

Índice de calidad y evaluación de un sistema PPG



Máster Universitario en
Ingeniería de Telecomunicación

Trabajo Fin de Máster

Noé San Fabián

Arantzazu Villanueva Larre

Pamplona, 18/09/2018

Índice

1. Estado del arte	2
1.1. PPG	2
1.2. Índices de calidad empleados	14
1.2.1. Conclusiones:.....	28
2. Índice de calidad propuesto	29
2.1. Introducción.....	29
2.2. Índice de calidad diseñado	30
2.2.1. Introducción	30
2.2.2. Acondicionamiento de la señal	31
2.2.3. Índice de calidad	34
3. Evaluación PulseCam	59
3.1. Introducción.....	59
3.2. Algoritmo PPG.....	59
3.2.1. La detección del área de interés (ROI).	61
3.2.2. La iluminación.....	61
3.2.3. Una cámara como sensor óptico.....	62
3.2.4. Extracción de los signos vitales	62
3.3. Evaluación empleando el índice de calidad	64
4. Conclusiones	77
4.1. Conclusiones del análisis del sistema PPG analizado	77
4.2. Conclusiones del índice de calidad diseñado.....	78
5. Bibliografía	79

1.Estado del arte

Para la correcta comprensión de este trabajo fin de Máster es importante entender los principios en los que se basa y la situación actual de los sistemas desarrollados. Para ello se introducirán a continuación los principales temas que toca este trabajo: los sistemas PPG (Photoplethysmography) y los índices de calidad empleados en la actualidad para evaluar la señal obtenida.

1.1. PPG

Photoplethysmography (PPG) es una tecnología óptica no invasiva empleada en la detección de cambios del flujo sanguíneo en el tejido micro vascular [1]. El tejido micro vascular está formado por una red de vasos sanguíneos con un diámetro menor o igual a 100 μm . Esta red se encarga de distribuir la sangre de forma homogénea por los tejidos.

El principio en el que se basa el PPG es el siguiente, la sangre arterial presenta una absorción de la luz diferente al resto de tejidos en ciertas longitudes de onda [2-4]. Fluctuaciones en la cantidad de proteínas y hemoglobina en la sangre, causadas por la circulación de sangre oxigenada, provocan variaciones en la absorción y dispersión de la luz en el tejido. Cuando el corazón bombea la sangre el flujo sanguíneo que llega a los capilares cercanos a la superficie de la piel aumenta, esto supone una mayor absorción de la luz. En cambio, cuando la sangre vuelve a corazón el flujo disminuye y la absorción de luz disminuye también.

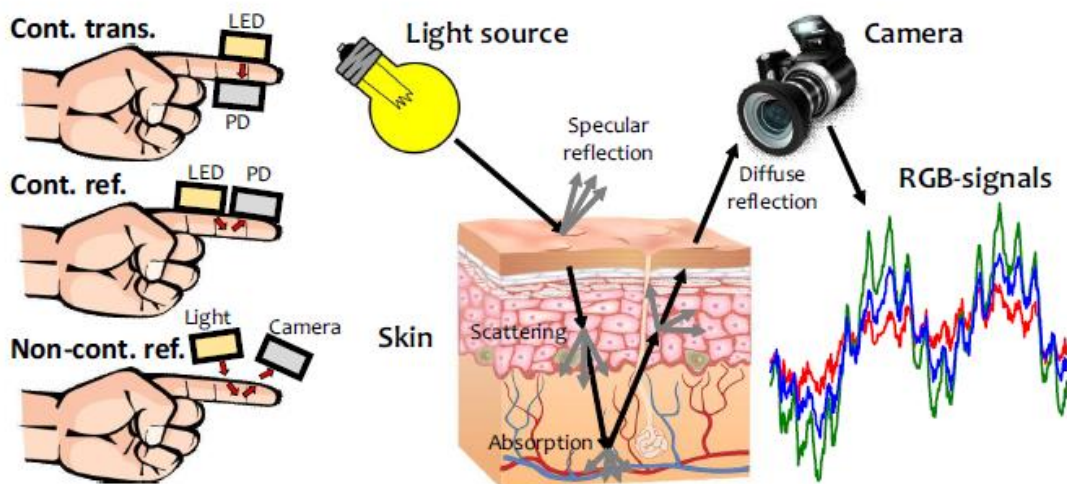


Ilustración 1: Principio básico de sistemas PPG [4]

En la ilustración 1 se representa el funcionamiento de un sistema PPG. Una fuente de iluminación emite luz hacia la piel, una parte de la luz atraviesa la piel y es absorbida mientras que otra parte se refleja en los tejidos sanguíneos y vuelve al exterior de la piel. La cantidad de sangre varía la cantidad de luz reflejada, un mayor flujo de sangre implica una mayor absorción y por lo tanto menos luz es reflejada. Analizando la luz reflejada se puede obtener una señal que contiene la información pulsátil.

En cuanto a la luz, si dividimos el espectro visible en tres componentes (rojo, verde y azul) cada una de las componentes presenta una respuesta diferente al interactuar con los tejidos.

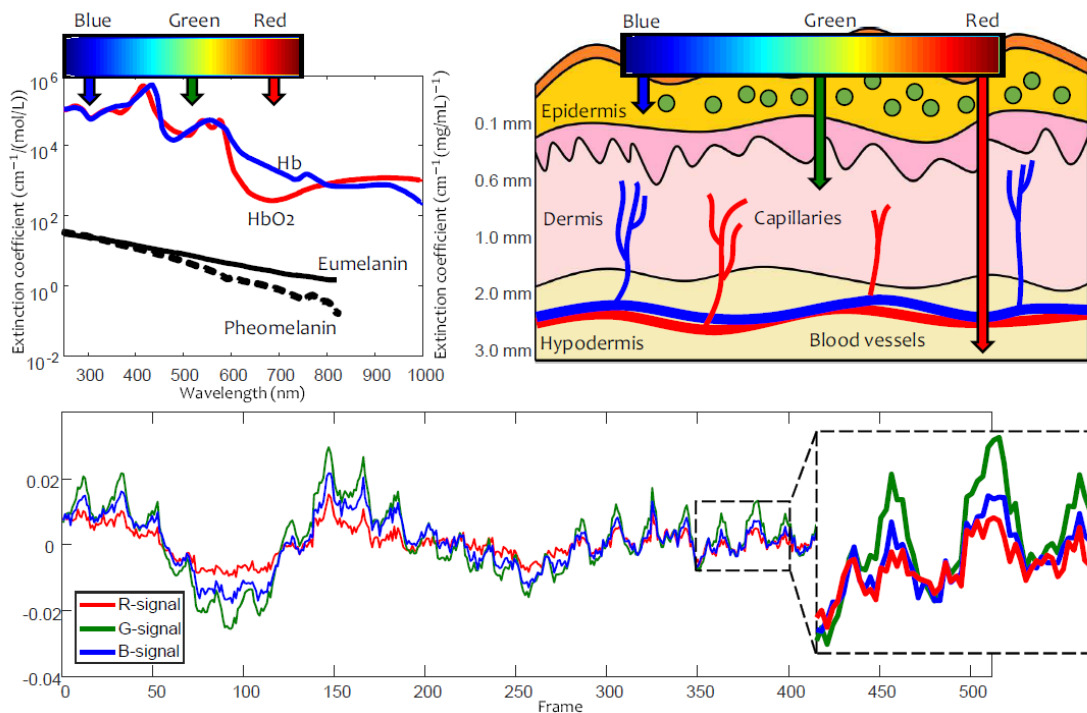


Ilustración 2: Comportamiento de las componentes RGB de la luz con la piel [4]. Arriba-izquierda se muestra el coeficiente de extinción, arriba-derecha se muestra el nivel de penetración en la piel de distintas longitudes de onda de la luz y abajo se muestra la señal pulsátil obtenida tras analizar la luz reflejada.

La componente azul presenta una señal débil, esto se debe a que no logra atravesar la epidermis y la señal reflejada contiene poca información acerca del flujo sanguíneo. La componente roja también presenta una señal débil, esto se debe a que es la señal que más tejidos atraviesa y gran parte de la señal se pierde debido a la absorción o dispersión en el camino de ida y vuelta. El espectro correspondiente al rojo (600-750 nm) se emplea en la medición del nivel de oxígeno en la sangre. La componente verde contiene la señal más clara del flujo sanguíneo, su amplitud varía mucho más que en el caso de la componente azul o roja. Logra atravesar la epidermis hasta llegar a los capilares donde el flujo de sangre se detecta correctamente, pero no atraviesa hasta los principales vasos sanguíneos donde la señal se debilita. Esta es la señal usada en

sistemas PPG monocanales, sin embargo, hay algoritmos que combinan los tres canales para obtener una mejor señal de salida.

Existen diversos métodos para la medición de signos vitales mediante PPG, pero en términos generales se pueden resumir en los siguientes dos [1]:

- Una fuente de luz (ej. un LED) ilumina y un fotodetector situado al otro lado recibe la luz que logra traspasar el tejido. Se emplea en zonas estrechas como un dedo y una oreja.
- Una fuente de luz (ej. un LED) ilumina y un fotodetector situado al lado recibe la luz que refleja el tejido superficial e interno [5]. Esta técnica permite el empleo de PPG en otras áreas como es la frente o la muñeca. En este caso la fuente de luz y el receptor pueden estar en contacto directo con la piel o no.

Hoy en día los sistemas basados en PPG se están desarrollando con fuentes de luz blanca o incluso sin fuente de luz artificial y como detector se emplean cámaras que capturan las imágenes que serán procesadas para extraer los signos vitales [4]. Este salto se debe al desarrollo de los sistemas basados en la reflexión de la luz sobre los tejidos sin necesitar el contacto directo. El simple hecho de evitar el contacto directo supone un gran cambio en el funcionamiento del sistema. Empleando el sistema de contacto directo la mayor parte de la luz recibida procede de tejidos internos (como arterias), sin embargo, al emplear el sistema sin contacto directo también se obtiene señal procedente de tejidos más superficiales [8 y 9]. Si bien es un avance en la búsqueda de sistemas cada vez menos intrusivos, también puede suponer un problema debido a reflexiones indeseadas que no contienen información sobre el flujo sanguíneo. Estos sistemas se conocen como Remote Photoplethysmography (rPPG).

Las variaciones detectadas empleando PPG son procesadas y se pueden llegar a extraer distintos parámetros como [1]:

- **Nivel de oxígeno en la sangre.** Es útil en diversos escenarios clínicos como son un hospital, medicina deportiva o clínica veterinaria entre otros. Para la obtención del nivel de oxígeno se emplea el espectro correspondiente al rojo (600-750 nm) y cercano al infrarrojo (850-1000 nm), que son los que logran atravesar más la piel y también son muy sensibles a cambios en el nivel de oxígeno en la sangre [6].
- **Ritmo cardiaco.** El ritmo cardiaco es un parámetro fundamental a la hora de monitorizar las constantes vitales de un paciente en un hospital o una ambulancia. La medición del ritmo cardiaco empleando PPG puede verse afectada por el movimiento o una arritmia cardiaca, por estos motivos es fundamental procesar la señal para detectar y eliminar estos artefactos. En la actualidad existen diversos algoritmos capaces de extraer el ritmo cardiaco de forma fiable.

- **Presión sanguínea.** La presión sanguínea es un parámetro a monitorizar fundamental de cara a detectar enfermedades o problemas cardiovasculares.
- **Respiración.** Este parámetro cobra especial relevancia en el cuidado de recién nacidos, estudios del sueño o con el empleo de ciertos fármacos que bloquean la sensibilidad o consciencia.

Además rPPG permite la monitorización de constantes vitales sin necesitar el contacto directo con el paciente, esta idea cobra especial relevancia en ciertos entornos [4]:

- Monitorización clínica de la salud, el uso de sistemas intrusivos puede suponer molestias o incluso dolor en algunos pacientes (ej. problemas en la piel o recién nacidos).
- Monitorización de la salud en el hogar, de esta manera se puede monitorizar las constantes vitales de pacientes desde su propia casa. En este escenario el uso del rPPG no tiene que estar directamente relacionado con problemas en la salud, también se puede utilizar mientras se hace ejercicio o mientras se duerme.
- Monitorización de conductores, esta aplicación toma cada vez más relevancia con el desarrollo de coches inteligentes. La posibilidad de extraer información del conductor y analizar el cansancio, la atención, si se está alterado o tranquilo, o desmayos es fundamental para amentar la seguridad en las carreteras y disminuir los riesgos.
- Monitorización durante el deporte, esta aplicación también ha tomado especial relevancia en los últimos años. Con el incremento de los wearables la monitorización del pulso y respiración entre otros parámetros se ha vuelto un proceso habitual durante la práctica de ejercicio. Sistemas rPPG integrados en dispositivos pueden ofrecer este tipo de información de forma cómoda sin afectar al movimiento del usuario.
- Vigilancia, la evolución de las cámaras de seguridad abre la puerta a la entrada del rPPG. Estas cámaras suelen estar conectadas a un centro encargado de procesar las imágenes para detectar movimiento, identificar caras o analizar el entorno entre otras cosas. La posibilidad de analizar las constantes vitales de una persona a distancia abre la puerta a detectar el momento en el que alguien va a tener una reacción violenta, o una emergencia.
- Sistemas de seguridad, en la actualidad existen muchos sistemas empleados para permitir o bloquear el acceso a algo (datos, dispositivos, lugares). La huella dactilar, el análisis de voz o el análisis facial entre otros. Estos sistemas se basan en analizar ciertos parámetros, sin embargo, estos sistemas se pueden falsear con elementos artificiales que cumplan con esos parámetros (caretas, imitación de huellas dactilares, fotos). Para evitar este problema se puede utilizar rPPG para buscar pulso en el tejido a analizar.

- Búsqueda y rescates, en este entorno el tener una herramienta capaz de analizar un área y buscar tejido con pulso puede suponer la diferencia entre la vida o la muerte. En derrumbamientos o tras avalanchas de nieve la búsqueda de supervivientes no es sencilla.
- Entretenimiento, el mundo del entretenimiento se basa en ofrecen una experiencia cada vez más inmersiva al usuario. La posibilidad de detectar el comportamiento del usuario puede dar lugar a sistemas de entretenimiento inteligentes capaces de adaptarse.

En este trabajo se analizará la respuesta de un sistema rPPG con cámara RGB, su funcionamiento es muy similar al que se observa en la Ilustración 3. La cámara captura imágenes del rostro de una persona y el algoritmo extrae los niveles RGB de una ROI (Region Of Interest), entonces de la combinación de esas señales RGB se extrae el parámetro deseado. En este caso el sistema está diseñado para extraer el pulso. Los algoritmos de este tipo suelen añadir procesamiento extra para eliminar artefactos indeseados, en este caso se filtra la distorsión causada por el movimiento de la cabeza.

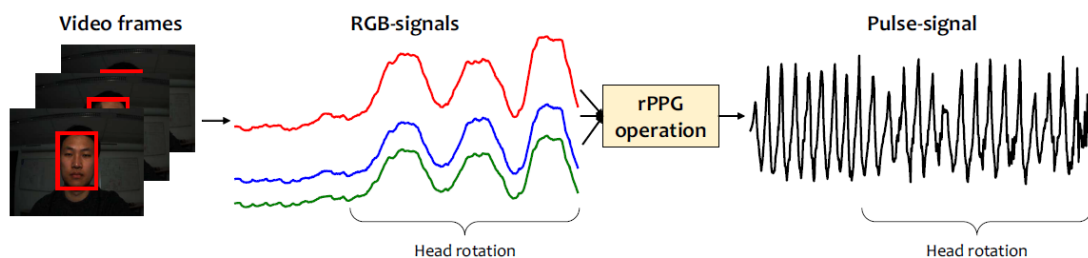


Ilustración 3: rPPG con cámara RGB [4].

Esta configuración tiene sus ventajas y desventajas en comparación con configuraciones de contacto directo [4]:

Ventajas:

- Como ya se vio con anterioridad al eliminar el contacto directo se permite la movilidad y elimina cualquier posible irritación/molestia que pueda causar el sensor con la piel.
- Mediciones multizona, la cámara captura imágenes de toda la cara y esto permite analizar varias zonas simultáneamente con un único sensor. La obtención del pulso en distintas zonas al mismo tiempo ayuda a obtener un resultado final más fiable y abre la puerta a nuevos análisis como la velocidad del flujo sanguíneo, tiempo de propagación del pulso y presión sanguínea entre otros.

- Mediciones multipaciente, la cámara puede capturar imágenes de varias personas a la vez y esto da la posibilidad de analizar a múltiples pacientes simultáneamente con un único sensor.
- Mediciones multilongitud de onda, las cámaras RGB capturan imágenes en todo el espectro visible e incluso pueden llegar a capturar señal en el infrarrojo cercano (entorno a los 800-900 nm). Esa posibilidad abre las puertas a la combinación de canales para la obtención de una señal de salida más robusta y se pueden emplear distintas longitudes de onda para extraer distintos parámetros simultáneamente, el canal verde presenta buen nivel de salida para la obtención del pulso mientras que el rojo o cercano al infrarrojo se emplea para la extracción del nivel de oxígeno en la sangre.

Desventajas

- Los sistemas ECG (electrocardiograma) miden el ritmo cardiaco mientras que los sistemas PPG miden los cambios en el flujo sanguíneo, esto puede ser un problema para ciertas enfermedades/condiciones donde el pulso del flujo sanguíneo no está sincronizado con el pulso del corazón. Los sistemas ECG son imprescindibles si se quiere obtener el ritmo cardiaco de forma fiable en personas con enfermedades relacionadas con el corazón. rPPG también presenta algunos problemas de fiabilidad en sujetos en movimiento, si se quiere medir con precisión a un sujeto mientras hace deporte un sistema ECG sigue siendo más fiable.
- Los sistemas PPG de contacto directo son fácilmente integrables en wearables y es factible su uso continuado, en cambio un sistema rPPG basado en una cámara es mucho más difícil de integrar y solo es capaz de monitorizar mientras el sujeto permanece en la imagen.

rPPG es una tecnología que todavía se está desarrollando y se enfrenta a varios desafíos: las cámaras, la iluminación y la piel.

Desde el punto de vista de las cámaras no es necesario recurrir a las caras, webcams básicas son suficientes para extraer el pulso [10]. Sin embargo, cuando se busca extraer otros parámetros, como el nivel de oxígeno en la sangre, se buscan cámaras que sean capaces de capturar señal en los 800-900 nm, o cámaras con alta captura de imágenes por segundo para monitorizar el flujo de sangre.

También es importante que los canales de la cámara se solapen lo menos posible para poder utilizar cada canal de forma independiente del resto. Todos los sistemas automáticos de ajustes de imagen deben estar apagados y el formato de la imagen de salida debe ser raw, de esta manera se evitan ajustes indeseados y la pérdida de información al aplicar compresión sobre la imagen. Un sistema de estabilización de imagen puede ser necesario en casos donde la cámara esté en constante movimiento (ej. un coche o un dron).

En cuanto a la iluminación, este es un parámetro crítico. Los sistemas PPG se basan en analizar la luz procedente de los tejidos de la piel. Si se quiere emplear un sistema basado en una cámara RGB es suficiente con la luz visible, sin embargo, si se emplea una cámara capaz de medir en el infrarrojo cercano podría ser necesario iluminar la zona con LEDs que emitan en la longitud de onda deseada.

La intensidad también es importante, debe ser lo suficientemente alta como para que penetre en la piel y la señal recibida tenga una amplitud lo suficientemente alta como para que sea detectada por el sensor. Además, la amplitud de la señal pulsátil obtenida es proporcional a la intensidad de la luz, mayor intensidad supone una señal más robusta [11]. Pero también hay que tener cuidado, un exceso de luz supone un problema dado que satura la imagen y hace imposible extraer el pulso de forma fiable. Una fuente de luz ideal sería aquella que ilumina de forma homogénea la zona a iluminar, ya que fuentes de luz que centren el haz en una zona pueden dar problemas con reflexiones. La luz debe ser constante, el parpadeo y variaciones similares dificultan la detección de las señales, especialmente si suceden a una frecuencia similar a la de la señal a analizar.

La piel también supone un desafío en la extracción de los signos vitales. La composición de la piel no es homogénea en todo el cuerpo, hay zonas con una mayor concentración de agua, zonas con más arterias y zonas con más grasa. La cara, las manos, las orejas o el pie muestran señales pulsátiles más fuertes que en zonas como los brazos, el cuello o el pecho [12].

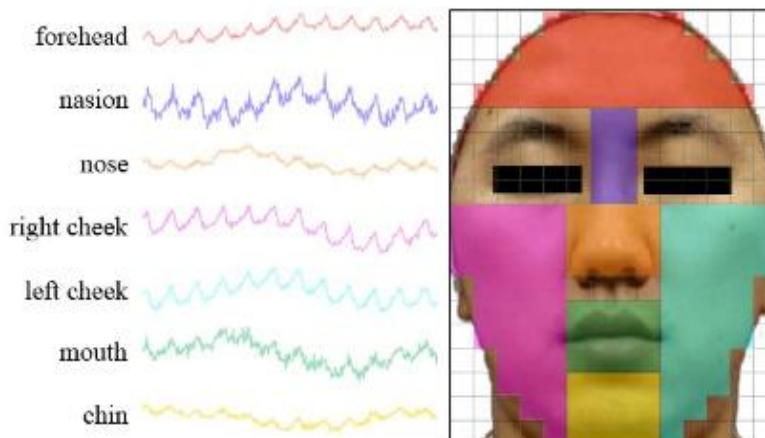


Ilustración 4: Señal pulsátil en diferentes partes de la cara [12].

En cuanto al tono de la piel, este también es un aspecto a tener en cuenta. Piel más oscuras con una concentración mayor de melanina pueden ser problemáticas, la melanina absorbe la luz reflejada en los tejidos internos y esto elimina señal útil en el receptor mientras que no reduce la reflexión producida en la superficie externa de la piel [11]. Estas dos situaciones dan como resultado una señal pulsátil débil con un nivel de SNR bajo.

Otro aspecto a tener en cuenta es el pelo, zonas cubiertas por pelo (como puede ser la frente o una barba cubriendo las mejillas) puede ser un problema. Lo mismo pasa con el maquillaje. Un aspecto fundamental en los algoritmos rPPG es su capacidad de detectar piel, si se es capaz de eliminar las zonas cubiertas por pelo o zonas como los ojos y boca se facilita mucho la detección de la señal pulsátil. A este aspecto hay que sumar la capacidad de filtrar movimientos tanto voluntarios (como la rotación de la cabeza) como involuntarios (como es la respiración).

Una vez se obtiene la imagen procesada con el área a analizar existen diversos algoritmos capaces de extraer la señal pulsátil. Todos ellos tratan de extraer la señal pulsátil a partir de los niveles RGB captados en la zona de interés. Para ello parten de [4]:

$$C_k(t) = I(t) \cdot (V_s(t) + V_d(t)) + V_n(t) \quad (1.1)$$

Donde $C_k(t)$ son los canales RGB del pixel k-ésimo (un $C_k(t)$ por cada canal), $I(t)$ es la intensidad luminosa, $V_s(t)$ es la reflexión especular, $V_d(t)$ es la reflexión difusa, $V_n(t)$ es el ruido de cuantificación de la cámara y t es el tiempo.

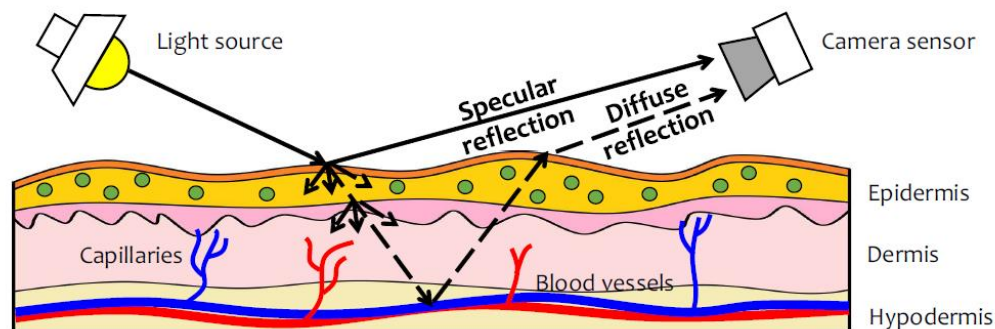


Ilustración 5: Interacción de la luz con la piel [4].

La reflexión especular es la que se produce en la superficie de la piel y no contiene ninguna información sobre el flujo sanguíneo. La reflexión difusa es la que se produce en los tejidos internos y varía con los cambios en el flujo sanguíneo.

El objetivo de los algoritmos PPG es el de extraer el pulso de las señales $C_k(t)$. La mayoría de los algoritmos emplean la media espacial de los valores RGB del área a analizar, de esta manera $C_k(t)$ pasa a ser una sola señal temporal $C(t)$. Además, si el número de píxeles es suficientemente alto, esto permite reducir el ruido de cuantificación $V_n(t)$ hasta volverse despreciable [4]. Por este motivo se debe buscar zonas de piel lo más amplias posible.

A continuación se presentarán e introducirán brevemente los principales algoritmos rPPG [4]:

- **Métodos basados en BSS (Blind Source Separation):**

BSS trata de separar una señal de una mezcla de señales sin emplear información (o muy poca) de las señales que forman la mezcla o del proceso de mezclado. En el caso de PPG se trata de extraer la señal pulsátil $p(t)$ de $C(t)$.

$$Y(t) = W \cdot C(t) \quad (1.2)$$

$Y(t)$ son las señales que componen $C(t)$, el pulso y el ruido, factorizadas. W es la matriz de “separación” obtenida mediante PCA [13] o ICA [14]. La diferencia entre emplear un método u otro depende de las suposiciones (ej. si las señales que forman $C(t)$ no son correladas o son independientes). El siguiente paso consiste en seleccionar la señal más periódica de $Y(t)$ y esta señal obtenida será la señal pulsátil [4].

La principal desventaja de este método es el problema que causa una señal periódica no relacionada con el pulso en la extracción del mismo. En casos donde el movimiento es periódico se puede confundir este con la señal pulsátil (ej. cuando se hace ejercicio o la misma respiración).

Estos métodos son soluciones generales en el procesado de señales que se aprovechan de ciertas características estadísticas que no contemplan todas las posibilidades o al menos las suficientes (como el suponer que la única señal periódica será la del pulso).

- **PBV/CHROM :**

Al contrario que los métodos basados en BSS, PVB y CHROM emplean las distintas características que ofrecen los tres canales de color RGB para extraer el pulso de $C(t)$. Los sistemas PPG tratan de detectar los cambios en el color de la piel debidos al flujo sanguíneo y de esta manera obtener la señal pulsátil. Además de estos cambios debidos a la sangre hay que tener en cuenta que la piel tiene un color propio que permanece constante y la luz en sí también tiene componentes que no varían.

Un primer paso de los métodos PBV y CHROM consiste en eliminar estos niveles de DC, dividiendo las señales RGB por sus medias temporales se consiguen eliminar las componentes constantes relativas a la propia luz o al color mismo de la piel [4]. A partir de este punto PBV y CHROM presentan estrategias distintas.

- **PBV**

Una vez obtenido la señal $C(t)$ sin las componentes en DC la señal pulsátil se obtiene de [4]:

$$\hat{p}(t) = \tilde{C}_n^T(t) \cdot z \quad (1.3)$$

$$\tilde{C}_n(t) = C_n(t) - 1 \quad (1.4)$$

Donde $\hat{p}(t)$ es una estimación de la señal pulsátil $p(t)$, $C_n(t)$ es $C(t)$ con media cero y z es [4]:

$$z \propto \tilde{\varepsilon}^{-1} \cdot u_{pbv} \quad (1.5)$$

$$\tilde{\varepsilon} = \{ \overline{\tilde{C}_n(t) \cdot \tilde{C}_n^T(t)} \} \quad (1.6)$$

Con u_{pbv} siendo un vector 3x1 que contiene el volumen de sangre del pulso y $\tilde{\varepsilon}$ la matriz 3x3 de la covarianza temporal.

La dificultad al emplear PBV no es $\tilde{\varepsilon}$, que se extrae del propio vídeo, es la obtención de u_{pbv} . Para ello es necesario tener información previa acerca del volumen de sangre.

- **CHROM**

A diferencia de PVB, CHROM busca reducir la dependencia del algoritmo con la necesidad de tener información previa. Para ello solo emplea información contenida en la señal de crominancia proyectándola de forma que se obtengan dos señales proyección donde las variaciones debidas al movimiento o al pulso se muestren en fase o en contrafase. De nuevo se calcula una predicción de la señal pulsátil [4]:

$$\hat{p}(t) = S_1(t) - \alpha \cdot S_2(t) \quad (1.7)$$

$$\alpha = \frac{\sigma(S_1)}{\sigma(S_2)} \quad (1.8)$$

$$S(t) = P_c \cdot M \cdot \tilde{C}_n(t) \quad (1.9)$$

Donde $S_i(t)$ es la i -ésima proyección de la crominancia, α es la división de las desviaciones estándar de $S_1(t)$ y $S_2(t)$ y $P_c \cdot M$ es una matriz con valores constantes calculados a partir de un vector de estandarización de color de piel [4 y 29].

De esta manera cuando la señal pulsátil es la que domina la señal $S_1(t)$ y $S_2(t)$ se suman de forma constructiva, en cambio, cuando es el movimiento el que domina la señal se cancelan las componentes relativas al movimiento. El único problema que tiene este sistema es cuando ni las variaciones debidas al pulso ni las variaciones debidas al movimiento dominan y existe un balance entre ambas, en ese momento la predicción obtenida $\hat{p}(t)$ no es lo suficientemente buena [4].

- **Método 2SR**

Este método crea un espacio que clasifica el color de la piel del sujeto y detecta cambios en el nivel de tono, en lugar de la intensidad, para extraer el pulso. Para ello calcula el tono en el área de piel a analizar mediante una distribución estadística. Este método se ha comprobado que funciona [16] pero no tiene en consideración aspectos fisiológicos y puede dar problemas si la zona a analizar no es buena o tiene mucho ruido.

- **Método del plano ortogonal a la piel**

POS (Plane Orthogonal to Skin) es el algoritmo en el que se basa el sistema rPPG empleado más adelante en el trabajo. Se basa en un principio similar al método CHROM, mediante proyecciones logra separar la señal pulsátil de otras.

En el caso de POS la señal " $C_n(t)$ se proyecta en un plano ortogonal a 1 en el espacio RGB temporalmente normalizado, que es un plano ortogonal al tono de piel temporalmente normalizada" [4]. La diferencia entre POS y CHROM es la siguiente, en el caso de CHROM el plano de proyección es ortogonal a la dirección de variación especular asumiendo un vector de tono de piel estandarizado, en cambio, en el caso de POS el plano de proyección es ortogonal a la dirección de variación de intensidad [4].

$$S(t) = P_p \cdot C_n(t) \quad (2.1)$$

Donde P_p es la matriz 2x3 de proyección.

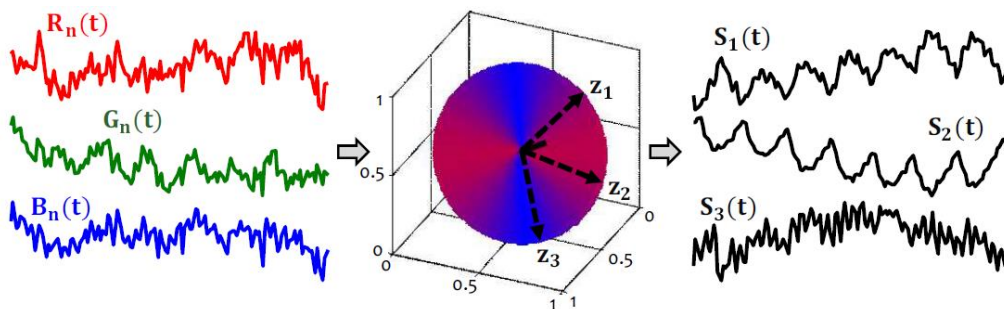


Ilustración 6: Distribución de la fuerza de la señal pulsátil en el plano ortogonal a 1 [4].

En la ilustración 6 se observa una esfera donde el color rojo señala las regiones con una mayor fuerza de la señal pulsátil y las zonas azules señalan las regiones con una señal pulsátil débil. En esta figura se ve también la dependencia de la dirección de la proyección con la fuerza de la señal pulsátil. Empleando Z_1 y Z_2 se obtienen señales pulsátiles claras pero en contrafase una respecto a la otra, sin embargo, Z_3 muestra una señal pulsátil débil con mucho ruido. Es por este motivo que se debe escoger cuidadosamente los ejes de la proyección.

$$P_p = \begin{pmatrix} 0 & 1 & -1 \\ -2 & 1 & 1 \end{pmatrix} \quad (2.2)$$

$$S(t) = \begin{pmatrix} S_1(t) \\ S_2(t) \end{pmatrix} = \begin{pmatrix} S_1(t) = G_n(t) - B_n(t) \\ S_2(t) = G_n(t) + B_n(t) - 2R_n(t) \end{pmatrix} \quad (2.3)$$

Tras escoger los valores de P_p y $S(t)$ anteriores [4].

$$h(t) = S_1(t) + \alpha \cdot S_2(t) \quad (2.4)$$

$$\alpha = \frac{\sigma(S_1(t))}{\sigma(S_2(t))} \quad (2.5)$$

Siendo σ la desviación estándar y $h(t)$ la señal pulsátil para un corto periodo de tiempo. De esta manera cuando la señal dominante es la del pulso $S_1(t)$ y $S_2(t)$ se suman en fase de forma constructiva y cuando la señal que domina es el movimiento se suman en contrafase y se cancelan las componentes relativas al movimiento.

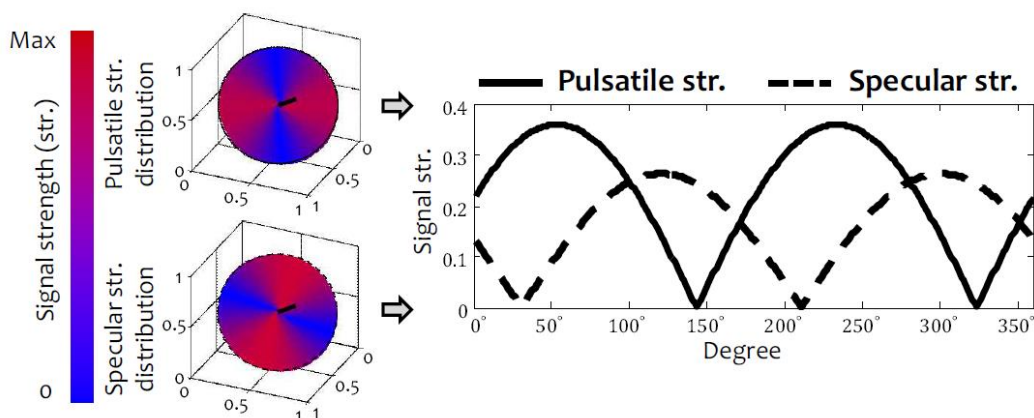


Ilustración 7: Distribución de una señal pulsátil fuerte y una señal con componente especular dominante [4].

Al igual que con el método CHROM cuando ni la señal pulsátil ni el movimiento dominan la predicción no es completamente fiable [4].

1.2. Índices de calidad empleados

En la actualidad existen numerosos sistemas para extraer los signos vitales de una persona, para valorar la efectividad de los métodos actuales y los que se están desarrollando es necesario emplear índices de calidad que los evalúen.

En este apartado se analizarán los resultados obtenidos utilizando algunos de los índices de calidad más comúnmente empleados en la actualidad para evaluar señales pulsátiles. La información sobre las distintas técnicas de análisis de señales PPG se ha obtenido de [17]. Para ello se emplearán tres vídeos, cada vídeo cuenta con una señal predicción (obtenida mediante el algoritmo explicado más adelante en el apartado Evaluación Pulsecam) y una referencia obtenida empleando un sistema profesional:

- Vídeo 25_5, se considera una predicción buena por la similitud de valores entre la predicción y la referencia y además la predicción es capaz de seguir a la referencia correctamente.
- Vídeo 25_3, se considera una predicción mala por la gran diferencia entre los valores de la predicción y la referencia. Además, la predicción no es capaz de seguir las variaciones de la referencia tampoco.
- Vídeo 30_9, esta predicción tiene zonas donde los valores entre la predicción y la referencia son similares y también tiene zonas donde la predicción sigue las variaciones de la referencia.

Una predicción buena:

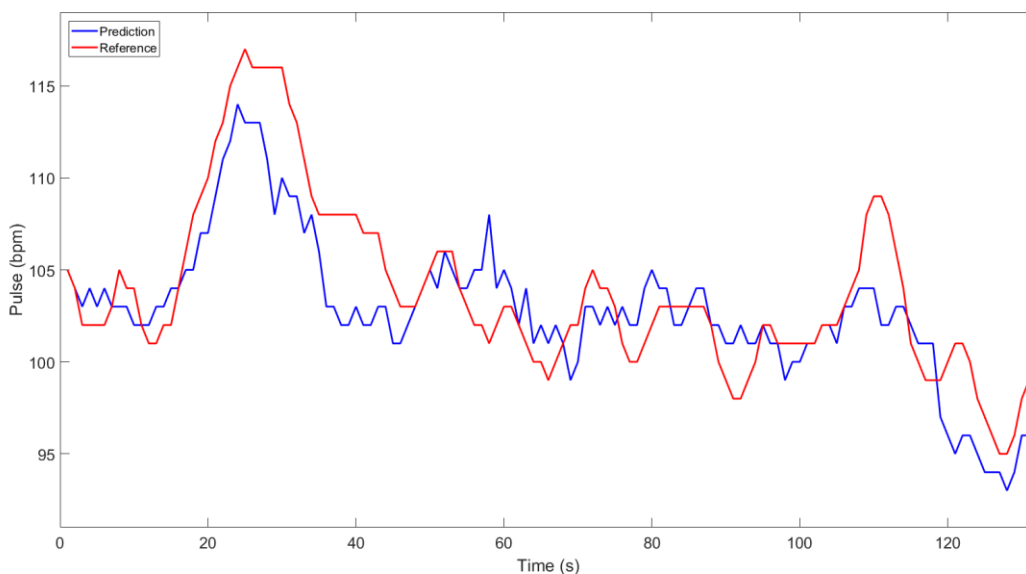


Ilustración 8: Vídeo 25_5, predicción buena.

Una predicción mala:

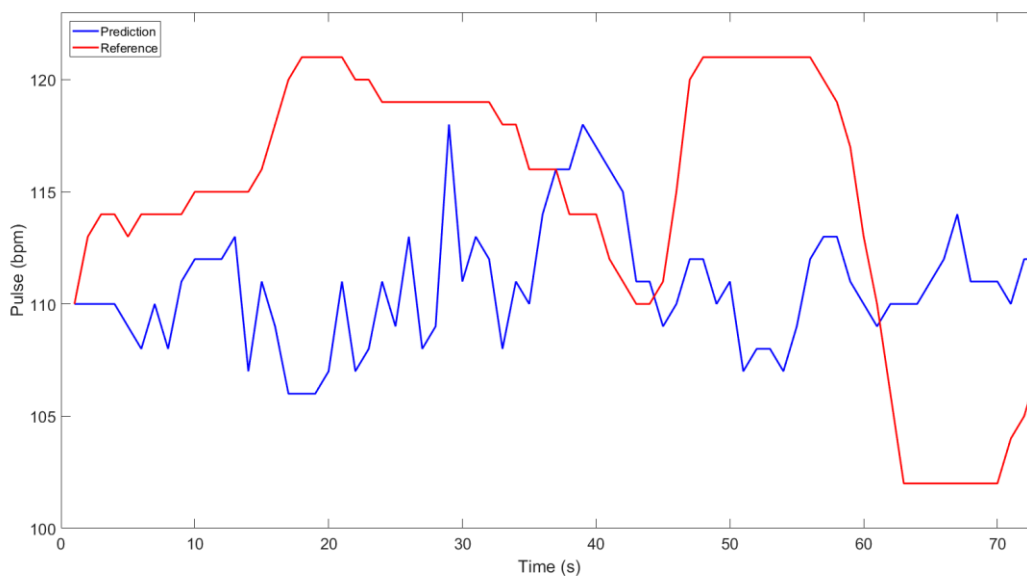


Ilustración 9: Vídeo 25_3, predicción mala.

Una predicción parcialmente buena y parcialmente mala:

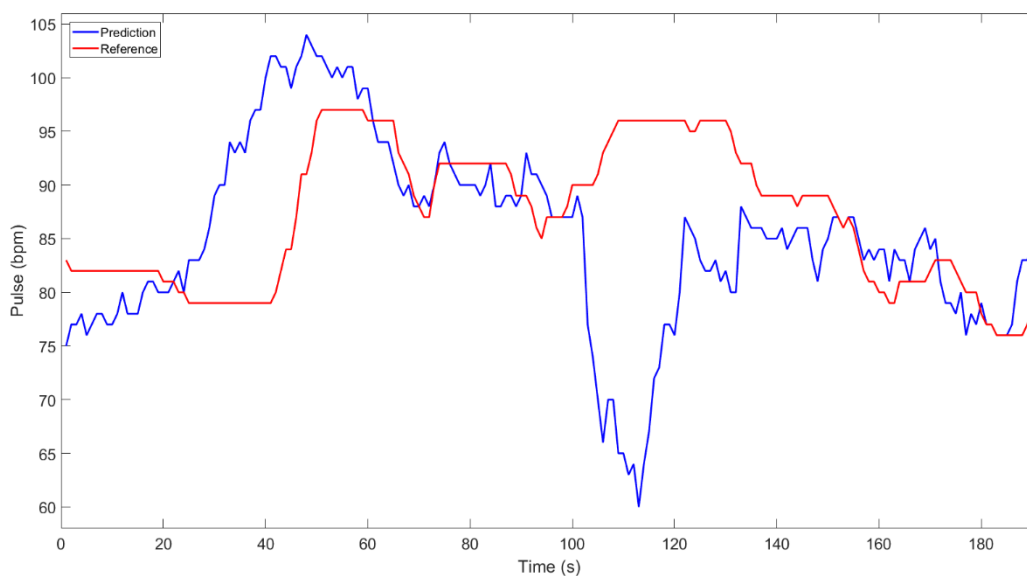


Ilustración 10: Vídeo 30_9, predicción parcialmente buena y parcialmente mala.

Para el cálculo de los siguientes índices de calidad es importante tener en cuenta los siguientes puntos:

- Señal PPG: señal que contiene las variaciones en el flujo sanguíneo. Dado que la predicción se obtuvo empleando un sistema basado en POS, la señal PPG es la señal $h(t)$ vista en el apartado anterior. Por cada segundo en el vídeo existe una señal PPG.
- La señal PPG suele filtrarse mediante un filtro paso banda para eliminar interferencias. La señal PPG sin procesar es aquella anterior al filtrado.

- **Perfusion (P_{SQI}):**

El primer índice de calidad a analizar será Perfusion (P_{SQI}). P_{SQI} mide la diferencia de luz absorbida, este parámetro se calcula empleando la siguiente fórmula:

$$P_{SQI} = \left[\frac{y_{max} - y_{min}}{|\bar{x}|} \right] \cdot 100 \quad (2.6)$$

Siendo $|\bar{x}|$ el valor absoluto de la media estadística de la señal PPG sin procesar y y_{min} y y_{max} los valores mínimo y máximo de la señal PPG tras el filtrado.

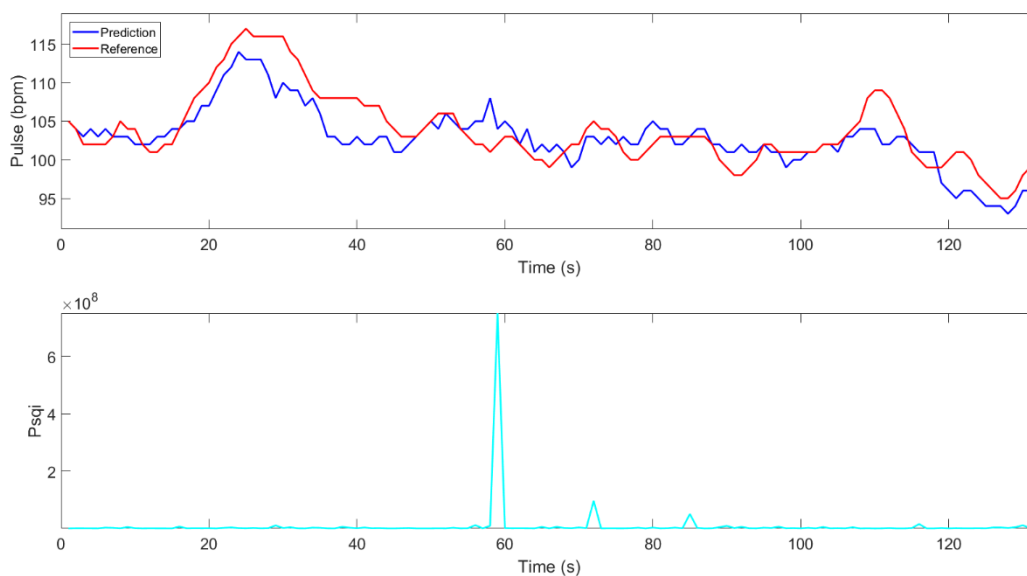


Ilustración 11: Vídeo 25_5, P_{SQI} .

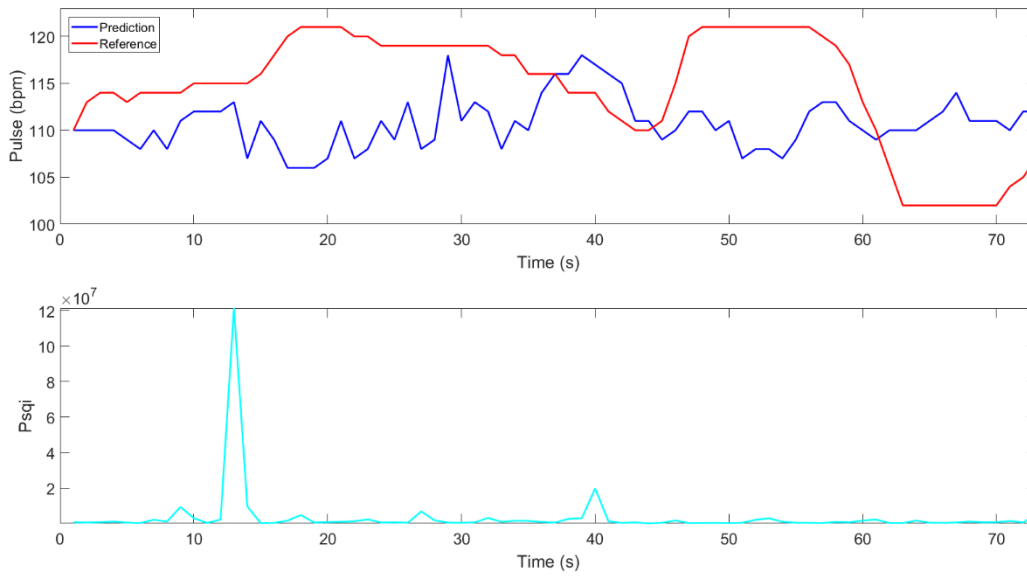


Ilustración 12: Vídeo 25_3, P_{sqi} .

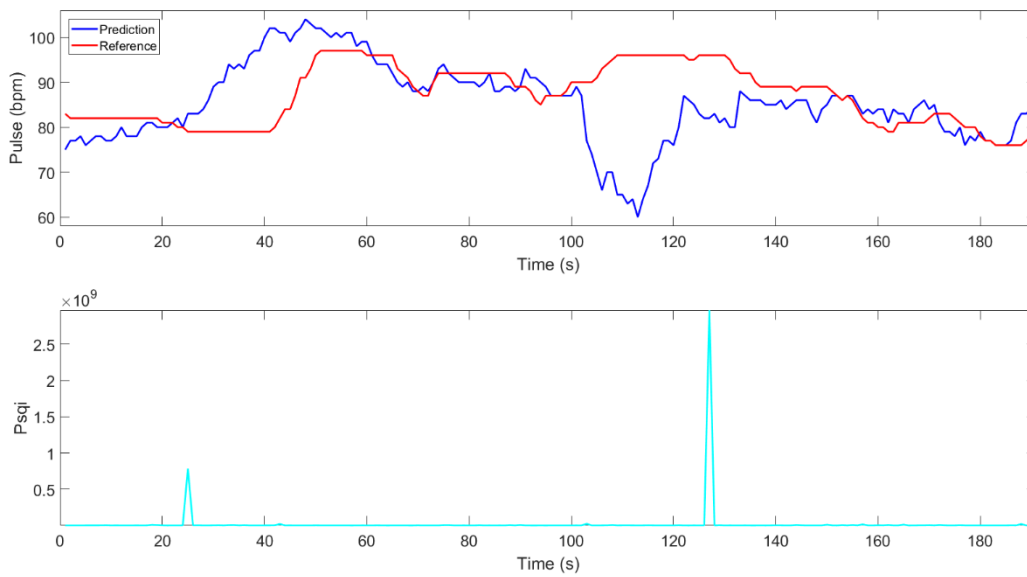


Ilustración 13: Vídeo 30_9, P_{sqi} .

Este índice es óptimo para sistemas con poco ruido donde interese detectar los puntos de la predicción donde la señal no es lo suficientemente intensa. De esta manera se puede valorar como de buena es la predicción en función de la fuerza de la señal PPG.

El inconveniente de este sistema es que no considera el ruido, un pico de ruido no filtrado altera considerablemente el resultado. Además, es fundamental no extraer la componente en DC de la señal PPG. El algoritmo empleado en este trabajo para la extracción de la señal PPG elimina la componente en DC y esto hace inservible este índice de calidad, dado que la media de la señal PPG es 0 o muy cercana a 0. Este es el motivo de la aparición de los picos en las gráficas relativas al P_{SQI} en las ilustraciones 11, 12 y 13 que realmente no señalan información útil.

- **Skewness (S_{SQI}):**

Este índice de calidad mide la simetría de la señal para detectar partes “corruptas”. Para su cálculo se emplea la siguiente fórmula:

$$S_{SQI} = \frac{1}{N} \sum_{i=1}^N [x_i - \hat{\mu}_x]^3 \quad (2.7)$$

Donde N es el número de muestras de la señal PPG, x_i es la señal PPG sin procesar y $\hat{\mu}_x$ y σ son la media y desviación estándar de x_i respectivamente.

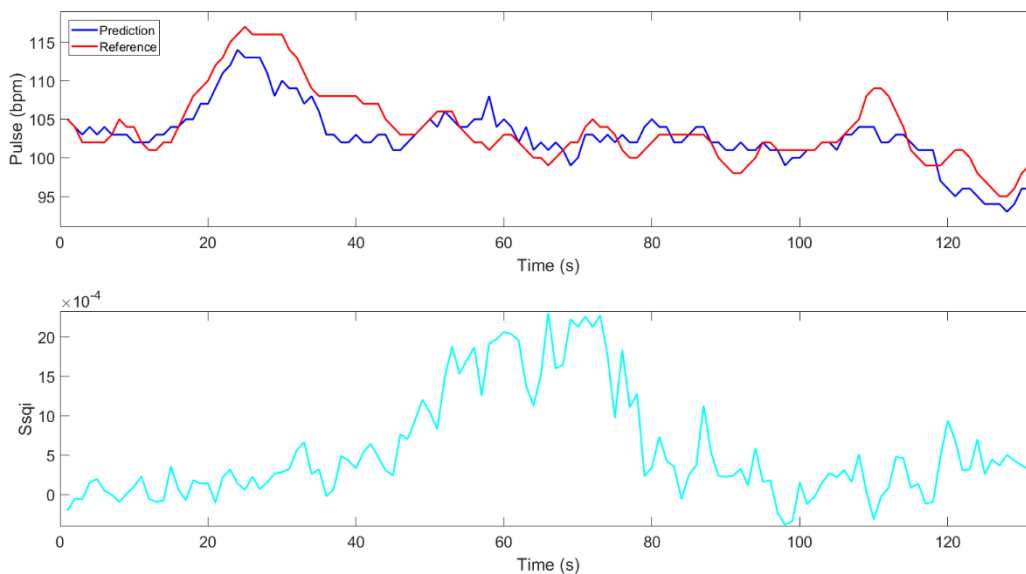


Ilustración 14: Vídeo 25_5, S_{SQI} .

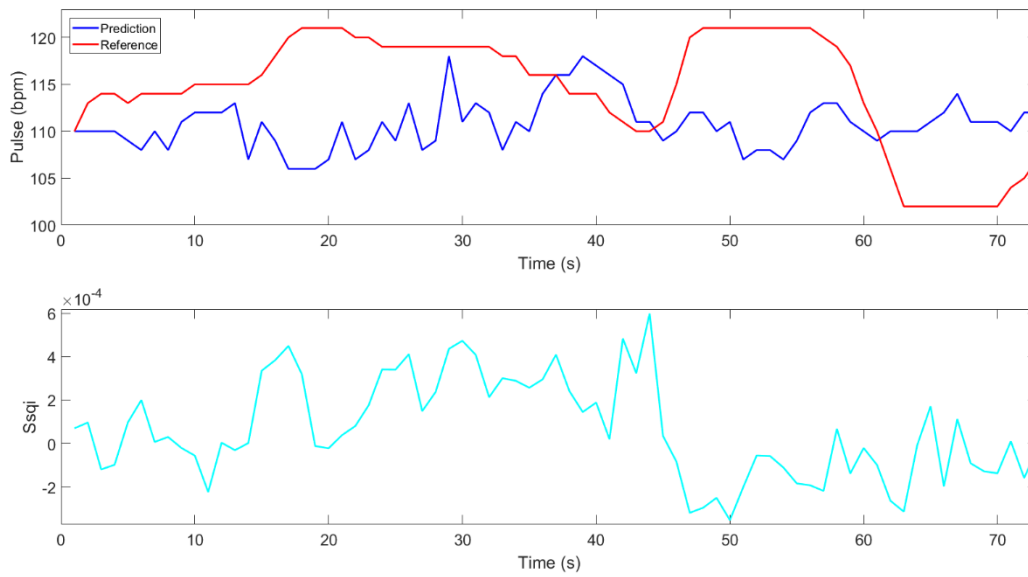


Ilustración 15: Vídeo 25_3, S_{sqj} .

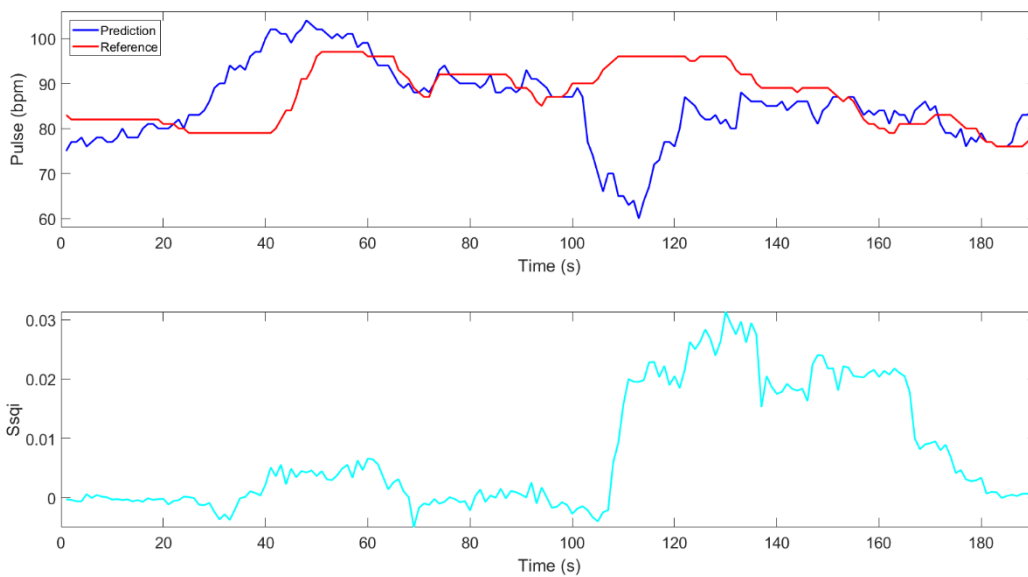


Ilustración 16: Vídeo 30_9, S_{sqj} .

De nuevo se trabaja sobre la señal PPG sin procesar, esto supone que el ruido o picos debidos a movimiento afectan al cálculo de calidad aun si se logran eliminar con el procesado posterior. Los valores de S_{sqj} obtenidos en los vídeos 25_3 y 25_5 son

muy similares entre sí, se encuentran entre $-3 \cdot 10^{-4}$ y $25 \cdot 10^{-4}$. En cambio, en el vídeo 30_9 los valores se encuentran entre -0.01 y 0.03, dos órdenes de magnitud por encima.

Los valores no son comparables en zonas de predicción buena entre vídeos y tampoco se distinguen zonas donde la predicción empeora o mejora

- **Kurtosis (K_{SQI}):**

Se emplea para describir la distribución de datos alrededor de la media. Su fórmula es muy similar a la de Skewness:

$$S_{SQI} = \frac{1}{N} \sum_{i=1}^N \left[x_i - \frac{\hat{\mu}_x}{\sigma} \right]^4 \quad (2.8)$$

Donde N es el número de muestras de la señal PPG, x_i es la señal PPG sin procesar y $\hat{\mu}_x$ y σ son la media y desviación estándar de x_i respectivamente.

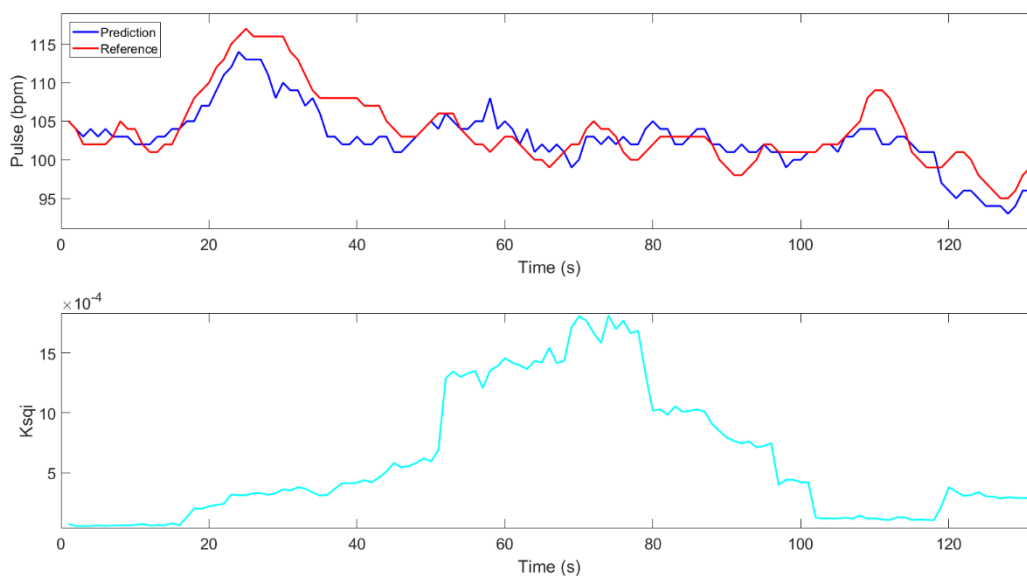


Ilustración 17: Video 25_5, K_{SQI} .

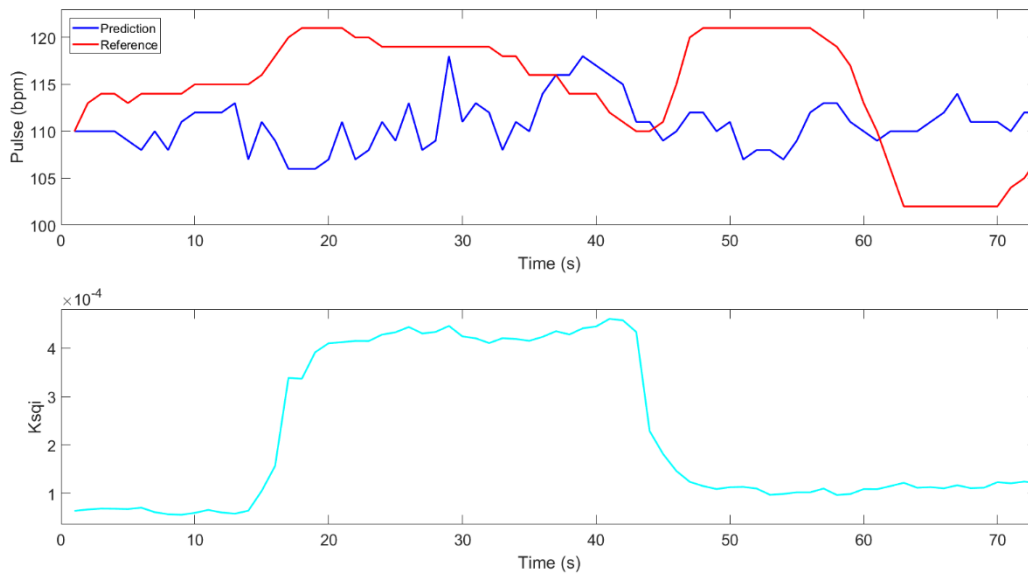


Ilustración 18: Video 25_3, K_{sqj} .

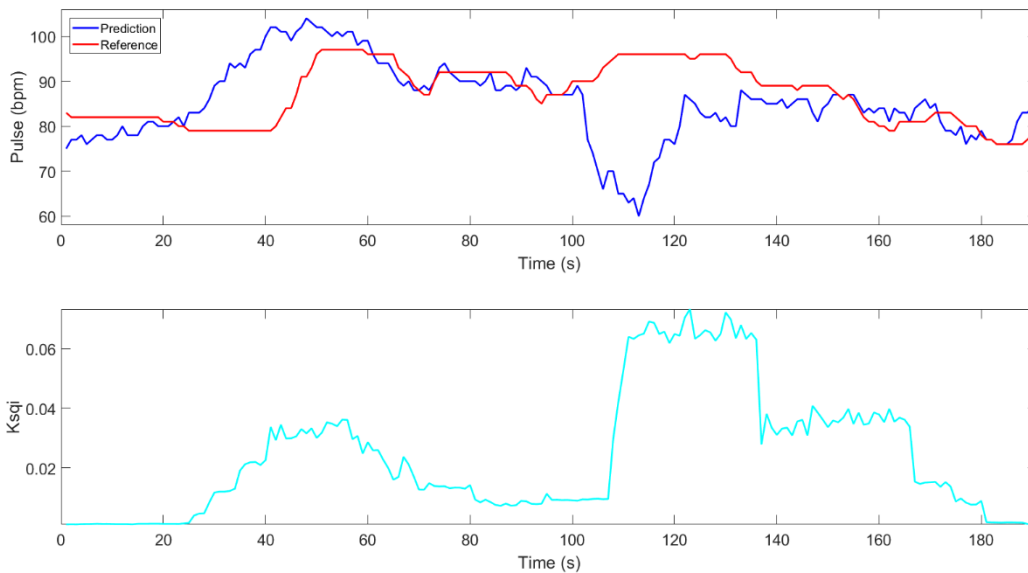


Ilustración 19: Video 30_9, K_{sqj} .

Este índice de calidad tampoco es consistente. Si bien en el video 30_9 es capaz de indicar las zonas donde la referencia y la predicción difieren, en los videos 25_3 y 25_5 el orden de magnitud es dos veces menor, esto los hace incomparables con el video 30_9, y no indica claramente los puntos donde la predicción es buena o mala.

- **Entropía (E_{sqi}):**

La entropía también se emplea en la evaluación de señales PPG, la fórmula empleada para su cálculo es:

$$E_{sqi} = -\sum_{i=1}^N x[i]^2 \cdot \log_e(x[i]^2) \quad (2.9)$$

Donde N es el número de muestras de la señal PPG y x la señal PPG sin procesar.

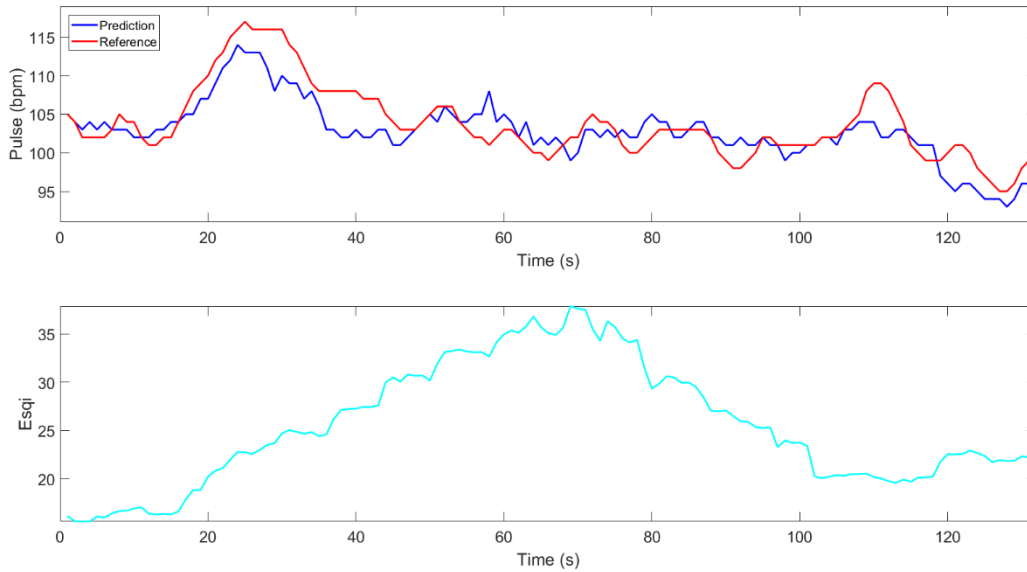


Ilustración 20: Video 25_5, E_{sqi} .

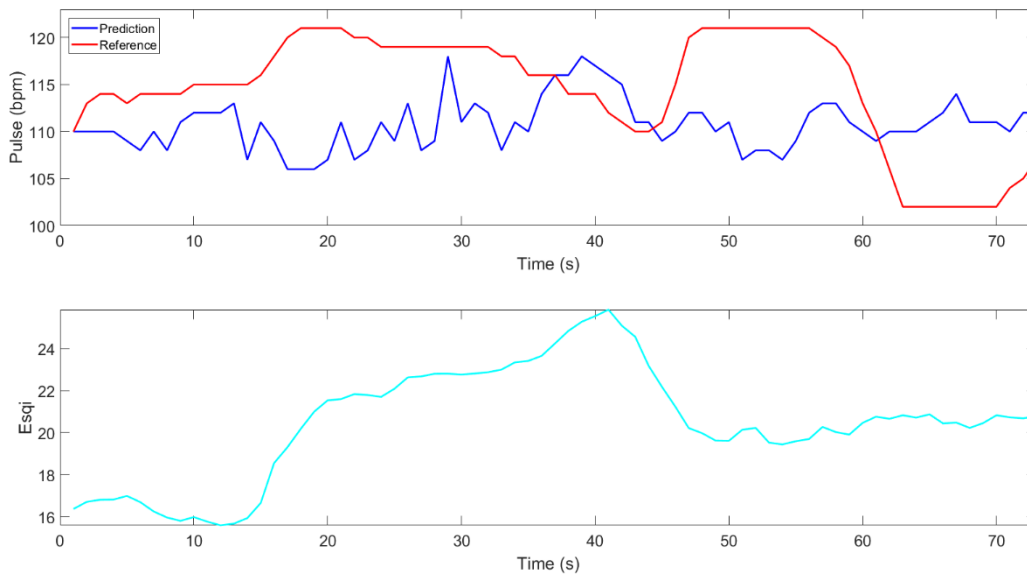


Ilustración 21: Video 25_3, E_{sqi} .

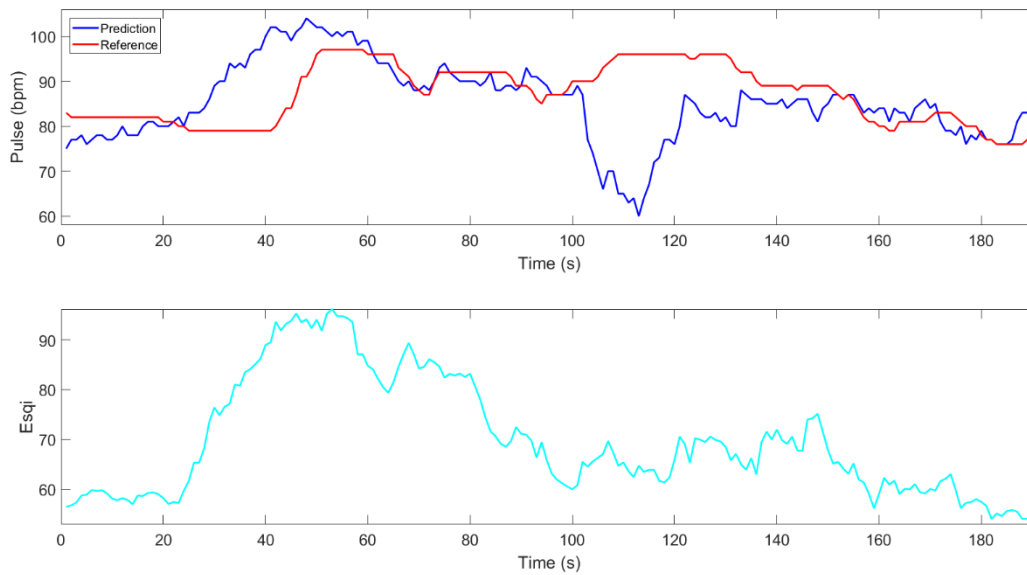


Ilustración 22: Vídeo 30_9, Esqi.

En este caso los valores obtenidos en los tres vídeos se encuentran entre 15 y 100, este parámetro si hace más comparables los distintos vídeos. Sin embargo, la gráfica de la entropía no resalta las zonas donde la predicción mejora o empeora y en el caso de los vídeos 25_5 y 25_3 la gráfica correspondiente a la entropía es similar cuando no debería.

- **Ratio de paso por cero (Z_{SQI}):**

Este índice mide la frecuencia con la que la señal PPG cambia de signo, para su cálculo se emplea la siguiente fórmula:

$$Z_{SQI} = \frac{1}{N} \cdot \sum_{n=1}^N I\{y < 0\} \quad (3.1)$$

Donde I es 1 si se cumple la condición y 0 en caso de no cumplirla y N es el número de muestras.

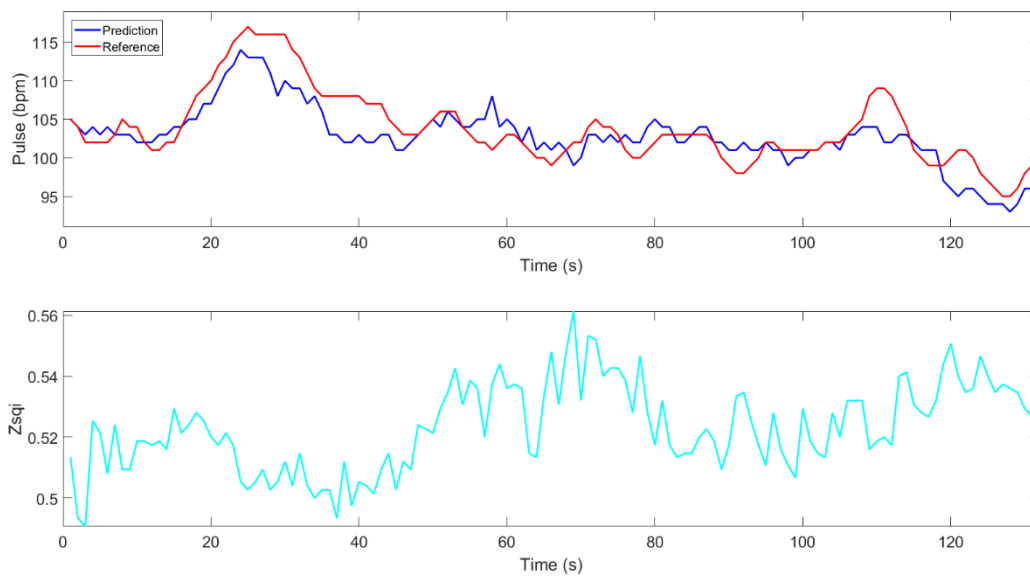


Ilustración 23: Vídeo 25_5, Z_{sqi} .

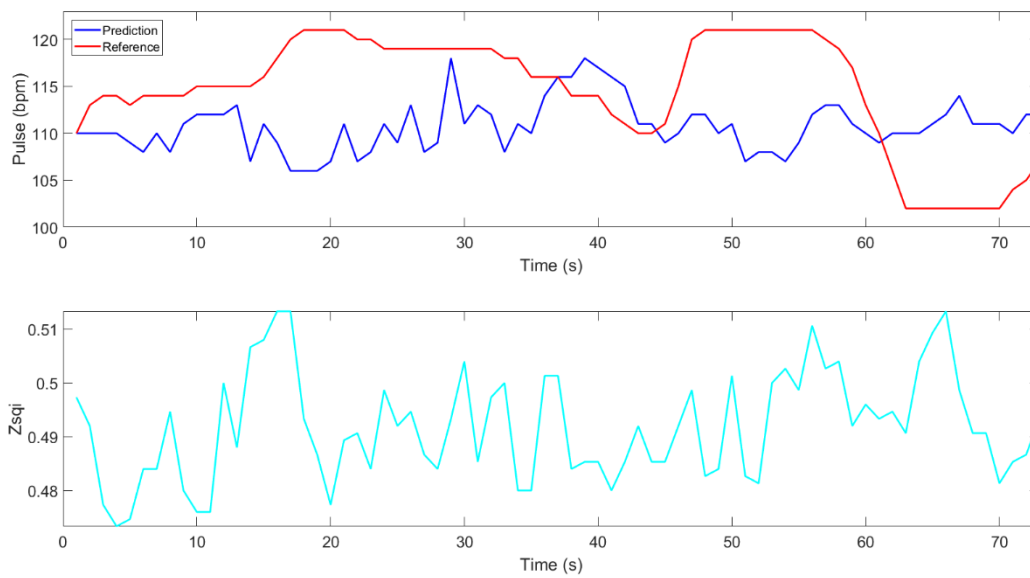


Ilustración 24: Vídeo 25_3, Z_{sqi} .

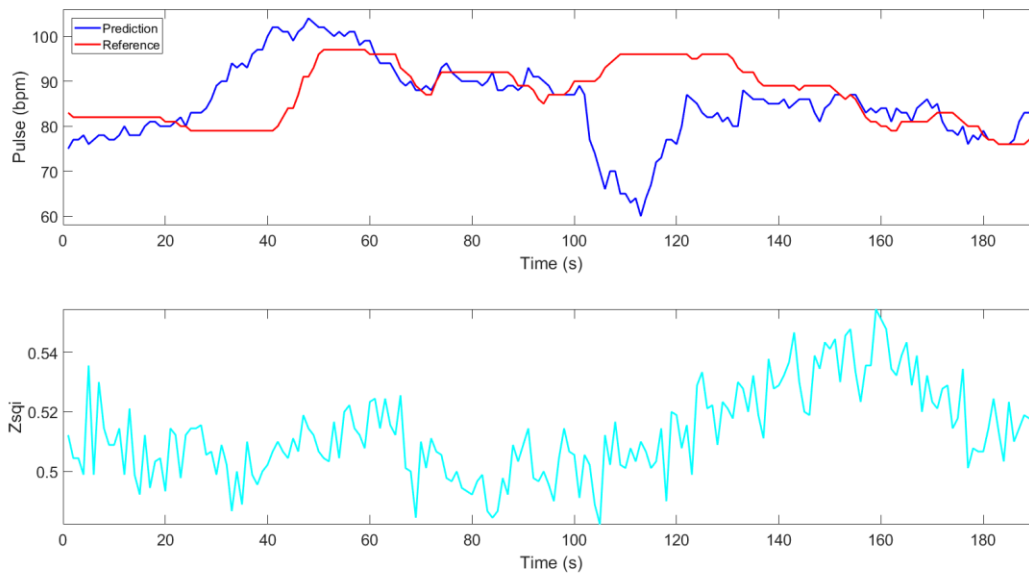


Ilustración 25: Vídeo 30_9, Z_{SQI} .

De nuevo a simple vista no se puede diferenciar las zonas donde la predicción es buena respecto a las zonas donde la predicción empeora. Sin embargo, si hay una clara diferencia entre la señal Z_{SQI} del vídeo 25_3, predicción mala durante todo el vídeo, y los otros dos vídeos donde la predicción es buena o parcialmente buena. En el vídeo 25_3 la señal Z_{SQI} muestra un valor inferior, se mueve en torno a 0.49, mientras que en los otros dos vídeos se mueven en torno a 0.52. Además, la frecuencia de oscilación en la señal Z_{SQI} del vídeo 25_3 es claramente inferior a la de los otros dos vídeos, los picos están más separados.

- **Relación señal a ruido (N_{SQI}):**

En este caso se emplea el nivel de SNR para evaluar la predicción, para ello se emplea la siguiente fórmula:

$$N_{SQI} = \frac{\sigma_{Señal}^2}{\sigma_{Ruido}^2} \quad (3.2)$$

Donde $\sigma_{Señal}$ es la desviación estándar del valor absoluto de la señal PPG filtrada y σ_{Ruido} es la desviación estándar de la señal PPG sin filtrar.

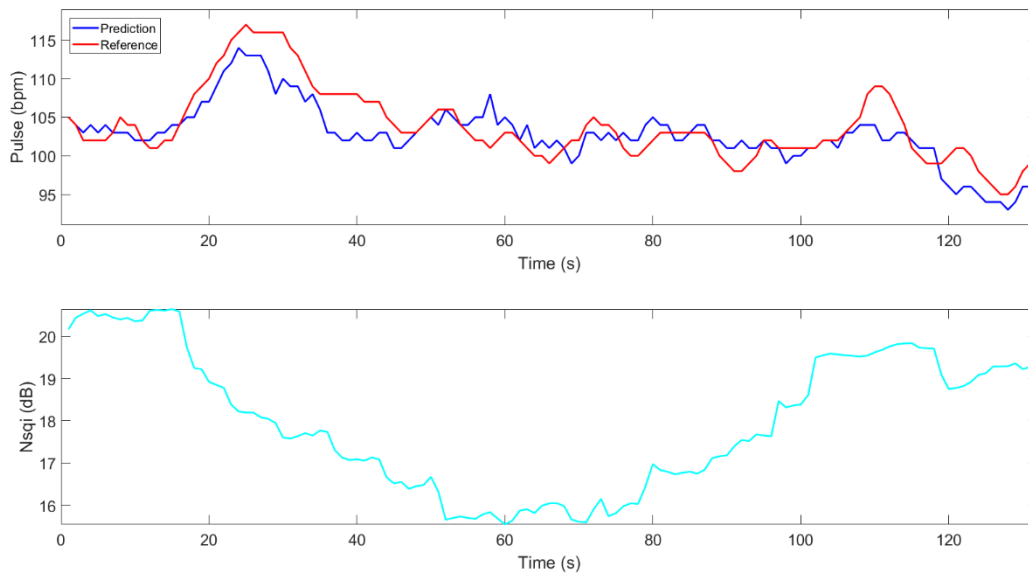


Ilustración 26; Vídeo 25_5, N_{sqi} .

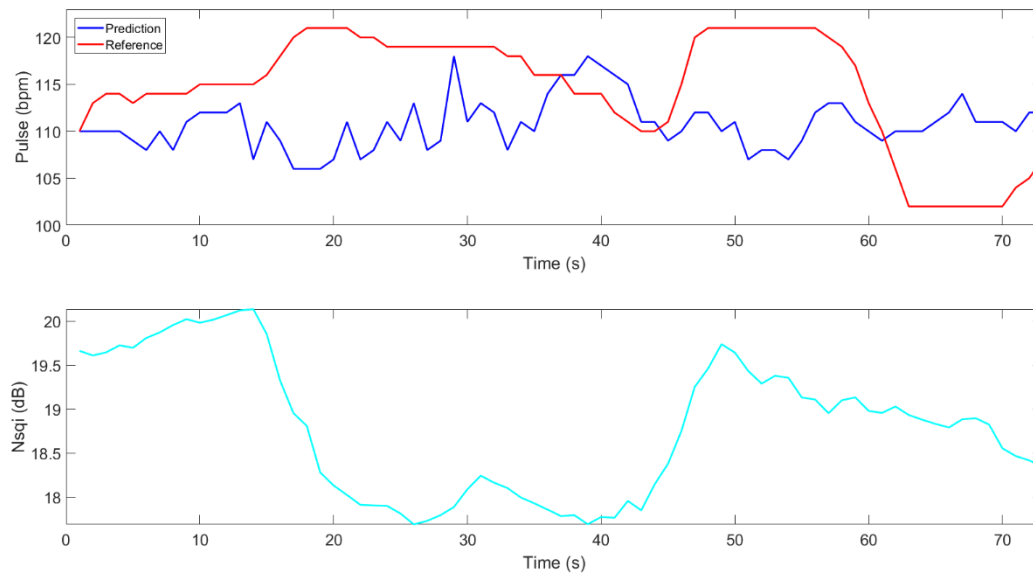


Ilustración 27: Vídeo 25_3, N_{sqi} .

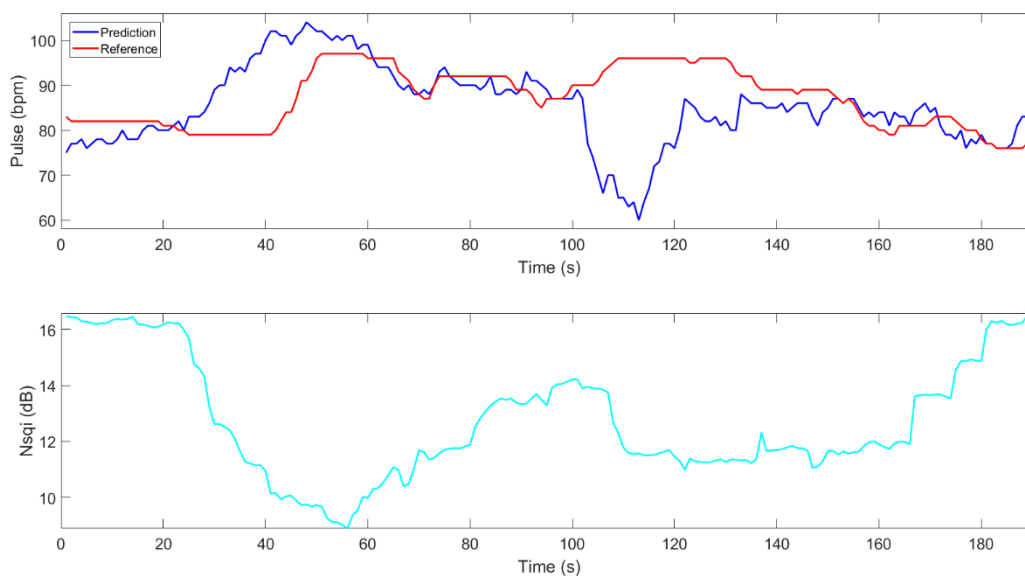


Ilustración 28: Vídeo 30_9, N_{sq} .

Si se compran los resultados obtenidos en los tres vídeos se ve que las tres señales SNR tiene una forma muy similar, esto se debe a que la persona que es grabada debe moverse en la mitad del vídeo y comienza y acaba el vídeo sin apenas mover la cara. Por este motivo el nivel de SNR es alto al principio y final mientras que a la mitad, cuando aparecen señales interferencia/ruido debido al movimiento, el nivel de SNR cae.

En cuanto a la capacidad de discernir cuando la predicción es buena o no, el nivel de SNR no da información útil. Los vídeos 25_3 y 25_5 tienen valores de SNR muy parecidos mientras que la predicción en el vídeo 25_5 es muy buena y la predicción en el vídeo 25_3 es bastante mala.

1.2.1. Conclusiones:

En este apartado se han analizado las respuestas obtenidas al emplear algunos de los índices de calidad más comúnmente empleados en sistemas PPG. Para cada índice se ha analizado la respuesta obtenida en tres vídeos: uno considerado predicción buena, uno considerado predicción mala y otro más considerado parcialmente buena. Comparando los resultados entre los tres vídeos se deberían detectar comportamientos diferentes del índice de calidad.

A primera vista parece que ninguno de los índices muestra una clara diferencia entre los videos de predicciones buenas y malas y tampoco es capaz de distinguir, en un mismo vídeo, las zonas donde la predicción mejora o empeora. Sin embargo, es importante señalar que la falta de eficacia al emplear estos índices se debe a que no encajan con el sistema de extracción de la señal PPG empleado en este trabajo. Los índices anteriormente nombrados se emplean en la actualidad y su funcionamiento está probado en diversos artículos. Es importante señalar que varios de estos índices trabajan con la señal PPG en bruto, sin procesar, en cambio la señal PPG “en bruto” empleada en este trabajo sí que ha pasado por algunas etapas de procesado (como la extracción del nivel de DC) y por este motivo es posible que las respuestas de los índices de calidad analizados disten de ser ideales.

Otro de los motivos de la falta de eficacia de los índices analizados es que están diseñados para sistemas completamente distintos. Algunos de los índices se centran en parámetros que llevan a error en el caso del algoritmo empleado en este trabajo. Un ejemplo claro es la Perfusión (P_{SQI}), este índice se centra en la fuerza de la señal pulsátil. Una señal fuerte es un indicador de predicción buena y una señal débil es un indicador de predicción mala, sin embargo, no considera problemas como puede ser ruido causado por el movimiento. Un pico en la señal pulsátil debido a una mala predicción causada por un movimiento brusco lleva a una evaluación errónea de la predicción.

También es cierto que las señales obtenidas de los índices de calidad no han sido profundamente analizadas, solo se han buscado diferencias en las zonas donde la predicción mejora o empeora. Un análisis más profundo de estas señales puede dar lugar a un resultado más adecuado.

Por estos motivos se han descartado estos índices de calidad y se propone el índice que se explicará en el siguiente apartado.

2. Índice de calidad propuesto

2.1. Introducción

El FZI (Forschungszentrum Informatik) es un centro de investigación localizado en Karlsruhe, Alemania, donde se están llevando a cabo varios proyectos de investigación relacionados con rPPG. El objetivo de estos proyectos no es otro que el de extraer de forma fiable ciertos signos vitales analizando y procesando imágenes capturadas por una cámara. Como ya se vio en el apartado PPG, dentro del estado del arte, los sistemas PPG, y más específicamente los sistemas rPPG, están cobrando especial relevancia en numerosos ámbitos de la actualidad como es la salud, la automoción, el deporte y la seguridad entre otros. En la búsqueda de sistemas cada vez menos intrusivos y fácilmente integrables la tecnología rPPG encaja perfectamente, es por esto que hoy en día existen varios proyectos de investigación en curso centrados en lograr un sistema lo más fiable posible. Sin embargo, para evaluar los resultados obtenidos por el sistema diseñado y detectar posibles problemas es necesario crear una herramienta capaz de analizar las predicciones obtenidas por el algoritmo rPPG.

El objetivo de este trabajo fin de Máster es el de crear una herramienta capaz de analizar una señal predicción y evaluar su similitud en relación a una señal de referencia en el entorno de los algoritmos rPPG. El índice de calidad diseñado tiene en cuenta varios parámetros de la señal y ofrece distintos índices de calidad en función de qué es lo que se quiere analizar. En algunas ocasiones se busca obtener la diferencia absoluta entre la predicción y la referencia, esto es cuando la exactitud del valor obtenido es el parámetro a medir. En otras ocasiones no interesa tanto la exactitud del valor obtenido en la predicción, sino que interesa saber hasta qué punto la predicción es capaz de seguir las variaciones.

Por último, señalar que el algoritmo rPPG empleado en este trabajo para la extracción de signos vitales, que se introducirá más adelante en el apartado Evaluación PulseCam, ha sido diseñado en el FZI. Además, los vídeos analizados también proceden de la base de datos del FZI.

2.2. Índice de calidad diseñado

2.2.1. Introducción

En vista del resultado obtenido con los índices de calidad actualmente existentes, en este apartado se propone una nueva herramienta para evaluar la predicción obtenida a partir de la señal PPG.

El índice de calidad diseñado trabaja con la señal del pulso ya extraído, no con la señal PPG como hacen los índices vistos anteriormente. Además, cuenta con la existencia de la señal de referencia en la evaluación de la predicción, esto no pasa con los índices analizados porque están diseñados para analizar la predicción sin necesidad de la referencia. La herramienta diseñada consta de dos partes fundamentales. La primera parte consiste en un acondicionamiento del pulso, esta señal tiene pequeñas variaciones constantes que en realidad no aportan información (no es relevante que el pulso sea 70 ahora, 71 un segundo después y 70 de nuevo a continuación). Tras el acondicionamiento se lleva a cabo el análisis de diversos parámetros de la señal predicción y de la señal referencia.

El resultado final son tres parámetros principales y una serie de parámetros secundarios. De los tres parámetros principales uno describe la precisión del valor obtenido en la predicción frente al de la referencia, otro la similitud entre las formas de la señal predicción y de la referencia, y el último es una combinación que da una idea general de la predicción. Los parámetros secundarios recogen información como el nivel de SNR o RMSE entre otros.

A continuación, se explicará en profundidad tanto la herramienta diseñada como los resultados obtenibles.

2.2.2. Acondicionamiento de la señal

Uno de los problemas observados en la señal de referencia y en la señal predicción es la variabilidad que presentan a lo largo del tiempo. Pequeñas variaciones continuas del pulso no contienen información y complican el análisis de las señales, no es relevante si el pulso fluctúa entre 70 y 74 bpm durante varios segundos.

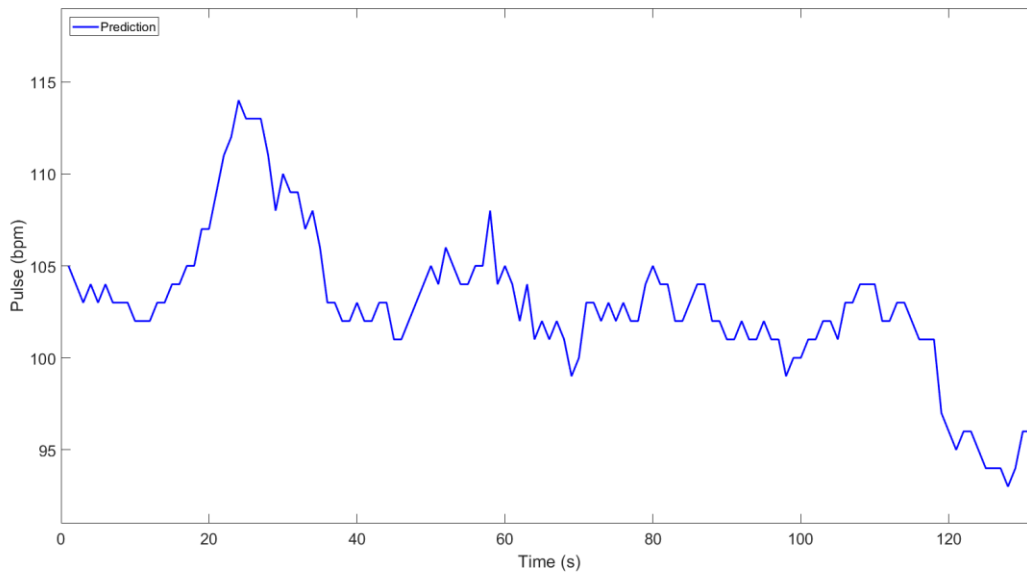


Ilustración 29: Señal sin acondicionar.

En la ilustración 29 se puede observar que hay muchas zonas donde el pulso fluctúa, hay variaciones relevantes que marcan un cambio importante en el pulso y variaciones que no aportan nada en el análisis de la señal. Este último caso se mostrará más claramente a continuación.

Un claro ejemplo de variaciones irrelevantes se ve en la ilustración 30 donde la señal varía entre 103 y 102 bpm desde el segundo 36 hasta el 46. Variaciones tan pequeñas en el pulso son normales, el cuerpo humano es una máquina muy compleja donde la respiración, un pequeño movimiento o la temperatura producen ligeros cambios en el pulso que no tienen más relevancia. De casos como este está repleta la señal de la ilustración 29. Este tipo de variaciones son las que se busca eliminar con el acondicionamiento de la señal.

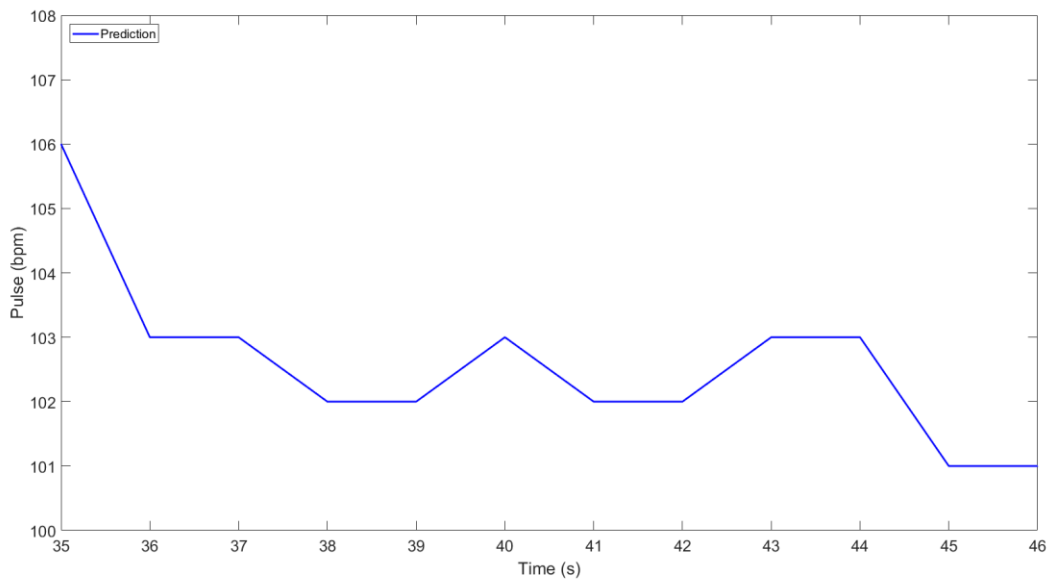


Ilustración 30: Señal sin acondicionar entre el segundo 35 y 46.

Para la eliminación de estas fluctuaciones se ha creado la siguiente herramienta. Su funcionamiento básico se puede ver en la siguiente ilustración.

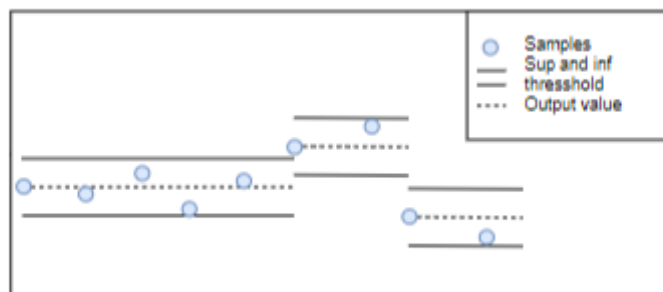


Ilustración 31: Acondicionamiento de la señal.

Dada una primera muestra a_n , esta pasa a ser el valor de salida y a su vez se crean un límite superior e inferior alrededor de este valor de salida. En la siguiente muestra, a_{n+1} , se comprueba si se encuentra dentro de los límites marcados por la muestra anterior a_n . Si a_{n+1} se encuentra dentro de los límites entonces el valor de salida sigue siendo a_n y los límites permanecen inalterados, en cambio, si a_{n+1} se encuentra fuera de los límites a_{n+1} pasa a ser el nuevo valor de salida y los límites se recalculan alrededor de a_{n+1} . De esta manera jugando con la separación de los límites situados alrededor del valor de salida se pueden eliminar las pequeñas fluctuaciones en la señal.

Si ahora acondicionamos la señal de la ilustración 29 se obtiene:

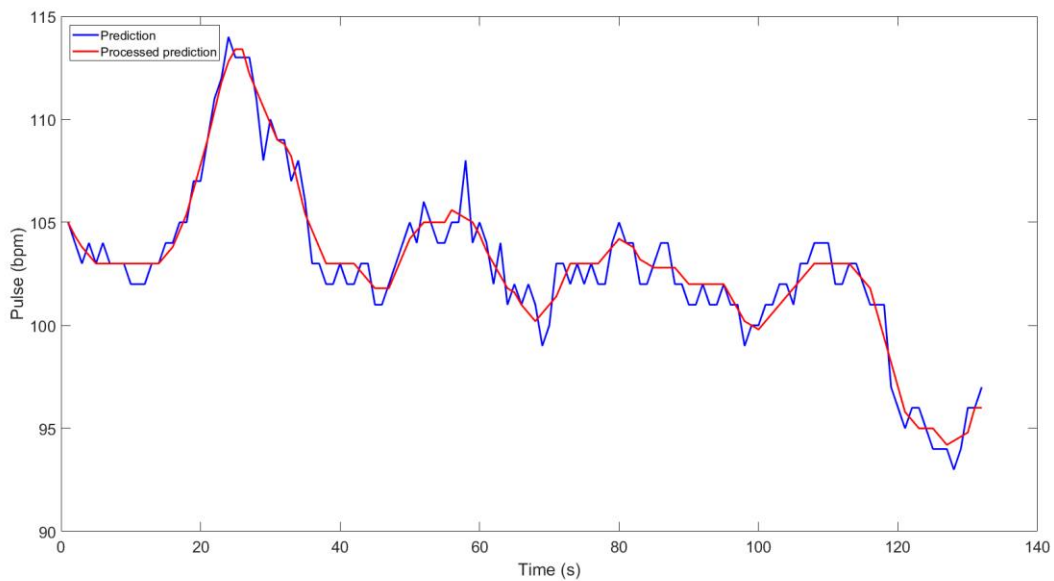


Ilustración 32: Comparación de la señal sin acondicionar y la señal acondicionada.

La señal resultante es notablemente más suave que la original y se logran hacer desaparecer las pequeñas variaciones que no contienen información.

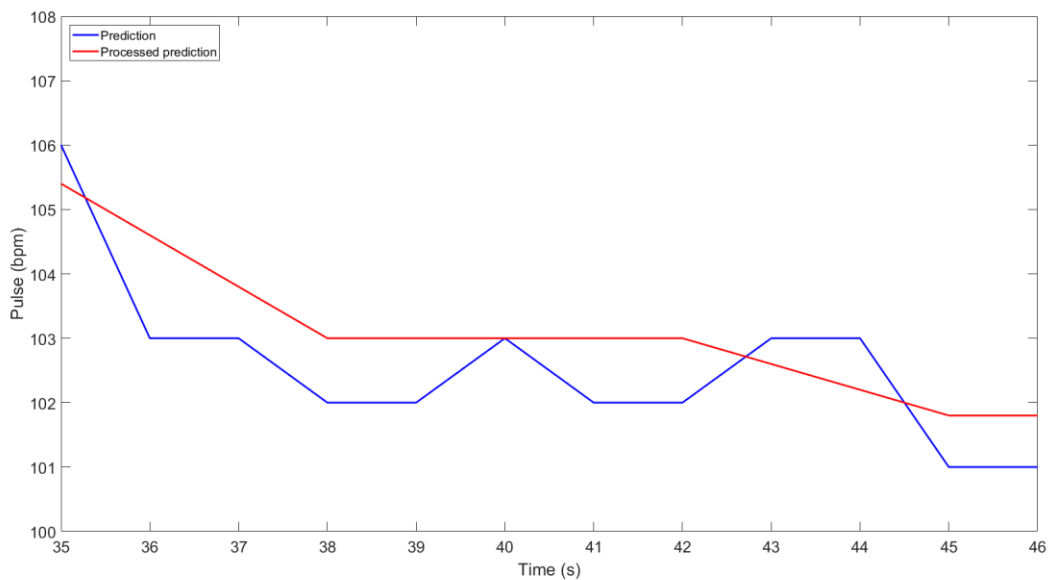


Ilustración 33: Comparación de la señal sin acondicionar y la señal acondicionada entre el segundo 35 y 46.

A partir de este punto todas las señales empleadas (salvo que se diga lo contrario), tanto la predicción como la referencia, serán el resultado de aplicar este código de suavizado.

Una vez introducido el código de suavizado se procederá a la explicación del índice de calidad diseñado.

2.2.3. Índice de calidad

El primer paso antes de analizar la señal pulsátil es el de acondicionar la señal tal y como se explicó en el apartado anterior, una vez procesada la señal se llevará a cabo la extracción de los parámetros empleados en el índice de calidad diseñado.

- **Porcentaje de seguimiento:**

Este parámetro evalúa el porcentaje de la señal donde la predicción es lo suficientemente cercana a la predicción como para considerarse buena. En este caso solo se tiene en cuenta el valor obtenido en la predicción y el valor obtenido en la referencia.

Para el cálculo de este parámetro se siguen los siguientes pasos:

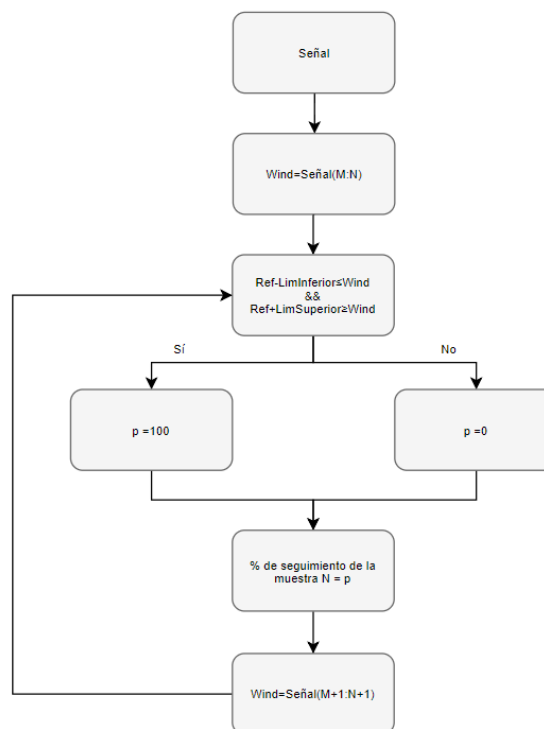


Ilustración 34: Esquema del porcentaje de seguimiento.

El primer paso consiste en analizar la señal empleando ventanas temporales. El principal motivo de analizar la señal empleando ventanas es el de evitar que zonas donde la predicción es buena en una muestra y mala en la siguiente, de forma salteada, dé un buen resultado. Haciendo cumplir una condición a toda la ventana se evita ese problema. El algoritmo a analizar busca obtener predicciones buenas de forma consistente y de esta manera se logra analizar esa consistencia también.

El tamaño de ventana se ha escogido considerando varios factores como son la duración de la señal, el número de muestras en la ventana y los objetivos del algoritmo. En relación a este último factor, se busca que el algoritmo a analizar sea capaz de obtener valores correctos durante un tiempo lo suficientemente largo para ser válido pero sin llegar a una situación que complique en exceso el sistema. Empleando ventanas de 15 segundos se obtienen 15 muestras por cada ventana, además, la evaluación empleando ventanas de 15 segundos da una buena idea sobre la precisión y consistencia de la predicción.

El análisis empleando ventanas se hace de la siguiente manera:

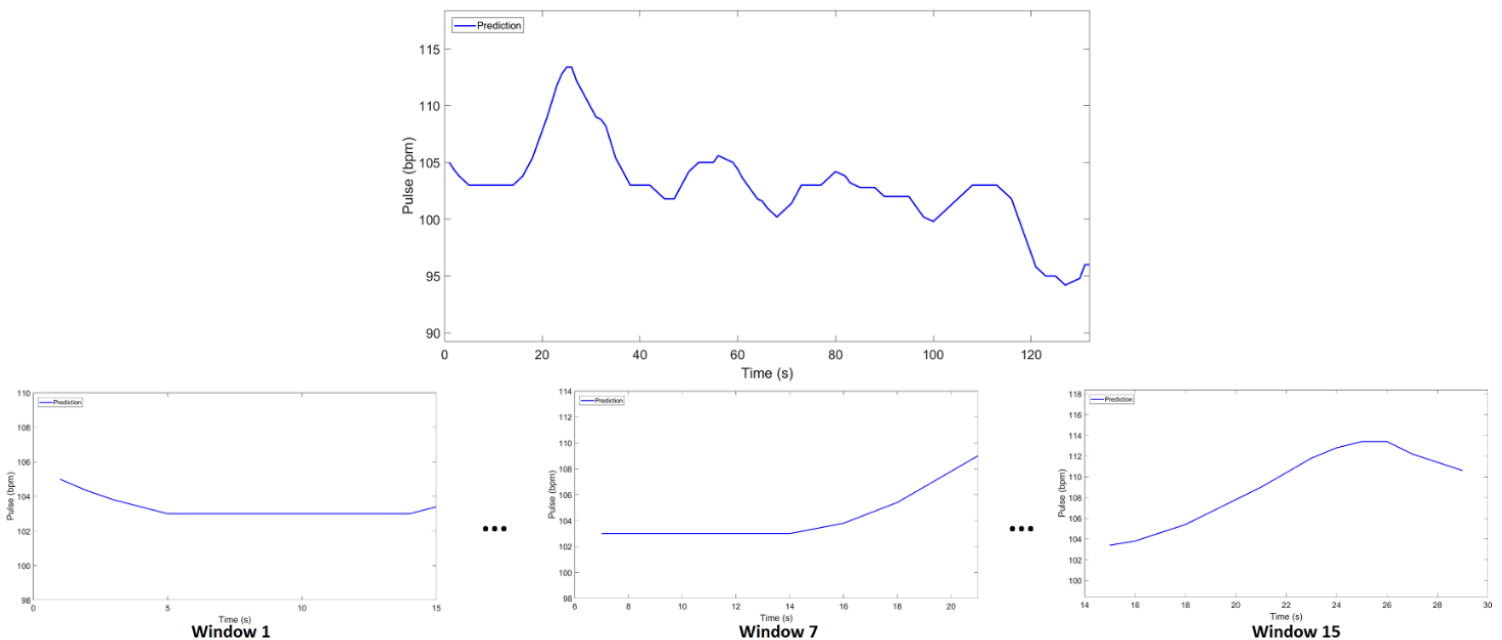


Ilustración 35: Análisis mediante ventanas temporales.

Se parte de una ventana inicial que contiene los primeros 15 segundos, tras ser analizada la siguiente ventana consiste en la ventana anterior tras eliminar la primera muestra y añadir una nueva al final. De esta manera se barre la señal segundo a segundo con ventanas deslizantes de 15 segundos.

Para cada ventana se analiza si todas las muestras se encuentran dentro de unos límites marcados por la referencia. Si esta condición no se cumple la muestra se marca con 0, si se cumple se marca con 100. Para las muestras analizadas en la primera ventana el porcentaje de seguimiento se calcula en un solo paso de forma directa para todas las muestras, a partir de esta ventana se analiza la señal segundo a segundo y el porcentaje de seguimiento calculado es el correspondiente a la muestra más reciente, es decir, la última muestra de la ventana.

En cuanto a los límites marcados por la referencia, que definen la condición a cumplir por la predicción, se emplea un límite superior y otro inferior alrededor de la referencia. De esta manera si se cumple la condición de la ecuación 3.3 el valor de la predicción se considera bueno y el porcentaje de seguimiento de esa muestra es 100.

$$\text{Predicción} \leq \text{Referencia} + \text{límite superior} \quad (3.3)$$

$$\text{Predicción} \geq \text{Referencia} - \text{límite inferior}$$

En cambio, si alguna de las muestras de la ventana analizada no cumple la condición de la ecuación 3.3 se considera que el porcentaje de seguimiento de esa muestra es 0.

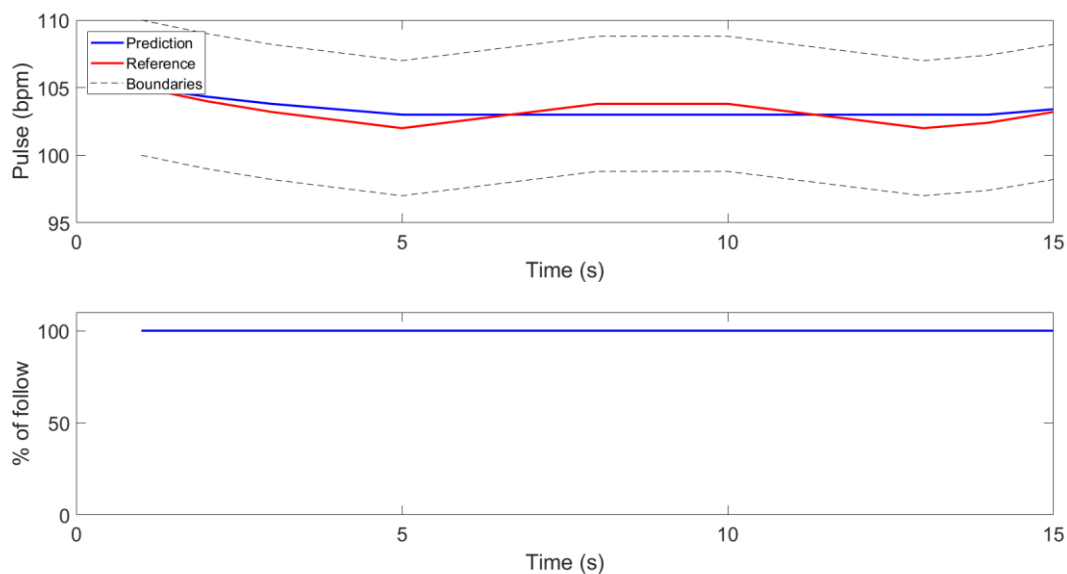


Ilustración 36: La predicción sigue a la referencia correctamente en la ventana analizada.

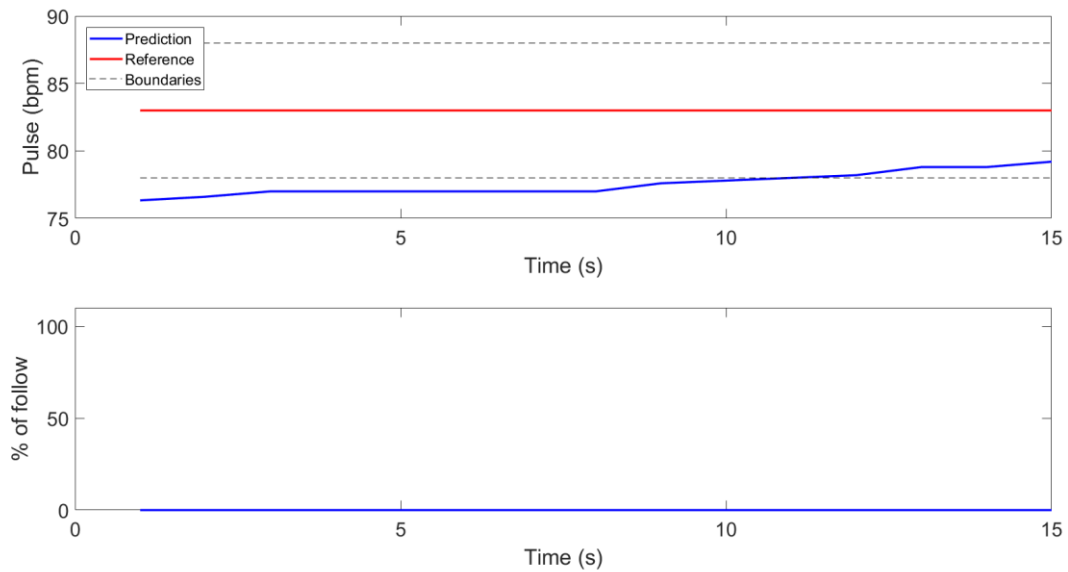


Ilustración 37: La predicción no cumple las condiciones marcadas por los límites en toda la ventana.

En la ilustración 36 se ve que la predicción se encuentra dentro de los límites calculados alrededor de la referencia y el porcentaje de seguimiento es del 100%. En el caso de la ilustración 37 la predicción no se mantiene dentro de los límites en toda la ventana y por ello el porcentaje de seguimiento de la última muestra es 0%. En este caso es importante señalar que el porcentaje de seguimiento es 0% en todas las muestras de la ventana y esto significa que tampoco las ventanas correspondientes a las muestras anteriores lograron cumplir con los límites.

Un caso intermedio es el que se ve en la siguiente ilustración:

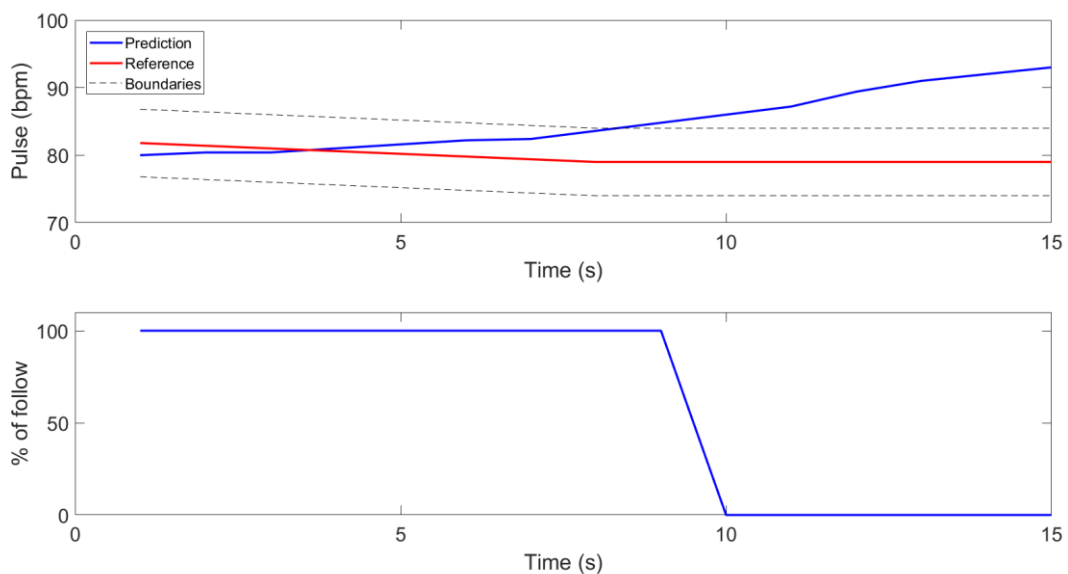


Ilustración 38: Caso intermedio.

En este caso las ventanas correspondientes a las primeras muestras si cumplen con los requisitos y por eso el porcentaje de seguimiento es 100% hasta que se llega a la muestra donde la ventana deja de cumplir, a partir de ese punto el porcentaje de seguimiento cae al 0%.

De esta manera el resultado final obtenido tras analizar la señal completa se puede ver en la siguiente figura, donde el color verde marca las zonas donde la predicción cumple con los requisitos y la zona azul donde no los cumple. El valor indicado en la parte superior de la gráfica señala el porcentaje de seguimiento de la señal, es decir, durante cuánto tiempo la predicción se considera buena dentro de un determinado rango y durante un tiempo mínimo.

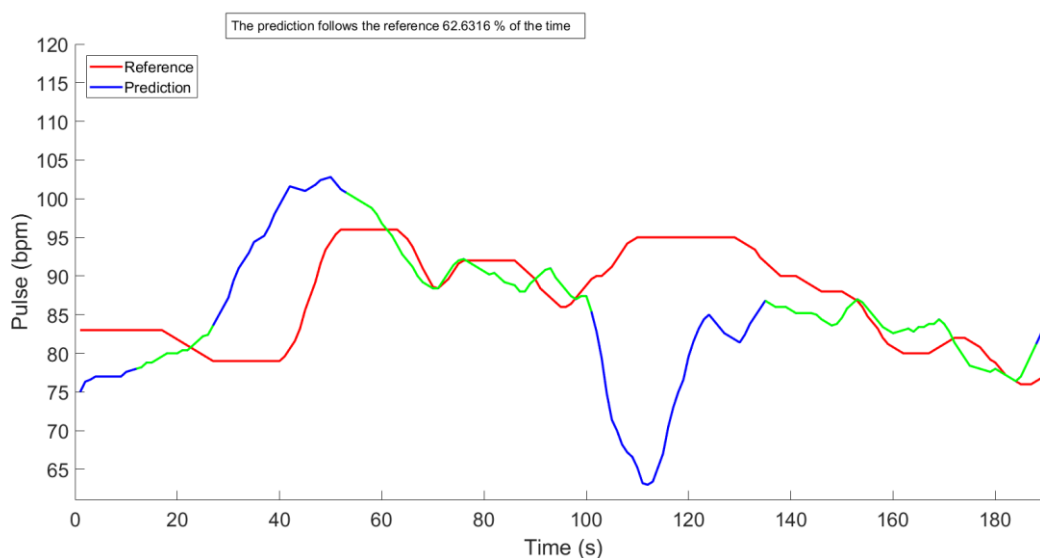


Ilustración 39: Porcentaje de seguimiento de la señal analizada.

A partir de este punto las partes de señal verdes, salvo que se indique lo contrario, muestran las zonas de la señal donde la predicción se encuentra dentro del área alrededor de la referencia más un cierto valor.

- **Delay:**

En el código empleado para evaluar el funcionamiento del algoritmo PPG hay un parámetro que es importante nombrar, aunque no se emplee directamente en el cálculo del índice de calidad. Este parámetro es el delay. El delay analiza si la predicción se encuentra temporalmente adelantada o retrasada en relación a la señal de referencia.

El primer paso consiste en analizar la señal empleando ventanas, el proceso es el mismo que con el cálculo del porcentaje de seguimiento. La ventana con la que se

trabajar se obtiene tanto de la señal predicción como de la señal referencia, ambas son necesarias.

Una vez obtenidas ambas ventanas hay dos procesos principales involucrados en el cálculo del delay, por un lado está el barrido y por otro el cálculo de la similitud.

Cálculo de similitud:

La obtención del delay se hace en el mismo código empleado para calcular el porcentaje de seguimiento. En la siguiente imagen se muestra un esquema general de cómo se calcula el delay:

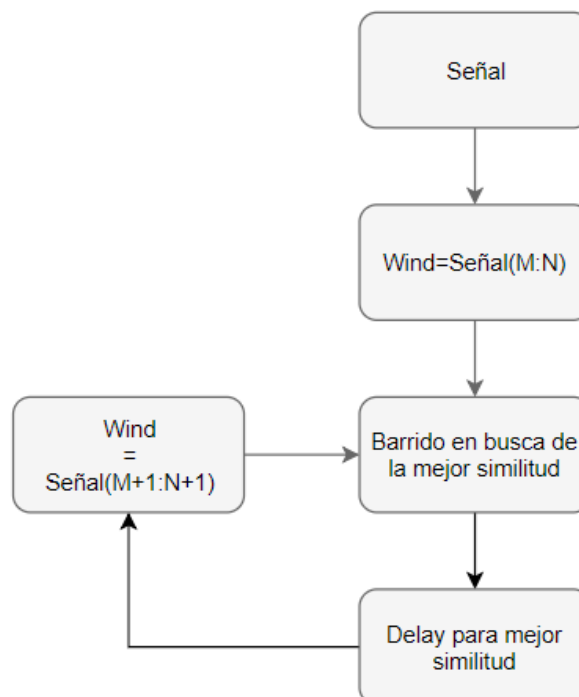


Ilustración 40: Esquema general del cálculo del delay.

El objetivo de la similitud es el de valorar cuánto se parecen la referencia y la predicción. Para ello se trabaja con la derivada de las señales predicción y referencia, para cada punto de la ventana correspondiente a la predicción y para cada punto de la ventana correspondiente a la referencia se calcula su derivada. De esta manera se obtiene la tendencia de la señal. A continuación se comparan las derivadas de la ventana correspondiente a la referencia y de la ventana correspondiente a la predicción punto por punto y se comprueba si se cumplen las siguientes condiciones:

$$\text{sign}(d\text{referencia}) = \text{sign}(d\text{predicción}) \quad (3.4)$$

$$d\text{referencia} = 0$$

$$d\text{predicción} = 0$$

Siendo $d\text{referencia}$ y $d\text{predicción}$ la derivada de la ventana referencia y de la ventana predicción respectivamente.

En los puntos que cumplen estas condiciones se transforma la pendiente en ángulo y se comparan los ángulos de la referencia con los de la predicción. En caso de que no se cumplan estas condiciones la similitud en ese punto será 0, esto pasa cuando la predicción y la referencia tienen pendientes opuestas. Para el cálculo del ángulo se emplea la arcotangente de la derivada y para comparar ambos ángulos se emplea la siguiente fórmula:

$$\text{Similitud}(\%) = \left(1 - \frac{|\text{GradRef} - \text{GradPred}|}{90}\right) \cdot 100 \quad (3.5)$$

Siendo GradRef el ángulo del punto a comparar de la referencia y GradPred el punto a comparar de la predicción.

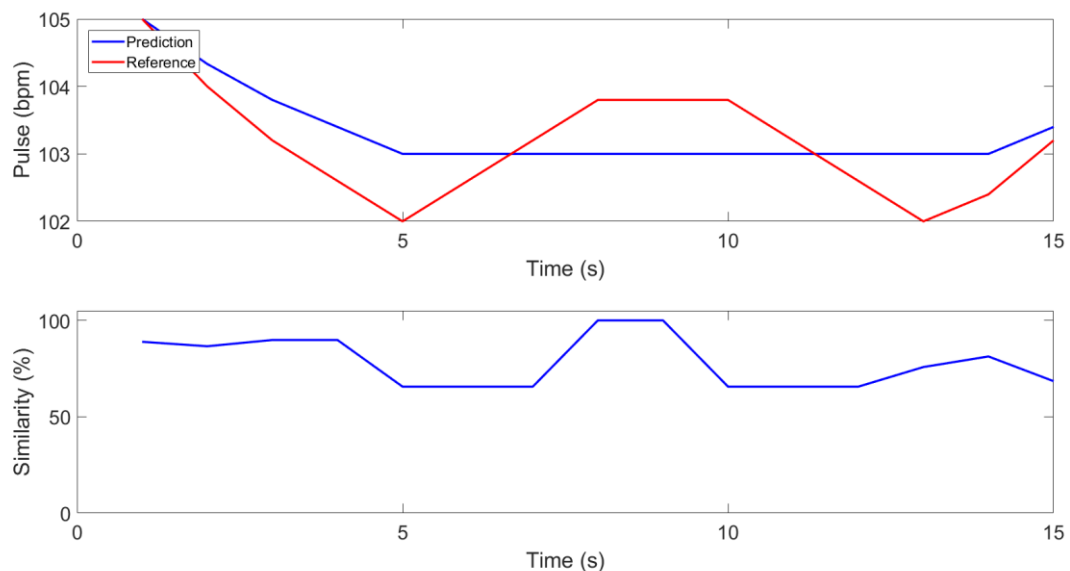


Ilustración 41: Similitud.

Empleando la fórmula anterior un ángulo de 0° y otro de 45° tendrían un 50 % de similitud y uno de 0° con otro de 90° tendrían 0 % de similitud. Con la media de la similitud, tras comparar todos los puntos de ambas ventanas, se evalúa como de parecida es la ventana predicción respecto a la ventana referencia.

Es importante señalar que el cálculo de la similitud solo se lleva a cabo en los puntos contenidos en una señal de al menos 15 s donde todos los puntos se encuentran dentro de los límites marcados alrededor de la referencia, tal y como se ve en el porcentaje de seguimiento en la ecuación 3.3. Los puntos que no cumplan esta condición no se tienen en cuenta en el cálculo de la similitud y tampoco en el del delay.

El barrido:

El objetivo del delay es el de detectar si la predicción se encuentra adelantada o retrasada temporalmente respecto a la señal de referencia. Para ello se calcula la similitud para una ventana referencia empleando varias ventanas predicción. Cada ventana predicción es una ventana adelantada o retrasada varios segundos respecto a la referencia. De esta manera se obtienen varios resultados de similitud dada una ventana de referencia, de todos estos valores el de mayor similitud será la ventana predicción más parecida y el delay de esta ventana se almacena.

Cada barrido comprueba la similitud de una ventana referencia dada con varias ventanas predicción. Dependiendo del delay máximo que se quiera comprobar se necesitarán más o menos barridos. El delay de la señal completa se obtiene tras los barridos de similitud, los delays óptimos de las ventanas predicción para cada ventana referencia se representan de la siguiente manera mostrando la tendencia de la predicción.

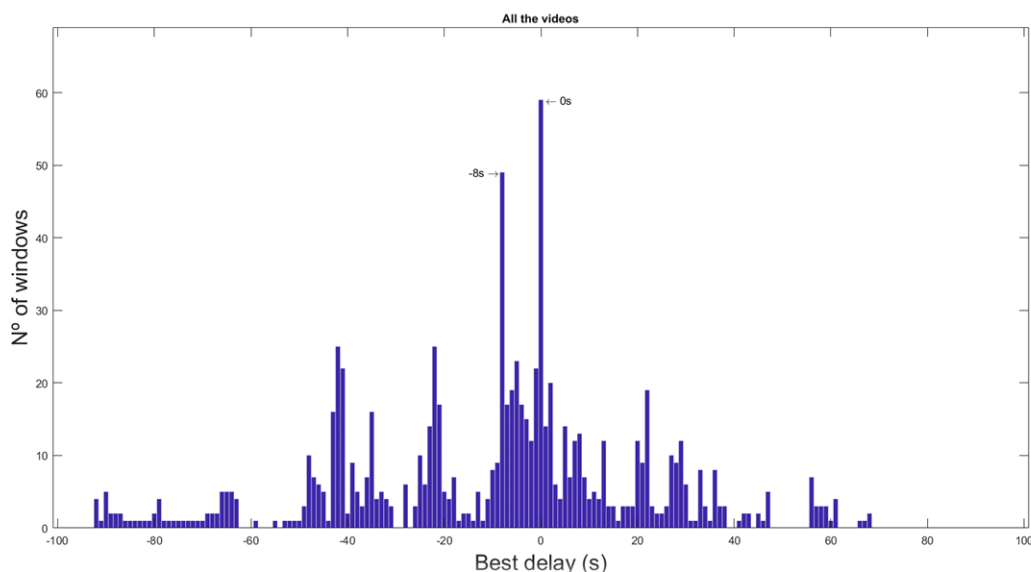


Ilustración 42: Delays de la señal completa.

Es importante señalar que el código da preferencia a los delays más cercanos a 0 segundos en caso de empate en cuanto a la similitud obtenida.

Otro parámetro disponible en el cálculo del delay es la tolerancia. La tolerancia no compara valores de similitud exactos sino que los compara considerando un cierto rango alrededor del valor exacto. La tolerancia se especifica en porcentaje, si el porcentaje de tolerancia es del 2 % el código considera iguales todos los valores de similitud situados en un rango del 2 % alrededor de la similitud a comparar. Si a esto le sumamos el hecho de que se da preferencia a los delays más cercanos a 0 segundos el resultado obtenido tiende a ser el siguiente:

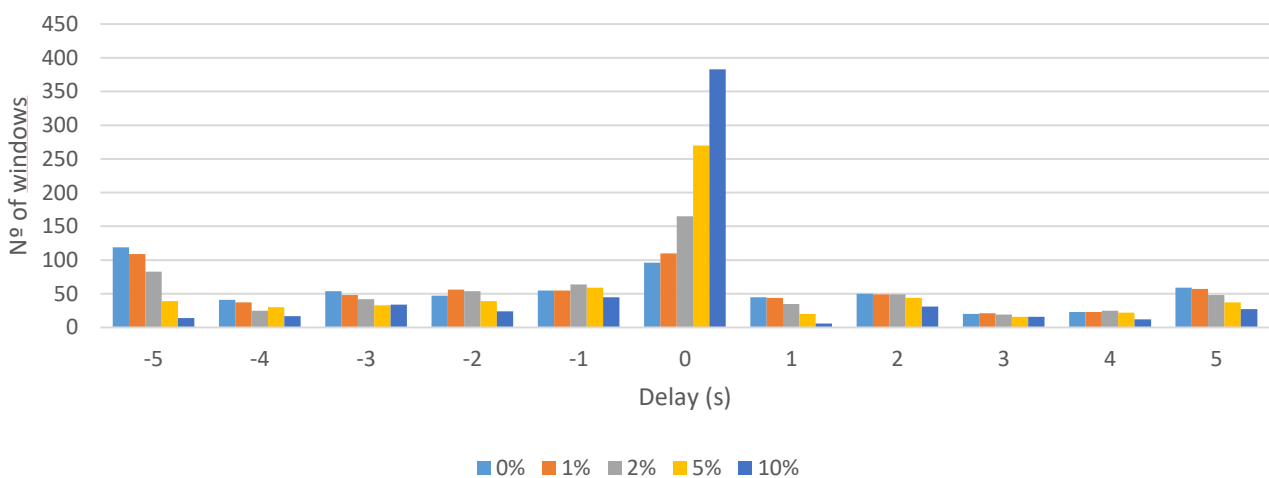


Ilustración 43: Tendencia del delay empleando varios porcentajes de tolerancia.

A medida que se aumenta la tolerancia, y teniendo en cuenta la prioridad de los delays cercanos a 0 segundos, la señal suele centrarse más alrededor de 0. De cara a utilizar esta herramienta hay que decidir qué porcentaje de tolerancia es aceptable para considerar valores de similitud iguales, si bien 10 puede ser excesivo emplear un 1% - 2% es recomendable y puede dar un valor más preciso del comportamiento real del algoritmo PPG a analizar.

- **Porcentaje de similitud:**

Este parámetro analiza la forma de la señal referencia y de la señal predicción e indica la semejanza entre ambas. El porcentaje de seguimiento se centra en analizar los valores de las señales, este parámetro es útil cuando se busca analizar la precisión lograda por el algoritmo a evaluar. En cambio, cuando lo que interesa es detectar cambios la precisión lograda en torno a la referencia no es tan relevante. Lo que se busca lograr con el porcentaje de similitud es analizar cómo responde la predicción frente a

los cambios en el pulso y evaluar de alguna manera la forma de las señales por encima de los valores obtenidos. Para lograr esto se siguen los siguientes pasos:

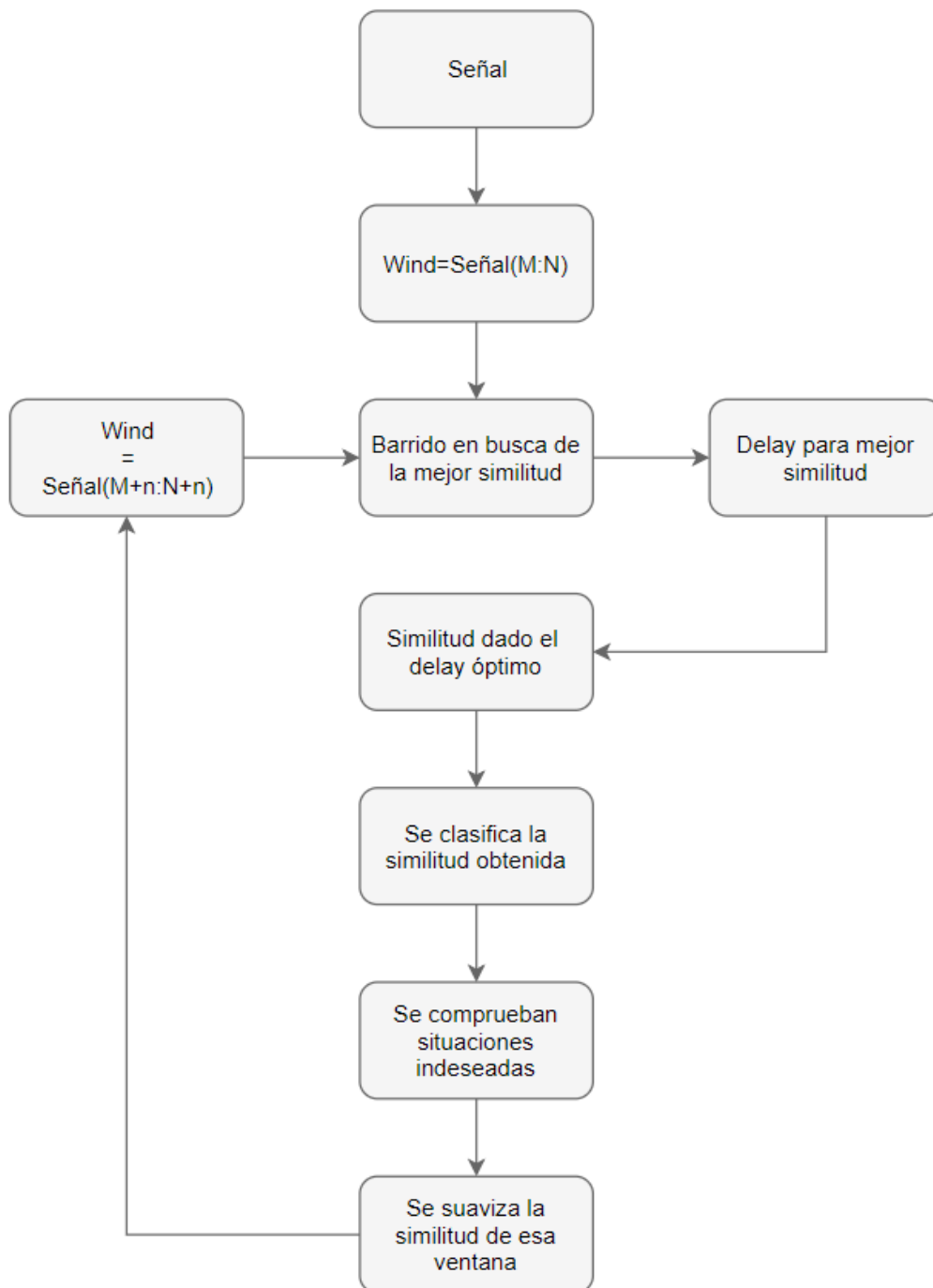


Ilustración 44: Esquema general del cálculo de la similitud.

Este parámetro es algo más complejo que el porcentaje de seguimiento, a continuación se explicarán los pasos llevados a cabo para su cálculo de forma más detallada incluyendo varios procesos no nombrados en el esquema general.

El primer paso, al igual que en el cálculo del porcentaje de seguimiento, es el análisis de la señal empleando ventanas temporales. En este caso se busca emplear ventanas lo suficientemente grandes como para tener una buena idea de la forma de la señal a analizar en esa ventana. Sin embargo también se quiere conseguir un número alto de ventanas, dado que cada ventana obtiene un valor de similitud y la similitud final se calculará empleando la similitud de cada ventana temporal analizada. Por lo tanto es necesario un compromiso entre ambos aspectos.

En este caso el análisis mediante ventanas no se hace segundo a segundo, como se busca tener una idea de la forma general de la señal la siguiente ventana a analizar avanza 10 segundos. De esta manera se superponen las ventanas para que no se pierda información de la forma de la señal pero no se avanza segundo a segundo acelerando el proceso.

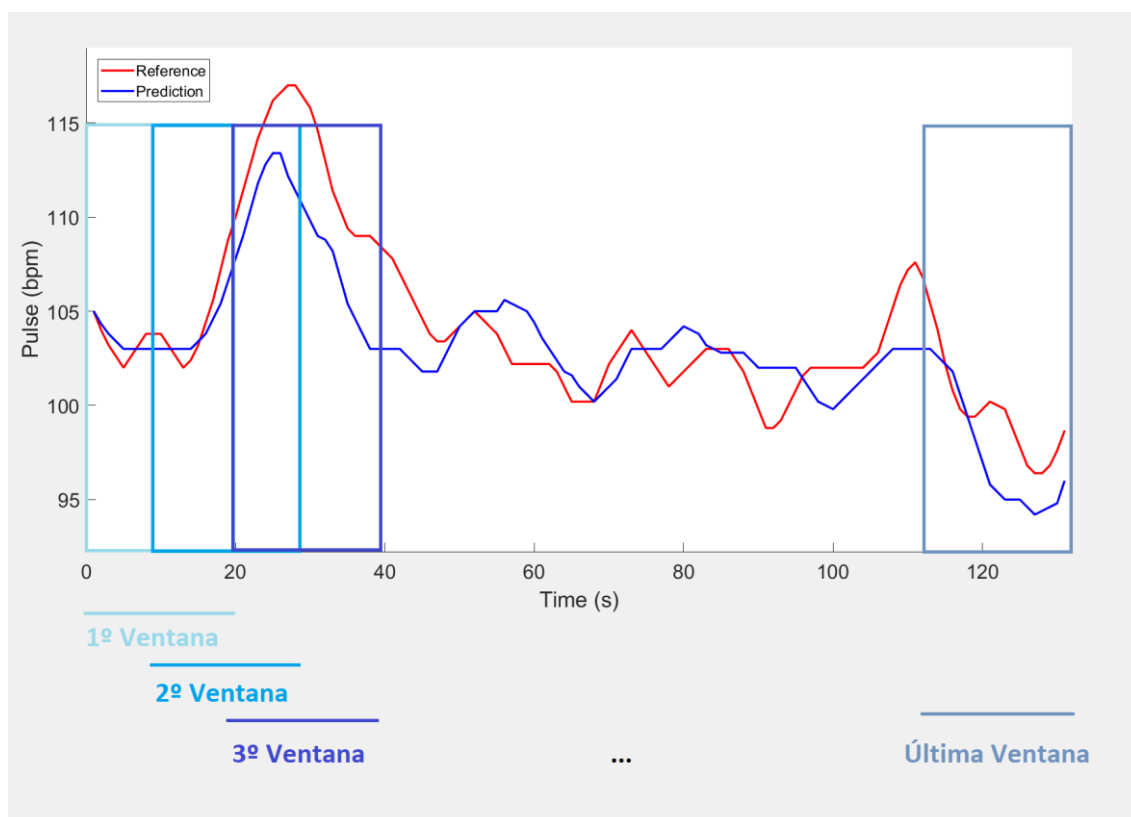


Ilustración 45: Análisis mediante ventanas.

Para el cálculo de la similitud se tiene en cuenta que la ventana óptima de la predicción pueda estar ligeramente adelantada o retrasada temporalmente respecto a la referencia, por este motivo se lleva a cabo un barrido para buscar el delay con el que

se consiga la ventana predicción óptima. El proceso es el mismo que el empleado en el delay, se barre empleando distintos delays y con una comparación de la mejor similitud lograda se obtiene el delay óptimo. En este caso el barrido de delays es menos amplio, se acota hasta un valor aceptable para la aplicación en la que se empleará el algoritmo PPG.

De nuevo existe la opción de definir una cierta tolerancia en la mejor similitud de la ventana de cara a obtener el delay óptimo.

Una vez obtenido el delay óptimo se recalcula la similitud punto por punto de la ventana tras comparar la ventana referencia con la venta predicción óptima. En este punto se tiene una señal similitud que abarca toda la ventana analizada y cuyos valores varían entre 0 y 100.

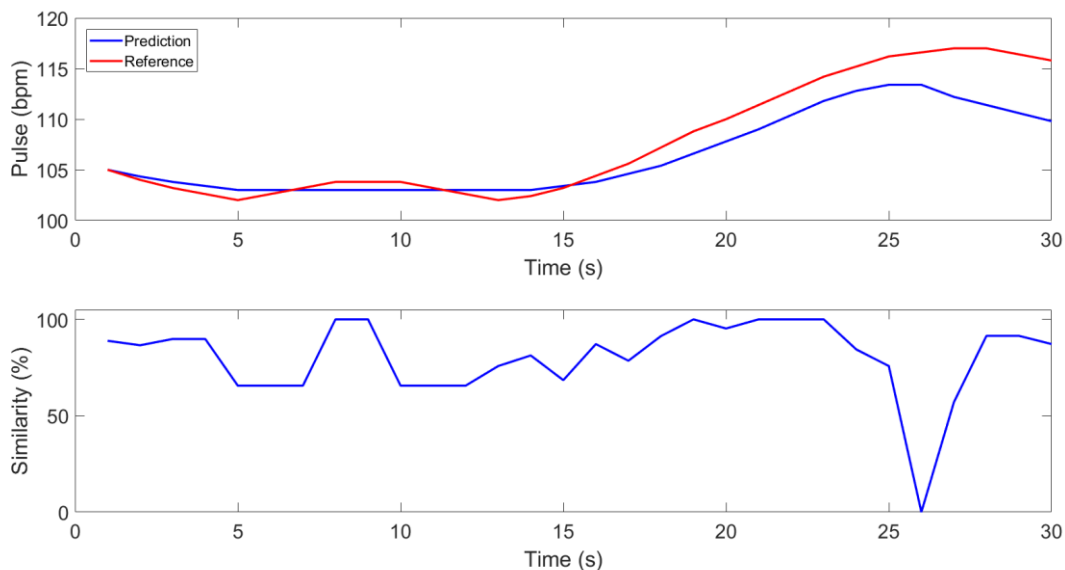


Ilustración 46: Similitud empleando la ventana predicción óptima.

El siguiente paso consiste en clasificar ese rango de similitud que varía entre 0 y 100 en cuatro clases:

- 100 % de similitud: valores mayores o iguales que 90.
- 80 % de similitud: valores entre 90 y 70, incluyendo 70.
- 60 % de similitud: valores entre 70 y 50.
- 0 % de similitud: valores menores o iguales a 50.

De esta manera la gráfica de similitud anterior pasa a ser:

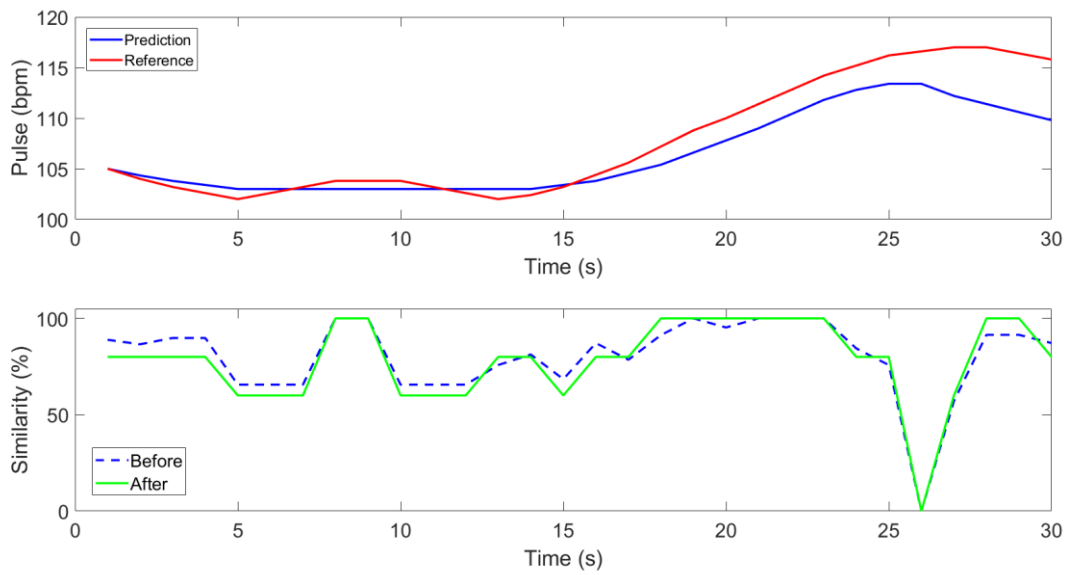


Ilustración 47: Similitud tras clasificar los valores obtenidos.

En azul se representa la similitud sin clasificar y en verde tras el clasificado, este paso supone una gran ayuda de cara a procesar la señal similitud a continuación.

El siguiente paso trata de eliminar un problema indeseado, la situación se puede observar en la siguiente gráfica:

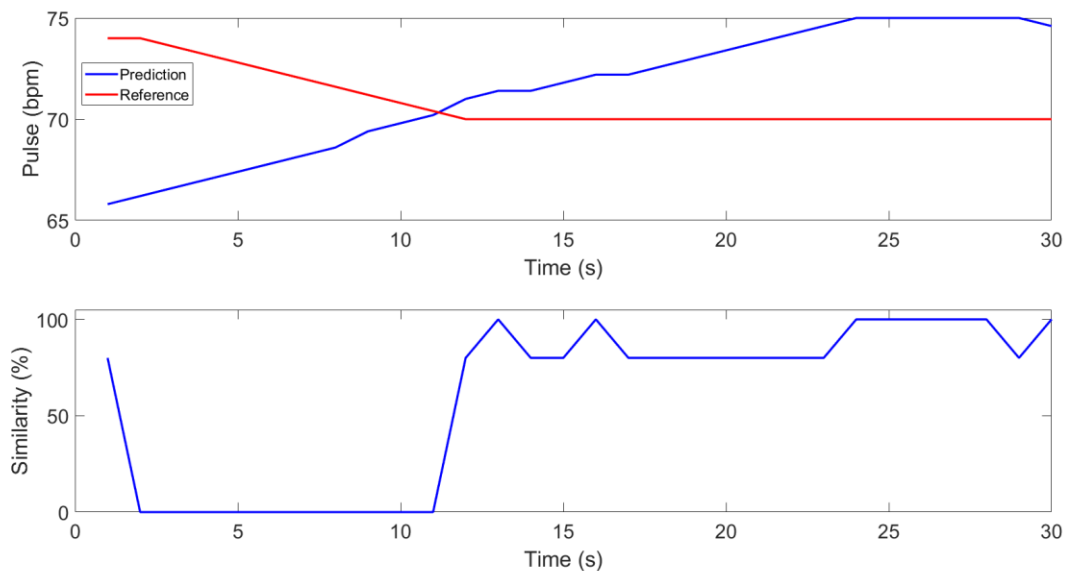


Ilustración 48: Crecimiento continuo frente a señal constante.

El problema es el incremento o decremento continuo de una de las señales frente a otra señal que permanece constante. La situación reflejada en la ilustración 48 muestra que la similitud obtenida es muy alta, esto sucede porque la pendiente con la que la señal crece o decrece es pequeña. Esta situación es indeseada porque a la larga, aunque la pendiente sea pequeña, la diferencia entre ambas señales es importante y una respuesta de este estilo durante un prolongado periodo de tiempo es una predicción errónea.

Para detectar estas situaciones se localizan las zonas donde una de las señales es constante y se buscan incrementos o decrementos continuos que duren al menos 10 segundos. Si se encuentra una situación como la que se ve en la imagen anterior la similitud de ese segmento baja a 0 %. Es importante señalar que también se tiene en cuenta la variación de la pendiente, si desde el inicio del incremento/decremento hasta pasados 10 segundos la variación en el pulso es inferior a 3 bpm la similitud se mantiene como estaba. También es importante señalar que no es necesario que la señal sea creciente o decreciente todo el tiempo para reducir la similitud, si se encuentran hasta como máximo 3 puntos por cada 10 segundos puntos donde la señal es constante también se reduce la similitud, de esta manera se evita que algún punto suelto suponga un problema.

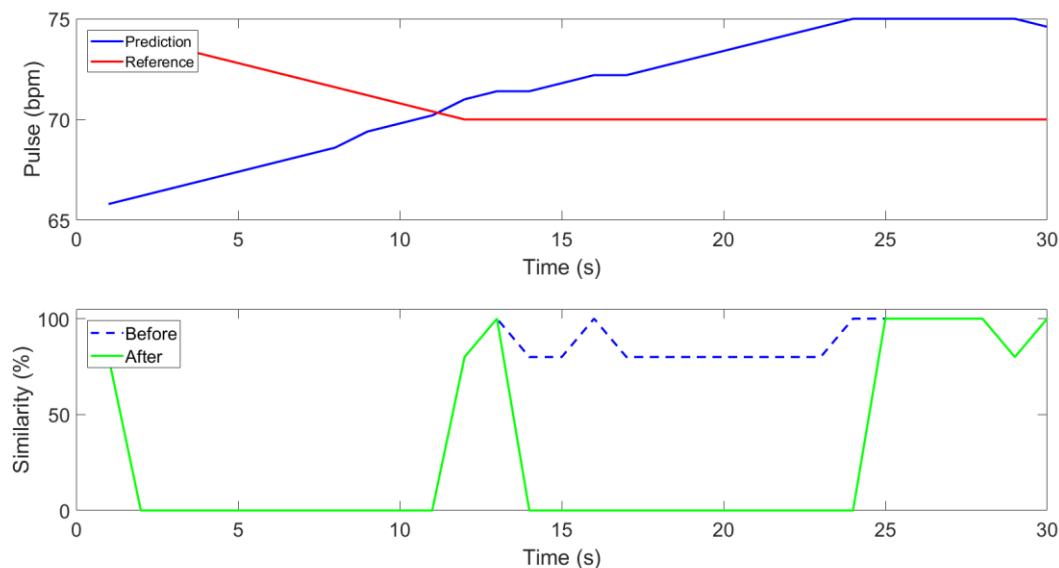


Ilustración 49: Similitud tras eliminar el crecimiento constante.

En la imagen superior se observa el resultado tras reducir la similitud en la zona donde el crecimiento de la predicción es continuo mientras que la referencia permanece constante.

En la imagen superior se ve una zona entre 10 s y 15 s donde la similitud no se ha reducido a 0 % a pesar de que a primera vista entra dentro de la zona de crecimiento continuo. Esto se debe a que desde el inicio del crecimiento hasta pasados 10 s, es decir, desde el segundo 12 hasta el segundo 21 la variación en pulso es inferior a 3 bpm, 2,8 bpm exactamente.

Otro aspecto que es importante señalar en la ilustración 49 es el final de la pendiente, donde la predicción se vuelve constante. El algoritmo reduce la similitud si detecta como máximo 3 puntos donde la señal es constante, este punto es el tercero si se cuentan los dos intermedios situados en 13 y 16 segundos y por eso reduce la similitud en ese punto a 0 %. Si bien esto podría considerarse un error, una sola muestra no debería afectar en exceso a la similitud total.

Una vez solucionados los problemas causados vistos anteriormente se procede con el siguiente paso, los puntos aislados. Los puntos aislados son aquellos que se encuentran rodeados por otros con una similitud completamente diferente, es decir, un punto con similitud 0% rodeado por puntos con similitud mayor de 50 % o un punto con alta similitud en medio de puntos con similitud nula.

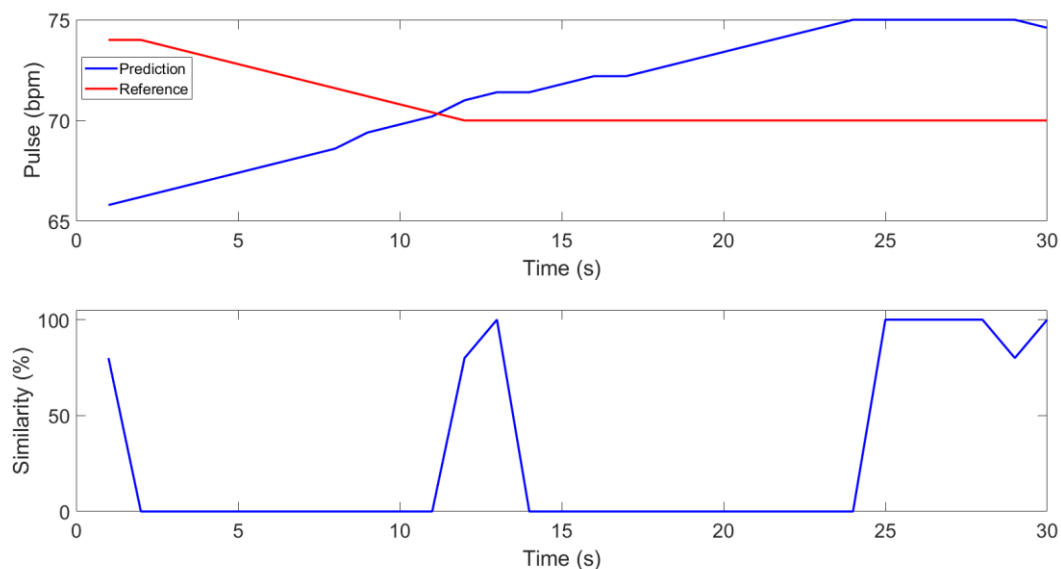


Ilustración 50: Puntos aislados de similitud alta rodeados por puntos de similitud nula.

En la imagen superior se observa una zona, entre los segundos 10 y 15, donde dos muestras obtienen una similitud alta en una zona donde todas las muestras que las rodean presentan una similitud nula. Esos dos puntos no aportan información útil e incluso aumentan la similitud total de la ventana dando una idea errónea de la similitud final de la ventana.

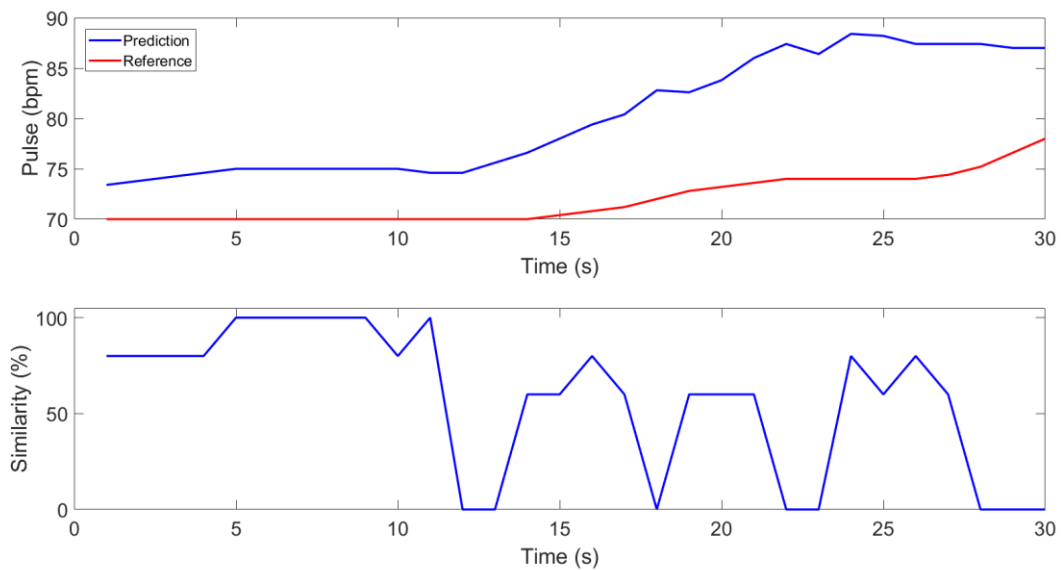


Ilustración 51: Puntos aislados de similitud nula rodeados por puntos de similitud alta.

En este otro caso se presenta la situación opuesta a la anterior. En la ilustración 51 se observan varios puntos cuya similitud es 0 %, debido al cambio del signo de la pendiente, rodeados de puntos donde la similitud es mayor que 50 %.

<i>Ventana de la ilustración</i>	<i>Similitud</i>
50	28,00 %
51	57,33 %

El objetivo de este punto es el de eliminar la influencia de puntos aislados en la similitud total, como se ve en las imágenes anteriores la similitud puede cambiar notablemente debido a la influencia de puntos aislados cuyo comportamiento difiere en gran medida con respecto a los que le rodean.

Para solucionar esta situación en primer lugar se localizan los puntos cuya similitud es mayor de 0 % y a continuación se mide la distancia entre esos puntos. Midiendo las distancias entre los puntos con similitud alta se pueden detectar dos casos:

- Puntos con similitud alta rodeados por puntos con similitud nula, es el caso de la ilustración 50. Si se da este caso esos puntos aislados pasan a tener 0 % de similitud. Para que esto pase los puntos aislados deben ser individuales o como máximo tres puntos de similitud alta seguidos, en el caso de ser cuatro puntos o más la similitud se mantiene.

- Puntos con similitud nula rodeados por puntos con similitud alta, es el caso de la ilustración 51. Si se da este caso los puntos aislados pasan a tener una similitud del 50 %, no sería justo que puntos con similitud nula pasen a tener una similitud más alta que otros puntos cuya forma se asemeja mucho más a la referencia. Esta situación se da para puntos aislados individuales, para dos o más la similitud se mantiene.

De esta manera las ilustraciones anteriores pasan a ser:

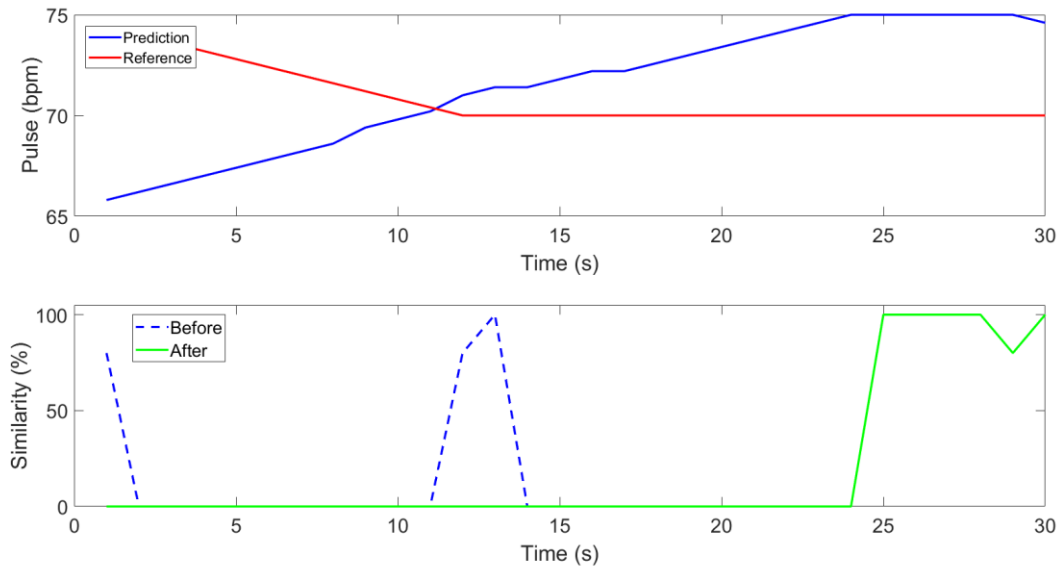


Ilustración 52: Similitud tras eliminar puntos aislados.

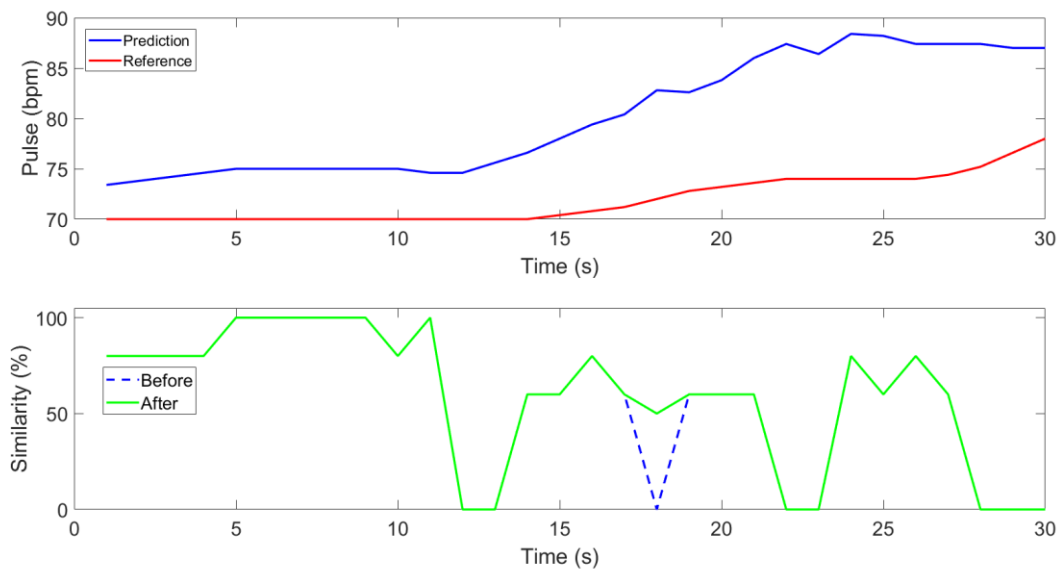


Ilustración 53: Similitud tras eliminar puntos aislados.

Ventana de la ilustración	Similitud antes	Similitud después
50	28,00 %	19,33 %
51	57,33 %	59,00 %

La similitud de la ilustración 50 pasa del 28 % al 19 %. Se reduce notablemente la similitud dado que la predicción tiene poca semejanza con la referencia en la mayoría de los puntos y los puntos aislados incrementaban la similitud en gran medida. En la ilustración 51 la similitud no se altera apenas, solo es un punto el que se “corrige”. El objetivo de este paso es el de reducir la influencia de los puntos aislados en la similitud de la ventana y se ha logrado.

Un caso que contiene ambas situaciones es el siguiente:

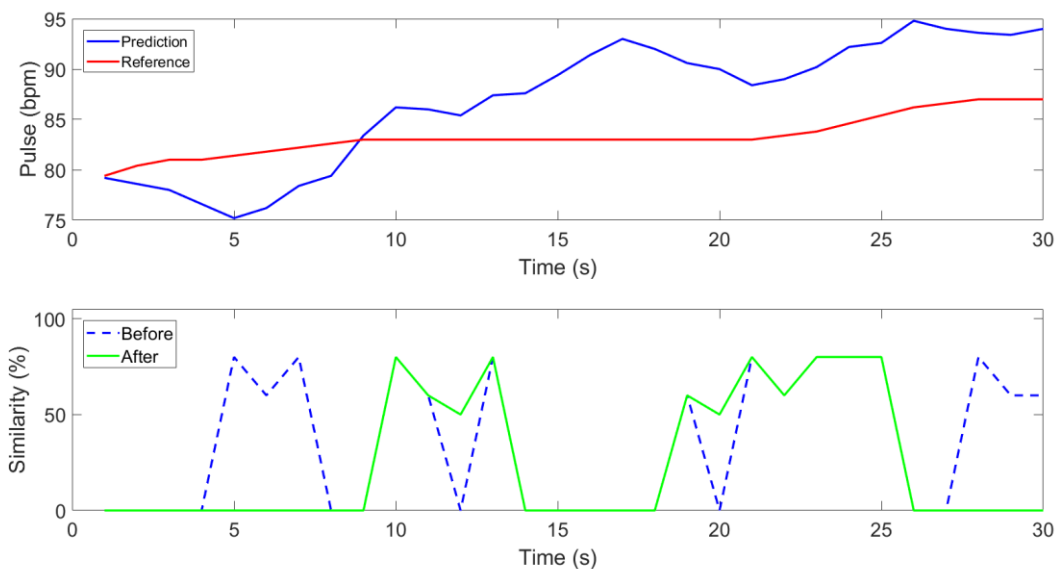


Ilustración 54: Ventana temporal con puntos aislados.

Entorno al segundo 5 un trozo de señal con similitud alta pasa a tener similitud nula y en los segundos 12 y 20 dos puntos con similitud baja pasan a tener 50 % de similitud. La similitud de la ventana pasa a ser del 36 % al 25 %.

Este es el paso final en el cálculo de la similitud, llegados a este punto y tras analizar la señal completa se obtienen las similitudes medias de cada una de las ventanas temporales analizadas. La similitud de la señal total es la media de la similitud de todas las ventanas.

- **Raíz del error cuadrático medio:**

La raíz del error cuadrático medio (RMSE) es un parámetro que indica el error cometido en toda la señal al comparar la predicción y la referencia. En este caso se trabaja con la señal pulso al completo una vez extraída de la señal PPG, sin ningún tipo de procesado como es el suavizado visto anteriormente. La fórmula empleada para su cálculo es:

$$RMSE = \sqrt{\overline{Error^2}} \quad (3.6)$$

Donde Error es la diferencia, punto por punto, de la predicción y la referencia y $\overline{Error^2}$ es la media de Error al cuadrado.

RMSE es un parámetro fiable que analiza la diferencia entre la predicción y la referencia, más adelante se empleará para afinar el índice de calidad combinado.

- **Análisis de ruido y de SNR**

En este punto se lleva a cabo el análisis del ruido de la relación señal/ruido de la predicción, para ello se seguirá el siguiente esquema:

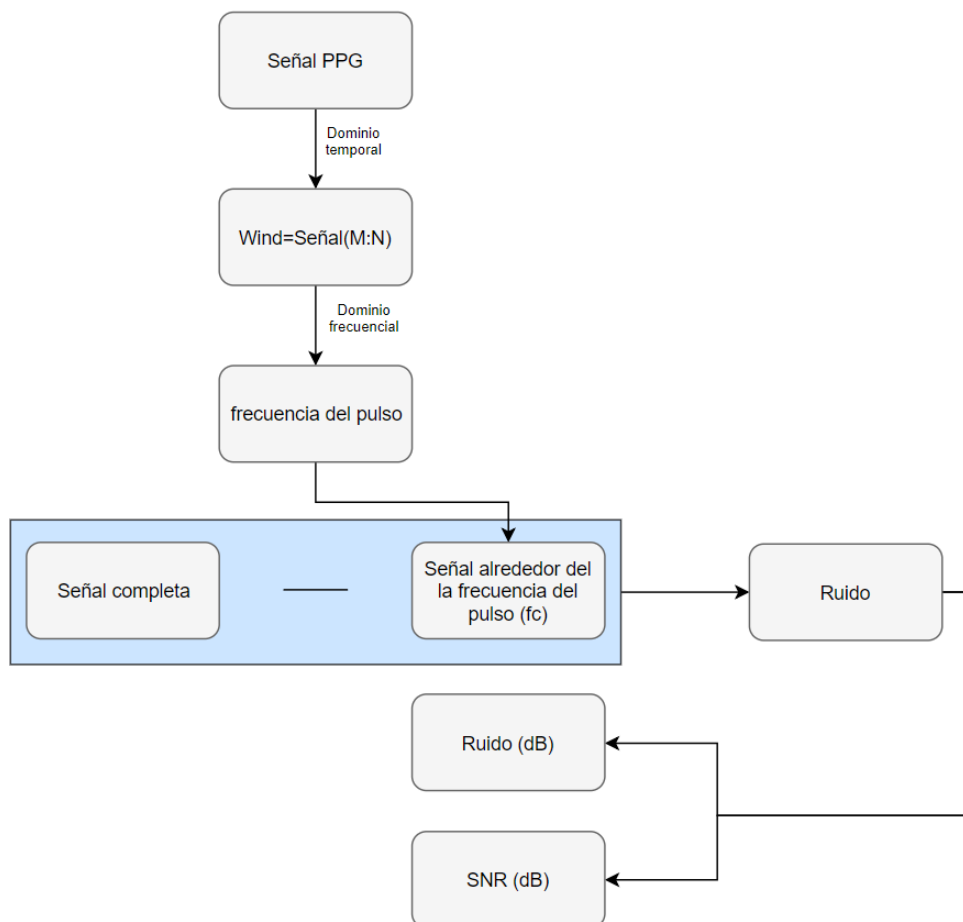


Ilustración 55: Esquema del cálculo del SNR y del ruido.

Para el cálculo del ruido y del SNR se emplea la señal PPG en lugar de la señal procesada que muestra el pulso. La señal PPG tiene esta forma:

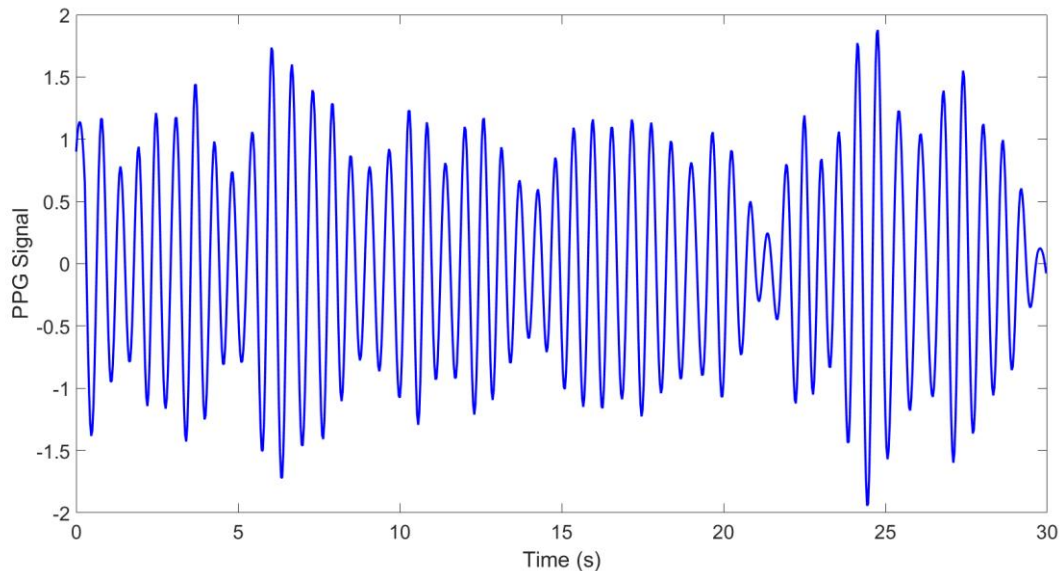


Ilustración 56: Señal PPG.

De la señal PPG se extraen los últimos 15 segundos para su análisis. Este paso se debe a que el algoritmo empleado para extraer la señal PPG trabaja con una señal PPG de 30 segundos que se va actualizando cada segundo. La muestra más vieja se elimina, la primera, y se añade una nueva muestra al final. De esta manera en cada segundo se mantienen los últimos 29 segundos en la señal PPG, el análisis del nivel de ruido y SNR se lleva a cabo en los 15 segundos más recientes. Se ha escogido 15 segundos como compromiso entre trabajar con varias muestras y emplear muestras recientes.

Una vez se tienen los últimos 15 segundos de señal PPG se procede al cálculo del ruido y SNR. El siguiente paso consiste en calcular la frecuencia “correcta” de la señal, para ello se emplea el pulso obtenido de la referencia de esos 15 segundos a analizar. En primer lugar se calcula la media del pulso en esos 15 segundos y a continuación se transforma el pulso, en latidos por segundo, a frecuencia, en Hz. Para la transformación se emplea la siguiente fórmula:

$$Hz = \frac{\text{Latidos por minuto}}{60 \text{ segundos}} \quad (3.7)$$

Una vez obtenida la frecuencia “correcta” de la señal se calcula el espectro de la señal PPG empleando la transformada rápida de Fourier (fft). De esta manera, los últimos 15 segundos de la señal PPG de la ilustración 56 representada en el dominio del tiempo pasa a tener la siguiente forma en el dominio frecuencial:

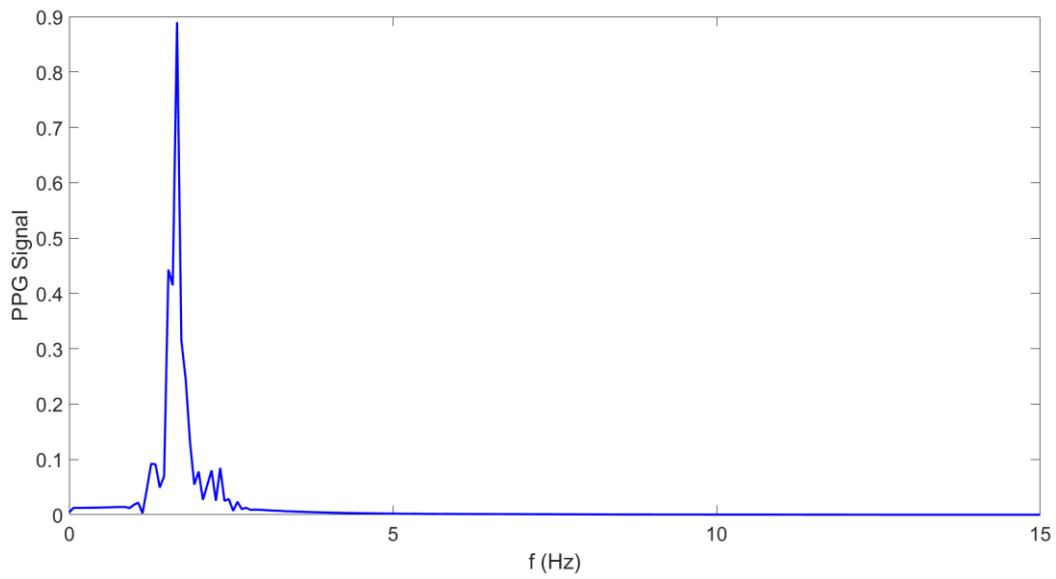


Ilustración 57: transformación frecuencial de la señal PPG.

El pico observado en la ilustración anterior se encuentra centrado en 1.667 Hz, que son unos 100 bpm. A continuación se busca en esta señal el punto correspondiente a la frecuencia “correcta” y el área de señal que corresponde a la frecuencia correcta más un pequeño rango por debajo y por encima, definido por el parámetro bins, será la parte considerada señal. El resto de la señal se considera ruido.

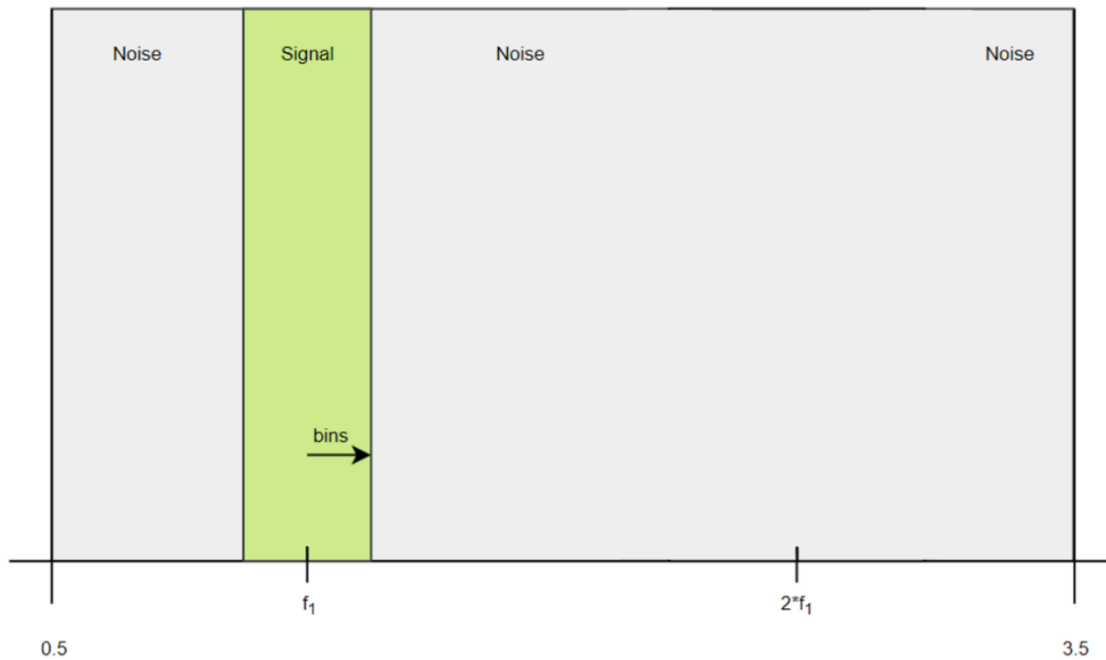


Ilustración 58: Obtención de la parte correspondiente a señal y la parte correspondiente al ruido.

En la ilustración 58 se ve señalado en el eje frecuencial el primer armónico, esto se debe a que en un principio se intentó trabajar con la señal PPG sin filtrar y se consideraba el primer armónico parte de la señal. Sin embargo solo en algunas señales se pudo distinguir señal entre el ruido entorno a la frecuencia del primer armónico y se descartó la idea. También se observa que el límite en el eje frecuencial varía de 0.5 Hz a 3.5 Hz, de 30 bpm a 210 bpm, estos se consideran los límites típicos entre los que se puede encontrar el pulso de una persona en reposo sin ninguna enfermedad que pueda afectar al pulso. De todos los vídeos analizados en ninguno se baja de 50 bpm y tampoco se superan los 120 bpm.

Llegados a este punto se tienen el área de la señal correspondiente al ruido y el área correspondiente a la señal. Solo queda hacer la conversión a dBs en el caso del ruido y dividir la parte correspondiente a la señal entre la parte correspondiente al ruido antes de la conversión en el caso del SNR.

En el cálculo del SNR y ruido se ha empleado la frecuencia obtenida de la señal referencia, por este motivo este nivel de ruido y SNR calculados se consideran centrados en la señal referencia. Otro cálculo que se lleva a cabo es el del nivel de ruido y SNR empleando la frecuencia obtenida de la señal pulso de la predicción. Estos niveles de ruido y SNR se consideran centrados en la predicción. De esta manera cuando se lleva a cabo el análisis del ruido y del nivel de SNR se indican ambos tal y como se indica en las siguientes ilustraciones:

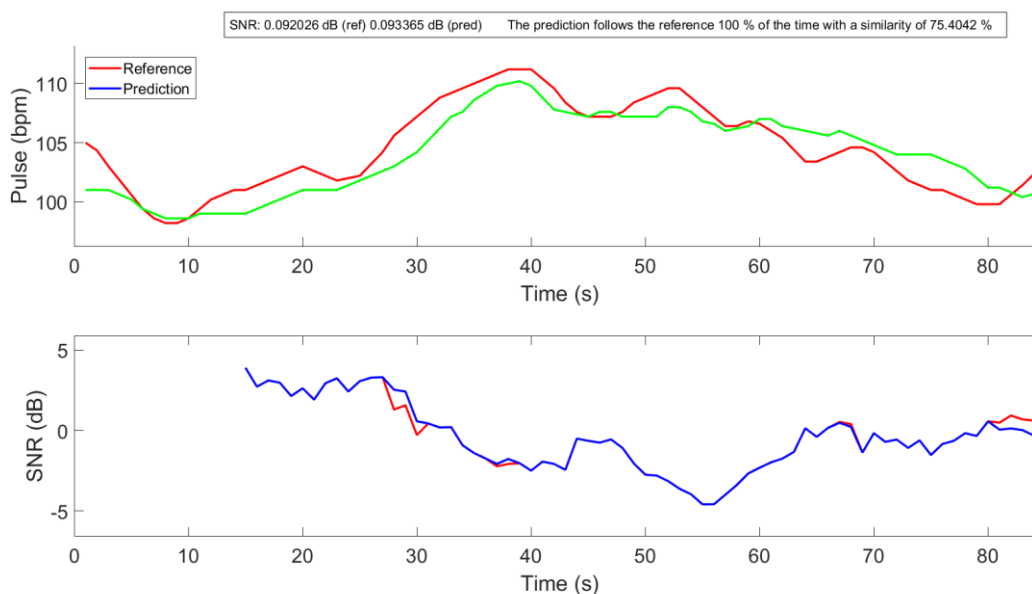


Ilustración 59: SNR de la señal.

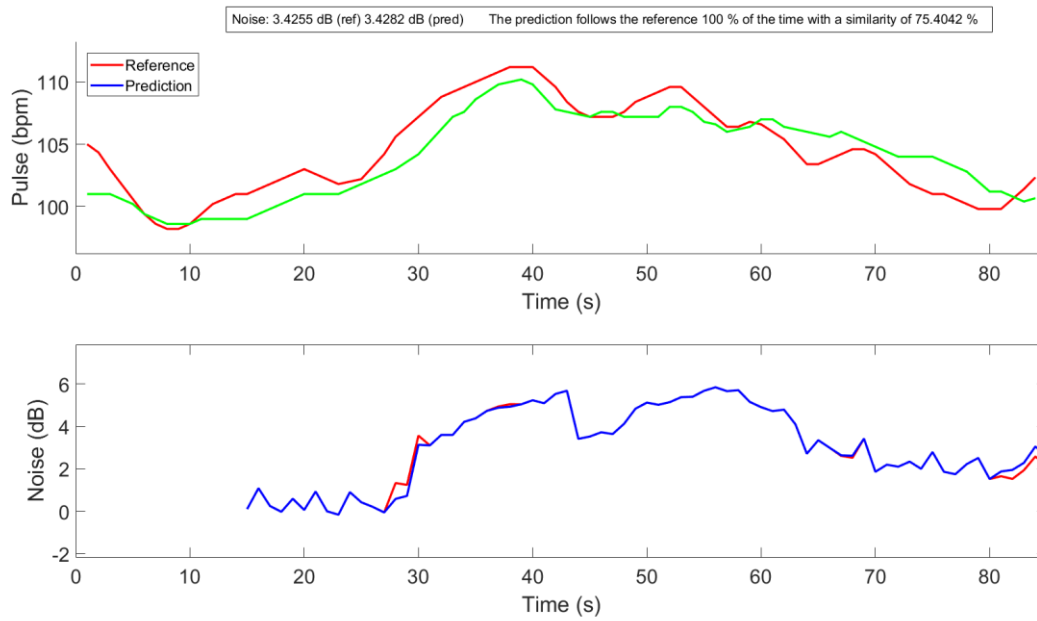


Ilustración 60: Ruido de la señal.

Tanto en la representación del ruido como del nivel de SNR la señal empieza en el segundo 15, esto se debe a que ambos parámetros se calculan ventana a ventana y la primera ventana abarca los primeros 15 segundos.

El nivel de ruido y el SNR son los últimos parámetros calculados en el análisis de la señal PPG, el resultado obtenido al realizar un análisis completo se representa en las tres imágenes siguientes:

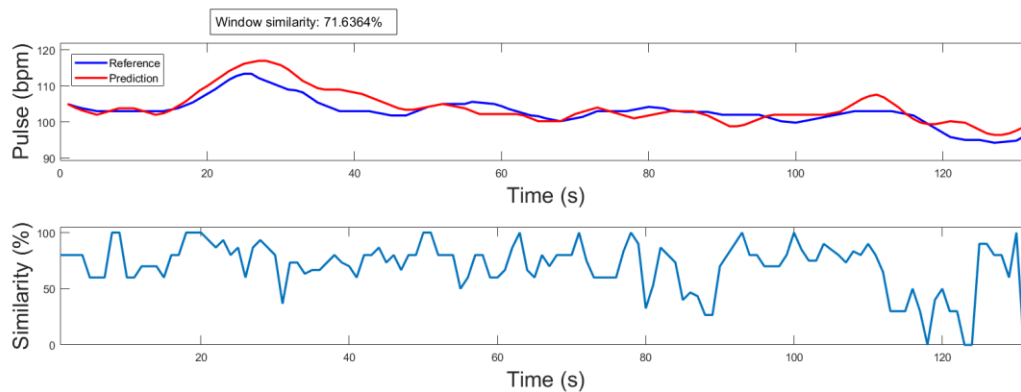


Ilustración 61: Similitud.

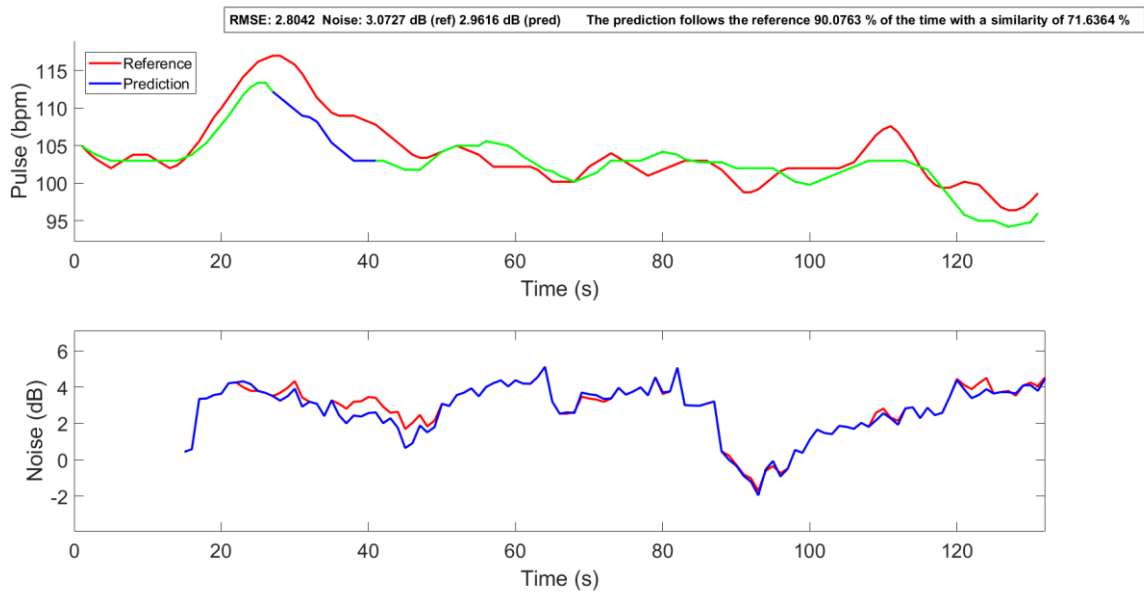


Ilustración 62: Análisis del ruido.

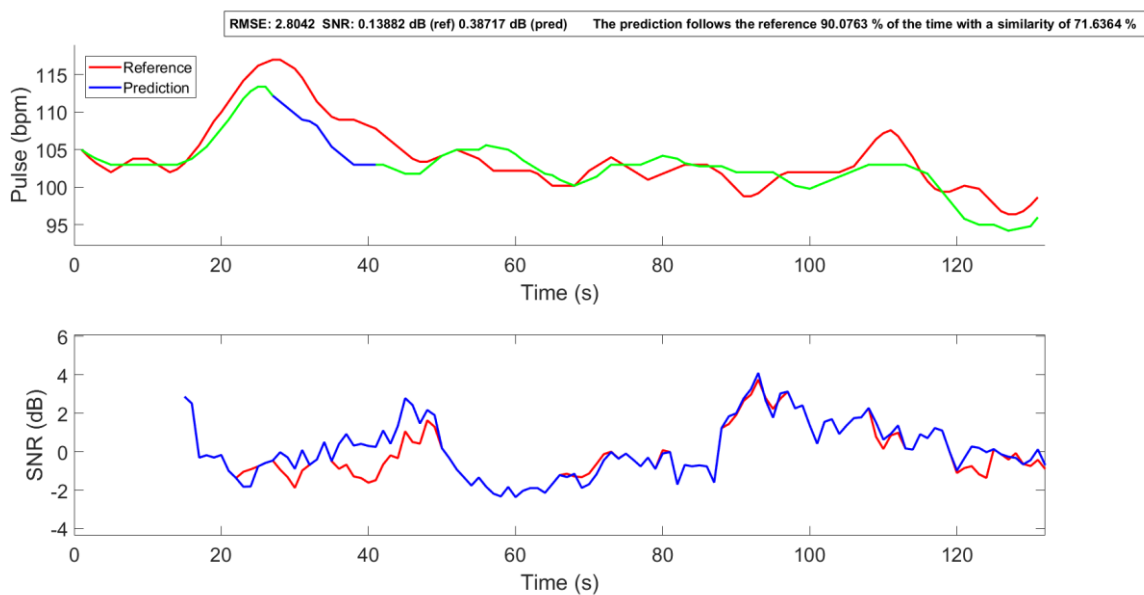


Ilustración 63: Análisis de SNR.

Por último solo falta explicar el índice de calidad combinado. Los análisis anteriores son útiles si se quiere entrar en los detalles de una señal específica, sin embargo cuando se trabaja con una gran cantidad de señales es necesario valorar la predicción sin entrar en tantos detalles. Por este motivo se ha diseñado un parámetro que combina los parámetros principales, este parámetro junto con la similitud y el

porcentaje de seguimiento es capaz de mostrar la validez de la predicción de forma rápida. El resultado final obtenido tras analizar un vídeo son los siguientes tres parámetros:

- Porcentaje de seguimiento, indica el porcentaje de la señal donde la predicción obtenida es buena, se considera una buena predicción si esta se encuentra dentro del rango marcado por la referencia $\pm 5\%$. Este parámetro viene a ser el porcentaje de éxito, si la precisión del sistema es fundamental este es el parámetro a tener en cuenta.
- Similitud, indica cómo de parecidas son la referencia y la predicción en cuanto a su forma. Evalúa cómo de capaz es la predicción de seguir el pulso. Si no interesa tanto el valor exacto de la predicción sino que se prefiere detectar las variaciones en el pulso este es el parámetro a tener en cuenta.
- Índice de calidad, este parámetro combina los dos anteriores junto con el RMSE para ajustar el valor. Es un parámetro que evalúa en general de la señal a analizar sin entrar en aspectos específicos.

El porcentaje de seguimiento y la similitud se obtienen directamente tal y como se ha visto en apartados anteriores, el índice de calidad se obtiene empleando la siguiente fórmula:

$$QI = \text{Seguimiento} * 0.75 + \text{Similitud} * 0.25 - \text{RMSE} \quad (3.8)$$

La similitud y el porcentaje de seguimiento son los principales valores empleados en la evaluación, el RMSE se emplea para ajustar el índice. Las señales con una buena predicción tienen un nivel bajo de RMSE y apenas se verán afectadas mientras que las malas predicciones obtienen un nivel más alto y reducirán considerablemente el valor obtenido.

Los parámetros similitud y porcentaje de seguimiento son valores acotados entre 0 % y 100 %, sin embargo el RMSE no está acotado, su rango va desde 0 a infinito. Esto puede parecer un problema a simple vista, una buena predicción puede sufrir una anomalía en un punto dando como resultado un RMSE lo suficientemente alto como para reducir considerablemente el índice de calidad. Pero en realidad bastaría con fijarse en la similitud y porcentaje de seguimiento por si solos para detectar que esa señal debería obtener un índice más alto. De esta manera se detectaría esta anomalía del comportamiento del algoritmo o de la señal a analizar. Por este motivo es importante mantener el seguimiento y la similitud como parámetros de apoyo al índice de calidad de cara a evaluar señales PPG.

Durante la evaluación del sistema PPG también se obtienen los niveles de SNR, este parámetro no se emplea en este caso porque diversos análisis, que se presentarán más adelante, muestran que no es un parámetro fiable en el análisis de este algoritmo PPG.

3. Evaluación PulseCam

3.1. Introducción

En este apartado se explicará el sistema empleado para la extracción de la señal PPG. El algoritmo empleado ha sido facilitado por el FZI. Este algoritmo se emplea en diversos proyectos relacionados con la extracción de signos vitales para uso médico y en vehículos principalmente.

A lo largo de este trabajo he tenido acceso completo al algoritmo de extracción de señales PPG y he trabajado en el análisis de numerosos vídeos. Todas las señales mostradas son los resultados que he obtenido tras aplicar el algoritmo sobre vídeos facilitados también por el FZI. Aunque he tenido acceso completo al algoritmo PPG, durante la explicación del funcionamiento no se entrará en detalles no publicados en artículos. El algoritmo todavía está en desarrollo y no está permitido revelar este tipo de detalles.

Tras la explicación se llevará a cabo una evaluación del sistema empleando el índice de calidad desarrollado en este trabajo.

3.2. Algoritmo PPG

En este apartado se explicará los fundamentos en los que se basa el sistema diseñado para la extracción de la señal PPG. Toda la información mostrada en este apartado se encuentra en el artículo *“An online PPGI approach for camera based heart rate monitoring using beat-to-beat detection”* [18], nombrado en la bibliografía.

Como se ha comentado al principio, cuando la luz alcanza una superficie se dan tres casos: atraviesa la superficie, se refleja o la superficie absorbe la luz. En este principio se basan numerosos sistemas PPG, la luz atraviesa los tejidos superficiales y se refleja en los tejidos internos. Una cámara captura la luz reflejada y esta señal es procesada por el algoritmo PPG para extraer los signos vitales.

La ilustración 64 muestra el comportamiento de la luz al interactuar con distintos tipos de tejidos. En gris se representa la interacción de la luz con huesos, la piel y otros tejidos que no presentan variación alguna con el tiempo. En este caso gran parte de la señal es absorbida y la pequeña porción de la luz que es reflejada no contiene información útil. Las señales azul y roja representan la interacción de la luz con las venas y arterias presentes en los tejidos. En este caso sí existe una variación de la señal respecto al tiempo, los latidos del corazón mueven la sangre por todo el cuerpo y con el

tiempo se detectan los cambios en el flujo sanguíneo. La señal reflejada sí contiene información útil.

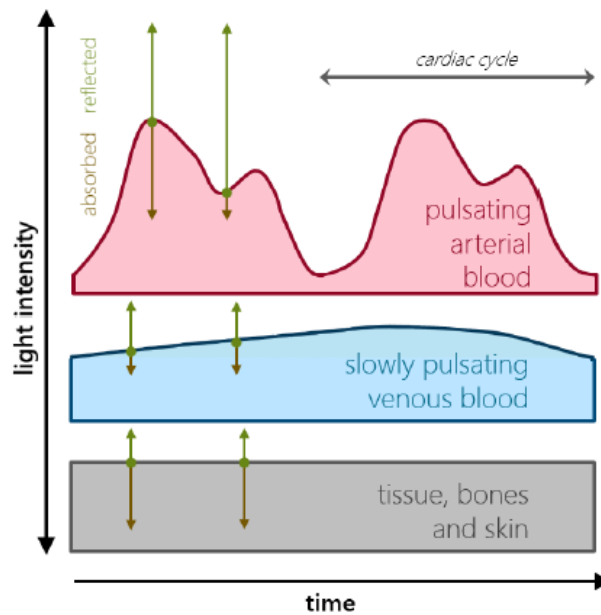


Ilustración 64: Interacción de la luz con tejidos del cuerpo humano [18].

La señal resultante está formada por componentes que interaccionan con los distintos tipos de tejidos. Las componentes obtenidas de la interacción de la luz con huesos y piel no presentan variación mientras que las componentes obtenidas de la interacción de la luz con el flujo sanguíneo son las que portan la información acerca de la señal pulsátil.

El sistema PPG se puede dividir en 4 elementos fundamentales:

- La detección del área de interés (ROI).
- La iluminación.
- Una cámara como sensor óptico.
- Extracción de los signos vitales

A continuación se explicarán en más detalle cada uno de estos aspectos fundamentales.

3.2.1. La detección del área de interés (ROI).

La zona a analizar es fundamental a la hora de obtener un sistema robusto. Las zonas ideales son aquella con una gran densidad de vasos sanguíneos lo más cerca posible del tejido externo de la piel. En este caso se escoge la cara como zona de análisis, contiene una gran densidad de vasos sanguíneos cerca de la piel y suele estar descubierta. Zonas los brazos pueden estar cubiertos por ropa y no siempre están disponibles para el análisis empleando cámaras.

La cara también presenta numerosas zonas donde la extracción del pulso es complicada o imposible. Los ojos, la nariz, cerca de la boca son algunas de las zonas que se deben evitar. En este caso la zona de extracción de signos vitales es la frente.

La frente es una de las zonas de la cara más estables en cuanto a los movimientos faciales, zonas como el entorno de la boca suelen moverse continuamente complicando la medición.



Ilustración 65: Localización de la frente [18].

Para la localización la región de interés se emplea un detector de caras entrenado capaz de seguir 66 puntos de interés. Dada la cara detectada se emplean los ojos para localizar la frente, que es la zona a analizar y por tanto la ROI. La localización de la cara también permite la obtención de puntos de interés que son empleados para corregir la inclinación de la cabeza.

3.2.2. La iluminación

El sistema empleado se basa en la interacción de la luz con los tejidos de la piel, es por este motivo que la luz desempeña un papel fundamental en el funcionamiento del sistema. Algunos de los sistemas PPG actuales emplean sistemas activos como son fuentes de luz artificial e incluso emiten luz en ciertas longitudes de onda específicas

como los 540 nm (longitud de onda del color verde, como se vio en el estado del arte es la longitud de onda que mejor respuesta obtiene). Sin importar el sistema empleado todos los algoritmos buscan una iluminación lo más homogénea posible que facilite la extracción de la señal pulsátil.

En este caso no se empleará ningún sistema de iluminación activo, la luz con la que se trabaja es luz ambiental. Uno de los motivos es que este sistema se pretende emplear en vehículos y la iluminación con luz del espectro visible podría dificultar la visibilidad del conductor.

3.2.3. Una cámara como sensor óptico

Existen diversos sistemas PPG basados en distintos tipos de sensores ópticos. Algunos sistemas se diseñaron para analizar la luz que logra atravesar los tejidos (empleados en las orejas o dedos), otros sistemas se basan en la luz reflejada, existen sistemas de contacto directo y sistemas sin contacto alguno con el tejido a analizar. Algunos de los sistemas son simples fotodiodos que analizan el nivel de luz recibido, sin embargo el sistema empleado en este trabajo emplea una cámara como detector óptico.

La cámara captura imágenes sobre al área a analizar y procesando estas imágenes se extrae la señal pulsátil. La cámara empleada no precisa de unas características técnicas difíciles de conseguir, cámaras webcam de bajo coste son suficientes. Los únicos parámetros fundamentales son:

- Las imágenes por segundo, define la velocidad de muestreo. En el caso de los vídeos analizados en este trabajo el framerate es de 25 o 30. Es importante que sea lo más estable posible.
- El tamaño del sensor también afecta al resultado, cuanto mayor es mejor para reducir el ruido de pixel.
- Formato de la imagen de salida, lo ideal es que el formato de la imagen de salida sea raw en lugar de otro donde se comprima la imagen. Al comprimir la imagen se pierde información.

En este caso se empleará una webcam Logitech C922.

3.2.4. Extracción de los signos vitales

El objetivo del sistema es la extracción de los signos vitales, para ello hace falta un algoritmo capaz de analizar la señal y extraer la información útil. Además de la extracción de la información útil también es muy importante la detección de artefactos en las imágenes que puedan afectar a la medición. Reflexiones indeseadas y movimiento

son algunos de los mayores problemas a los que se enfrentan los sistemas PPG basados en cámaras.

El código empleado ha sido desarrollado en Matlab con vistas a convertirlo a otros lenguajes más comunes en entornos industriales. El esquema general que muestra el funcionamiento del algoritmo se puede ver en la siguiente ilustración.

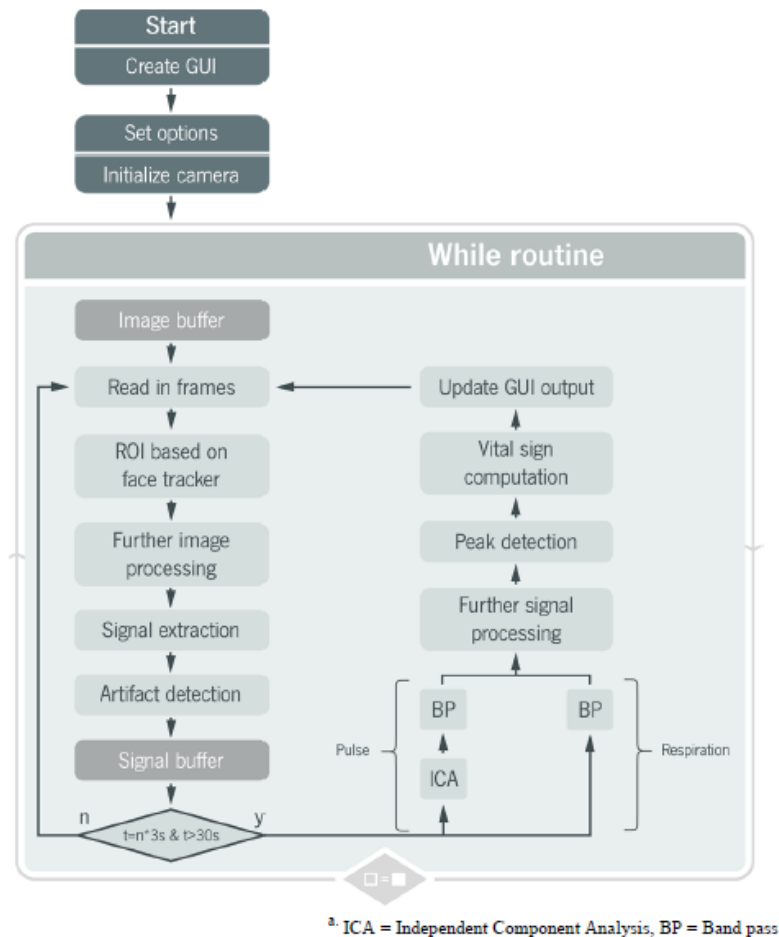


Ilustración 66: Algoritmo PPG empleado [18].

Uno de los primeros pasos consiste en definir la ROI, para ello se emplea un detector entrenado que permite localizar la frente de una persona. A continuación se lleva a cabo procesado de la imagen, como la aplicación de un filtro espacial, para facilitar la extracción de los signos vitales.

La extracción de la señal se lleva a cabo calculando la media, en la zona de interés, de los píxeles correspondientes al verde, al rojo y al azul. De esta manera se obtiene una señal con la media de las tres componentes para cada imagen analizada. Después se lleva a cabo la detección de artefactos que puedan afectar negativamente a la señal.

Los siguientes pasos se llevan a cabo pasados 30 segundos del inicio del análisis, este periodo de tiempo inicial se emplea para almacenar suficiente información como para obtener una predicción del pulso lo más precisa posible. Una vez pasados estos 30 segundos iniciales se procede al análisis de la señal obtenida en el paso anterior y la extracción de los signos vitales. Esta parte del algoritmo requiere más tiempo para completarse y no se ejecuta continuamente, sino que se ejecuta en cortos intervalos de tiempo.

Para la extracción del pulso y la frecuencia de respiración se aplican filtros paso banda adecuados a la señal a extraer. A continuación el pulso o la frecuencia de respiración se obtiene detectando los picos de la señal obtenida en cada caso.

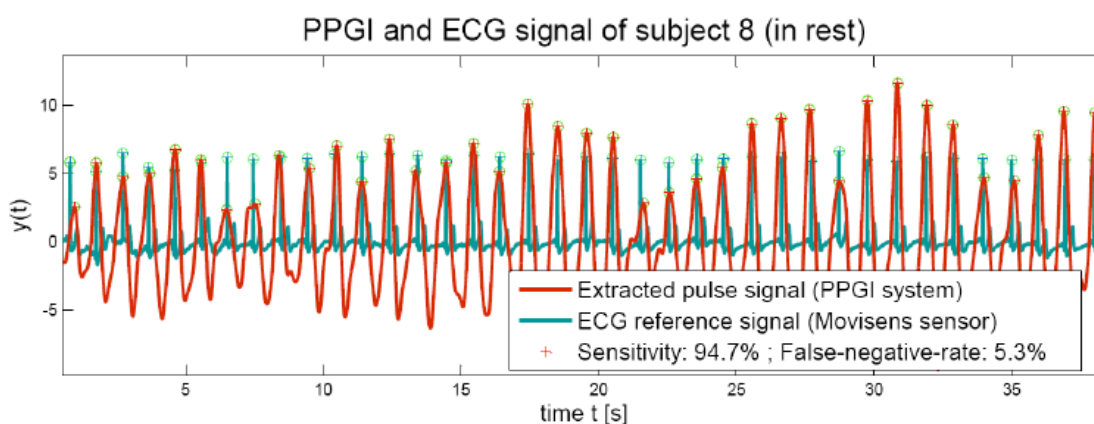


Ilustración 67: Señal pulsátil [18].

En cuanto a la detección de los artefactos, el sistema monitoriza tres aspectos fundamentales:

- El número de puntos de interés detectados en la cara.
- Cambios rápidos del ángulo de la cara.
- Que las amplitudes de las señales obtenidas no superen ciertos límites impuestos.

Los primeros dos elementos a monitorizar se obtienen directamente del detector de la cara mientras que el último se obtiene del análisis de las señales. Cuando se detecta un artefacto se localiza el espacio temporal en el que sucedió y se elimina de la extracción de los signos vitales.

3.3. Evaluación empleando el índice de calidad

En el apartado anterior se explican los fundamentos del algoritmo empleado para extraer la señal PPG de las imágenes.

En el apartado relativo a la explicación del índice de calidad diseñado se ha hablado de distintos parámetros cuyos valores debían definirse para el análisis del sistema PPG, de esta manera el primer paso es seleccionar los valores empleados en el análisis.

En el porcentaje de seguimiento se debe especificar el rango alrededor de la referencia dentro del cual la predicción se considera válida y durante cuánto tiempo debe mantenerse como mínimo.

Rango: ± 5 bpm. Rango suficiente como para ignorar pequeñas variaciones en el pulso sin ignorar cambios importantes.

Tiempo: 15 segundos. La predicción debe mantenerse dentro del rango durante al menos 15 segundos, de esta manera se asegura cierta estabilidad por parte de la predicción.

En el porcentaje de similitud se define un delay máximo y con el que se buscará la ventana de análisis óptima. Este delay se acota en -10 y 10 segundos, los resultados obtenidos al analizar el delay general de la señal muestra que gran parte de los casos óptimos se encuentran dentro de -10 y 10 segundos y un delay de 10 segundos es aceptable en este caso. En cuanto a la tolerancia, se ha decidido emplear un 0 %. No se ve necesario emplear este parámetro.

Por último solo queda definir el número de bins empleado en el cálculo del SNR y del ruido. Este parámetro decide cuántas muestras alrededor de la frecuencia central se consideran señal y a partir de cuál se considera ruido. Para seleccionar el valor se realizaron análisis empleando varios valores de bins, desde 1 hasta 5. Para cada señal se calculó la diferencia de SNR empleando la referencia como frecuencia central y SNR empleando la predicción como frecuencia central, de esta manera se obtuvo la diferencia de SNRs para cada valor de bins. La diferencia de SNRs resultó ser menor en el caso de 3 bins. Además un análisis posterior del espectro mostró que valores muy bajos como 1, y en algunos casos 2, no eran suficientes para contener el pico de la señal y valores altos como 5 contenían mucho ruido. Por estos motivos se escogió 3 como el valor adecuado de bins.

Estos son los parámetros empleados en el siguiente análisis del algoritmo PPG. A continuación se llevará a cabo el análisis del sistema PPG empleando el índice de calidad propuesto.

	% of the signal followed	SIMILARITY (%)	RMSE	SNR (dB) around Reference	SNR (dB) around Prediction	Quality Index
Proband_1	100	81.72	1.94	-2.02	-2.03	88.92
Proband_2	100	96.56	0.88	-0.45	-0.49	97.4
Proband_3	100	99.44	0.71	-1.45	-1.45	99.01
Proband_4	100	66.61	1.79	-1.45	-1.45	81.52
Proband_5	100	69.72	2.06	-2.18	-2.19	82.8
Proband_6	93.25	78.17	4.29	-2.53	-2.56	81.42
Proband_7	86.51	79.72	3.07	-3.51	-3.53	80.05
Proband_8	100	81	1.35	-2.06	-2.05	89.15
Proband_9	100	90.67	1.16	-0.78	-0.82	94.17
Proband_10	100	90.44	1.42	-3.16	-3.1	93.8
Proband_11	35.95	92.11	5.22	-2.26	-2.15	58.81
Proband_12	100	95.22	1.06	-2.16	-2.15	96.55
Proband_13	100	99.67	0.75	-0.51	-0.54	99.09
Proband_14	94.38	67.89	3.14	-1.78	-1.83	78
Proband_15	100	96.56	0.92	-1.48	-1.57	97.36
Proband_16	70.78	58.89	4.86	-3.5	-3.43	59.98
Proband_17	85.39	81.72	3.14	-1.3	-1.33	80.41
Proband_18	100	74.67	1.82	-0.61	-0.62	85.51
Proband_19	100	94	1.17	-2.18	-2.17	95.83
Proband_20	100	87	1.63	-1.64	-1.66	91.87
Proband_21	0	41.72	9.13	-5.5	-5.55	11.74
Proband_22	100	90.11	1.85	-0.55	-0.67	93.21
Proband_23	91.01	81.44	8.10	-0.86	-0.9	78.13
Proband_24	100	95.56	1.44	-2.26	-2.27	96.34
Proband_25	100	90.11	1.13	-3.16	-3.17	93.93
Proband_26	100	95	0.99	-1.08	-1.16	96.51

El primer análisis se lleva a cabo sobre vídeos con condiciones ideales como son poco movimiento y buena luz. En la tabla superior se han marcado en verde aquellas señales con un porcentaje de seguimiento superior al 40 %, son las predicciones consideradas buenas, y en rojo aquellas con un porcentaje inferior al 40 %, son las predicciones consideradas malas.

De las 26 señales analizadas 24 obtienen un porcentaje de seguimiento superior al 70 %. La media del porcentaje de seguimiento es 87,16 %. Esto quiere decir que de media un 87,16 % del tiempo el valor obtenido en la predicción se encuentra dentro del rango referencia ± 5 bpm.

En cuanto a la similitud, solo 5 vídeos obtienen un valor inferior al 70 % y la media es de 80,54 %. En media la señal referencia y la predicción tienen una forma muy parecida.

El RMSE es un buen indicador de la señal, todos los vídeos menos uno considerados buenas predicciones obtienen un valor del RMSE inferior a 5. Solo hay un caso donde el RMSE es alto siendo una predicción buena, el vídeo 23. Si nos fijamos en las predicciones malas los valores de RMSE son de 5, 22 y de 9,13 en los vídeos 11 y 21 respectivamente. Si analizamos el vídeo 11 vemos que el porcentaje de seguimiento es realmente bajo, 34 %, pero la similitud es bastante elevada, 92 %, dependiendo los criterios empleados se podría considerar una predicción buena si solo es importante la forma de la predicción en relación con la forma de la referencia. En este caso solo se obtienen dos resultados y no son suficientes para obtener conclusiones, más adelante se considerará de nuevo el RMSE de las predicciones malas y se extraerán conclusiones mejor respaldadas.

En cuanto al nivel de SNR, la mayoría de los valores obtenidos se encuentran por debajo de 0 dB. La diferencia entre las predicciones buenas y malas no muestra una clara diferencia en cuanto al SNR. Es cierto que en el caso del vídeo 21 el valor de SNR obtenido es el más bajo de todos, sin embargo el video 11 obtiene un valor de SNR de 2,26 dB que es muy similar al nivel de SNR de las predicciones buenas e incluso es superior al de algunos casos como el vídeo 7 y el vídeo 16, ambos con valores entorno a los 3,5 dB.

Finalmente, si nos fijamos en el índice de calidad la media obtenida es de 84,67 %. A continuación se mostrarán los resultados de algunos vídeos para ilustrar la respuesta del sistema PPG.

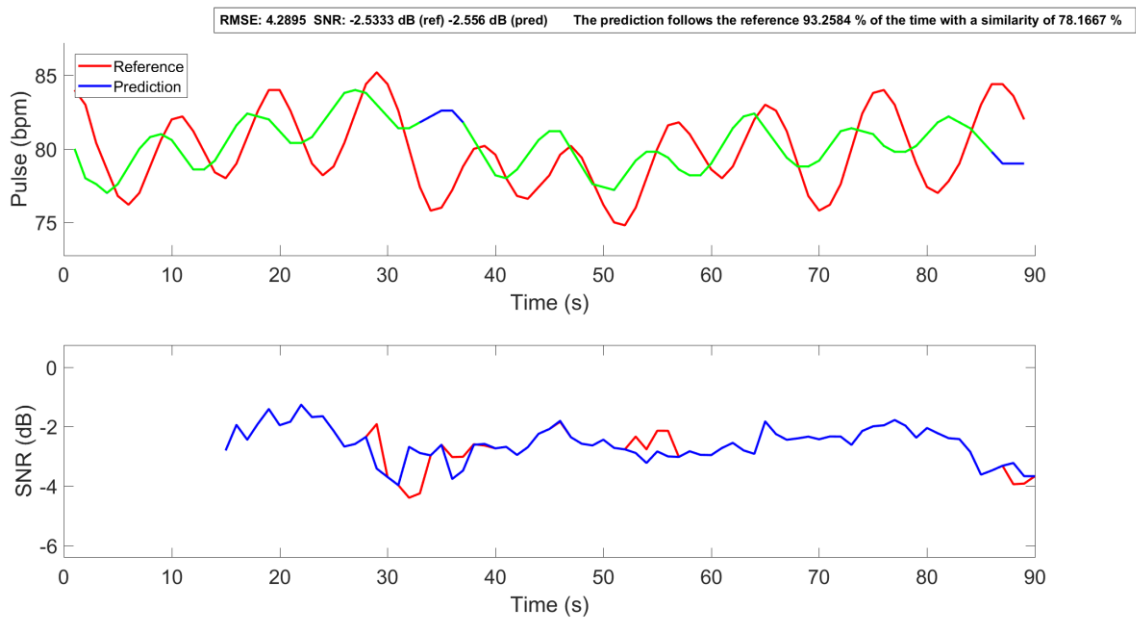


Ilustración 68: Proband 6.

En la ilustración 68 se muestra el resultado obtenido del vídeo 6. La señal de referencia muestra una variación continua del pulso que la predicción es capaz de seguir, aunque es cierto que la predicción se encuentra ligeramente desfasada en ciertas ocasiones. Un 93 % del tiempo la predicción obtiene un valor cercano a la referencia y la similitud es del 78 %. El índice de calidad es un 81 % en este caso.

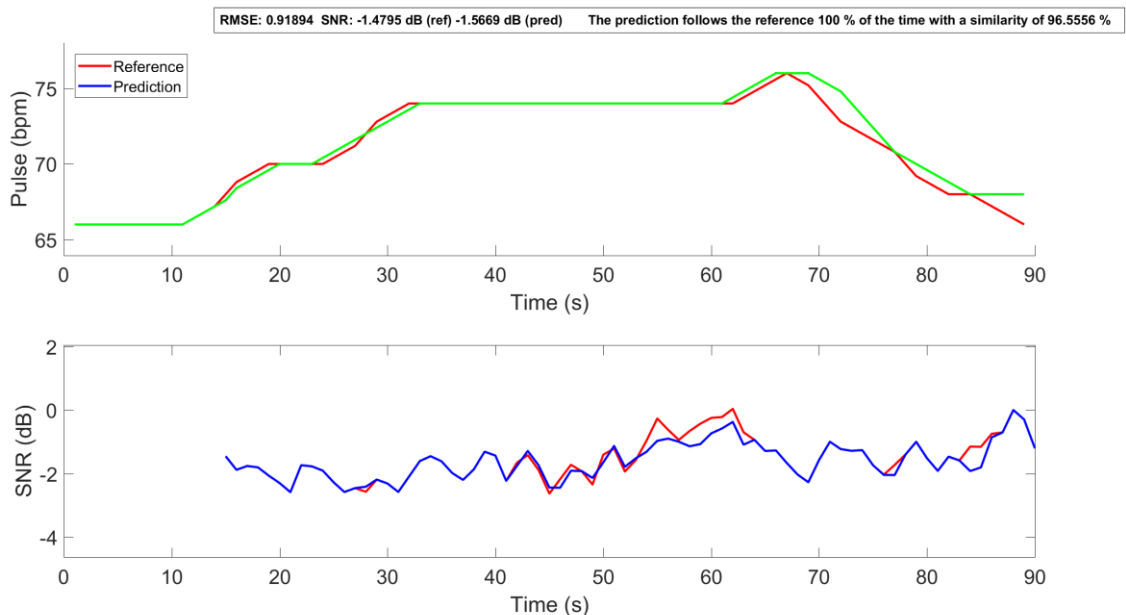


Ilustración 69: Proband 15.

La imagen anterior muestra el resultado del vídeo 15. La predicción se mantiene durante todo el vídeo muy cerca de la referencia. Este es prácticamente un caso ideal, por este motivo obtiene un 97,36 % en el índice de calidad.

La siguiente imagen muestra el peor caso medido, el vídeo 21. El porcentaje de seguimiento es del 0 %, los valores obtenidos en la predicción no son lo suficientemente cercanos a la referencia con la constancia requerida. En cuanto a la similitud este vídeo obtiene un 41 %, durante la primera mitad del vídeo la predicción detecta erróneamente variaciones que no se producen en la referencia. Esto puede ser debido al reflejo de la luz o movimientos bruscos. El índice de calidad obtenido es del 11 %.

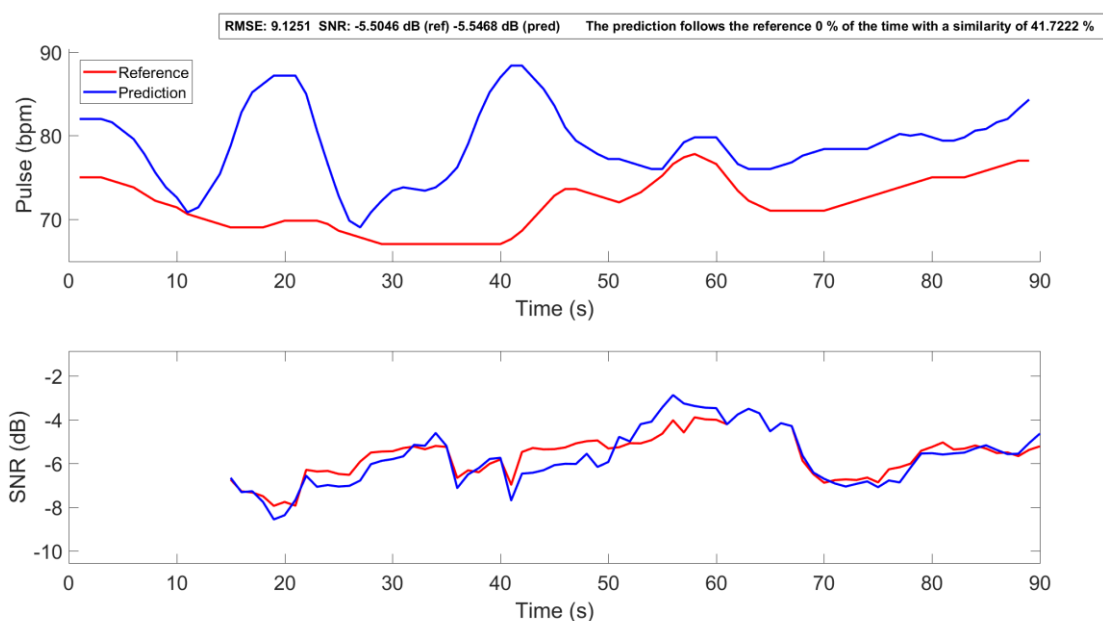


Ilustración 70: Proband 21.

En el siguiente análisis se han procesado vídeos donde la persona a analizar se mueve más y de forma más brusca, aunque manteniendo la naturalidad. En este caso se quiere estudiar cómo reacciona el algoritmo PPG frente a una situación habitual donde la persona que es analizada mueve la cabeza con naturalidad. Esta evaluación es importante para valorar el sistema de cara a aplicarlo en situaciones donde la persona a analizar mueve la cabeza para hablar o mirar en una dirección determinada.

	% of the signal followed	SIMILARITY (%)	RMSE	SNR (dB) around Reference	SNR (dB) around Prediction	Quality Index
Proband_1	17.98	51.67	7.93	-4.73	-4.75	26.89
Proband_2	53.93	73.11	8.09	-3.77	-3.77	55.43
Proband_3	94.38	64.06	2.21	-2.75	-2.7	77.01
Proband_4	0	34.28	11.31	-5.23	-5.41	5.82
Proband_5	-	-	-	-	-	-
Proband_6	41.57	65.61	7.62	-5.36	-5.31	45.97
Proband_7	55.05	65.67	7.75	-5.33	-5.88	52.61
Proband_8	68.54	71.94	3.69	-3.2	-3.33	66.55
Proband_9	100	85.28	1.75	-2.87	-2.87	90.89
Proband_10	21.35	71.44	8.32	-5.36	-5.56	38.08
Proband_11	61.80	65.94	4.79	-3.89	-4.04	59.08
Proband_12	58.42	52.61	8.41	-5.24	-4.55	47.11
Proband_13	80.90	60.83	3.65	-2.09	-2.1	67.22
Proband_14	50.56	68.94	4.68	-5.58	-5.6	55.07
Proband_15	61.80	67.33	4.50	-5.87	-5.93	60.06
Proband_16	28.09	54.11	6.39	-5.38	-5.28	34.71
Proband_17	40.45	60.50	5.19	-3.42	-3.75	45.28
Proband_18	70.78	79.50	4.67	-2.71	-2.74	70.47
Proband_19	62.92	60.22	5.25	-3.93	-3.75	56.32
Proband_20	92.13	79.44	3.18	-5.1	-4.9	82.60
Proband_21	0	45.83	18.74	-6.45	-5.78	4.18
Proband_22	76.40	67	3.04	-4.94	-4.98	68.66
Proband_23	59.55	63.56	4.61	-3.45	-3.56	56.94
Proband_24	89.88	76.72	2.65	-3.06	-3.14	80.65
Proband_25	34.83	53.06	9.11	-5.03	-5.36	34.83
Proband_26	41.53	45.22	8.87	-5.13	-5.13	34.52

Este análisis es fundamental, uno de los objetivos de los sistemas PPG es el de interferir lo menos posible con la persona a analizar de forma que no sea molesto. El objetivo del análisis en estas condiciones es el de valorar cómo de preparado está el algoritmo frente a movimientos y si es necesario mejorar su funcionamiento.

De nuevo los vídeos marcados en rojo son aquellos considerados una mala predicción, porcentaje de seguimiento inferior al 40 %, los vídeos en verde son considerados buenas predicciones, porcentaje de seguimiento superior al 70 %, y por último en gris están los vídeos intermedios, con porcentaje de seguimiento entre 40 % y 70 %. El vídeo 5 no se tiene en cuenta debido a problemas con la referencia, en un momento dado la referencia deja de capturar el pulso.

Ocho de los vídeos se consideran buenas predicciones, en cuanto a las malas predicciones en este caso hay ocho, y por último las señales consideradas intermedias son once. En el caso anterior había pocos vídeos que podrían considerarse puntos intermedios entre una buena predicción y una mala y por este motivo no se ve necesario clasificarlos como tal, sin embargo, en este caso si hay una cantidad suficiente como para diferenciar el caso intermedio.

Con el movimiento de la persona a analizar el porcentaje de seguimiento baja considerablemente. La media es de 54 %, un 33 % inferior al caso anterior. En cuanto a la similitud, la media, en un entorno donde la persona a analizar se mueve, es del 63 %. La similitud cae un 17 % respecto al caso ideal.

Si nos fijamos en el RMSE se observa que en los vídeos considerados buenas predicciones el valor obtenido se mantiene por debajo de 5 en todos los casos, en cuanto a las malas predicciones el nivel de RMSE más bajo es 6,39 y el resto de casos se mantienen muy cerca o por encima de 8, por último los casos intermedios varían entre 4,5 y 9.

De nuevo el nivel de SNR no es un claro indicador, hay casos como los vídeos 20 y 22 con nivel de SNR de -5 dB que se consideran buenas predicciones mientras que existen casos varios casos considerados malas predicciones un nivel de SNR muy cercano a los -5 dB y también hay varios casos intermedios con mejor nivel de SNR o muy parecido a las buenas predicciones.

Como es de esperar el índice de calidad obtenido también se reduce notablemente. El índice de calidad obtenido es del 52 %, un 28 % menor que en el caso anterior. Las predicciones consideradas buenas obtienen un índice da calidad superior al 65 %, las predicciones consideradas malas obtienen un índice inferior al 40 % y las intermedias varía entre 60 % y 35 %.

De nuevo se mostrarán algunos de los casos a continuación:

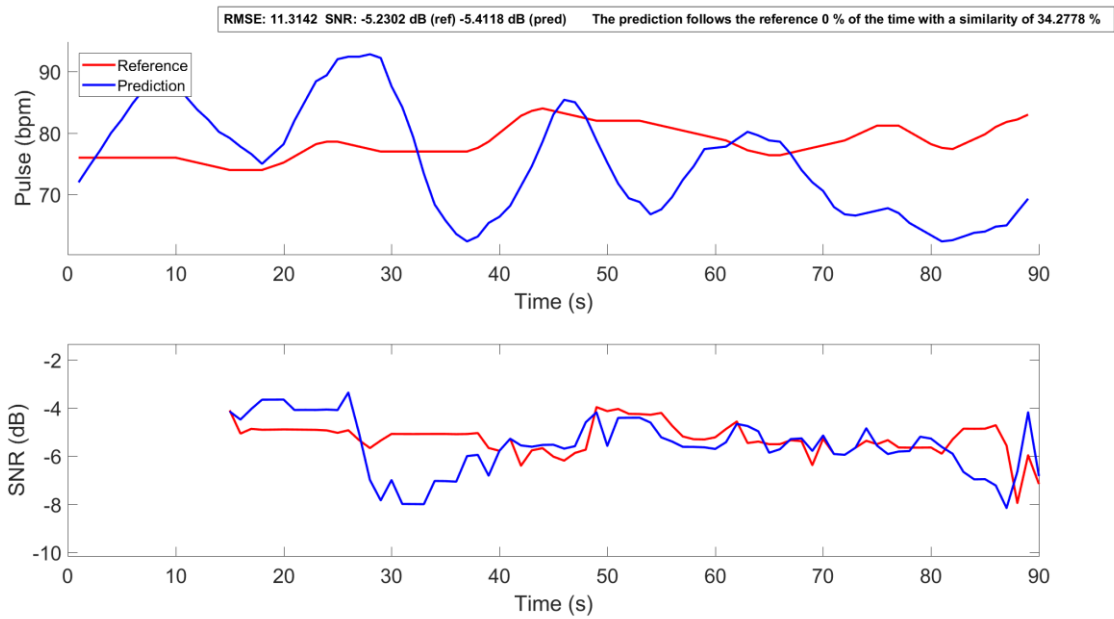


Ilustración 71: Proband 4.

La ilustración 71 muestra un claro ejemplo de una mala predicción, la predicción detecta variaciones en el pulso erróneas debido al movimiento. En este caso las variaciones detectadas tienen una gran amplitud y evitan que la predicción obtenga valores cercanos a la referencia de forma constante, por este motivo el porcentaje de seguimiento es del 0 %. Además estas variaciones deforman por completo la señal en relación a la señal de referencia, por este motivo la similitud obtenida es del 34 %.

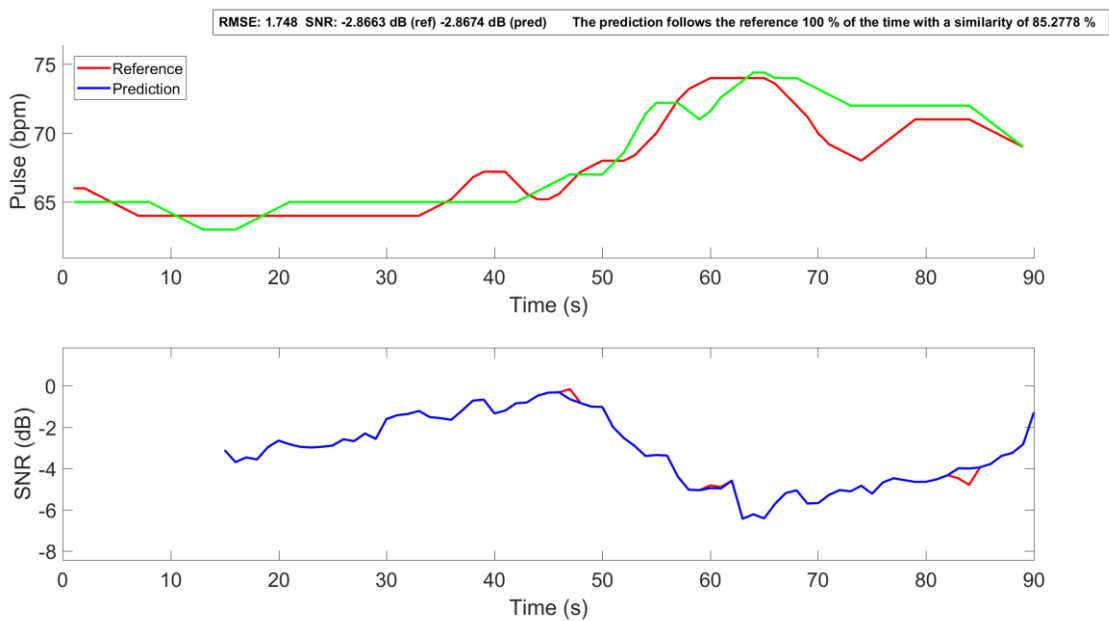


Ilustración 72: Proband 9.

La imagen anterior muestra una buena predicción en un entorno donde el sujeto se mueve, el porcentaje de seguimiento es del 100 %. En cuanto a la forma de la señal, la similitud detectada es un 85 %.

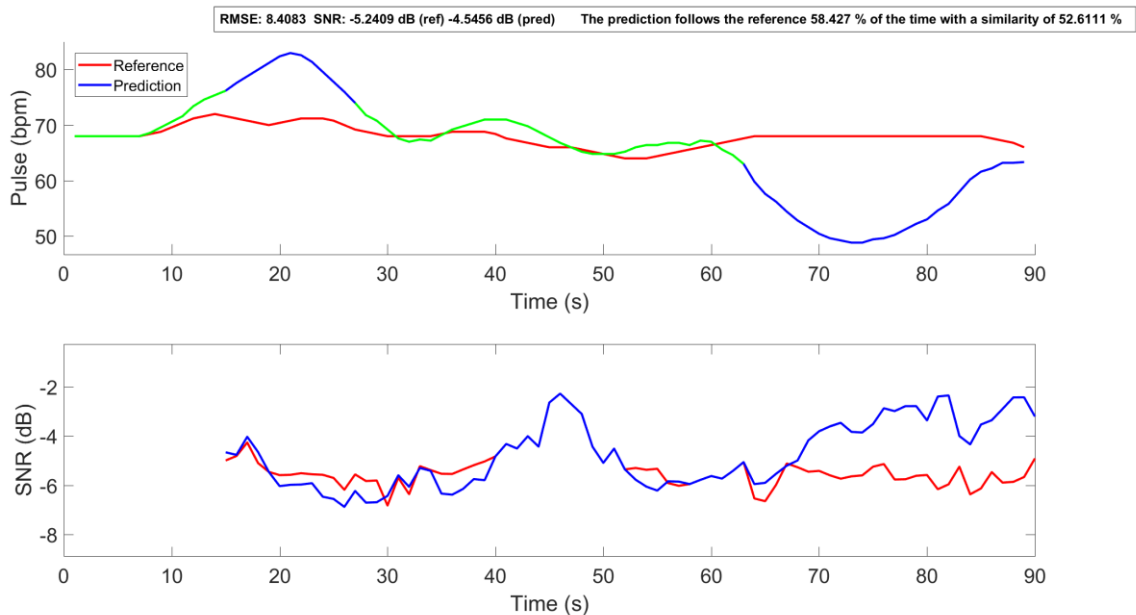


Ilustración 73: Proband 12.

El último caso muestra el análisis de un vídeo considerado intermedio entre una buena predicción y una mala predicción. En este caso hay dos zonas fácilmente identificables a simple vista donde la predicción empeora considerablemente. Los picos en el pulso erróneamente detectados por el algoritmo reducen la similitud y el porcentaje de seguimiento, sin embargo durante el resto de la señal la predicción es realmente buena. Por este motivo esta señal obtiene un índice de calidad del 47 %, prácticamente la mitad de la señal se ve afectada por la predicción errónea.

El último análisis del sistema es el del delay, en las siguientes imágenes se muestran los resultados obtenidos tras analizar varios vídeos empleando los delays temporales óptimos en cada ventana, de esta manera se puede saber si la predicción se encuentra adelantada, retrasada o en el mismo instante que la referencia.

La siguiente ilustración muestra los resultados obtenidos tras realizar barridos de -100 a 100 segundos de delay y con un 0 % de tolerancia.

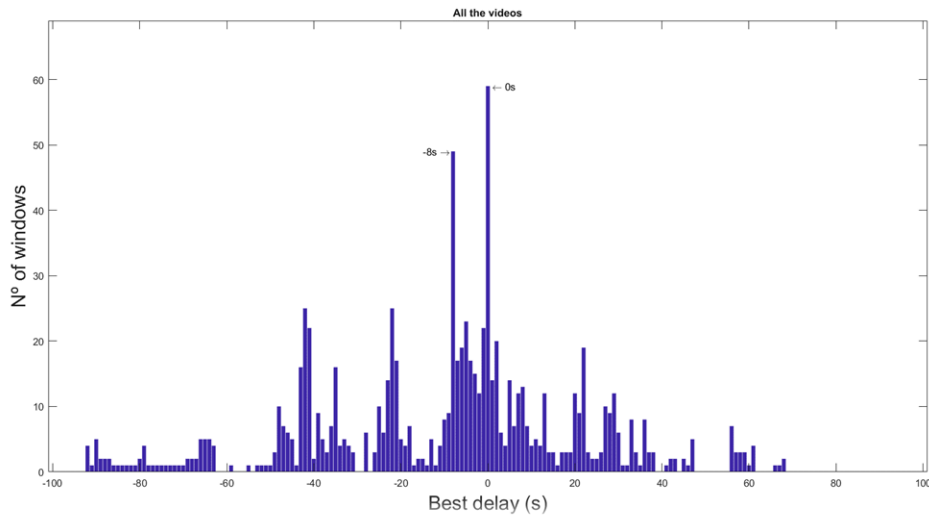


Ilustración 74: Barrido con 0 % de tolerancia.

La ilustración 74 muestra que el mejor delay se encuentra en 0 segundos seguido de -8 segundos. También se observa que una gran parte de los delays obtenidos se centran en el lado negativo del eje, es decir, que la predicción es ligeramente más rápida que la referencia. Para comprobar hasta qué punto estas dos situaciones representan el comportamiento real del algoritmo se llevaron a cabo análisis de delay aplicando un porcentaje de tolerancia.

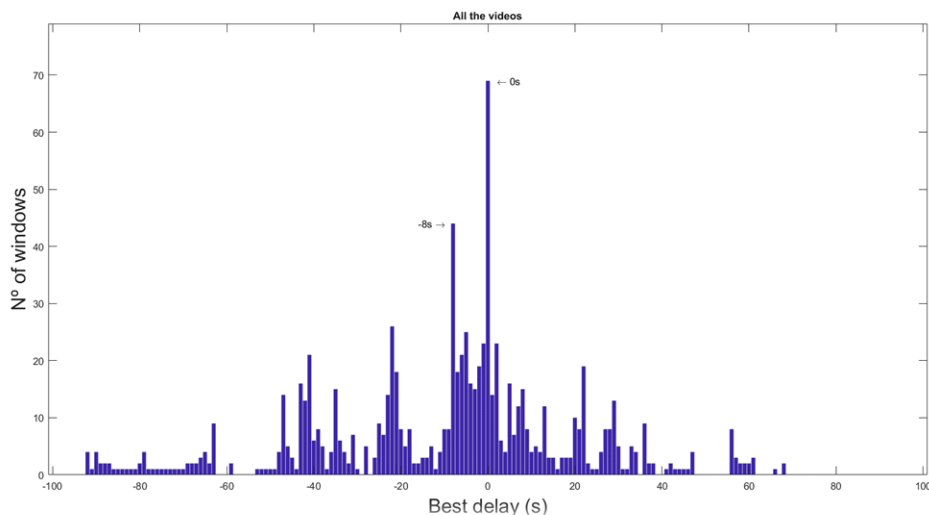


Ilustración 75: Barrido con 1 % de tolerancia.

La imagen anterior representa el delay tras aplicar un 1 % de tolerancia. De nuevo los delays óptimos son 0 y -8 segundos, sin embargo la diferencia entre ambos ha aumentado a favor de los 0 segundos. En cuanto a la tendencia, todavía se observa claramente una mayor densidad entorno a los valores negativos del delay.

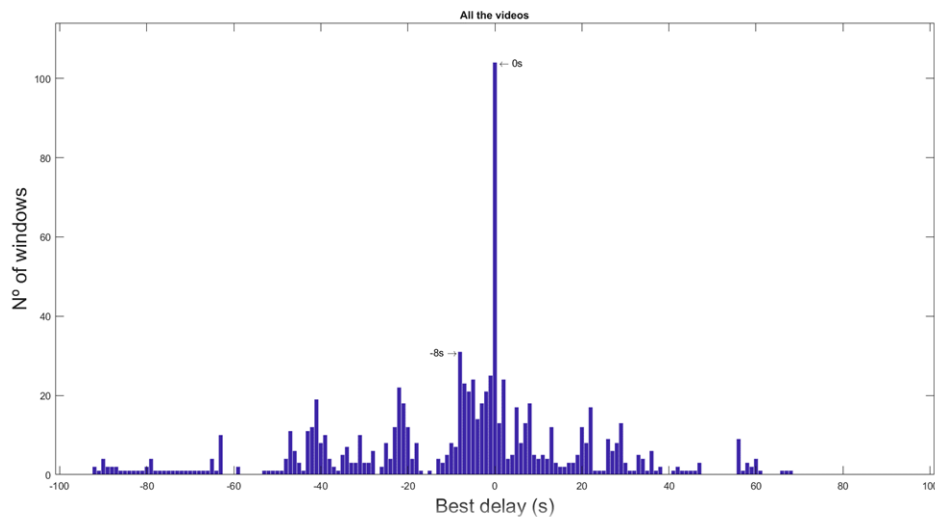


Ilustración 76: Barrido con 2 % de tolerancia.

Tras aplicar un 2 % de tolerancia en el análisis del delay se empiezan a notar cambios importantes. Sin duda alguna el delay óptimo es 0 segundos, el número de casos donde -8 segundos es el delay óptimo ha disminuido considerablemente. Si bien los delays óptimos distintos de 0 segundos han disminuido, se observa mayor densidad en los valores negativos.

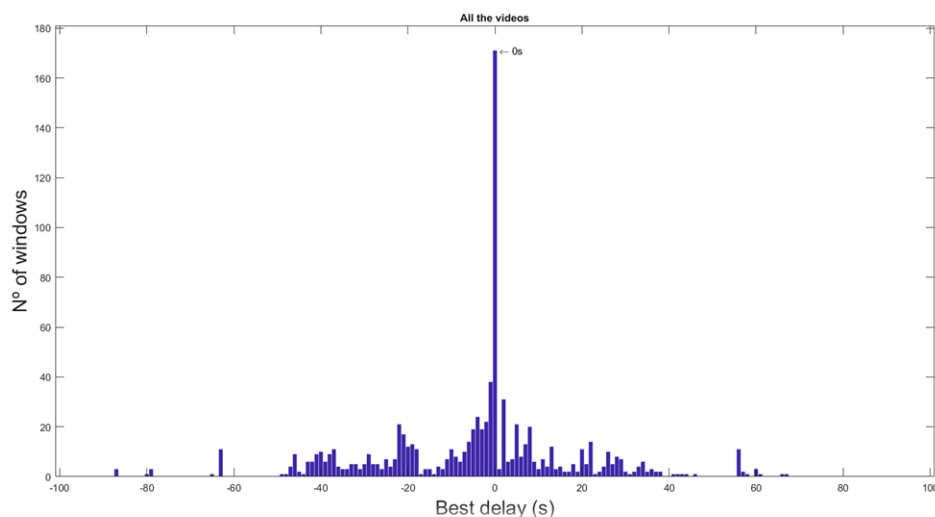


Ilustración 77: Barrido con 5 % de tolerancia.

El último caso es el análisis del delay aplicando un 5 % de tolerancia. El delay óptimo es 0 segundos y aunque existe una mayor densidad de valores en el lado negativo el número de casos distintos de 0 segundos no es tan significativo como es al aplicar 0 o 1 % de tolerancia.

4. Conclusiones

En este apartado se expondrán las conclusiones del índice de calidad diseñado y del análisis del sistema PPG.

4.1. Conclusiones del análisis del sistema PPG analizado

Una vez concluido el análisis del sistema PPG se ha observado que el algoritmo es capaz de obtener el pulso correcto y de forma precisa en condiciones ideales. Esto demuestra que el algoritmo de extracción del pulso funciona correctamente. La predicción es precisa en prácticamente toda la señal y la forma de la predicción es muy parecida a la referencia. En cuanto al RMSE este se mantiene bajo en prácticamente todos los casos. El análisis del SNR demuestra que el sistema no es dependiente de este parámetro, hay vídeos con niveles altos de SNR y otros con niveles bajos sin embargo el SNR alrededor de la referencia y alrededor de la predicción son prácticamente iguales en todos los casos.

Una vez se analizan vídeos donde las condiciones no son perfectas y existe movimiento los resultados obtenidos empeoran. En general la predicción pierde precisión y en diversos casos solo se mantiene en valores lo suficientemente cercanos a la referencia un 50 % del tiempo, aunque también existen varios casos donde la predicción es precisa durante más de un 70 % del tiempo. En cuanto a la similitud, esta también baja pero se mantiene en torno al 60 %. En este caso sí se notan diferencias de SNR al comparar el SNR obtenido alrededor de la referencia y el obtenido alrededor de la predicción, especialmente en vídeos donde la predicción no es muy buena. Este parámetro, la diferencia de SNR, no se tiene en cuenta en este estudio aunque se plantea la posibilidad de ser añadido en el cálculo del índice de calidad. En vista de los resultados, y tras un análisis más profundo de los vídeos donde la predicción empeora, se observa que el algoritmo sí es capaz de trabajar en situaciones como la que se plantea en este punto. El movimiento afecta a la predicción de forma clara y aun así el algoritmo es capaz de obtener un resultado bueno durante más de un 60 % del tiempo. Dependiendo de la aplicación y la precisión necesaria este resultado puede ser suficiente o no, lo que sí está claro es que el algoritmo sí es capaz de trabajar en condiciones más cercanas a la realidad. En cuanto al delay, se ve una clara tendencia si se analizan los resultados obtenidos al aplicar los distintos porcentajes de tolerancia. La predicción óptima se obtiene prácticamente en el mismo momento que la referencia, es decir, el sistema no es ni más rápido ni más lento que los sistemas más comúnmente utilizados en la actualidad. En cuanto a la tendencia en la zona negativa del delay óptimo que podría indicar que la predicción es ligeramente más rápida, esta se reduce

considerablemente al aplicar una tolerancia del 2 % e incluso más al aplicar la tolerancia del 5 %. Esto quiere decir que la mejora obtenida de esas muestras ligeramente adelantadas a la referencia es mínima.

El algoritmo analizado funciona correctamente, en condiciones ideales la respuesta es realmente buena mientras que en condiciones más cercanas a la realidad mantiene unos resultados prometedores. Actualmente se está trabajando en mejorar la respuesta del algoritmo frente al movimiento y reflexiones de manera que se aumente su fiabilidad en situaciones con movimiento e iluminación indeseada. Si se logra este avance el sistema diseñado tiene futuro en la implementación de diversas aplicaciones relacionadas con la salud y la automoción.

4.2. Conclusiones del índice de calidad diseñado

El índice de calidad diseñado está basado en el sistema PPG analizado, sin embargo se puede emplear con otros sistemas PPG cuyos principios sean similares. Esto no es muy difícil dado que muchos de los algoritmos diseñados se basan en los mismos principios.

El análisis del porcentaje de seguimiento da una clara indicación de la precisión obtenida a lo largo de la extracción del pulso. En cuanto a la similitud, este parámetro tiene utilidad si la aplicación del sistema requiere cierta capacidad de análisis de las variaciones del pulso. Al nivel de SNR no se le ha dado mucho peso en el análisis del sistema PPG debido a los resultados obtenidos, sin embargo puede ser un parámetro muy útil en otros casos. Además el análisis de las diferencias del SNR calculado alrededor de la referencia y alrededor de la predicción puede ser un parámetro interesante a estudiar. En el caso estudiado se han notado diferencias al respecto en vídeos con predicciones buenas y vídeos con predicciones malas o intermedias, este es un parámetro que se está analizando actualmente. El RMSE si está claro que es un índice fiable, por este motivo se emplea para afinar el índice de calidad.

El índice de calidad diseñado es un parámetro que combina varios y por ello muestra una valoración completa. Acompañado del porcentaje de seguimiento y de la similitud es suficiente para tener una clara idea de la predicción obtenida.

El índice de calidad funciona correctamente y es capaz de analizar satisfactoriamente predicciones sin embargo todavía se están realizando pequeños ajustes y mejoras, especialmente en cuanto a la similitud. Existen ciertas situaciones donde el sistema no se comporta como se espera, aunque apenas afecta al resultado obtenido por ser muy puntuales y poco habituales.

5. Bibliografía

- [1] J. Allen, “Photoplethysmography and its application in clinical physiological measurement,” *Physiological Measurement*, vol. 28, no. 3, pp. R1–R39, 2007.
- [2] H. Ugnell and P. A. Oberg, “The time-variable photoplethysmographic signal; dependence of the heart synchronous signal on wavelength and sample volume,” *Medical Engineering and Physics*, vol. 17, no. 8, pp. 571–578, 1995.
- [3] W. Cui et al., “In vivo reflectance of blood and tissue as a function of light wavelength,” *IEEE Transactions on Biomedical Engineering*, vol. 37, no. 6, pp. 632–639, June 1990.
- [4] Wang, W., “Robust and automatic remote photoplethysmography”, Eindhoven: Technische Universiteit, 23/10/2017.
-
- [5] S. Vogel et al., “In-ear vital signs monitoring using a novel microoptic reflective sensor,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 6, pp. 882–889, 2009.
- [6] P. Mannheimer, “The light-tissue interaction of pulse oximetry,” *Anesthesia and Analgesia*, vol. 105, no. 6, pp. S10–S17, 2007.
- [7] W. Verkruysse, M. Bartula, E. Bresch, M. Rocque, M. Meftah, and I. Kirenko, “Calibration of contactless pulse oximetry.” *Anesthesia and analgesia*, May 2016.
- [8] T. J. Farrell et al., “A diffusion theory model of spatially resolved, steady-state diffuse reflectance for the noninvasive determination of tissue optical properties in vivo,” *Medical Physics*, vol. 19, no. 4, pp. 879–888, 1992.
- [9] F. Corral et al., “A photoplethysmographic imaging system with supplementary capabilities,” *Optica Applicata*, vol. 44, no. 2, 2014.
- [10] Y. Sun et al., “Comparison of scientific cmos camera and webcam for monitoring cardiac pulse after exercise,” *Proc. SPIE*, vol. 8135, pp. 813 506– 813 506–7, 2011.
- [11] W. Wang et al., “Algorithmic principles of remote-ppg,” *IEEE Trans. Biomed. Eng.*, vol. 64, no. 7, pp. 1479–1491, July 2017.
- [12] S. Kwon et al., “Roi analysis for remote photoplethysmography on facial video,” in *Proceedings of 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2015, pp. 4938–4941.

- [13] M. Lewandowska et al., “Measuring pulse rate with a webcam - a noncontact method for evaluating cardiac activity,” in Computer Science and Information Systems (FedCSIS), Federated Conference on, Sept. 2011, pp. 405–410.
- [14] M.-Z. Poh et al., “Advancements in noncontact, multiparameter physiological measurements using a webcam,” Biomedical Engineering, IEEE Transactions on, vol. 58, no. 1, pp. 7–11, Jan. 2011.
- [15] G. de Haan and V. Jeanne, “Robust pulse rate from chrominance-based rppg,” Biomedical Engineering, IEEE Transactions on, vol. 60, no. 10, pp. 2878–2886, Oct. 2013.
- [16] G. R. Tsouri and Z. Li, “On the benefits of alternative color spaces for noncontact heart rate measurements using standard red-green-blue cameras,” J. Biomed. Opt., vol. 20, no. 4, p. 048002, Apr. 2015.
- [17] Mohamed Elgendi, “Optimal Signal Quality Index for Photoplethysmogram Signals”, *Bioengineering*, vol. 3, n. 4, 2016.
- [18] T. Blöcher, J. Schneider, M. Schinle and W. Stork, "An online PPGI approach for camera based heart rate monitoring using beat-to-beat detection," 2017 IEEE Sensors Applications Symposium (SAS), Glassboro, NJ, 2017, pp. 1-6.