



# FRECUENCIA RESPIRATORIA EN IMÁGENES

**Daniel Estrada Calderón, Valentina Perea Torres, Juan Manuel López López**

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

## **Resumen**

La frecuencia respiratoria (FR) es la cantidad de respiraciones por minuto que realiza un ser vivo y está regulada por el sistema nervioso y algunas otras características fisiológicas o patológicas. La FR se puede extraer por medio de Fotopleetismografía de imagen (IPPG), la cual es una técnica que consiste en extraer diferentes variables fisiológicas del video de una persona, de manera remota y sin contacto. Para esta investigación se decidió estudiar la FR debido a su importancia en el contexto clínico, ya que refleja algunas funciones esenciales del cuerpo. Puede ser de utilidad, por ejemplo, para detectar enfermedades respiratorias y/o cardiacas. En este proyecto, se estudiaron un total de 13 participantes, cuyas edades variaban entre los 18 y 50 años. Para la adquisición de los videos, se utilizaron las cámaras con las que contaba cada participante, como la cámara integrada de un computador o la cámara de un celular. El ejercicio consistía en la grabación de un video por cada participante de alrededor de 3 minutos, donde se le indicaba cómo respirar para así controlar su FR y después compararla con los resultados. La medición de la FR se hizo por medio de la aplicación "Paced Breathing" ya usada en investigaciones anteriores. Para la extracción de la señal se realizó primero la localización de una región de interés (ROI) en el rostro, con el algoritmo de Viola-Jones, para luego aplicar una etapa de remoción de artefactos. A través de técnicas de procesamiento de imágenes y procesamiento estadístico de señales multivariantes se pudo extraer la FR de cada sujeto con error promedio de 3.3% a través de análisis con transformada Wavelet y 2.65% con análisis de Fourier. Se utilizaron herramientas de preprocesamiento como filtros frecuenciales, remoción de tendencias lineales y filtros de media móvil. Los métodos propuestos en este trabajo son novedosos y presentan una alternativa para la extracción de diferentes variables fisiológicas. Las condiciones de grabación de los participantes no están bajo muchas restricciones, por lo que podría ser utilizado en varios campos de la salud sin necesidad de condiciones exigentes. Además, muestra un desempeño similar al estado del arte y provee una forma no invasiva para la medición de una variable fisiológica de importancia clínica y de investigación.

**Palabras clave:** procesamiento digital de imágenes; fotopleletismografía de imagen; frecuencia respiratoria

### **Abstract**

*Respiratory rate (FR) is the number of breaths per minute performed by a living being and is regulated by the nervous system and some other physiological or pathological characteristics. FR can be extracted using image photoplethysmography (IPPG); which is a technique for extracting different physiological variables from a person's video; remotely and contactless. For this research it was decided to study FR due to its importance in the clinical context; as it reflects some essential functions of the body. It may be useful; for example; to detect respiratory and/or cardiac diseases. In this project; a total of 13 participants were studied; the ages ranged from 18 to 50 years. For the acquisition of the videos; the cameras available to each participant were used; such as the integrated camera of a computer or the camera of a cell phone. The exercise consisted of recording a video by each participant for about 3 minutes; where he was told how to breathe to control his FR and then compare it with the results. The measurement of FR was done through the application "Paced Breathing" already used in previous research. For the extraction of the signal; the location of a region of interest (ROI) on the face is found; with the algorithm of Viola-Jones; and then; an artifact removal stage was applied. Through image processing techniques and statistical processing of multivariate signals; the FR of each subject was obtained; with an average error of 3.3% for a method using Wavelet transform and 2.65% for a Fourier analysis method. Preprocessing tools such as frequency filters; linear trend removal; and moving average filters were used. The methods proposed in this work are novel and presents an alternative for the extraction of different physiological variables. Participants' recording conditions are not under many restrictions; so it could be used in various contexts without the need for sophisticated devices. In addition; the proposed method shows similar performance to the state of the art and provides a noninvasive way for measuring a physiological variable of clinical and research importance.*

**Keywords:** digital image processing; image photoplethysmography; respiratory rate

## **1. Introducción**

A lo largo del tiempo se han utilizado diferentes instrumentos y métodos para la extracción de la frecuencia respiratoria; en su mayoría; métodos que constan de instrumentos biomédicos no tan comunes en nuestra vida cotidiana; como puede ser el electrocardiógrafo; para tener señales de electrocardiografía o el espirómetro para medir volúmenes y capacidades del pulmón; o simplemente métodos por observación que no suelen ser tan precisos. Una gran desventaja de estos métodos es el uso limitado fuera de instalaciones médicas o centros de investigación. La Fotopleletismografía de imagen (IPPG) es una técnica para extraer señales fisiológicas del video de una persona y proviene de la Fotopleletismografía (PPG) convencional; que es utilizada tanto en diferentes aplicaciones clínicas; como cotidianas. Dichas señales pueden ser usadas para el diagnóstico de diferentes condiciones; por ejemplo; hipoxia; taquicardia; bradicardia. Estos



videos pueden ser tomados en diversas regiones del cuerpo [1]. A través de diferentes técnicas aplicadas a las señales de los videos de cada uno de los participantes se puede llegar a la señal de IPPG y desde aquí; a la extracción de alguna variable fisiológica específica de interés. En este caso; la frecuencia respiratoria; la cual se puede relacionar con estados fisiológicos o patológicos de las personas. La FR es la cantidad de respiraciones por minuto que tiene una persona. Esta debe estar entre las 12 y 20 respiraciones por minuto en un adulto sano y en estado de reposo; y cuyo estado es el que se estudia en esta investigación. Para la evaluación de las señales fotopleletismográficas se decidió crear una base de datos de videos de diferentes personas entre los 18 y 50 años que; a su conocimiento; no tuviesen ningún problema respiratorio ni cardiaco; realizando un video de aproximadamente 3 minutos frente a una cámara común y controlando sus respiraciones con ayuda de una aplicación para celular. Esta base de datos será explicada con mayor detalle más adelante. De igual manera; el lector se encontrará una sección detallada de los métodos usados para la extracción de cada una de las señales y la discusión de estos resultados El presente trabajo es un derivado del proyecto "Monitoreo Remoto de Frecuencia Cardiaca a través de Fotopleletismografía de Imagen (IPPG)" (Sastoque; in al; 2020).

## **2. Método y Materiales**

### **2.1 Participantes**

Se tomaron un total de 17 videos de una muestra no aleatoria de 13 diferentes participantes voluntarios hombres y mujeres; cuya edad oscilaba entre los 18 y 50 años. Los criterios de exclusión fueron: Tener enfermedades respiratorias y/o cardiacas; fumadores activos o el consumo medicamentos que pudiesen haber afectado la frecuencia respiratoria o cardiaca (como jarabes para la tos; medicamentos para la presión arterial; entre otros); haber sido diagnosticado de enfermedades respiratorias agudas en los últimos 6 meses. También se les comentó a los participantes que los videos debían ser grabados con buena iluminación natural; es decir; que su rostro estuviera totalmente iluminado uniformemente y sin estar a contraluz.

### **2.2 Procedimiento experimental**

Para la grabación de los videos se lo pidió a cada participante se ubicará frente a la cámara y que cumpliera con los parámetros anteriormente mencionados. El procedimiento consistió en la realización de un video con cualquier dispositivo que contara con cámara; como por ejemplo el celular o computador. El video constaba de aproximadamente tres minutos de grabación; donde la persona estuviese mirando a la cámara sin realizar movimientos bruscos o que pudiesen interferir con la grabación del rostro del participante. A su vez; cada participante descargo la aplicación "Paced Breathing" ya usada en anteriores estudios mostrando resultados de confianza (Alafeef; in al; 2020); esta aplicación ayudó al estudio ya que pudimos tener un FR real estimada por la aplicación. Lo que hacia la aplicación; dependiendo de la configuración dada; era decirle al participante en qué momento exhalar e inhalar; y así se logró controlar la cantidad de respiraciones por minutos del participante.



### 2.3 Procesamiento de las señales

La detección de rostro se hizo a través del algoritmo de Viola-Jones; y luego de su detección; se aplicó un algoritmo de seguimiento del rostro de la persona una vez este se haya identificado. Una vez hecho esto se seleccionó una región de interés (ROI) de la cual se extrajeron las señales fotopletiométricas más adelante. La ROI seleccionada para cada participante fue la frente; cómo se puede observar en la figura 1. La ROI es adaptativa al video y rostro detectado. Se escogió la frente como ROI debido a que es un lugar del rostro en donde hay pocos movimientos involuntarios y por lo tanto es más sencillo de trabajar con ella (Sastoque; in al; 2020).

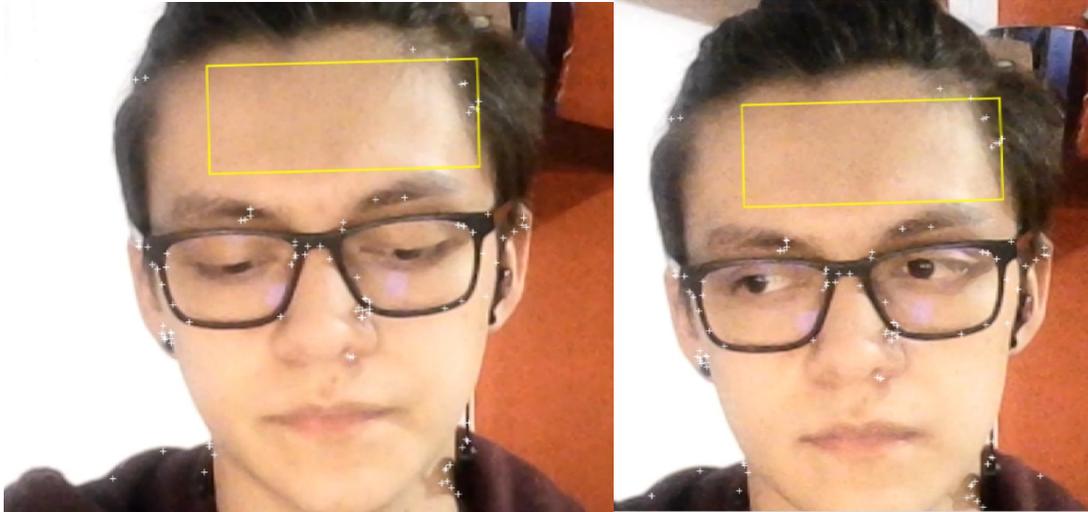


Figura 1. Detección del rostro con su debida región de interés

Para la extracción de la señal cruda del vídeo de cada persona; lo que se hizo después de tener almacenados los datos de la cantidad de recuadros del vídeo; matriz RGB de cada recuadro y coordenadas de la ROI cuadro por cuadro fue separar las matrices R;G y B del vídeo e ir promediando todos los pixeles de cada matriz formando así 3 señales en el dominio del tiempo; R; G y B; correspondientemente. De igual forma; en seguida se normalizó cada señal para poder analizar mejor cada una de estas como se puede ver en la figura 2.1; cada una de estas 3 señales representan los pixeles promediados de cada señal.

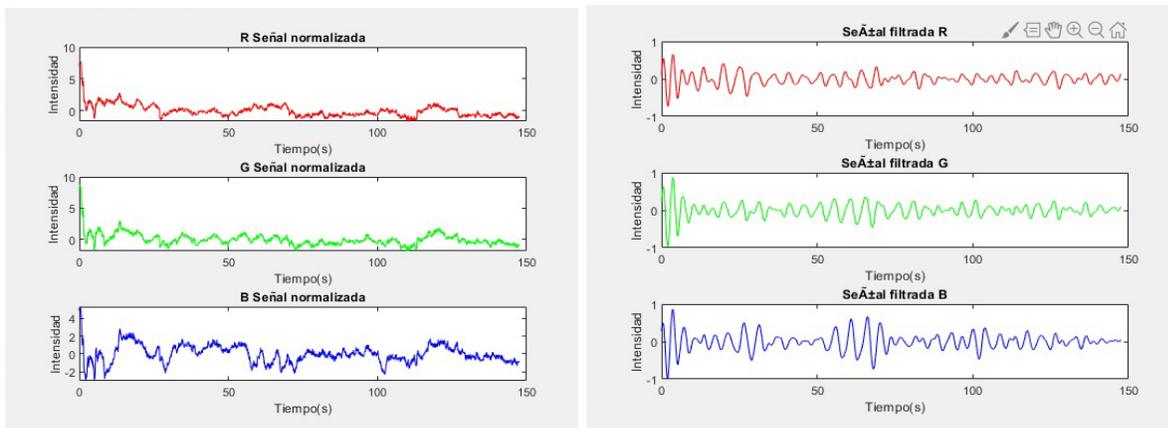


Figura 2.1. Señales R; G y B normalizadas después haber promediado los pixeles por cada cuadro. Figura 2.2. Señales R; G y B después por pasar por técnicas de procesamiento para eliminar artefactos no deseados.



Luego de tener las señales crudas de cada canal del video se aplicaron diferentes métodos de filtrado y procesamiento para dar paso a una señal más clara y precisa; estos métodos se explicarán a continuación.

Primero; se aplicó un filtro pasa altas tipo IIR Butterworth de orden 9; con frecuencia de corte de 0.18 Hz y atenuación de 40 dB en la banda de rechazo; para eliminar cualquier artefacto no deseado por debajo de la frecuencia de corte ya mencionada.

En la siguiente etapa de procesamiento; con el mismo fin que el primer filtro; se utilizó un filtro pasa-bajas de tipo FIR; con una frecuencia de corte de 6 Hz y orden 40 para eliminar cualquier frecuencia por encima de 6 Hz.

Las frecuencias de corte utilizadas para los filtros se eligieron debido a que la frecuencia respiratoria en individuo sano mayor a 18 años y en estado de reposo esta entre 0.2 y 0.3 Hz; fuera de estos rangos no se encontró la variable fisiológica deseada en este estudio (Barret; in al; 2012).

Finalmente; se realizó una etapa de filtrado adicional; donde se utilizó un filtro FIR de media móvil con un ancho de ventana de 40 muestras; para suavizar cualquier dato ruidoso que estuviese en las señales.

Posterior a los filtros de frecuencia; se utilizó el método de *remoción de tendencias* (*detrending* en inglés) (Unakafov; 2018) para eliminar cualquier tendencia lineal que pudiese tener nuestras señales y evitar contaminar nuestra señal final; ya que el componente pulsátil de IPPG tiene una amplitud menor que la línea de b.

### 3. Resultados

Con las tres señales procesadas podemos dar paso a los métodos seleccionados en el estudio para la extracción de la frecuencia respiratoria. Estos son por medio de la transformada de Fourier y por medio de la transformada Wavelet (Bousefsaf et al.; 2013) Para la primera técnica; se aplicó la transformada rápida de Fourier a cada una de las tres señales; para luego poder detectar la frecuencia dominante en cada señal a través de una búsqueda de máximos (picos) en la señal. La frecuencia del pico de mayor potencia será la estimación de la frecuencia respiratoria; como se muestra en las figuras 3.1 y 3.2.

Por otro lado; al usar transformada Wavelet se tienen las gráficas de dos diferentes videos con sus correspondientes estimaciones en las figuras 4.1 y 4.2.

Para la validación del algoritmo y el análisis de los resultados; este se probó con 17 videos de 13 diferentes participantes; donde 3 videos no funcionaron y para 1 video el algoritmo de detección de rostro no funcionó. De los videos 13 videos restantes para los cuales el algoritmo dio resultados; 11 contenían una información de frecuencia respiratoria de 12 respiraciones por



minuto y 2 videos que contaban con información de 15 respiraciones por minuto. La tabla 1 reúne la información correspondiente a los datos reales y a los datos obtenidos por el algoritmo. El error relativo de la FR para los videos correspondientes a 0.2 Hz o 12 respiraciones por minuto; por el método de Fourier es 0.5% y con transformada Wavelet de 1.8% y para los videos con FR igual a 2.5hz o 15 respiraciones por minuto el error relativo con Fourier fue de 4.8% y con Wavelet de 4.8%.

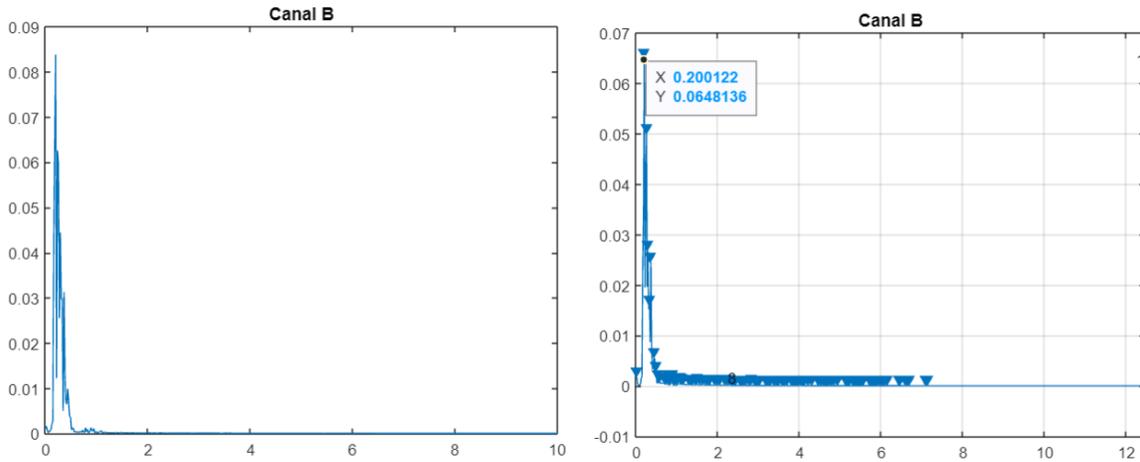


Figura 3.1. Transformada de Fourier del canal B con una estimación de FR de 0.2hz con y sin la función findpeaks

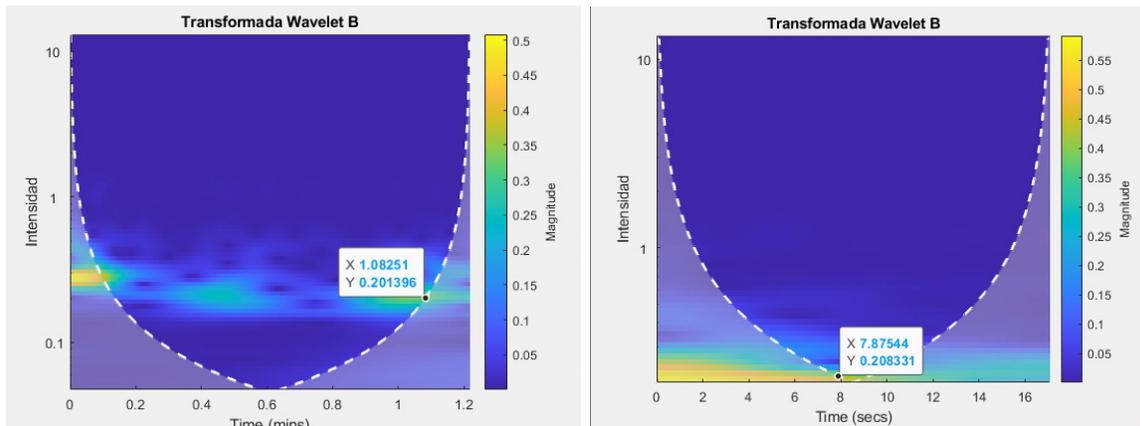


Figura 4.1 y 4.2 Estimaciones de la FR de dos diferentes videos y participantes.



No. Participante	FR real (Hz)	FR estimada (Hz)	
		Fourier	Wavelet
1	0.2	0.183	0.2
2	0.2	0.195	0.2
3	0.2	0.2	0.2
4	0.2	No hubo ROI	No hubo ROI
5	0.2	0.2087	0.2
6	0.2	0.18	0.18
7	0.2	0.229	0.22
8	0.2	0.1925	0.2
9	0.2	0.2109	0.211
	0.25	No hubo video	No hubo video
10	0.2	No hubo video	No hubo video
	0.25	No hubo video	No hubo video
11	0.2	0.1981	0.204
	0.25	0.2573	0.246
12	0.2	0.1932	0.205
	0.25	0.225	0.23
13	0.2	0.1985	0.22

Tabla 1. Recopilación de datos de los resultados del algoritmo para cada participante

#### 4. Discusión y Conclusiones

Primero; el algoritmo de detección de rostros presentó buenos resultados en todos los casos a excepción del participante 4; debido al ángulo de la cámara con respecto al rostro de la persona. El segundo video del participante 9 y los dos videos del participante 10 fueron descartados debido a incompatibilidad de formato. Con respecto a los demás videos no hubo problema alguno con la detección y seguimiento del rostro. La duración de los videos era de alrededor de 3 minutos; sin embargo; analizando los videos con diferentes duraciones entre 30 segundos y 3 minutos; se podían obtener buenos resultados cualquiera que fuera el tiempo dentro de este rango. Otro factor para tener en cuenta es el hecho de que los videos fueron grabados bajo luz artificial algunos y otros bajo luz ambiente; el algoritmo funcionó bastante bien con ambos tipos de videos.

Para la estimación de la FR con transformada de Fourier se toma el pico de mayor potencia en cada canal; normalmente el valor era el mismo para cada canal; pero solía suceder que en dos canales fueran el mismo valor y en el tercero diferente; para este caso se tomaba el valor presente en dos de los tres canales. Con transformada Wavelet pasaba algo muy similar a la técnica anterior; solo que en este caso los valores eran diferentes en cada canal; pero muy cercanos por lo que se promedió cada resultado y así se obtuvo la FR del video de una persona. Las dos técnicas para la estimación de la frecuencia respiratoria obtuvieron un desempeño bastante bueno; sin embargo; con la transformada de Fourier podemos tener un valor más conciso de la FR ya que con ondas wavelets puede ser un poco más impreciso a la hora de



analizar las gráficas; ya que se ven muchos artefactos de más que pueden confundir la extracción de la FR.

En comparación de muchos otros métodos de extracción de la FR; este es un método sin contacto y sin muchas limitaciones a la hora de grabar los videos; donde el algoritmo logró desenvolverse muy bien a lo largo del estudio. Muchos otros métodos sin contacto han logrado una estimación de la FR con buenos resultados; pero con ambientes experimentales limitando así; la aplicación del algoritmo en entornos más cotidianos

## Referencias

- Cobos Torres; J. C.; & Abderrahim; M. (2017). Measuring Heart and Breath Rates by Image Photoplethysmography using Wavelets Technique. IEEE Latin America Transactions; 15(10); 1864-1868. <https://doi.org/10.1109/TLA.2017.8071228>
- Sastoque; S.; & López López; J. M. (2020). MONITOREO REMOTO DE FRECUENCIA CARDÍACA A TRAVÉS DE FOTOPLETISMOGRAFÍA DE IMAGEN (IPPG). Encuentro Internacional De Educación En Ingeniería. Recuperado a partir de <https://acofipapers.org/index.php/eiei/article/view/795>
- Unakafov; A. M. (2018). Pulse rate estimation using imaging photoplethysmography: Generic framework and comparison of methods on a publicly available dataset. Biomedical Physics & Engineering Express; 4(4); 045001. <https://doi.org/10.1088/2057-1976/aabd09>
- Alafeef; M.; & Fraiwan; M. (2020). Smartphone-based respiratory rate estimation using photoplethysmographic imaging and discrete wavelet transform. Journal of Ambient Intelligence and Humanized Computing; 11(2); 693-703. <https://doi.org/10.1007/s12652-019-01339-6>
- Barrett; Kim E.; Barman; Susan M.; Boitano; Scott; Brooks; Heddwen (2012-04-05). Ganong's Review of Medical Physiology (24 ed.). p. 619. ISBN 978-0071780032.
- Bousefsaf; F.; Maaoui; C.; & Pruski; A. (2013). Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate. Biomedical Signal Processing and Control; 8(6); 568-574. <https://doi.org/10.1016/j.bspc.2013.05.010>

## Sobre los autores

- **Daniel Estrada Calderón:** Estudiante de Ingeniería Biomédica. Escuela Colombiana de Ingeniería Julio Garavito – Universidad del Rosario. [daniel.estrada@mail.escuelaing.edu.co](mailto:daniel.estrada@mail.escuelaing.edu.co)
- **Valentina Perea Torres:** Estudiante de Ingeniería Biomédica. Escuela Colombiana de Ingeniería Julio Garavito – Universidad del Rosario. [valentina.perea@mail.escuelaing.edu.co](mailto:valentina.perea@mail.escuelaing.edu.co)
- **Juan Manuel López López:** Ingeniero Electrónico; Magister en Ingeniería Electrónica. Doctor en Ingeniería. Profesor Asistente Programa de Ingeniería Biomédica – Escuela Colombiana de Ingeniería Julio Garavito. [juan.lopezl@escuelaing.edu.co](mailto: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 © 2021 Asociación Colombiana de Facultades de Ingeniería (ACOFI)

