



LA FORMACIÓN DE INGENIEROS:
UN COMPROMISO PARA EL
DESARROLLO Y LA SOSTENIBILIDAD

15 al 18
DE SEPTIEMBRE

20
20

www.acofi.edu.co/eiei2020

MONITOREO REMOTO DE FRECUENCIA CARDÍACA A TRAVÉS DE FOTOPLETISMOGRAFÍA DE IMAGEN (IPPG)

Santiago Sastoque Granados, Juan Manuel López López

**Escuela Colombiana de Ingeniería Julio Garavito
Bogotá, Colombia**

Resumen

La frecuencia cardíaca se regula mediante el sistema nervioso autónomo (SNA) y algunos factores fisiológicos o patológicos afectan esta variable; por ejemplo, la fiebre y la respiración. La fotopleletismografía de imagen (IPPG) es una técnica para extraer señales fisiológicas del video de una persona. En este proyecto se implementa y analiza este método de detección de la frecuencia cardíaca y funciona mediante una cámara web. Debido a que en la sangre se encuentran componentes dinámicos de alta absorbancia, es posible extraer del video los cambios de densidad de la sangre que están sincronizados con el ciclo cardíaco. La ventaja de este método sobre la técnica tradicional de fotopleletismografía (PPG) radica en la posibilidad de hacer monitoreo remoto y sin contacto; adicionalmente, se evita la deformación de la pared arterial al no haber presión sobre la piel. Se utilizaron dos modalidades de adquisición de video distintas. En la primera, se adquieren todos los videos con la misma cámara y resolución, y además se procesa la información en tiempo real. En la segunda modalidad, se utilizan videos tomados con diferentes dispositivos, de forma remota. Se contó con la participación de 10 personas para la experimentación, con videos de una duración promedio de 50 s, con una resolución mínima de 264x480 pixeles y máxima de 1920x1080 pixeles; todos con una tasa de refresco de 30 cuadros por segundo (fps). Cada video es acompañado del registro de la frecuencia cardíaca, tomado con dispositivos digitales. En el conjunto de videos se obtuvo un error promedio de 1.97% en la frecuencia cardíaca y una desviación estándar de 0.73%; es decir, el sistema de detección de frecuencia cardíaca es preciso y funcional, en comparación con sistemas actuales de medición de esta variable.

Palabras clave: fotopleletismografía de imagen; frecuencia cardíaca; procesamiento digital de imágenes

Abstract

The heart rate (HR) is regulated by the Autonomous Nervous System (ANS) and some physiological and pathological factors may affect this variable, for instance, fever and breathing. The image photoplethysmography (IPPG) is a technique for the extraction of physiological signals from video. This project implements and analyzes this method of HR detection and works through a webcam. Due to the fact that the blood contains high dynamic components of absorbance, it is possible to obtain from videos the changes in the blood density, which is synchronized with the person's heart cycle. The main advantage of this technique over the traditional photoplethysmography (PPG) is the possibility of non-contact, remote monitoring. In addition, the arterial wall deformation is avoided since no device is making pressure in the skin. Two video acquisition modes were used. The first one was an online acquisition with the same camera and resolution for the participants. In the second one, remote videos from different devices were obtained. 10 people were enrolled in the experiment, providing videos with an average length of 50 s each, with a minimal resolution of 264x480 pixels and maximum of 1920x1080 pixels; all of them with a frame rate of 30 frames per second (fps). Each video is accompanied by the recording of the heart rate, taken with digital devices. The average error over all videos was 1.97% in the HR with a standard deviation of 0.73%; proving that the HR measurement system is accurate and functional, in comparison with other current measurement devices.

Keywords: digital image processing; imaging photoplethysmography; heart rate

1. Introducción

Los métodos convencionales para la detección de la frecuencia cardíaca, en su mayoría o son invasivos o requieren dispositivos biomédicos especializados, como es el caso de la electrocardiografía (que requiere del posicionamiento de electrodos para capturar la actividad eléctrica cardíaca) o la fotoplethysmografía (que utiliza una fuente de luz con longitudes de onda específicas y un receptor adecuado para ese espectro). La necesidad de estos dispositivos trae consigo una limitación en la adquisición y requieren del contacto con el tejido lo que resulta en artefactos por movimiento en la señal. Finalmente, se pueden entender como métodos de monitoreo localizado, despreciando posibles cambios de la señal en otros lugares (Morales, *et al.*, 2018), (Nakajima, *et al.*, 1996). Por otro lado, condiciones de aislamiento como la que estamos viviendo actualmente a raíz del CoVID-19, llevan a pensar en la posibilidad de realizar mediciones remotas, con equipos más asequibles. Un sensor de luz se encuentra en un cámara de video convencional, como la que tiene la mayoría de teléfonos móviles o computadores. De aquí surge la búsqueda de técnicas de procesamiento digital de video que permitan medir variables fisiológicas relevantes como la frecuencia cardíaca. A la técnica de extracción de frecuencia cardíaca a partir de videos se le conoce como fotoplethysmografía de imagen (IPPG) (Blöcher, *et al.*, 2017), (Hassan, *et al.*, 2018). Las señales de IPPG se obtienen de zonas específicas de los videos, llamadas regiones de interés (ROI); siendo el rostro el área más estudiada. Las señales de las ROI contienen información de la cantidad de luz absorbida por los vasos sanguíneos en el instante en que se tomó el cuadro de video. En este documento se explicará la implementación detallada de un método de IPPG, explicando la formación de la base de datos de trabajo, el esquema de procesamiento de los

videos y la validación de resultados. Finalmente, el lector podrá encontrar una discusión de los resultados, contrastándolos con métodos de medición de frecuencia cardíaca actuales y explorando los posibles usos de esta técnica.

2. Materiales y métodos

2.1. Participantes

Se toma una muestra por conveniencia sobre la base de participantes que voluntariamente quisieran hacer parte de este proyecto. Los criterios de inclusión fueron: tener entre 18 y 30 años de edad y estar sanos. Los criterios de exclusión fueron tener alguna enfermedad diagnosticada, crónica o aguda, que afectara su actividad cardíaca y condiciones insuficientes de iluminación durante la toma de los videos. Un total de 10 personas participaron en este proyecto, con una edad promedio de 25.2 años. La muestra se conformó por 5 hombres y 5 mujeres.

2.1 Protocolo experimental

Se planteó un protocolo experimental. Se propusieron dos modalidades distintas de adquisición: tiempo real y análisis fuera de línea. Para los videos tomados en tiempo real se utilizaron videos RGB tomados con una cámara Intel Real Sense D435i, con una resolución de 1920x1080 pixeles a una tasa de refresco de 30 fps y con una distancia entre el participante y la cámara de aproximadamente 40 cm a 1 m. Los videos tienen una duración promedio de 50 s. El montaje para la adquisición en tiempo real consta únicamente de la cámara y la unidad de procesamiento. Adicionalmente, los participantes contaron con un reloj inteligente en la muñeca que se utilizó para obtener un valor de referencia para la frecuencia cardíaca (Blöcher, *et al.*, 2017). Una vez posicionados los participantes, se les pidió que abrieran una aplicación de cámara y revisaran que su rostro fuera completamente visible en el cuadro de la imagen. Luego, se les pidió que grabaran un video con una duración mínima de 40 segundos, durante el cual debían permanecer relajados, observando a la cámara y sin realizar movimientos rápidos. En total, se procesaron 6 videos para esta modalidad, de 3 participantes. Por otro lado, para la modalidad de adquisición fuera de línea, los videos fueron tomados por los participantes y procesados posteriormente. Se utilizaron cámaras de teléfonos celulares a la misma tasa de refresco que para la condición anterior (30 fps), y a la misma distancia de la cámara. En esta modalidad, los participantes podían estar de pie o sentados, en una posición cómoda, observando a la cámara. Las resoluciones de los videos van desde 264x480 hasta 1920x1080. En la tabla 1 se muestran las especificaciones detalladas de los videos:

Tabla 1. Videos por sujeto y por modalidad

Sujeto	Resolución	Duración (s)	Sujeto	Resolución	Duración (s)
1	1920x1080	52	1	480 x 272	53
1	1920x1080	50	2	568 x 320	50
2	1920x1080	48	3	480 x 264	56
2	1920x1080	53	4	480 x 272	45
3	1920x1080	51	5	568 x 320	47
3	1920x1080	50	6	568 x 320	46
			7	480 x 272	46
			8	480 x 264	50
			9	480 x 272	48
			10	480 x 264	52

2.2 Procesamiento

El esquema de procesamiento propuesto, independiente de la modalidad se compone de tres etapas principales: Detección de rostro y extracción de ROI, preprocesamiento y análisis de las señales y extracción de la frecuencia cardíaca. Para el desarrollo de esta etapa se utilizó la herramienta computacional MATLAB 2019a.

2.2.1. Detección de rostro y extracción de ROI

Para la detección del rostro se utilizó el algoritmo de Viola-Jones (Viola, *et al.*, 2001), implementado en MATLAB. Una vez detectado el rostro en un cuadro, se procede a hacer seguimiento del rostro utilizando el algoritmo de KLT (Tomasi, *et al.*, 1991). Si en un cuadro el seguimiento falla debido a movimientos rápidos o cambios de iluminación, se vuelve a hacer primero el proceso de detección. Con el rostro detectado, se extraen las 5 ROI mostradas en la figura 1, provenientes de estudios previos (Blackford, *et al.*, 2015), (Hassan, *et al.*, 2018), (Sun, *et al.*, 2012). Las coordenadas de cada ROI se calculan de la siguiente manera: a partir del proceso de detección, se tiene un marco sobre todo el rostro, este marco constituye la primera ROI (figura 1.A). Las siguientes ROI se obtienen dividiendo este marco de forma horizontal o vertical.

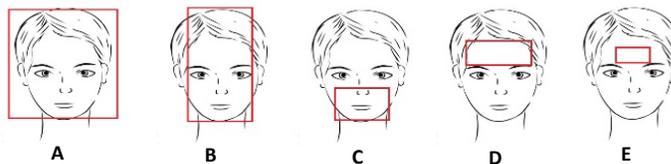
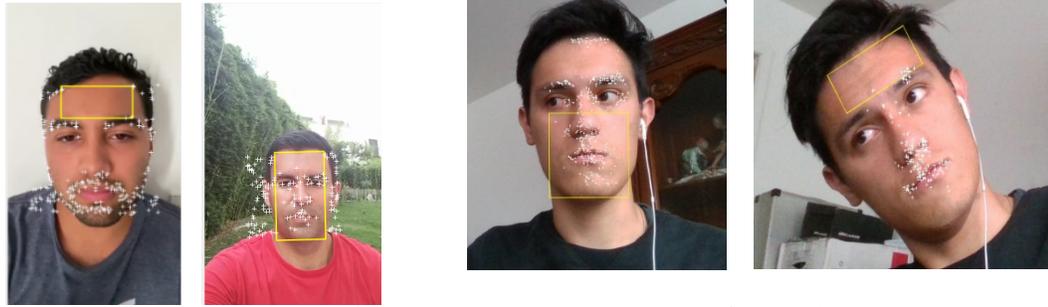


Figura 1. ROI para el análisis de la frecuencia cardíaca.

La figura 2 muestra ejemplos de diferentes ROI tomadas con las dos modalidades de adquisición.



A B

Figura 2. Ejemplos de regiones de interés. A) Diferentes ROI en modalidad: Análisis fuera de línea. B) Ejemplos regiones de interés y seguimiento facial en modalidad: Tiempo real

2.2.2. Preprocesamiento

La función principal del preprocesamiento es extraer una señal unidimensional de las señales adquiridas que son bidimensionales, es decir, terminada esta etapa se tendrá la señal de IPPG cruda. Para extraer la señal se descompone cada imagen del video en sus tres planos de color RGB y para cada plano de color se seleccionan los datos que están dentro de la ROI. Esto se hace despreciando los 10 primeros cuadros. Cuando se tiene el segmento de imagen extraído, se realiza el promedio espacial de las intensidades de todos los pixeles que hacen parte de la ROI; es decir, para cada imagen del video se tendrán tres puntos y de esta manera, concatenando la información de todos los cuadros, se construyen 3 señales. Luego de extraídas las señales, se normalizan de tal forma que su media sea cero y su desviación estándar sea uno (Poh, *et al.*, 2010), (Akar, *et al.*, 2013).

En la figura 3.A. se pueden apreciar componentes DC y de bajas frecuencias que pueden aparecer debido a cambios en las presiones venosas y el movimiento por la respiración (Rubins, *et al.*, 2016). Por este motivo es importante un primer filtro tipo FIR, pasa altas, con una frecuencia de corte de 0.6 Hz, la cual corresponde a una frecuencia cardíaca de 36 lpm y posteriormente dependiendo de la longitud del video, se realiza un filtro FIR pasa bajas, con frecuencia de corte de 11 Hz como se muestra en la figura 3.B. (Blackford, *et al.*, 2015), (Rubins, *et al.*, 2011), (Liang, *et al.*, 2018).

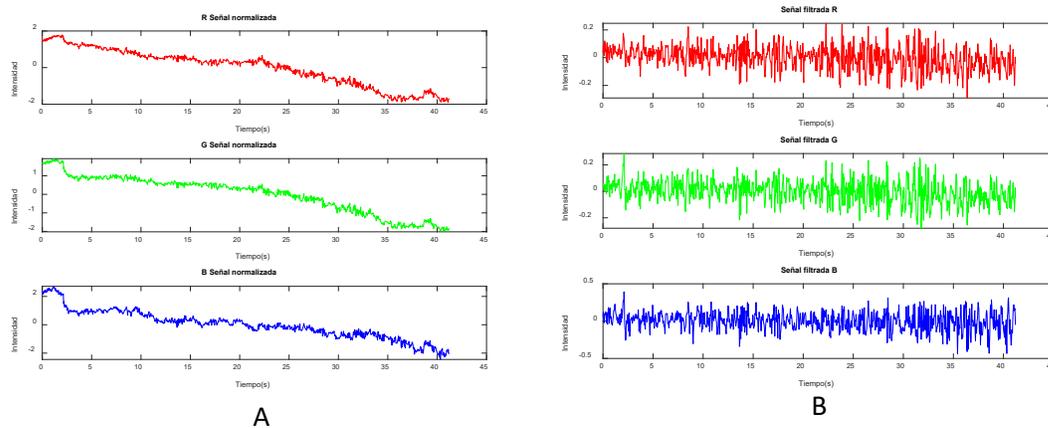


Figura 3. Señales obtenidas de las ROI. A. Señales crudas de los planos de color. De arriba hacia abajo: plano rojo, verde y azul. B. Señales filtradas y normalizadas.

2.2.3. Análisis de las señales y extracción de la frecuencia cardíaca

El reto principal de esta etapa es extraer la señal de interés, debido a que las señales que se obtuvieron son una mezcla de varios procesos físicos ocurriendo al mismo tiempo. Para este fin se hace uso de la técnica de Análisis de Componentes Independientes (ICA). El algoritmo está motivado en la idea de que las variaciones en la información son generadas por diversos factores, entonces, cada variable observada se puede ver como una combinación de estos factores (Poh, *et al.*, 2010). Las señales de entrada para ICA serán las tres señales RGB filtradas y normalizadas. Las señales obtenidas con esta técnica se filtran con un filtro pasa banda entre 0.6 Hz y 5 Hz, que corresponden a frecuencias cardíacas de 36 lpm y 300 lpm respectivamente. Por último, las señales se interpolan usando una interpolación cúbica (Akima, *et al.*, 1970) para lograr una frecuencia de muestreo de 1200 Hz; de esta forma se mejora la resolución temporal y frecuencial de las señales (Nordhausen, *et al.*, 2017), (Mannapperuma, *et al.*, 2015).

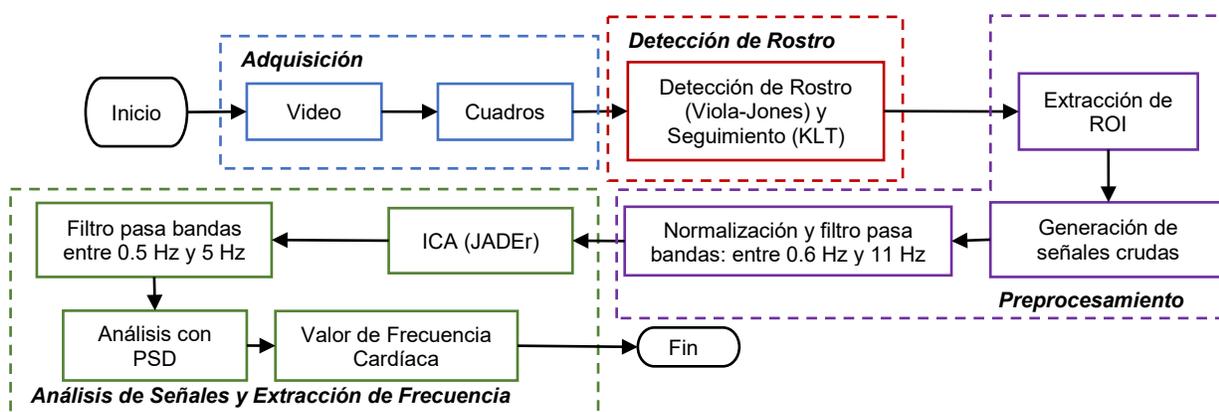
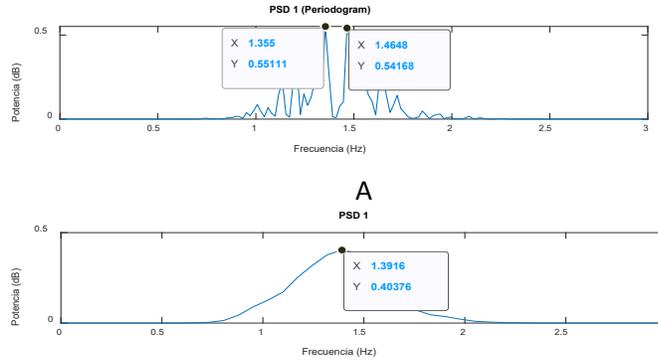


Figura 4. Esquema de procesamiento

De estas tres señales, solo una corresponde a las variaciones de absorbancia causadas por el flujo de sangre en los vasos sanguíneos. Por esta razón, son analizadas en frecuencia a través de su densidad espectral de potencia (PSD). Se utilizaron dos métodos de estimación de PSD no paramétricos: periodograma de Barlett y periodograma de Welch (Semmlow, 2004). Los mejores resultados se obtuvieron con el segundo método. La figura 4 muestra todo el esquema de procesamiento de videos expuesto en esta sección.

3. Resultados

Los resultados de la PSD para un participante de la modalidad de adquisición no presencial se presentan a continuación:



B

Figura 5. Resultados de análisis en frecuencia. A PSD por el método de periodograma. B PSD por el método de Welch.

Como se puede identificar en ambos métodos la segunda señal es la que presenta el pico de frecuencia de mayor potencia, pero está muy cercana a otro pico, cuyas frecuencias cardiacas serían 72.48 lpm y 78.5 lpm, ahora para determinar qué frecuencia tomar, es útil el segundo método para hallar la PSD (método de Welch), en la que se puede ver que el componente de mayor potencia se encuentra en la segunda señal con una frecuencia cardiaca de 74.7 lpm, y si se analiza la figura 5.A, es posible notar que el valor del pico por el método de Welch es el promedio de los picos por el método de periodograma con un error del 1.05%, y la frecuencia cardiaca de la persona a la que se le están analizando los datos es de 76 lpm. El error con respecto al pico de mayor potencia en la PSD es del 1.710%.

Existen diferentes maneras de analizar los datos y diferentes combinaciones de parámetros, como lo son: diferentes ROI, tiempos de los videos, métodos de análisis en frecuencia, frecuencias de corte de los filtros, que señal escoger, entre otras. Debido a esto, se probaron diferentes combinaciones y las que obtuvieron menor error se utilizaron para el esquema de procesamiento. Para la determinación de la ROI más eficiente se resumen los resultados en la figura 6, en el que se muestran resultados de 13 participantes de modalidad de adquisición no presencial. Cada video se procesó con 4 diferentes ROI y se muestran sus respectivos errores relativos.

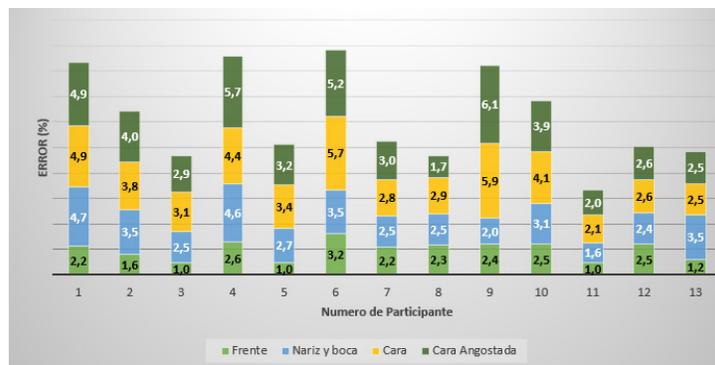


Figura 6. Diagrama de barras del error para las diferentes ROI

Cada color representa una región diferente, y la longitud del color para cada participante representa el error que obtuvo dicha ROI.

4. Discusión y conclusiones

Primeramente, con respecto a las diferentes ROI, según el diagrama de barras, se puede observar que la ROI con el porcentaje de error más bajo es la frente, mismos resultados que investigaciones anteriores (Blöcher, *et al.*, 2017). Esto se puede deber a que las otras ROI ocupan áreas de la cara que pueden tener más movimiento durante la adquisición (como el movimiento de la nariz al respirar, o los ojos al parpadear) y provocar componentes no deseados en la señal. Por otro lado, la frente presenta menos movimiento y únicamente abarca píxeles del rostro, mientras que otras ROI podían algunas veces contener píxeles del fondo del video. Empíricamente, el tiempo de grabación de los videos fue suficiente (50 s) para obtener un valor preciso de la frecuencia cardíaca. Algunas veces los videos externos tienen artefactos, que en la mayoría de los casos se pueden retirar para frecuencias por encima 11 Hz. Sin embargo, se encontró en la mayoría de los videos componentes armónicos de la frecuencia cardíaca, de alta potencia. Para estos casos, la medida se tomó con base no en la potencia sino en la ubicación de los picos de la PSD. En el análisis con PSD, en repetidas ocasiones se encontraron dos picos de potencias similares en frecuencias diferentes pero cercanas, esto se puede deber variaciones de la frecuencia cardíaca durante la adquisición del video, pues se presentan mayormente en los videos de larga duración, para solucionar esto se podría implementar un procesamiento para tiempos más cortos en la señal de IPPG, como en (Blöcher, *et al.*, 2017); Aunque hacer un promedio entre los picos resultó ser una buena solución. Además, se encontró que los cambios en la uniformidad de iluminación provocan componentes de altas frecuencias, pero dentro del ancho de banda de interés. Por otro lado, con las señales de ICA se pueden obtener los tacogramas necesarios para analizar la HRV y con este tener información del SNA, únicamente con videos. Finalmente, según (Hassan, *et al.*, 2018), el método de detección de frecuencia cardíaca presentado obtuvo resultados similares a otras investigaciones anteriores en IPPG, en varios casos con porcentajes de errores menores. El método se puede implementar para un monitoreo remoto de la frecuencia cardíaca y posteriormente utilizar esta variable para encontrar otras señales fisiológicas como HRV y patrón respiratorio. Por otro lado, existen relojes inteligentes y teléfonos móviles que tienen implementados funciones para hallar la frecuencia cardíaca, pero son menos asequibles se limitan únicamente a esta función, es decir, no se puede utilizar la señal resultante para explorar otras variables fisiológicas.

5. Referencias

Artículos de revistas

1. Akar, S. A., Kara, S., Latifoğlu, F. and Bilgic, V. (2013). Spectral analysis of photoplethysmographic signals: The importance of preprocessing, Vol. 8, No. 1, pp 16-22.
2. Akima, Hiroshi. (1970). A new method of interpolation and smooth curve fitting based on local procedures. Journal of the ACM (JACM), 17.4, pp. 589-602.

3. Blackford, E. and Estep, J. (2015). Effects of frame rate and image resolution on pulse rate measured using multiple camera imaging photoplethysmography, Vol 9417, No. 9417, pp 1- 14.
4. Hassan, M. A., Malik, A. S., Fofi, D., Saad, N., Karasfi, B. and Mériaudeau, F. (2017). Heart rate estimation using facial video: A review, Vol. 38, pp. 346-360.
5. Liang, Y., Elgendi, M. and Chen, Z. (2018). An optimal filter for short photoplethysmogram signals, Vol 5, No. 1, pp 1-180076.
6. Mannapperuma, K., Holton, B., Lesniewski, P. and Thomas, J. (2015). Performance limits of ICA-based heart rate identification techniques in imaging photoplethysmography, Vol. 36, No. 1, pp. 1-83.
7. Miettinen, J., Nordhausen, K. and Taskinen, S. (2017). Blind Source Separation Based on Joint Diagonalization in R: The Packages JADE and BSSasyp. Journal of Statistical Software, Vol. 76, No. 2, pp. 1-31.
8. Moraes, J.L., Rocha, M.X., Vasconcelos, G.G., Vasconcelos Filho J.E., De Albuquerque V.H.C. and Alexandria A.R. (2018). Advances in Photoplethysmography Signal Analysis for Biomedical Applications. Vol. 18, No. 6.
9. Nakajima, K., Tamura, and T. Miike, H. (1996). Monitoring of heart and respiratory rates by photoplethysmography using a digital filtering technique, Vol. 18, No. 5, pp. 365-372.
10. Poh, M., McDuff, D. J. and Picard, R. W. (2010). Non-contact, automated cardiac pulse measurements using video imaging and blind source separation, Vol. 18, No. 10, pp. 10762-10774.
11. Rubīns, U., Spīgulis, J. and Miščuks, A. (2016). Photoplethysmography imaging algorithm for continuous monitoring of regional anesthesia, pp 67–71.
12. Rubins, U., Upmalis, V., Rubenis, O., Jakovels, D. and Spigulis, J. (2011). Real-Time Photoplethysmography Imaging System. Vol 16, No. 26, pp.183-186.
13. Sun, Y. and Thakor, N. (2017). Photoplethysmography Revisited: From Contact to Noncontact, From Point to Imaging. Vol. 63, No. 3, pp. 463-477.
14. T. Blöcher, J. Schneider, M. Schinle and W. Stork. (2017). An online PPGI approach for camera based heart rate monitoring using beat-to-beat detection, pp. 1-6.
15. Tomasi, C., Kanade, T. (1991). Detection and Tracking of Point Features. Carnegie Mellon University Technical Report CMU-CS-91-132.
16. Viola, Paul A. and Jones, Michael J. (2001). Rapid Object Detection using a Boosted Cascade of Simple Features, IEEE CVPR.
17. Yu Sun, Sijung H., Azorin-Peris, V. Kalawsky, R. and Greenwald, S. E. (2012). Noncontact imaging photoplethysmography to effectively access pulse rate variability, Vol. 10, No. 6.

Libros

1. Semmlow, J. (2004). Biosignal and Biomedical Image Processing. Marcel Dekker, Nueva York, pp. 282-292.

Sobre los autores

- **Santiago Sastoque Granados:** Estudiante de Ingeniería Biomédica. Integrante del Semillero de Investigación en Procesamiento de Imágenes y Señales PROMISE. Escuela

Colombiana de Ingeniería Julio Garavito - Universidad del Rosario.
santiago.sastoque@mail.escuelaing.edu.co

- **Juan Manuel López López:** Ing. Electrónico, Magister en Ingeniería. Área: Electrónica y de Computadores, Doctor en Ingeniería, Profesor Asistente del Programa de Ingeniería Biomédica de la Escuela Colombiana de Ingeniería Julio Garavito.
juan.lopezl@escuelaing.edu.co.

Los puntos de vista expresados en este artículo no reflejan necesariamente la opinión de la Asociación Colombiana de Facultades de Ingeniería.

Copyright © 2020 Asociación Colombiana de Facultades de Ingeniería (ACOFI)