. Signal Process. Control.*, 55.
- Ghassemlooy, Z., Ng, W. P., of Northumbria at Newcastle., U., Engineering, and of Northumbria at Newcastle. School of Computing, I. S. U. (2010). *CSNDSP 2010 : 2010 7th International Symposium on Communication Systems, Networks Digital Signal Processing : 21-23 July 2010, Northumbria University, Newcastle upon Tyne, United Kingdom*. IEEE.
- Janssen, R., Wang, W., Moço, A., and de Haan, G. (2016). Video-based respiration monitoring with automatic region of interest detection. *Physiological Measurement*, 37(1):100.

- Koolen, N., Decroupet, O., Dereymaeker, A., Jansen, K., Vervisch, J., Matic, V., Vanrumste, B., Naulaers, G., Huffel, S. V., and Vos, M. D. (2015). Automated respiration detection from neonatal video data. In Marsico, M. D., Figueiredo, M. A. T., and Fred, A. L. N., editors, *ICPRAM 2015 - Proceedings of the International Conference on Pattern Recognition Applications and Methods, Volume 2, Lisbon, Portugal, 10-12 January, 2015*, pages 164–169. SciTePress.
- Kouamou, G., Ferrari, G., Lofino, F., Raheli, R., and Pisani, F. (2011). Extraction of video features for real-time detection of neonatal seizures. In *12th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks, WOWMOM 2011, Lucca, Italy, 20-24 June, 2011*, pages 1–6. IEEE Computer Society.
- Liu, H., Allen, J., Zheng, D., and Chen, F. (2019). Recent development of respiratory rate measurement technologies.
- Lukac, T., Pucik, J., and Chrenko, L. (2014). Contactless recognition of respiration phases using web camera.
- Martens, J.-B. (1990). The Hermite Transform-Theory. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38(9):1595–1606.
- Morley, C. J., Thornton, A. J., Fowler, M. A., Cole, T. J., Hewson, P. H., Fowler, C. J. M. A. J. T. M. A., and Hewson, A. P. H. (1990). Respiratory rate and severity of illness in babies under 6 months old 1 1 4 7 +1+4 content crying sleeping content crying sleeping.
- Ntonfo, G. M. K. (2014). Monitoring and diagnosing neonatal seizures by video signal processing. *ELCVIA Electronic Letters on Computer Vision and Image Analysis*, 13(2):58–60.
- Ntonfo, G. M. K., Ferrari, G., Raheli, R., and Pisani, F. (2012). Low-complexity image processing for real-time detection of neonatal clonic seizures. *IEEE Trans. Inf. Technol. Biomed.*, 16(3):375–382.

- Pisani, F., Spagnoli, C., Pavlidis, E., Facini, C., Ntonfo, G. M. K., Ferrari, G., and Raheli, R. (2014). Real-time automated detection of clonic seizures in newborns. *Clinical neurophysiology*, 125(8):1533–1540.
- Sun, Y., Kommers, D., Wang, W., Joshi, R., Shan, C., Tan, T., Aarts, R. M., van Pul, C., Andriessen, P., and de With, P. H. N. (2019). Automatic and continuous discomfort detection for premature infants in a NICU using video-based motion analysis. In *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2019, Berlin, Germany, July 23-27, 2019*, pages 5995–5999. IEEE.
- Tarassenko, L., Villarroel, M., Guazzi, A., Jorge, J., Clifton, D. A., and Pugh, C. (2014). Non-contact video-based vital sign monitoring using ambient light and auto-regressive models. *Physiological Measurement*, 35(5):807.
- Villarroel, M., Guazzi, A., Jorge, J., Davis, S., Watkinson, P., Green, G., Shenvi, A., McCormick, K., and Tarassenko, L. (2014). Continuous non-contact vital sign monitoring in neonatal intensive care unit. *Healthcare technology letters*, 1(3):87–91.
- Wu, H.-Y., Rubinstein, M., Shih, E., Guttag, J., Durand, F., and Freeman, W. T. (2012). Web page: Eulerian Video Magnification for Revealing Subtle Changes in the World. <http://people.csail.mit.edu/mrub/evm/>. [accessed 13.05.2015].
- Zhao, F., Li, M., Qian, Y., and Tsien, J. Z. (2013). Remote measurements of heart and respiration rates for telemedicine. *PLoS ONE*, 8(10):e71384.