

Estimación Óptica Remota de la Actividad y Frecuencia Respiratoria durante Diversas Maniobras Respiratorias

J. A. Vazquez-Segura^{1,*}, A. R. Mejía-Rodríguez¹, B. A. Reyes¹

¹ Facultad de Ciencias, Universidad Autónoma de San Luis Potosí, S.L.P., México.

* j.agustin_vazquez@hotmail.com

Resumen— El monitoreo de la frecuencia respiratoria es relevante para predecir oportunamente eventos adversos, y tradicionalmente se realiza con dispositivos especializados que son difícilmente trasladables fuera de ambientes clínicos y de investigación. En este trabajo se exploran algoritmos de procesamiento de imágenes y señales para estimar la frecuencia respiratoria instantánea (FRI) a lo largo de diversas maniobras respiratorias en estado de reposo, mediante la estimación óptica remota (EOR). La señal de respiración fue extraída de una secuencia de video obtenida de un smartphone y comparada con la señal de respiración obtenida con un termistor conectado al Biopac. Se encontró que ambas señales están altamente correlacionadas linealmente ($\rho > 0.8$) para la mayoría de las maniobras. En cuanto a la FRI, el análisis de Bland-Altman reveló un sesgo de -0.05 rpm y límites de concordancia entre -4.25 y 4.15 rpm entre la estimación con la EOR y el termistor.

Palabras clave— estimación óptica remota, frecuencia respiratoria instantánea, análisis tiempo-frecuencia.

I. INTRODUCCIÓN

La respiración es el proceso de entrada y salida del aire en los pulmones, y para valorarla es necesario verificar la frecuencia, profundidad, y eficacia de las respiraciones. La frecuencia respiratoria (FR), uno de los signos vitales, se mide contando el número de respiraciones por minuto (rpm), y en condiciones normales es de 12 a 20 rpm en una persona adulta [1]. Se ha demostrado que una FR anormal (mayor a 27 rpm) puede ser predictor importante de acontecimientos adversos, como un paro cardíaco [2]. Sin embargo, a pesar de su relevancia clínica, la FR a menudo no es registrada, incluso cuando el problema primario de los pacientes es una condición respiratoria [2].

Existen diversos métodos para estimar la FR. Por ejemplo, suele emplearse la medición de la impedancia en el tórax, pero pueden obtenerse mediciones inexactas debido al movimiento del paciente y la obstrucción de vías respiratorias [3]. Otro método para obtener la FR es por medio de la capnografía, que cuenta con la desventaja de estar implementada solo en monitores de signos vitales de unidades de cuidados intensivos y quirófono [3], lo que dificulta poder monitorear a los pacientes después de que fueron intervenidos quirúrgicamente, y que pueden seguir bajo el efecto de la anestesia.

Un enfoque más reciente para el monitoreo cardiorrespiratorio de fácil acceso y bajo costo es el uso de

smartphones, donde se ha demostrado que es posible estimar a través de la cámara de un celular las frecuencias respiratorias y cardíacas promedio [4]. En particular, se ha empleado una estimación óptica remota (EOR) para medir cambios en la intensidad de la luz reflejada debido a los movimientos respiratorios de las cavidades torácicas durante la inhalación y exhalación de una persona en estado de reposo [4]. Es importante mencionar que la mayoría de los esfuerzos realizados involucran únicamente la estimación de la FR promedio.

En este trabajo, se proponen los siguientes dos objetivos:

- 1) mostrar que es factible obtener un subrogado de la señal respiratoria, adquirida con un sensor especializado, usando una cámara de celular y técnicas de procesamiento de imágenes, y 2) mostrar que es factible estimar la FR instantánea para dar seguimiento a la misma en diferentes maniobras respiratorias, que incluyen respiraciones espontáneas, a frecuencias fijas y forzadas.

II. METODOLOGÍA

A. Protocolo experimental.

Se adquirieron datos de nueve ($N=9$) sujetos voluntarios (5 hombres y 4 mujeres), de entre 21 y 25 años, (media \pm desviación estándar: 22.4 ± 1.4 años), 66.6 ± 12.4 kg, y 163.9 ± 6.4 cm, sanos y sin enfermedades cardiorrespiratorias. Previo a la adquisición de los datos, se obtuvo su consentimiento informado por escrito de acuerdo con los Protocolos de Helsinki.

Durante el protocolo se adquirió la señal de temperatura (termistor Vernier, STS-BTA) a una frecuencia de muestreo (f_s) de 1 kHz, mediante el empleo del hardware BIOPAC MP150®. Además, se adquirieron secuencias de video utilizando la videocámara de un smartphone (Samsung Galaxy S3, Samsung Electronics, Seúl, Corea del Sur) a partir de las cuales se estimó la actividad respiratoria mediante la EOR. La resolución empleada fue de 1920×1080 píxeles @ 30 cuadros-por-segundo (fps), como se muestra en la Fig. 1.

Cada sujeto voluntario realizó la siguiente secuencia de maniobras respiratorias:

- 1) *Respiraciones espontáneas*; por un periodo de 2 minutos.
- 2) *Respiraciones controladas a una frecuencia fija* por un periodo de 2 minutos; en un rango de 6 a 36 rpm en intervalos de 6 rpm.

- 3) *Maniobra respiratoria controlada*; cada sujeto inició con respiraciones tenues a una FR alta, y gradualmente fue más profunda a la par que se disminuía la FR, para finalmente volver a decrecer progresivamente la profundidad de las respiraciones e incrementar la FR. La duración de esta maniobra dependió del ritmo y profundidad de la respiración del voluntario.

B. Procesamiento de las señales.

Antes de realizar el procesamiento de las señales, se realizó la edición de los videos, con el software VideoPad Video Editor (v6.10, NCH Software, Canberra, Australia), con la finalidad de alinear los tiempos de adquisición entre la señal obtenida por el termistor, y los videos en cada una de las maniobras. El procesamiento de las señales se realizó en MATLAB® (R2017a, The Mathworks, Inc., MA, USA).

B.1 Obtención de la señal de respiración de los videos y de la señal de referencia

Para obtener la señal de respiración de los videos, se definió una región de interés (ROI) rectangular enfocada en la sección alta del tórax de cada sujeto. En esta ROI se calculó la intensidad promedio de los pixeles en los canales rojo, verde y azul (RGB, por sus siglas en inglés) fotograma por fotograma. Como resultado del análisis de la intensidad, se obtuvo la forma de onda de la señal de respiración correspondiente a cada maniobra, y se calculó el promedio (i_{Resp}) de las tres intensidades RGB por cada fotograma del video, i.e., en cada instante de tiempo (t), de acuerdo con la siguiente ecuación:

$$i_{Resp}(t) = \left(\frac{1}{3D}\right) \left(\sum_{\{m,n \in ROI\}} i_R(m,n,t) + i_G(m,n,t) + i_B(m,n,t) \right) \quad (1)$$

donde $i_r(m,n,t)$ es la intensidad evaluada en la m -ésima fila y n -ésima columna del canal rojo (R), verde (G) o azul (B), dentro de la ROI que contiene un total de D pixeles [5].

Cabe mencionar que también se adquirieron y analizaron señales de respiración obtenidas con un sensor piezoeléctrico. Sin embargo, se observó que dicha señal es muy susceptible a los movimientos corporales no relacionados con la respiración [6]. Así, la señal obtenida con el termistor se consideró como la señal respiratoria de referencia ($Resp_{term}$).

B.2. Remuestreo y filtrado

Se uniformizó la frecuencia de muestreo de ambas señales (i_{Resp} y $Resp_{term}$) a 25 Hz, debido a que las señales obtenidas por el sistema BIOPAC® fueron muestreadas a una frecuencia de 1 kHz, y los videos estaban muestreados no uniformemente entre los 15 Hz y los 30 Hz.

Posteriormente, se filtraron ambas señales empleando un filtro digital pasabandas con respuesta al impulso finita (FIR)

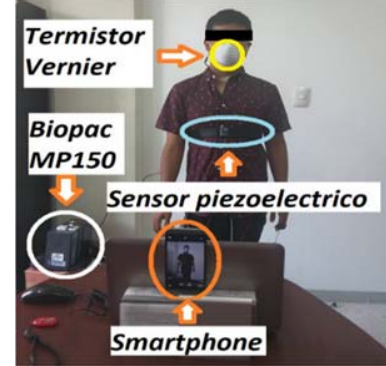


Fig. 1. Protocolo de adquisición.

de orden 20, con frecuencias de corte de 0.05 y 2.1 Hz, esto después de analizar el espectro en frecuencia de las señales respiratorias de referencia. Finalmente, se removió el valor medio de cada una de las señales respiratorias.

B.3. Alineación de señales y eliminación de tendencia

La alineación de las señales del termistor y de la videocámara se realizó con el propósito de corregir los distintos tiempos de inicio de ambas señales. Para ello, se empleó la función de correlación cruzada entre ellas, ajustando de acuerdo con el retraso o adelanto correspondiente al valor máximo de correlación.

Para eliminar la tendencia de las señales, se empleó la técnica de reconstrucción empírica de modos (EMD, Empirical Mode Decomposition), caracterizada por identificar los modos de oscilación intrínsecos (IMF) en los datos en sus escalas de tiempo [7]. Así, una señal arbitraria $s(t)$ puede representarse como

$$s(t) = \sum_{k=1}^K IMF_k(t) + r_k(t) \quad (2)$$

donde K es el número total de IMFs, y $r_k(t)$ es la señal de residuo. En nuestro caso, dicho residuo o tendencia fue removida de ambas señales para facilitar su posterior análisis.

Finalmente, para minimizar el efecto de artefactos de movimiento externos a la maniobra se descartaron los extremos, y se empleó la porción central donde eran más estables.

C. Similitud entre i_{Resp} y $Resp_{term}$.

Se calculó el coeficiente de correlación (ρ) entre ambas señales (i_{Resp} y $Resp_{term}$) de acuerdo con la ecuación

$$\rho = \frac{Cov(Resp_{term}, i_{Resp})}{\sqrt{Var(Resp_{term})} * \sqrt{Var(i_{Resp})}} \quad (3)$$

Además, se aplicó el análisis de regresión lineal en cada una de las maniobras respiratorias.

D. Estimación de la frecuencia respiratoria instantánea

La estimación de la FR a lo largo de la maniobra se realizó empleando una representación tiempo-frecuencia, es decir, el espectrograma (SP) definido como la magnitud al cuadrado de la transformada de Fourier de tiempo corto (STFT) [8] de acuerdo a la ecuación

$$STFT = \left| \frac{1}{\sqrt{2\pi}} \int s(\tau) h(\tau - t) e^{-j\omega t} d\tau \right|^2 \quad (4)$$

donde $h(t)$ representa la ventana de análisis empleada para enfocarse una zona temporal y espectral de la señal $s(t)$. Así, primero se obtuvo el espectro de potencia de la señal para extraer su ancho de banda, de tal forma que la FR instantánea (FRI) fue extraída de la imagen SP como la frecuencia que corresponde al valor de intensidad máximo para cada instante de tiempo, dentro de los límites en frecuencia establecidos, resultando en una nueva serie de tiempo con la evolución temporal de la FR a lo largo de la maniobra.

Se empleó el error cuadrático medio normalizado (NRMSE) entre ambas señales FRI (obtenidas del termistor y del smartphone). Además, se calculó el índice de correlación 2D entre ambas imágenes SP, así como un análisis Bland-Altman para todas las maniobras.

III. RESULTADOS Y DISCUSIÓN

Las señales adquiridas con ambos sistemas (smartphone: *iResp* y termistor: *Respterm*), presentaban cierta tendencia no lineal, como se muestra en la Fig. 2. Por ello resulta útil el empleo de la técnica EMD para eliminar dicha tendencia. Por su parte, *iResp* (obtenida como el promedio RGB) presentó mejores resultados que las señales de los canales individuales ya que obtiene características morfológicas de las tres intensidades, que por sí solas no tienen. Adicionalmente, dicha definición de *iResp* permitió obtener resultados adecuados independientemente del color de la playera utilizada por los voluntarios. En la Fig. 3 se pueden observar las señales preprocesadas, percibiendo la eliminación de la tendencia, y una alta similitud al compararlas entre sí en cuanto a sus ciclos respiratorios a lo largo del tiempo.

Con respecto a la similitud entre ambas señales, en la Tabla I se muestran los resultados de la regresión lineal para cada maniobra (media \pm desviación estándar) en términos de la pendiente m , intercepto b y coeficiente de determinación r^2 . Los valores de r^2 mayores a 0.8 indican que señal *iResp* se correlaciona fuertemente con la señal *Respterm*. Dichos resultados se corroboran con la distribución del índice de correlación lineal entre las señales *iResp* y *Respterm* por maniobra, mostrados en la Fig. 4. Cabe mencionar que para las maniobras con respiración a una frecuencia variable y espontánea se obtuvo el desempeño más bajo, donde los sujetos pudieron perder el cuidado de mantenerse sin movimientos ante la cámara, y perder el estado de reposo.

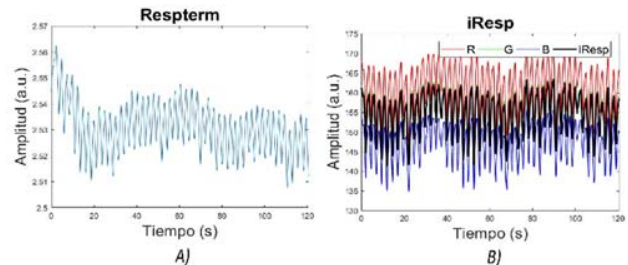


Fig. 2. Señales adquiridas. A) *Respterm* (termistor). B) *iResp* (smartphone).

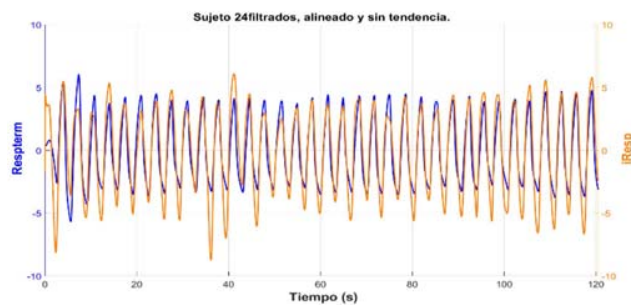


Fig. 3. Señales de respiración *iResp* y *Respterm* procesadas (filtradas, alineadas y sin tendencia).

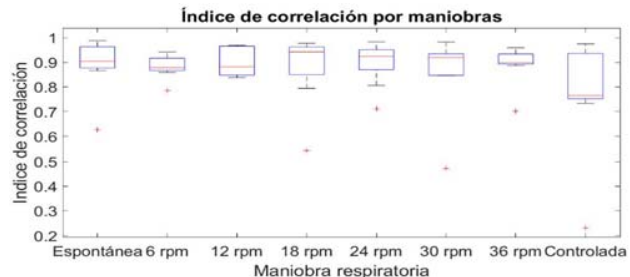


Fig. 4. Distribución del índice de correlación lineal (ρ) entre las señales de respiración *iResp* y *Respterm*

TABLA I. RESULTADOS DE LA REGRESIÓN LINEAL ENTRE LAS SEÑALES IRESP Y RESPTERM POR MANIOBRA. N=9.

Maniobra Respiratoria	M	b	r^2
Esponánea	1.0 \pm 0.7	0.0 \pm 0.0	0.6 \pm 0.3
Fija a 6 rpm	1.4 \pm 1.4	0.2 \pm 0.3	0.8 \pm 0.1
Fija a 12 rpm	1.5 \pm 1.6	-0.2 \pm 0.3	0.8 \pm 0.1
Fija a 18 rpm	2.0 \pm 2.2	0.0 \pm 0.1	0.8 \pm 0.2
Fija a 24 rpm	1.8 \pm 1.7	0.0 \pm 0.1	0.8 \pm 0.2
Fija a 30 rpm	2.0 \pm 2.4	0.0 \pm 0.1	0.8 \pm 0.2
Fija a 36 rpm	2.2 \pm 2.4	0.0 \pm 0.1	0.8 \pm 0.1
Controlada	2.0 \pm 2.2	0.1 \pm 0.5	0.8 \pm 0.2

Para el valor la pendiente m se encuentran valores muy diversos que pueden estar influenciados por factores como la distancia, la región utilizada para obtener la señal *iResp*, la inclinación de la cámara, y las condiciones de iluminación variables.

Con respecto a la estimación de la FRI, en la Fig. 5 se muestran los espectrogramas obtenidos a partir de las señales *Respterm* e *iResp*, en donde es posible observar que existe una alta similitud entre ambas representaciones tiempo-frecuencia (valores altos de correlación 2D), la cual se ve reflejada en los bajos valores de NRMSE de las frecuencias instantáneas extraídas (ver Tabla 2). Dichos valores muestran que la estimación de la FRI por medio de la cámara del smartphone (*iResp*) es factible, ya que indica un favorable seguimiento a la FRI medida por la señal del termistor (*Respterm*). Con respecto al NRMSE se observa un valor alto en la maniobra de 6 rpm, pudiendo deberse a la dificultad de mantener respiraciones a frecuencias bajas por parte de los sujetos voluntarios, mismo que se va reduciendo conforme aumenta la FR. En la maniobra espontánea se encontraron resultados individuales que son muy altos, debido a que los sujetos, al no tener su atención en el metrónomo que marcó el paso de la respiración, tendieron a realizar movimientos y no mantenerse en reposo, mismo que afectó al cálculo del error. Finalmente, para el análisis Bland-Altman se obtienen resultados aceptables con un sesgo de -0.05 rpm y límites de concordancia entre -4.25 y 4.15 rpm entre ambos sistemas (ver Fig. 6). La regresión lineal entre todas las FRI calculadas, resultó en la ecuación $FRI_{smartphone} = 0.98FRI_{termistor} + 0.371$, y el valor de $r^2=0.9561$, en un total de $n=67537$ puntos.

IV. CONCLUSIONES

A partir de los resultados obtenidos, se corroboró que la señal de respiración obtenida por medio de la técnica de estimación óptica remota con un smartphone es un subrogado de la señal de respiración adquirida por el termistor, tomando en cuenta la alta similitud entre ambas señales en las diferentes maniobras respiratorias realizadas en este trabajo. Además, también se obtuvo un seguimiento adecuado de la frecuencia respiratoria a lo largo de las diferentes maniobras respiratorias y no sólo de su valor promedio.

A pesar de los resultados obtenidos, aún se tienen varias limitantes, por ejemplo, degradación del desempeño del

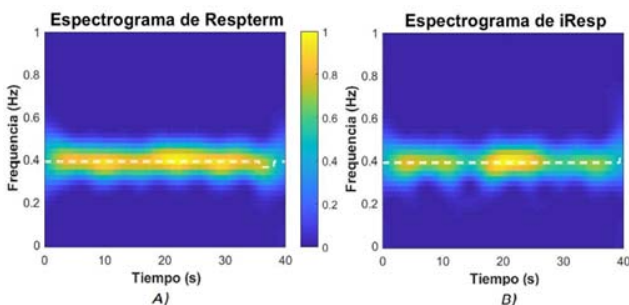


Fig 5. Espectrograma de las señales de respiración obtenidas. A) *Respterm* (termistor). B) *iResp* (smartphone).

TABLA 2. ÍNDICES DE DESEMPEÑO DE LA ESTIMACIÓN DE LA FRECUENCIA RESPIRATORIA INSTANTÁNEA CON LA TÉCNICA ÓPTICA REMOTA EN COMPARACIÓN CON EL TERMISTOR COMO REFERENCIA.

Maniobra	Correlación 2D	NRMSE
Espontánea	0.8±0.2	18.0±27.9
Fija a 6 rpm	0.9±0.1	29.2±9.1
Fija a 12 rpm	0.9±0.1	8.3±6.6
Fija a 18 rpm	0.9±0.2	8.1±9.2
Fija a 24 rpm	0.9±0.1	2.7±1.2
Fija a 30 rpm	0.8±0.1	4.2±5.3
Fija a 36 rpm	0.9±0.1	2.3±1.0
Controlada	0.9±0.2	7.4±4.6

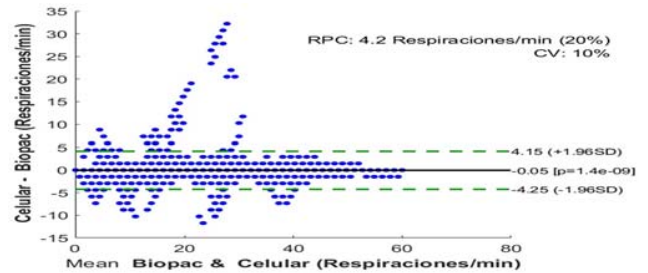


Fig. 6. Representación Bland-Altman de todas las frecuencias respiratorias instantáneas obtenidas para las maniobras

estimador debido a artefactos de movimiento. Así, como trabajo futuro se plantea la incorporación de algoritmos de procesamiento de imágenes, donde se realice un seguimiento en la zona de análisis del sujeto (ROI) conforme este realiza algún movimiento, los cuales permitan mejorar la frecuencia respiratoria estimada en cada instante.

BIBLIOGRAFÍA

- [1] B. Kozier, et al. "Fundamentos de enfermería. Conceptos, procesos y practica", Ed. McGraw-Hill, ch. 27, pp 539-561.
- [2] Cretikos M, et al. Respiratory rate: the neglected vital sign. The medical journal of Australia. 188,657. 2008.
- [3] Smith I, et al. Respiratory rate measurement: a comparison of methods. British Journal of Healthcare Assistants. 18, 19. 2011.
- [4] Nam Y, Youngsun K, Reyes B, Natasa R, and Ki Ch. "Monitoring of HR and RR Using Dual Cameras on a Smartphone" 2016. Plos One. DOI: 10.1371/journal.pone.0151013
- [5] Reyes B, Natasa R, Youngsun K, Yunyoung N, and Ki Ch. "Tidal Volume and Instantaneous Respiration Rate Estimation Using a Smartphone Camera" 2016. IEEE Journal of Biomedical and Health Informatics. DOI: 10.1109/JBHI.2016.2532876.
- [6] Lozano L, Areny P, "Detección oculta de la frecuencia cardiaca y la frecuencia respiratoria en personas sentadas", Congreso Anual de la Sociedad Española de Ingeniería Biomédica, 2009.
- [7] Huang N. 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, vol. 454, pp. 903–995, 1998.
- [8] Cohen L, "Time-frequency analysis", IEEE Xplore Digital Library, DOI: 10.1109/5.30749, 1989.