

Trabajo Fin de Grado

Grado en Ingeniería de las Tecnologías de
Telecomunicación

Estimación del ritmo cardíaco mediante análisis de
secuencias de vídeo

Autor: Ángel González Sánchez

Tutor: Rubén Martín Clemente

Dep. Teoría de la Señal y Comunicaciones
Escuela Técnica Superior de Ingeniería
Universidad de Sevilla

Sevilla, 2019



Trabajo Fin de Grado
Grado en Ingeniería de las Tecnologías de Telecomunicación

Estimación del ritmo cardíaco mediante análisis de secuencias de vídeo

Autor:

Ángel González Sánchez

Tutor:

Rubén Martín Clemente

Profesor titular

Dep. de Teoría de la Señal y Comunicaciones

Escuela Técnica Superior de Ingeniería

Universidad de Sevilla

Sevilla, 2019

Trabajo Fin de Grado: Estimación del ritmo cardíaco mediante análisis de secuencias de vídeo

Autor: Ángel González Sánchez

Tutor: Rubén Martín Clemente

El tribunal nombrado para juzgar el Trabajo arriba indicado, compuesto por los siguientes miembros:

Presidente:

Vocales:

Secretario:

Acuerdan otorgarle la calificación de:

Sevilla, 2019

El Secretario del Tribunal

A mi familia

A mis amigos

A mis maestros

Agradecimientos

Estoy agradecido a toda aquella persona que, de cualquier forma, ha participado en mi vida sirviendo de alguna ayuda. Ya sea debido a su apoyo, sus enseñanzas, su motivación en mí, o simplemente por su cercana compañía. Desde profesores y compañeros, hasta amigos y familiares.

Agradecer sobre todo a mis padres, ya que han sido la base fundamental sobre la que me he apoyado para poder crecer, aprender y llegar hasta aquí superando grandes dificultades por el camino.

A mi tutor Rubén por haberme dado la oportunidad de hacer este trabajo que tanto me ha gustado y con el que he aprendido con cada dificultad encontrada.

Un agradecimiento especial para mi compañera durante estos años, gracias por apoyarme siempre cuando me venía abajo y por animarme a continuar adelante, gracias Sara.

Ángel González Sánchez

Sevilla, 2019

Resumen

La estimación y el seguimiento de la frecuencia cardíaca han sido esenciales en los campos de la medicina y la ingeniería biomédica a lo largo de los años. En la actualidad, existe un gran interés por extraerla sin contacto, por lo que la estimación de la frecuencia cardíaca mediante el uso de la cámara digital es un área de investigación en rápido crecimiento debido a su bajo costo y su naturaleza no invasiva. El objetivo de este trabajo es diseñar y desarrollar un algoritmo que nos permita detectar el ritmo cardíaco de un individuo mediante el análisis de secuencias de vídeo de su cara. Se ha diseñado e implementado una nueva aproximación, basada en un algoritmo base fundamentado en el trabajo previo en el que nos apoyamos.

En primer lugar se realiza la detección facial mediante el algoritmo de Viola-Jones para localizar el área de la imagen que contiene la cara. A continuación, se extrae el movimiento de la cabeza usando el método de Lucas-Kanade y se aísla el movimiento correspondiente al pulso mediante un filtrado. Posteriormente se realiza el análisis de componentes principales (PCA) para seleccionar la componente que mejor se corresponda con los latidos del corazón en función de su espectro de frecuencia temporal. Finalmente, se analiza el movimiento proyectado en esta componente y se estima el ritmo cardíaco como la frecuencia (en bpm) con mayor potencia.

Los experimentos muestran que, si los vídeos están grabados bajo unas buenas condiciones, podemos detectar el ritmo cardíaco de una persona obteniendo unas muy bajas tasas de error.

Abstract

The estimation and monitoring of heart rate have been essential in the fields of medicine and biomedical engineering over the years. At present, there is a great interest to extract it without contact, so the estimation of the heart rate by using the digital camera is a research area in rapid growth due to its low cost and non-invasive nature. The objective of this work is to design and develop an algorithm that allows us to detect the heart rhythm of an individual by analyzing video sequences of his face. A new approach has been designed and implemented, starting from a base algorithm based on previous work in which we rely.

First, facial detection is performed using the Viola-Jones algorithm to locate the area of the image that contains the face. Then the motion of the head is extracted using the Lucas-Kanade method and the motion corresponding to the pulse is isolated by filtering. Afterwards, the main components analysis (PCA) is performed to select the component that best corresponds to the heartbeat according to its temporal frequency spectrum. Finally, the motion projected to this component is analyzed and the heart rate is estimated as the frequency (in bpm) with greater power.

The experiments show that, if the videos are recorded under good conditions, we can detect the heart rate of a person getting low error rates.

Índice

Agradecimientos	ix
Resumen	xi
Abstract	xiii
Índice	xv
Índice de Tablas	xvii
Índice de Figuras	xix
Notación	xxi
1 Introducción	1
1.1 <i>Motivación del trabajo</i>	1
1.2 <i>Objetivos y enfoque</i>	2
1.3 <i>Organización de la memoria</i>	2
2 Estado del arte	3
2.1 <i>Métodos de obtención del ritmo cardíaco por análisis de vídeo</i>	3
2.1.1 <i>Mediante cambios de color</i>	3
2.1.2 <i>Mediante movimientos</i>	6
3 El corazón y el electrocardiograma	9
3.1 <i>Fisiología cardíaca</i>	9
3.2 <i>El electrocardiograma (ECG)</i>	11
3.2.1 <i>Componentes del ECG</i>	12
3.2.2 <i>Frecuencia cardíaca</i>	12
3.2.3 <i>Principales trastornos del ritmo cardíaco</i>	14
4 Marco teórico	15
4.1 <i>Detección facial</i>	15
4.1.1 <i>Algoritmo de Viola-Jones</i>	15
4.2 <i>Estimación del flujo óptico</i>	17
4.2.1 <i>Algoritmo de Lucas-Kanade</i>	19
4.3 <i>Reducción de la dimensionalidad</i>	22
4.3.1 <i>Análisis de Componentes Principales (PCA)</i>	22
5 Implementación del sistema	25
5.1 <i>Entorno de desarrollo</i>	25

5.1.1	Lenguaje de programación	25
5.1.2	Entorno de trabajo	25
5.1.3	Librerías y extensiones	25
5.2	<i>Algoritmo propuesto</i>	26
5.2.1	Detección facial y selección de ROI	26
5.2.2	Selección y seguimiento de puntos de interés	28
5.2.3	Filtrado temporal	29
5.2.4	Reducción de componentes	29
5.2.5	Obtención de la frecuencia cardíaca	29
6	Experimentos realizados y resultados	31
6.1	<i>Evaluación de la detección</i>	31
6.2	<i>Evaluación de la estimación en vídeos grabados</i>	32
6.3	<i>Evaluación de la estimación en vídeos en tiempo real</i>	35
7	Conclusiones y líneas futuras	39
7.1	<i>Conclusiones</i>	39
7.2	<i>Líneas futuras</i>	40
	Referencias	41

ÍNDICE DE TABLAS

Tabla 2-1. Resumen de algoritmos basados en cambios de color	6
Tabla 3-1. Pulsaciones normales en reposo en mujeres.	13
Tabla 3-2. Pulsaciones normales en reposo en hombres.	13
Tabla 6-1. Resultados de los algoritmos para cada cara	33
Tabla 6-2. Resultados del algoritmo para cada caso	38

ÍNDICE DE FIGURAS

Figura 2-1. Apps basadas en cambios de color	4
Figura 2-2. Ejemplo de algoritmo basado en cambios de color con separación en bandas de color y uso de ICA	5
Figura 2-3. Balistocardiografo y balistocardiogramas del Dr. Starr	6
Figura 2-4. Movimiento involuntario debido a la circulación de la sangre desde el corazón a la cabeza	7
Figura 3-1. Anatomía del corazón y flujo de sangre	10
Figura 3-2. Sistema de conducción cardíaco	11
Figura 3-3. Representación de un ECG	12
Figura 3-4. Ejemplo de medición del pulso en la muñeca	13
Figura 3-5. ECG de distintas afecciones cardíacas	14
Figura 4-1. a) Características tipo <i>Haar</i> . b) Ejemplos de características <i>Haar</i> comunes de un detector de cara	16
Figura 4-2. Detección de la región de la cara	17
Figura 4-3. Efecto <i>barber pole</i> : ¿En qué sentido se mueven las líneas negras? Su movimiento aparente es hacia la izquierda y arriba, pero en realidad podrían estar moviéndose de infinitas otras maneras.	17
Figura 4-4. Problema de apertura: En esta aproximación, el cambio de brillo es lineal según dirección y magnitud de $\nabla I_x, y$. Debido a esto, existen infinitos vectores v que son solución de (4.6)	19
Figura 4-5. Esquema descriptivo de la obtención del flujo óptico	19
Figura 4-6. Vecindad alrededor del elemento q_5 de tamaño 3×3	20
Figura 4-7. Pirámide de resolución: cada nivel tiene la mitad de resolución que el nivel inferior.	21
Figura 4-8. Ejemplo de reducción de dimensionalidad aplicando PCA	23
Figura 5-1. Diagrama de las fases del algoritmo propuesto	26
Figura 5-2. Detección facial inicial	27
Figura 5-3. Selección final de las ROIs	27
Figura 5-4. Representación de los puntos de interés	28
Figura 5-5. Ejemplo de estimación del pulso a partir de la PSD	30

Figura 5-6. Estimación final del ritmo cardíaco	30
Figura 6-1. Detección correcta	31
Figura 6-2. Vídeos del MIT. a) 54 bpm b) 55 bpm c) 152 bpm	32
Figura 6-3. Estimación final del ritmo cardíaco del caso 55bpm	33
Figura 6-4. Variación de la posición de los puntos seleccionados en el vídeo	34
Figura 6-5. Componentes principales, destacando la más periódica	34
Figura 6-6. Densidad espectral de la componente principal	35
Figura 6-7. Estimación del ritmo cardíaco en reposo (pulsómetro 1)	36
Figura 6-8. Estimación del ritmo cardíaco en reposo (pulsómetro 2)	36
Figura 6-9. Estimación del ritmo cardíaco acelerado (pulsómetro 1)	37
Figura 6-10. Estimación del ritmo cardíaco acelerado (pulsómetro 2)	37

Notación

(x, y, t)	Coordenadas x, y de un píxel en el tiempo t
I	Intensidad del píxel
I^T	Traspuesta
I^{-1}	Inversa
$\sum$	Sumatorio
Δx	Desplazamiento en la dirección x
$\frac{\partial I}{\partial x} \equiv I_x$	Derivada parcial de I respecto a x
$\mathbf{v}$	Velocidad del píxel
$\frac{\Delta x}{\Delta t} \equiv v_x$	Componente x de la velocidad
q_i	Puntos en el parche
$\mathbf{A}, \mathbf{b}$	Matrices utilizadas para emplear el método de mínimos cuadrados
$\mathbf{X}$	Conjunto de datos
m	Dimensión del conjunto $\mathbf{X}$
$\mathbf{P}$	Matriz de representación
$\mathbf{Y}$	Conjunto de datos transformado
$m' (m' < m)$	Dimensión del conjunto $\mathbf{Y}$
$\mathbf{C}_X$	Matriz de covarianza de $\mathbf{X}$
$\mathbf{U}, \mathbf{S}, \mathbf{V}$	Matrices resultado de la descomposición SVD

1 INTRODUCCIÓN

1.1 Motivación del trabajo

La estimación de la frecuencia cardíaca es un componente esencial para determinar el estado fisiológico y patológico de una persona o individuo. El corazón es uno de los órganos más importantes del cuerpo humano. La función del corazón no es otra que la de bombear sangre por todo el cuerpo latiendo aproximadamente de 60 a 100 latidos por minuto (bpm). La frecuencia cardíaca por encima de 100 bpm, se conoce como taquicardia, y la frecuencia cardíaca por debajo de 60 bpm, se conoce como bradicardia. Las anomalías en la frecuencia cardíaca son causadas por factores fisiológicos y patológicos que interrumpen el impulso eléctrico normal que controla la acción de bombeo del corazón.

La estimación y el seguimiento de la frecuencia cardíaca han sido esenciales en los campos de la medicina y la ingeniería biomédica. Los investigadores han desarrollado varios métodos para estimar la frecuencia cardíaca, desde sensores de contacto hasta sensores sin contacto. Los métodos de medición de la frecuencia cardíaca estándar y establecida están en contacto con el cuerpo, tales como: electrocardiograma y fotopletimografía, que se basan en electrodos conductores y fototransistores. Los métodos experimentales sin contacto son la medición a distancia por microondas/ultrasonido, las imágenes térmicas y las imágenes RGB. El objetivo común de cualquiera de estos métodos es la extracción de parámetros fisiológicos para estimar y monitorizar la frecuencia cardíaca de forma fiable.

Existe un interés creciente por extraerlo sin contacto, especialmente para poblaciones como los recién nacidos prematuros y las personas de edad avanzada, para quienes la piel es frágil y vulnerable a los sensores tradicionales. Además, a medida que la población envejece, la monitorización continua o al menos frecuente fuera de los entornos clínicos puede proporcionar a los médicos no sólo muestras oportunas, sino también tendencias a largo plazo y análisis estadísticos.

La estimación de la frecuencia cardíaca mediante el uso de sensores de cámara digital es un área de investigación en rápido crecimiento debido a su bajo coste y a su naturaleza no invasiva para la medición de la frecuencia cardíaca. La estimación de la frecuencia cardíaca utilizando un sensor de cámara digital se realiza principalmente sobre la base de los conceptos de fotopletimografía y balistocardiografía. Los métodos de extracción de frecuencia cardíaca basados en cámaras digitales han evolucionado a partir del uso de dispositivos de acoplamiento de carga de alta definición (CCD) y sensores de cámara de semiconductores de óxido metálico (CMOS) para teléfonos móviles y cámaras webcam. Por lo tanto, la medición de la frecuencia cardíaca a partir de vídeos faciales con sensores de cámara digital se ha convertido en una de las direcciones motivadoras para extraer señales fisiológicas de forma no invasiva. El progreso de la medición de la frecuencia cardíaca mediante vídeos faciales ha dado un nuevo rumbo hacia la monitorización de la salud en los campos de la asistencia sanitaria, la telemedicina, la rehabilitación, los deportes y la ergonomía.

En este trabajo se busca diseñar, desarrollar y evaluar un algoritmo, que nos permita captar la actividad cardíaca mediante el análisis del movimiento inducido en la cabeza. El movimiento cíclico de la sangre del corazón a la cabeza a través de la aorta abdominal y las arterias carótidas hace que la cabeza se mueva en un movimiento periódico.

1.2 Objetivos y enfoque

El objetivo principal de este trabajo es el estudio, implementación y evaluación de un sistema de obtención del pulso cardíaco a partir de vídeos faciales. No se crearán nuevos algoritmos, sino que se implementará uno de los métodos ya existentes estudiados en el estado del arte, en este caso el método de obtención del pulso cardíaco en función del movimiento de la cabeza, que es una de las técnicas menos conocida y a su vez cada vez más utilizada en la actualidad.

El usuario dispondrá de un sistema que le permitirá obtener el ritmo cardíaco tanto en vídeos grabados como en vídeos en tiempo real. Una vez implementado el algoritmo, se evaluarán los resultados comparándolos con las mediciones reales hechas por los aparatos de medición convencionales.

1.3 Organización de la memoria

El presente documento se divide en siete capítulos:

- **Capítulo 1: Introducción.** Se trata del capítulo actual, en el cual se exponen las razones que han motivado a la realización del trabajo así como los objetivos perseguidos con el mismo.
- **Capítulo 2: Estado del arte.** En primer lugar se explican los fundamentos de la obtención del ritmo cardíaco por análisis de vídeo añadiendo un trasfondo histórico exponiendo los hitos más relevantes hasta el presente. Por último, se detallan las dos técnicas que existen en la actualidad para la obtención del pulso cardíaco y se exponen algunas de las investigaciones que se han hecho tanto de una técnica como de otra.
- **Capítulo 3: El corazón y el electrocardiograma.** En este capítulo se expone la fisiología y anatomía del corazón, pues es el órgano afectado en los pacientes del estudio. Seguidamente se explica el electrocardiograma, el cual aporta la señal de interés para analizar en este trabajo.
- **Capítulo 4: Marco teórico.** Se explica todo el fundamento teórico de las técnicas empleadas en cada una de las fases de la implementación del sistema.
- **Capítulo 5: Implementación del sistema.** En este capítulo se detallan los pasos seguidos para llevar a cabo la implementación del sistema. La explicación se divide según las diferentes fases vistas en el capítulo 4 entre otras. También se expone el entorno de trabajo utilizado para llevar a cabo dicha implementación.
- **Capítulo 6: Experimentos realizados y resultados.** Muestra los experimentos que se llevan a cabo para evaluar el funcionamiento del sistema, contrastando los resultados con mediciones reales hechas por los aparatos de medición convencionales.
- **Capítulo 7: Conclusiones y líneas futuras.** En este capítulo se resume el trabajo realizado en el proyecto extrayendo las conclusiones más importantes y se proponen líneas de trabajo futuras que hagan posible la continuación y mejora del trabajo aquí expuesto.

2 ESTADO DEL ARTE

2.1 Métodos de obtención del ritmo cardíaco por análisis de vídeo

La detección de ritmo cardíaco mediante análisis de vídeo es un campo en el que se lleva trabajando desde el año 2002 [1] con el uso de secuencias obtenidas mediante cámaras térmicas. En los últimos años ha experimentado un mayor crecimiento y han aparecido nuevas líneas de trabajo, como el uso de vídeos obtenidos mediante cámaras convencionales. Aunque ya se trabajaba con secuencias de color desde el año 2007 [2], es a partir de la publicación de un trabajo del Instituto Tecnológico de Massachusetts (MIT) en el año 2012 [3], donde se mostraba que había una variación de color en el rostro producida por el flujo sanguíneo, cuando empieza a producirse este crecimiento.

Durante este estudio veremos que existen diversas técnicas para la obtención del ritmo cardíaco mediante el análisis de vídeo, pudiéndose agrupar en dos grandes tipos: mediante variaciones de color y mediante movimiento. Los métodos basados en variaciones en el color emplean los cambios en la tonalidad de la piel (por ejemplo, de la cara o de un dedo) provocados por la circulación de la sangre y los métodos basados en movimiento aprovechan los movimientos que produce la sangre en la cabeza al pasar por la aorta. Tanto los cambios de color como de movimiento inducidos por el flujo de la sangre son imperceptibles para el ojo humano, pero se han hecho visibles al ojo humano mediante técnicas de amplificación de vídeo [3] aplicadas a imperceptibles cambios de color y movimiento.

2.1.1 Mediante cambios de color

Este método, también llamado fotoplethysmografía, es el más usado en la actualidad. Durante el estudio del estado del arte se han consultado referencias relativas a trabajos relacionados con esta aproximación. Esta técnica se basa en obtener las variaciones de color de los píxeles de la región facial según la cantidad de luz reflejada, debidas a la circulación sanguínea, para posteriormente filtrarlas en la banda apropiada. A través de estas señales, analizadas directamente o procesadas, se obtiene una estimación del ritmo cardíaco. Para ello se busca la frecuencia a la que tenemos la máxima energía en un espacio transformado. En la actualidad existen sistemas comerciales basados en esta tecnología, como por ejemplo, la aplicación Heart Rate Monitor [4] o la aplicación móvil Cardiio [5], como se muestran en la figura 2-1, disponibles para descargar en las principales tiendas de apps.



(a) Cardio App



(b) Heart Rate Monitor App

Figura 2-1. Apps basadas en cambios de color

Estos sistemas se basan en la selección de una región de interés (ROI), normalmente de todo el rostro, aunque ciertos investigadores delimitan la región a la frente, mejillas y labios en busca de unas señales más robustas. Estas regiones de la piel comprenden una mayor proporción de capilares que generarían una mayor potencia de la señal debido a la mayor absorción de luz en comparación con otras regiones de la cara.

Sin embargo, muchos investigadores han seleccionado la región de la frente por tener una menor proporción de capilares debajo de la piel. Esto se debe a que la frente es menos propensa a los movimientos musculares en comparación con otras áreas de la cara. Los movimientos musculares dentro de la cara servirían para debilitar la fuerza de la señal al agregar ruido en forma de artefactos de movimiento. En general, se observa que la región del ojo se excluye de la ROI, ya que no consta de un componente pulsátil relacionado con el cambio de volumen de la sangre y, además, el parpadeo del ojo provocaría artefactos de movimiento que debilitarían la señal. Se ha demostrado que la extracción de la señal con múltiples regiones mejora la potencia general de la señal [6].

Dentro de esta metodología hay distintas variantes para medir los cambios de coloración:

- Separar frames en R, G y B, y normalizarlos, para posteriormente aplicar Análisis de Componentes Independientes (ICA - Independent Component Analysis) o Análisis de Componentes Principales (PCA - Principal Component Analysis) [7] [8] [9] [10].
- Obtención de la media de la intensidad o uso de banda G del rostro para después aplicar análisis espectral auto regresivo [11].
- Mediante el cálculo de las diferencias de brillo [2].
- Uso de descomposición piramidal y análisis de la componente R [12].

También tenemos distintas variantes a la hora de obtener la frecuencia dominante, siendo el cálculo de la Transformada Rápida de Fourier (FFT - Fast Fourier Transform) [7] y el uso de la Densidad Espectral de Potencia (PSD - Power Spectral Density) [8] [9] [10] las dos técnicas más usadas.

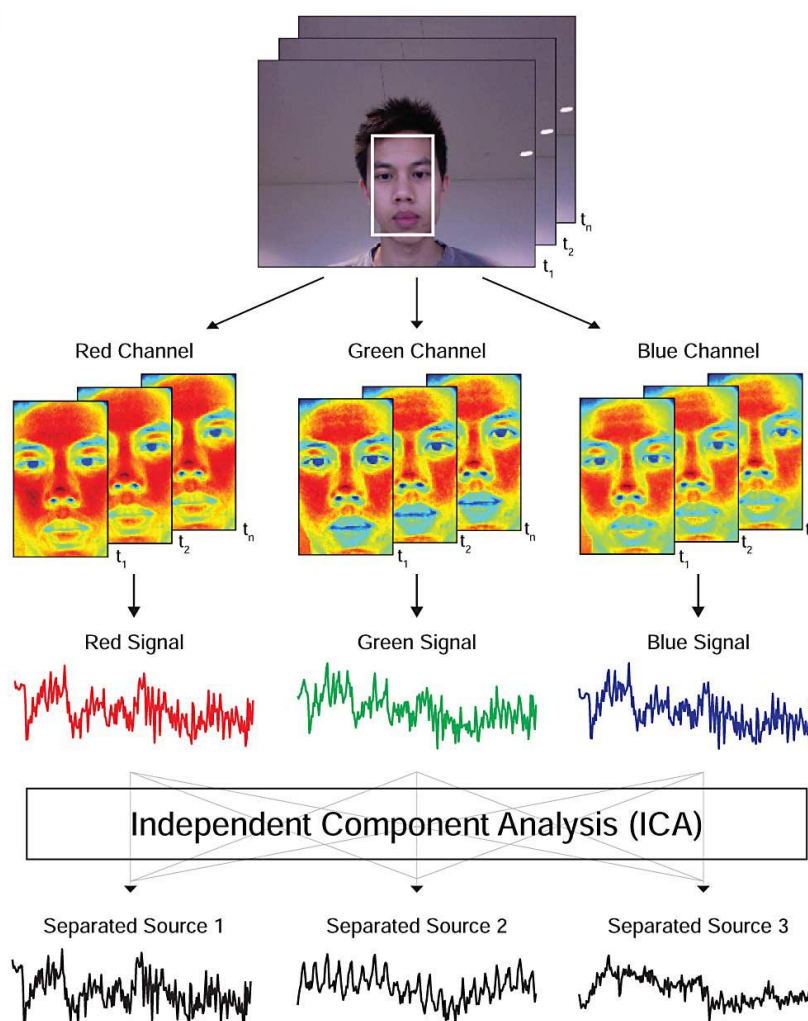


Figura 2-2. Ejemplo de algoritmo basado en cambios de color con separación en bandas de color y uso de ICA

Estos sistemas necesitan que la piel de la zona de interés sea visible para la cámara, ya que el rostro humano no es simétrico por naturaleza, lo que conduce a una dispersión de luz proporcional a la superficie de la piel. La posición de la fuente de luz y la orientación de la cara causan variaciones en la iluminación dando como resultado diferentes intensidades de color en la piel.

En la siguiente tabla se muestra un resumen de los principales estudios de algoritmos basados en cambios de color:

Algoritmo	Detección facial	Cálculo cambio de color	Obtención frecuencia dominante
Poh et al. [7] [8]	Viola Jones	Separación de componentes RGB	ICA + FFT/PSD
Tarassenko et al. [11]	Everingham et al.	Solo banda G	Análisis espectral auto regresivo
Lewandowska et al. [9]	Viola Jones	Separación de componentes RGB	PCA/ICA + PSD/Zero crossing
Aarts et al. [10]	Manual	Separación de componentes RGB	Fast ICA + PSD
Carvalho et al. [12]	Viola Jones	Descomposición piramidal y análisis banda R	Detección de picos
Takano et al. [2]	Manual	Cambios de brillo	Análisis espectral auto regresivo

Tabla 2-1. Resumen de algoritmos basados en cambios de color

2.1.2 Mediante movimientos

La balistocardiografía es un método para medir la frecuencia cardíaca mediante la estimación del movimiento generado por el bombeo de la sangre desde el corazón en cada ciclo cardíaco. Es uno de los muchos métodos que se basan en el movimiento mecánico de los sistemas cardiovasculares, como la apex-cardiografía, quinetocardiografía, fonocardiografía y sismo-cardiografía. La balistocardiografía (BCG) se descubrió inicialmente en el siglo XIX, aunque fue en 1939 cuando Starr elaboró un balistocardiografo (ver Figura 2-3), que inscribía las ondas en papel fotográfico. Se convirtió en el foco de interés en la investigación desde la década de 1940 a la de 1980, después de lo cual el método desapareció.

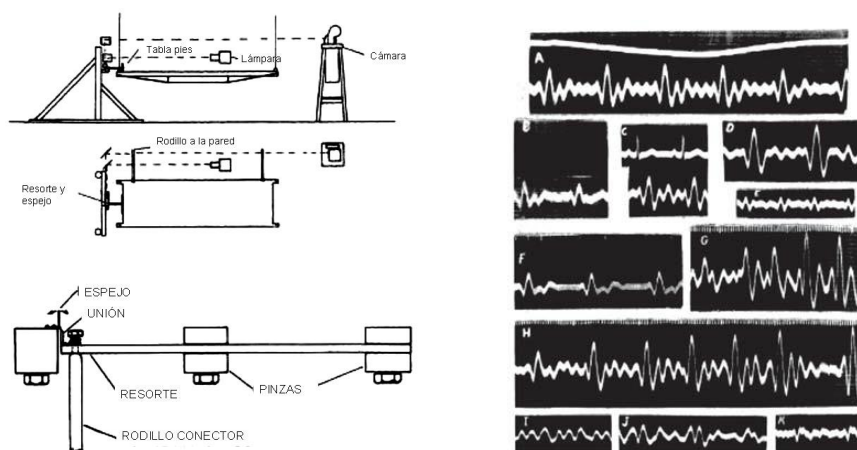


Figura 2-3. Balistocardiografo y balistocardiogramas del Dr. Starr

El balistocardiograma es el registro del movimiento del cuerpo que se origina por la energía transmitida desde el corazón en cada ciclo cardíaco, el ventrículo izquierdo se contrae y expulsa sangre a una presión alta al arco aórtico. Este flujo de sangre en cada ciclo pasa a través de las arterias carótidas en ambos lados del cuello, generando una fuerza en la cabeza (ver Figura 2-4). Por lo tanto, según la Tercera Ley de fuerza de Newton, esta fuerza creada por el flujo de sangre en la cabeza, es igual a la fuerza de la cabeza que actúa sobre el flujo de sangre causando un movimiento reaccionario de cabeza cíclica. Las flechas blancas ilustran esto en la Figura 2-4.



Figura 2-4. Movimiento involuntario debido a la circulación de la sangre desde el corazón a la cabeza

Los métodos de extracción de BCG más típicos implican que el sujeto se acueste en una cama plana o sentado en una silla con un movimiento voluntario e involuntario mínimo [13]. Este movimiento de cabeza es demasiado pequeño para ser observado a simple vista. Sin embargo, el trabajo presentado en [14], mediante el uso de la amplificación de vídeo, demostró que la cabeza se mueve periódicamente al movimiento de la frecuencia cardíaca a una amplitud menor.

En 2013, Balakrishnan et al. [14], propuso un método de estimación de la frecuencia cardíaca utilizando un vídeo facial basado en BCG, el cual implementaremos en el presente trabajo y será explicado más en profundidad en los siguientes capítulos. Este método se centró en extraer el movimiento involuntario de la cabeza en el vídeo. Los puntos característicos de la cabeza se extrajeron con una combinación de regiones sobre la frente y la región debajo de la línea del ojo que incluye las mejillas y la parte superior de los labios. La justificación para la selección de ROI es que están seleccionados y contienen gran parte de los capilares que se ramifican desde las arterias carótidas. La línea ocular se excluyó porque el parpadeo del ojo provocaría artefactos de movimiento en la señal de BCG extraída.

Los autores utilizaron el detector de rostro Viola-Jones [15] para detectar el rostro y extraer la ROI. Los puntos característicos de la ROI fueron rastreados usando el algoritmo de seguimiento de cara de Kanade Lucas Tomasi (KLT) [16]. Los puntos rastreados se filtraron temporalmente usando un filtro de paso de banda que operaba de 0.75 Hz a 2 Hz. La frecuencia operativa se eligió para asemejarse a la frecuencia cardíaca dentro de 45 a 120 bpm. Se usó el análisis de componentes principales (PCA) para eliminar artefactos y extraer la señal periódica como BCG. En consecuencia, el ritmo cardíaco del sujeto se determinó mediante un algoritmo de detección de picos simple.

Shan et al. [17] propuso un método de extracción de BCG basado en [14]. La región de la frente solo se usó para extraer los puntos característicos ya que el área debajo de la línea del ojo (es decir, las mejillas y el labio superior) era mucho más propensa a los artefactos de movimiento, causados por movimientos no rígidos del músculo facial debido a las emociones de hablar o expresar. Además, se utilizaron análisis de componentes independientes (ICA) en lugar de PCA para extraer la señal BCG. Finalmente, la frecuencia cardíaca se mide extrayendo la densidad espectral de potencia (PSD) de la señal BCG. Aunque no hubo mucha diferencia entre los resultados obtenidos con los métodos ICA o PCA, el significado principal del enfoque basado en BCG fue que la metodología en sí misma era invariable a la variación de la iluminación y las variaciones de cada piel.

Posteriormente en 2014, en la Universidad de Aalborg, Irani et al. [18] replicaban el algoritmo realizado en el MIT y proponían modificaciones para obtener mejoras en grabaciones con movimiento voluntario de la cabeza. Los cambios introducidos en este trabajo fueron el uso de un filtro de suavizado después del filtrado de la señal, y el uso de la Transformada Discreta del Coseno.

3 EL CORAZÓN Y EL ELECTROCARDIOGRAMA

Las tecnologías de la información y de la comunicación han penetrado en la ciencia y se han hecho un hueco, llegando a especialidades tan importantes para el ser humano como es la Medicina. Entre las enfermedades más importantes que causan muerte están las cardiovasculares, pasando la cifra de los 15 millones anuales mundiales.

Según la Sociedad Española de Cardiología (SEC), las enfermedades cardiovasculares se dividen en: reumáticas, hipertensivas, cardiopatías isquémicas, insuficiencia cardíaca, cerebro vasculares y arritmias. El electrocardiograma (ECG) [19] es la representación gráfica de la actividad eléctrica del corazón, que se obtiene de un electrocardiógrafo de forma continua.

3.1 Fisiología cardíaca

El corazón es el órgano responsable de bombear la sangre mediante contracciones rítmicas para que ésta sea distribuida por el cuerpo a través de los vasos sanguíneos. Se encuentra situado a la izquierda del tórax, entre los dos pulmones, reposando sobre el diafragma. La sangre contiene los nutrientes que son requeridos por el organismo para realizar las funciones metabólicas necesarias, además de cumplir muchas otras funciones como la regulación de la temperatura corporal o el transporte de los excedentes metabólicos [20].

Respecto a la estructura del corazón, podemos diferenciar tres capas. La más externa es el pericardio, que se encarga de proteger el corazón y permite el movimiento durante la contracción cardíaca. A continuación está el músculo cardíaco denominado miocardio, el cual está formado por células musculares que permiten que el impulso eléctrico se transmita por el corazón. Este impulso es el que provoca la contracción del músculo cardíaco. Finalmente, la capa más interna es el endocardio, que es la que está en contacto con la sangre.

Además, el corazón presenta cuatro cavidades musculares [21], las dos superiores son las aurículas y las dos inferiores los ventrículos. La zona izquierda de la derecha está separada por una membrana denominada tabique. El corazón también presenta cuatro válvulas cardíacas mediante las cuales se asegura el flujo sanguíneo unidireccional. Las situadas entre las aurículas y los ventrículos se denominan válvulas auriculoventriculares, mientras que las que impiden el reflujo de las arterias hacia los ventrículos son las válvulas semilunares, según se muestra en la Figura 3-1.

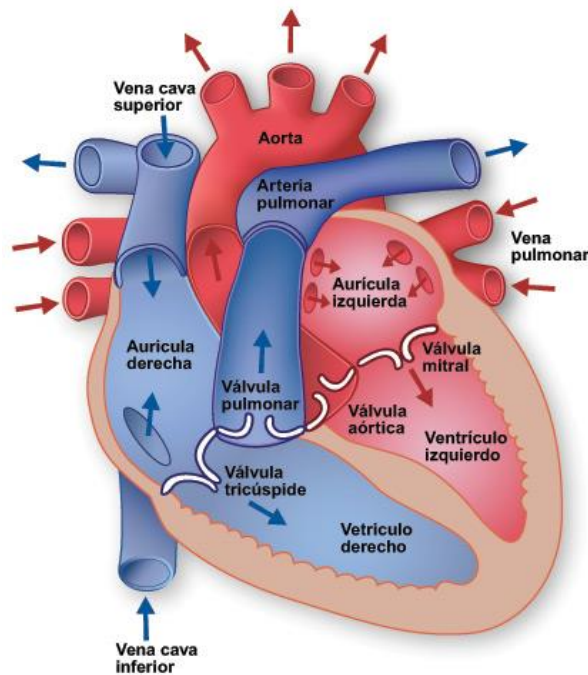


Figura 3-1. Anatomía del corazón y flujo de sangre

Sistema circulatorio

El circuito sanguíneo en el corazón empieza en las aurículas, que son las cámaras que reciben la sangre de las venas y la pasan a los ventrículos a través de las válvulas auriculoventriculares. La válvula que regula el flujo en la zona derecha del corazón es la tricúspide, mientras que la de la zona izquierda es la válvula mitral.

A continuación, los ventrículos se contraen para bombear la sangre hacia las arterias, que son las encargadas de distribuirla por todo el organismo. En este caso, las válvulas pulmonar y aórtica son las que impiden el reflujo de sangre en los ventrículos desde la arteria pulmonar y la aorta respectivamente. Todo este proceso se conoce como ciclo cardíaco.

Además, la circulación está organizada de modo que el lado derecho bombea la sangre a través de los pulmones para oxigenarla y el lado izquierdo la bombea hacia el resto del organismo. Debido a esta necesidad de impulsar la sangre para que ésta proporcione oxígeno a todo el cuerpo, el músculo cardíaco del lado izquierdo requiera más potencia y, por lo tanto, está más desarrollado.

Sistema de conducción

Las contracciones rítmicas del miocardio son las que permiten el bombeo de la sangre y se consiguen mediante la estimulación eléctrica. El estímulo se inicia al variar el potencial de membrana de una célula cardíaca, y esto se conoce como despolarización. Esta variación se produce gracias al transporte de iones de potasio, sodio y calcio a través de la membrana celular y provoca que las células vecinas también se despolaricen transmitiendo así el impulso eléctrico por todas las células miocárdicas.

Este impulso se difunde por rutas especiales que se conocen como el sistema de conducción, el cual está compuesto por células especializadas que presentan un menor umbral de excitabilidad.

Este estímulo se inicia y se regula desde el nodo sinusal (SA), que se encuentra cerca del orificio de entrada de la vena cava a la aurícula derecha. Además, la frecuencia del nodo SA puede ajustarse en función de los estímulos de dos nervios, el vago y el simpático. A pesar de esta regulación, los valores de frecuencia habituales varían de 60 a 100 estímulos por minuto.

Por lo tanto, el impulso eléctrico que proviene del nodo SA estimula las células de las dos aurículas a través de tres conductos y alcanza el nódulo auriculoventricular (AV). En el AV se produce un retardo de la señal eléctrica para que la contracción ventricular sea posterior a la auricular. A continuación, el impulso eléctrico alcanza el haz de His, que es un conducto que ramifica hacia los dos ventrículos. Estas dos ramas, a su vez, se dividen en forma de red en las fibras de Purkinje, que son las encargadas de estimular los ventrículos casi de forma simultánea. Esta estimulación es la que contrae el ventrículo para que la sangre pase a las arterias con la potencia suficiente para que se distribuya por todo el cuerpo.

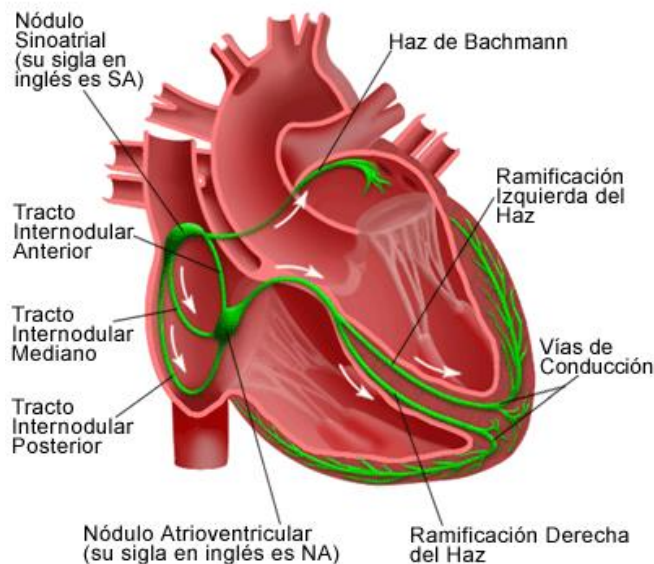


Figura 3-2. Sistema de conducción cardíaco

3.2 El electrocardiograma (ECG)

El electrocardiograma (en adelante ECG), es una prueba no invasiva que arroja como resultado una gráfica a través de la cual se puede obtener información acerca del estado del corazón [22]. El ECG es una herramienta básica para la detección y posterior diagnóstico de enfermedades cardíacas. Además, no supone ningún riesgo para el paciente, ni requiere de preparación previa alguna. Sí que es necesario que, en el momento de la realización, el sujeto que se somete a la prueba esté lo más inmóvil posible, puesto que un leve movimiento puede distorsionar sobremedida el ECG.

En su realización, se pueden colocar hasta 12 parches por el tórax del paciente, obteniéndose así hasta 12 puntos de observación distintos. Hay enfermedades que no pueden ser detectadas desde algunos puntos de observación y sí desde otros.

3.2.1 Componentes del ECG

El ECG está formado por múltiples componentes que se corresponden con las distintas etapas del ciclo cardíaco y que presentan unas características determinadas que permiten diagnosticar anomalías cardíacas. Por lo tanto, podemos comentar los siguientes elementos del ECG que se observan en la figura 3-3 y algunas de sus características morfológicas:

- Onda P: se corresponde con la despolarización de las dos aurículas. La amplitud normal es inferior a 300 μV , y la duración no supera los 120 ms. Además, esta onda normalmente presenta unas características espectrales de baja frecuencia (entre 10 y 15 Hz) y una morfología poco definida que dificulta su estudio en los registros con mucho ruido [22].
- Intervalo PR: tiempo de inactividad eléctrica correspondiente al retraso producido en el nodo AV.
- Complejo QRS: refleja la despolarización de los ventrículos y en las personas sanas dura de 70 a 110 ms. Este complejo está formado por tres ondas independientes y su morfología varía significativamente en función del latido.
- Segmento ST: representa el tiempo que los ventrículos se mantienen despolarizados.

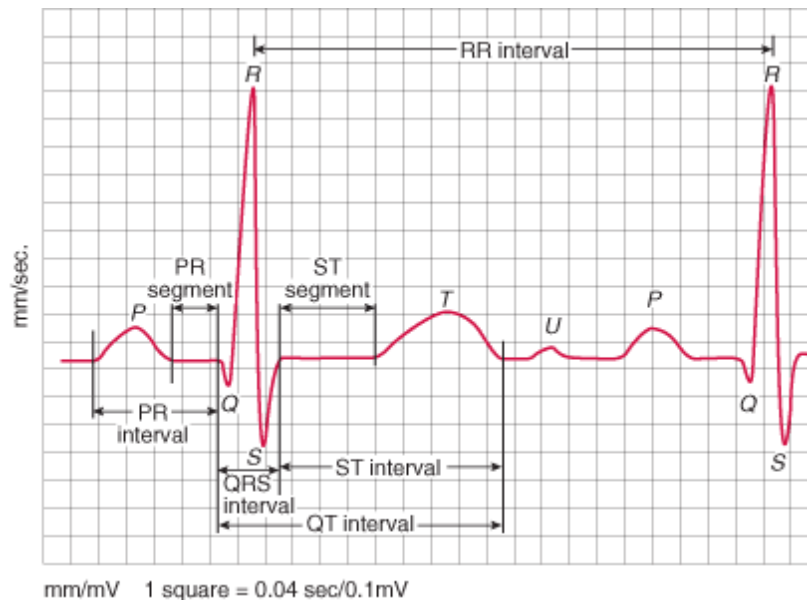


Figura 3-3. Representación de un ECG

- Onda T: se corresponde con la repolarización ventricular. Su duración promedio es de 200 ms.
- Intervalo QT: representa el tiempo desde la despolarización hasta la repolarización ventricular. Este intervalo varía significativamente en función del ritmo cardíaco.
- Intervalo RR: representa el tiempo de un latido cardíaco medido entre dos ondas R sucesivas. Este intervalo es el que se utiliza para estudiar la variabilidad del ritmo cardíaco.

3.2.2 Frecuencia cardíaca

Se conoce como frecuencia cardíaca (o pulso) a la cantidad de veces que late el corazón por minuto y se puede medir en las arterias periféricas. Es uno de los parámetros considerados como signos vitales, junto a la presión arterial, la temperatura corporal y la frecuencia respiratoria. El pulso puede ser medido fácilmente en las muñecas, el cuello, parte posterior de las rodillas, la sien, etc, aunque cada vez existen dispositivos más innovadores para realizar la medición [23].

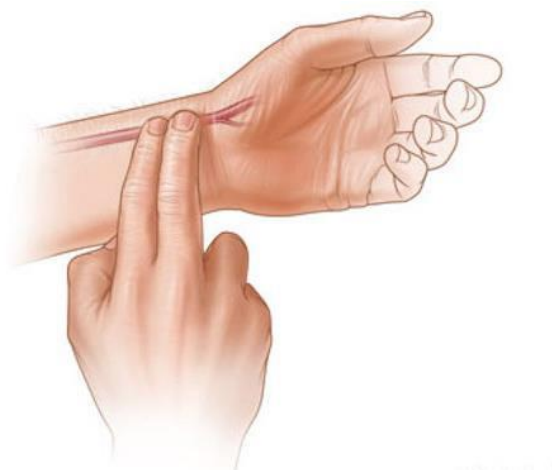


Figura 3-4. Ejemplo de medición del pulso en la muñeca

Para calcular la frecuencia cardíaca se suele hacer en reposo, ya que esta es fácilmente alterada por el sueño, el ejercicio, el estrés y otras situaciones. En las siguientes tablas se recogen los valores de frecuencia cardíaca considerados normales para personas adultas en función del sexo y la edad (siempre teniendo en cuenta situación de reposo) [24].

Para mujeres:

Edad	Inadecuado (PPM)	Normal (PPM)	Bueno (PPM)	Excelente (PPM)
20-29	96 o más	78-94	72-76	70 o menos
30-39	98 o más	80-96	72-78	70 o menos
40-49	100 o más	80-98	74-78	72 o menos
50 o más	104 o más	84-102	76-82	74 o menos

Tabla 3-1. Pulsaciones normales en reposo en mujeres.

Para hombres:

Edad	Inadecuado (PPM)	Normal (PPM)	Bueno (PPM)	Excelente (PPM)
20-29	86 o más	70-84	62-68	60 o menos
30-39	86 o más	72-84	64-70	62 o menos
40-49	90 o más	74-88	66-72	64 o menos
50 o más	90 o más	76-88	68-74	66 o menos

Tabla 3-2. Pulsaciones normales en reposo en hombres.

3.2.3 Principales trastornos del ritmo cardíaco

Los resultados anormales en un ECG normalmente son sinónimo de diversos tipos de patologías cardíacas. La alteración de las distintas ondas que conforman la gráfica puede llevar a unos tipos u otros de enfermedades del corazón. Por ejemplo, si nos encontramos con una onda T negativa puede ser sinónimo de enfermedad. Una onda Q demasiado ancha puede indicar un infarto de miocardio, y las irregularidades presentadas en el complejo QRS pueden ser indicios de anomalías como taquicardia ventricular, bloqueo de rama, hipertrofia ventricular u otras similares.

Las principales afecciones cardíacas que se dan en la sociedad actual y que pueden afectar al ritmo cardíaco son [25]:

- **Taquicardia:** Se conoce como taquicardia un tipo frecuente de trastornos en el ritmo cardíaco (arritmia), que produce un aumento de las pulsaciones aun en estado de reposo debido a la presencia de erróneas señales eléctricas que aceleran el latido del corazón. Además durante una taquicardia, el corazón puede latir más rápido en las cavidades superiores, inferiores o en ambas [26].
- **Braquicardias o bradiarritmias:** Al contrario que la taquicardia, la braquicardia se caracteriza por un descenso de frecuencia cardíaca normal en estado de reposo. En algunos casos puede suponer un grave problema si el corazón no bombea suficiente sangre rica en oxígeno al organismo, pudiendo llegar a ser necesario un implante de un marcapasos para corregir esta anomalía y que el corazón mantenga un ritmo adecuado.
- **Arritmias supraventriculares:** Como su nombre indican, son arritmias (modificaciones del ritmo normal) con frecuencia superiores a 100 bpm, que tienen su origen en las aurículas o en el nodo auriculoventricular.
- **Arritmias ventriculares:** Similares a las anteriormente expuestas, modificaciones del ritmo cardíaco que, en este caso concreto, tienen su origen en los ventrículos.

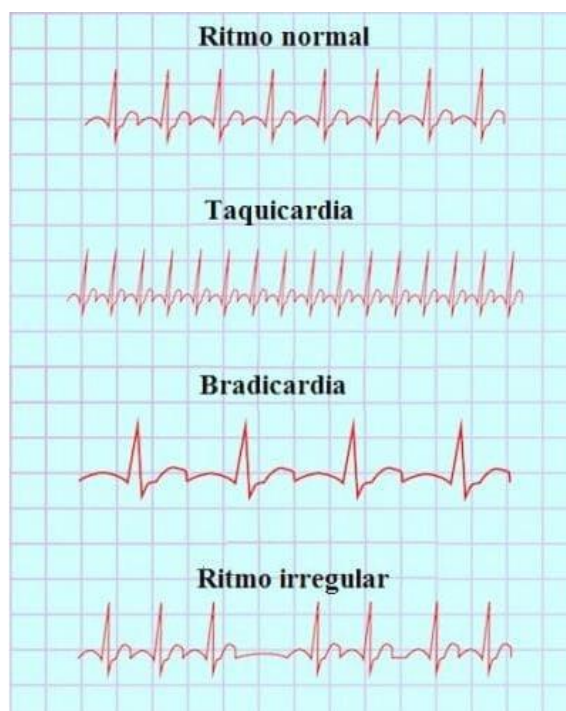


Figura 3-5. ECG de distintas afecciones cardíacas

4 MARCO TEÓRICO

Tras haber llevado a cabo la fase de búsqueda de información y técnicas empleadas para la estimación del ritmo cardíaco, es necesario decidir qué técnica va a desarrollarse. Se decide implementar la técnica que utiliza el movimiento de la cabeza para la estimación.

El presente capítulo presenta el trasfondo teórico de los algoritmos que se utilizarán más adelante en la implementación del sistema, pues de esta forma cuando sean utilizados en la implementación se comprenderá mucho mejor lo que aportará cada uno de ellos al sistema y el porqué de su elección.

4.1 Detección facial

La fase de detección consiste en encontrar áreas de la imagen que contengan una cara para aislarlas del resto. Este es uno de los objetivos del trabajo, y una de las fases más importante de la aplicación ya que una mala detección conllevaría un error en el resto de las fases.

En la actualidad, existen distintos algoritmos que pueden ser implementados para una detección facial, muchos de ellos un tanto complejos o específicos según el área donde se desean emplear. Estos pueden llegar a tratarse de redes neuronales, grafos, entre otros. Sin embargo, en este trabajo se va a comprender uno de los métodos más populares y eficientes que permiten detectar objetos en tiempo real.

4.1.1 Algoritmo de Viola-Jones

El algoritmo de Viola-Jones [15] es un método de detección de objetos que destaca por su bajo coste computacional, lo que permite que sea empleado en tiempo real. Su desarrollo fue motivado por el problema de la detección de caras, donde sigue siendo ampliamente utilizado, pero puede aplicarse a otras clases de objetos que, como las caras, estén caracterizados por patrones típicos de iluminación.

El algoritmo se basa en una serie de clasificadores débiles denominados *Haar-like features* que se pueden calcular eficientemente a partir de una imagen integral. Estos clasificadores, que por sí mismos tienen una probabilidad de acertar solo ligeramente superior a la del azar, se agrupan en una cascada empleando un algoritmo de aprendizaje basado en AdaBoost para conseguir un alto rendimiento en la detección así como una alta capacidad discriminativa en las primeras etapas.

Haar-like features

Las *Haar-like features* son los elementos básicos con los que se realiza la detección. Reciben este nombre por similitud a los *wavelets de Haar* [27]. Estos clasificadores son características muy simples que se buscan en las imágenes y que consisten en la diferencia de intensidades luminosas entre regiones rectangulares adyacentes. Las características quedan por tanto definidas por unos rectángulos y su posición relativa a la ventana de búsqueda y adquieren un valor numérico resultado de la comparación que evalúan.

En el trabajo presentado por Viola-Jones existen tres tipos de características, representadas en la Figura 4-1. Este algoritmo considera regiones rectangulares en una ventana de detección, suma las intensidades de los píxeles en cada región y calcula la diferencia entre estas sumas. En la Figura 4-1 pueden verse algunos ejemplos de características *Haar* comunes de un detector de caras.

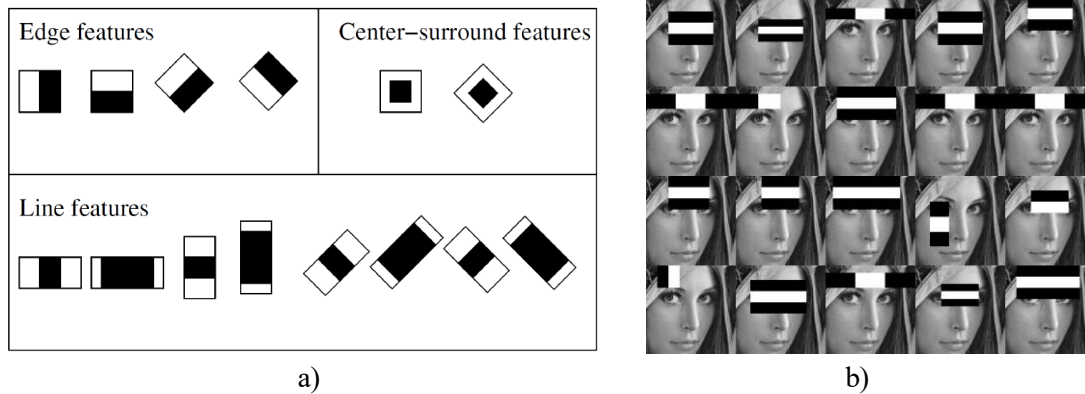


Figura 4-1. a) Características tipo *Haar*. b) Ejemplos de características *Haar* comunes de un detector de cara

Imagen Integral

La suma de los píxeles de un rectángulo puede ser calculada de manera muy eficiente empleando una representación intermedia denominada imagen integral. La imagen integral en un punto contiene la suma de todos los píxeles que están arriba y hacia la izquierda de ese punto en la imagen original. La imagen integral total se puede calcular fácilmente en un solo barrido de la imagen.

Proceso de aprendizaje

Es necesario realizar un proceso de entrenamiento supervisado para crear el clasificador mediante una gran cantidad de imágenes positivas (caras) e imágenes negativas (imágenes sin caras). Este proceso se realiza mediante un algoritmo basado en AdaBoost [28], un algoritmo adaptativo de machine learning cuyo nombre es una abreviatura de adaptive boosting.

Cascada de clasificadores

En vez de construir un único clasificador mediante el proceso anterior, se construyen clasificadores más pequeños y eficientes. Estos clasificadores más simples se utilizan para rechazar la mayoría de las ventanas de búsqueda y solo en aquellas en las que hay mayores probabilidades de encontrar caras se llama a clasificadores más complejos. Se obtiene así una cascada de clasificadores, cada uno de los cuales es entrenado con AdaBoost.

Detección de caras

Empleando el algoritmo de Viola-Jones explicado en el apartado anterior y mediante un entrenamiento correcto, se hace posible un detector de caras robusto basado en un clasificador en cascada como se muestra en la Figura 4-2.

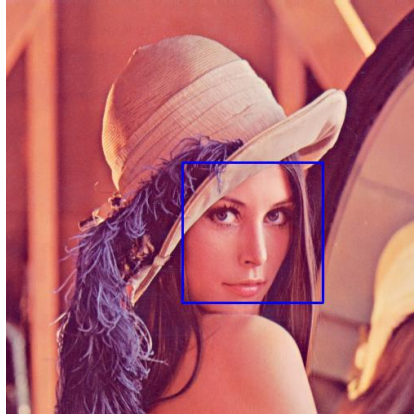


Figura 4-2. Detección de la región de la cara

4.2 Estimación del flujo óptico

El flujo óptico se entiende como el patrón aparente de movimiento existente en los objetos de una escena, provocado por el cambio de intensidad de luz y que es dependiente de la distancia que existe entre el observador y la escena [29]. Se le denomina “aparente” ya que puede diferir del movimiento real que realizaban los objetos a la hora de captar la escena. Un ejemplo clásico de este fenómeno es el del *barber pole*, que se puede ver en la Figura 4-3. En general, esta diferencia se debe al hecho de que la cámara esté captando la escena desde un punto de vista en particular, conjugado con otros factores como rotaciones y cambios de iluminación.

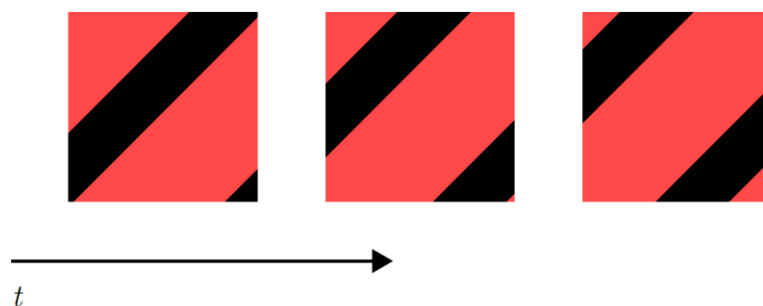


Figura 4-3. Efecto *barber pole*: ¿En qué sentido se mueven las líneas negras? Su movimiento aparente es hacia la izquierda y arriba, pero en realidad podrían estar moviéndose de infinitas otras maneras.

Entonces podemos decir que el flujo óptico estudia el cambio de intensidades de los píxeles que conforman una secuencia de imágenes, obteniendo de este análisis las velocidades y el sentido con las que los objetos se mueven dentro de un vídeo.

La mayoría de las técnicas existentes para la estimación del flujo óptico se pueden clasificar en 4 categorías [30]: las basadas en gradientes espacio-temporales (diferenciales), las basadas en comparación de regiones, las basadas en fase y las basadas en energía.

Además, los algoritmos de flujo óptico dentro de los métodos diferenciales, también se clasifican en locales y globales, categorizados de esta manera por su diferente forma de ser calculados, donde el método local calcula el flujo en la vecindad del punto de interés [31], mientras que el método global calcula el movimiento en toda la imagen. Cada uno de ellos tiene ventajas específicas sobre el otro, por ejemplo, los métodos locales son más robustos ante el ruido, y los métodos globales producen campos de flujo densos [31]. De la clasificación anterior, son el de Lucas-Kanade [32] (método local) y el de Horn-Schunck [33] (método global), los más usados y mayormente estudiados y analizados por la eficiencia de los algoritmos.

En todas las estrategias de estimación de flujo óptico se parte de la hipótesis de que los niveles de gris permanecen constantes ante movimientos espaciales en un tiempo dado y que el objeto en la escena es plano, de manera tal que al moverse su brillo aparente no cambia. Dicha hipótesis es conocida como la *condición de brillo constante* y da lugar a la ecuación general de flujo óptico:

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t) \quad (4.1)$$

Supongamos que un pixel en la posición (x, y) se movió Δx y Δy en un espacio de tiempo Δt . Asumiendo un movimiento pequeño podemos realizar la siguiente aproximación por una serie de Taylor de primer orden:

$$I(x + \Delta x, y + \Delta y, t + \Delta t) \approx I(x, y, t) + \frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t \quad (4.2)$$

Combinando las ecuaciones (4.1) y (4.2) llegamos a

$$\frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t = 0 \quad (4.3)$$

y por último dividimos por Δt , obteniendo

$$\frac{\partial I}{\partial x} \frac{\Delta x}{\Delta t} + \frac{\partial I}{\partial y} \frac{\Delta y}{\Delta t} + \frac{\partial I}{\partial t} = 0 \quad (4.4)$$

Las derivadas parciales $\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial I}{\partial t}$ son los respectivos componentes de la imagen gradiente en el punto (x, y) en tiempo t y las podemos escribir como I_x, I_y e I_t . Notar que las dos primeras tienen información de los alrededores del punto (x, y) en el tiempo t mientras que la última tiene información de cómo cambia el brillo de los pixeles que pasan por la posición fija (x, y) a lo largo del tiempo. Los factores $\frac{\Delta x}{\Delta t}$ y $\frac{\Delta y}{\Delta t}$ son los componentes de la velocidad $\mathbf{v} = (v_x, v_y)^T$ del píxel, los cuales se convierten en nuestras incógnitas:

$$I_x v_x + I_y v_y = -I_t \quad (4.5)$$

Llegamos así a una ecuación lineal con dos incógnitas, por lo que el sistema está subespecificado. Esto es conocido como el problema de apertura y es lo que obliga a poner más condiciones sobre cómo se modela el flujo óptico para poder llegar a una solución. Para entender mejor el problema de apertura, reescribamos nuestra ecuación usando los vectores gradiente y velocidad:

$$\nabla I \cdot \mathbf{v} = \nabla I^T \mathbf{v} = -I_t \quad (4.6)$$

Esta ecuación nos dice que el componente de la velocidad en la dirección del gradiente es directamente proporcional al cambio de brillo en el tiempo. Visto de otra manera, estamos aproximando la apariencia local de la imagen de manera lineal conforme al sentido, dirección y magnitud del gradiente y al movimiento local como una traslación de esta aproximación de forma tal que el brillo en el punto cambia proporcionalmente a $-I_t$. Una representación visual de esto se puede ver en la Figura 4-4.

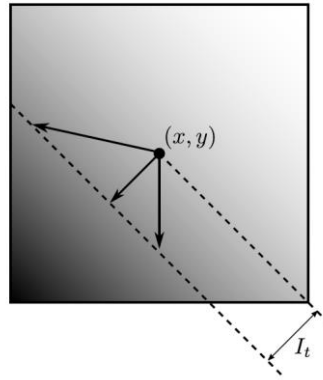


Figura 4-4. Problema de apertura: En esta aproximación, el cambio de brillo es lineal según dirección y magnitud de $\nabla I(x, y)$. Debido a esto, existen infinitos vectores $\mathbf{v}$ que son solución de (4.6)

Sin embargo, para un cambio de brillo dado, existen infinitas traslaciones que lo causan, correspondientes a la línea dentro la aproximación a la imagen que une todos los puntos con el mismo valor de brillo deseado. Por lo tanto, el grado de libertad que tenemos para resolver la ecuación puede interpretarse como la componente de la velocidad en la dirección de tal línea. Este problema es análogo al ejemplo del barber pole de la Figura 4-1. El hueco o “apertura” a través del cual vemos el barber pole no nos deja ver todo el contexto necesario para dilucidar de manera precisa el movimiento del objeto.

De la misma manera, la falta información en las ecuaciones de flujo óptico no permite desambiguar la forma exacta del vector $\mathbf{v}$: podría estar apuntando en dirección del gradiente o de muchas otras maneras. El problema de apertura es evidencia de que no podemos trackear un solo punto en la imagen. Una forma de lidiar con esto es introduciendo más información a la especificación del problema.

4.2.1 Algoritmo de Lucas-Kanade

Una de las opciones para la obtención del flujo óptico es el llamado método de Lucas Kanade con el cual se puede calcular una aproximación bastante buena. En la figura 4-5 podemos apreciar, de manera general, las operaciones que hay que llevar a cabo para el cálculo del campo de velocidades. De una secuencia de imágenes se seleccionan dos consecutivas a las cuales se les realiza un tratamiento visual y una serie de operaciones propias del método LK para obtener el resultado final y mostrarlo si fuese necesario.

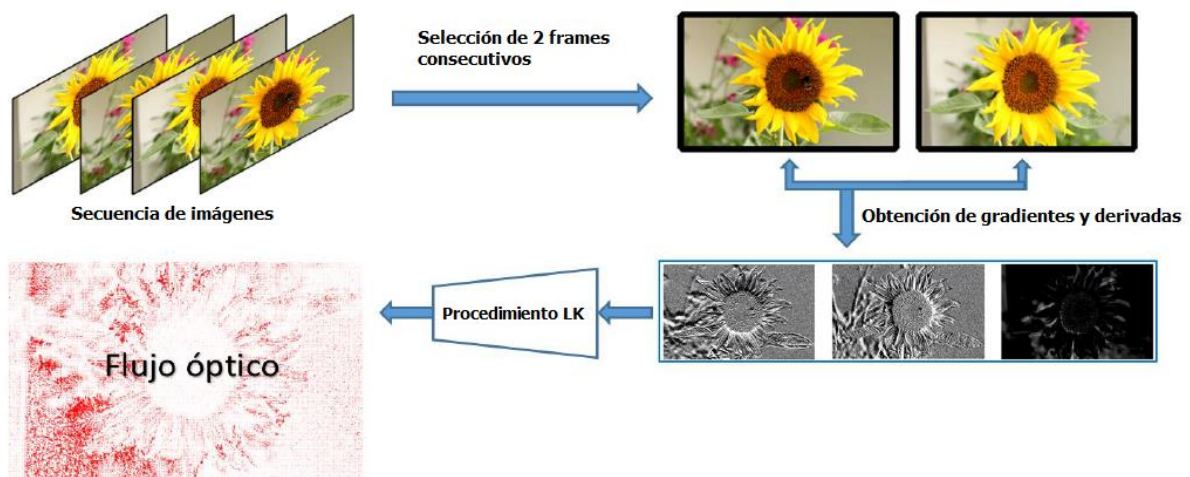


Figura 4-5. Esquema descriptivo de la obtención del flujo óptico

Bruce D. Lucas y Takeo Kanade [34, 35] implementaron, en 1981, una aproximación de mínimos cuadrados, para resolver el problema de restricción que presenta la ecuación del flujo óptico. El algoritmo de Lucas-Kanade extiende las condiciones de flujo óptico discutidas en la sección anterior a todos los píxeles dentro de una ventana rectangular centrada en un punto. Este recorte de la imagen original es también conocido como un *parche* de la imagen. Se obtiene así un sistema de ecuaciones de flujo óptico como el siguiente, donde $q_1, q_2, \dots, q_n$ son los puntos en el parche y donde extendemos I_x, I_y e I_t a todos ellos:

$$\begin{aligned} I_x(q_1)v_x + I_y(q_1)v_y &= -I_t(q_1) \\ I_x(q_2)v_x + I_y(q_2)v_y &= -I_t(q_2) \\ &\vdots \\ I_x(q_n)v_x + I_y(q_n)v_y &= -I_t(q_n) \end{aligned} \quad (4.7)$$

q_1	q_2	q_3
	q_5	
		q_9

Figura 4-6. Vecindad alrededor del elemento q_5 de tamaño 3×3

La idea detrás de esta extensión es que los píxeles del parche, por estar en la cercanía del punto central, tienen un vector velocidad aproximadamente similar y conservan la intensidad de la imagen. Es deseable que la ventana sea suficientemente grande como para que las gradientes de la imagen dentro del parche tengan información suficiente para poder resolver estas ecuaciones. Por el contrario, mientras más grande sea el parche contendrá píxeles más alejados del centro, por lo cual la aproximación lineal de la velocidad será peor e introducirá más error en los cálculos, disminuyendo la precisión de nuestra solución.

El método propuesto por Lucas Kanade consiste en resolver las ecuaciones anteriores con el fin de obtener el desplazamiento o velocidad $\mathbf{v}$. Estas ecuaciones se pueden escribir en la forma matricial $\mathbf{A}\mathbf{v} = \mathbf{b}$, donde:

$$\mathbf{A} = \begin{bmatrix} I_x(q_1) & I_y(q_1) \\ I_x(q_2) & I_y(q_2) \\ \vdots & \vdots \\ I_x(q_n) & I_y(q_n) \end{bmatrix} \quad \mathbf{v} = \begin{bmatrix} v_x \\ v_y \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} -I_t(q_1) \\ -I_t(q_2) \\ \vdots \\ -I_t(q_n) \end{bmatrix} \quad (4.8)$$

Puesto que el sistema presentado está sobredeterminado la solución buscada será una solución de compromiso mediante el método de mínimos cuadrados. Anticipándonos a una futura implementación del algoritmo se tiende a buscar expresiones susceptibles de ser programadas de manera sencilla, para ello:

$$\mathbf{A}^T \mathbf{A} \mathbf{v} = \mathbf{A}^T \mathbf{b} \rightarrow \mathbf{v} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} \quad (4.9)$$

Resolviendo el sistema de ecuaciones anterior se obtiene:

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} \sum_i I_x(q_i)^2 & \sum_i I_x(q_i)I_y(q_i) \\ \sum_i I_x(q_i)I_y(q_i) & \sum_i I_y(q_i)^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum_i I_x(q_i)I_t(q_i) \\ -\sum_i I_y(q_i)I_t(q_i) \end{bmatrix} \quad (4.10)$$

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \frac{1}{\sum_i I_{xi}^2 \sum_i I_{yi}^2 - (\sum_i I_{xi} I_{yi})^2} \begin{bmatrix} \sum_i I_{xi}^2 & \sum_i I_{xi} I_{yi} \\ \sum_i I_{xi} I_{yi} & \sum_i I_{yi}^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum_i I_{xi} I_{ti} \\ -\sum_i I_{yi} I_{ti} \end{bmatrix} \quad (4.11)$$

Realizando las operaciones pertinentes sobre la ecuación anterior llegamos al resultado final que puede ser implementado fácilmente en cualquier lenguaje de programación.

$$v_x = \frac{-\sum_i I_{xi}^2 \sum_i I_{xi} I_{ti} + \sum_i I_{xi} I_{yi} \sum_i I_{yi} I_{ti}}{\sum_i I_{xi}^2 \sum_i I_{yi}^2 - (\sum_i I_{xi} I_{yi})^2} \quad (4.12)$$

$$v_y = \frac{\sum_i I_{xi} I_{ti} \sum_i I_{xi} I_{yi} - \sum_i I_{xi}^2 \sum_i I_{yi} I_{ti}}{\sum_i I_{xi}^2 \sum_i I_{yi}^2 - (\sum_i I_{xi} I_{yi})^2}$$

Debido a que este método está diseñado para calcular el flujo óptico cuando se presentan pequeños desplazamientos de los píxeles entre dos cuadros, se puede recurrir a imágenes piramidales, es decir, podemos encontrar una aproximación del flujo óptico de la imagen original aplicando el algoritmo al mismo cuadro con menor resolución. Cada vez que le quitamos resolución a la imagen original, generamos un nivel de la pirámide, que además de optimizar el cálculo del flujo óptico, calcula desplazamientos de píxeles más grandes (Figura 4-7).

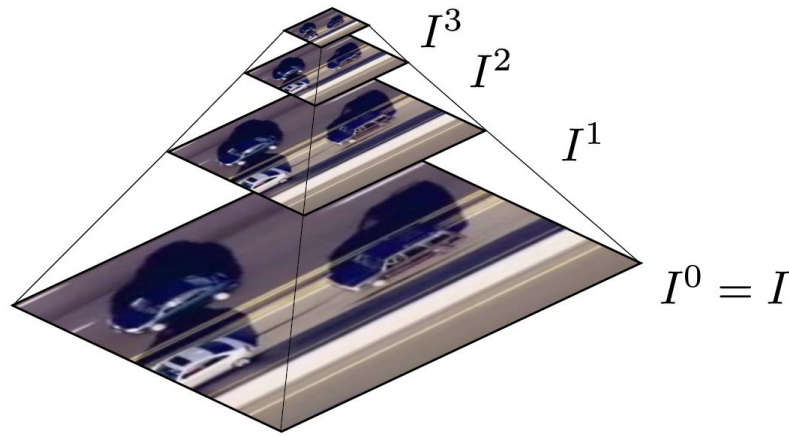


Figura 4-7. Pirámide de resolución: cada nivel tiene la mitad de resolución que el nivel inferior.

4.3 Reducción de la dimensionalidad

En la actualidad, el creciente volumen de información generado por sistemas de información y comunicación demanda nuevas técnicas de manipulación de datos con el objetivo de extraer información no trivial que reside, de manera implícita, para facilitar la obtención de patrones y su análisis.

Los métodos de reducción de dimensionalidad (MDR) son algoritmos que mapean el conjunto de los datos a subespacios derivados del espacio original, de menor dimensión, que permiten hacer una descripción de los datos a un menor costo [36]. Es decir, transforman un conjunto de datos $\mathbf{X}$ con dimensionalidad m a un nuevo conjunto $\mathbf{Y}$ con dimensionalidad m' ($m' < m$), de forma tal que conserve la mayor información intrínseca posible. Como resultado de los MRD, se favorece la compresión, eliminación de redundancia del conjunto de datos y permite mejorar procesos de clasificación y visualización de los datos a un menor costo computacional. Por su importancia, son ampliamente usados en procesos asociados a aprendizaje de máquina.

Uno de los métodos de reducción de dimensionalidad más popular hoy día es el Análisis de Componentes Principales (PCA) [37], técnica lineal que se utiliza para la eliminación de la redundancia de los datos.

4.3.1 Análisis de Componentes Principales (PCA)

PCA permite un cambio de base a una de menor dimensionalidad sobre $\mathbf{X}$ a través de la ecuación de transformación $\mathbf{Y} = \mathbf{P}\mathbf{X}$, donde $\mathbf{P}$ es una matriz ortogonal denominada matriz de representación. El objetivo es determinar la matriz $\mathbf{P}$ que permita que la nube de datos pueda ser proyectada a un espacio de menor dimensión. La estrategia es buscar $\mathbf{P}$ de forma que se garantice la no correlación entre vectores de $\mathbf{Y}$, ya que si la correlación entre las distintas muestras es nula, se elimina la redundancia y el subespacio de datos puede ser descrito por $\mathbf{P}$.

El algoritmo para hallar la matriz de representación $\mathbf{P}$ es el siguiente:

1. Se inicia con el centrado y estandarizado de los datos $\mathbf{X}$.
2. Luego, se calcula la matriz de covarianza de $\mathbf{X}$ (4.13), que es simétrica y diagonalizable, y que cuantifica la covarianza entre las mediciones.

$$\mathbf{C}_X = \frac{1}{n} \mathbf{X}\mathbf{X}^T \quad (4.13)$$

3. Una vez calculada la matriz de covarianza, el siguiente paso consiste en obtener sus autovectores, que serán las columnas de la matriz $\mathbf{P}$ buscada y que servirán de nuevas coordenadas del sistema donde es maximizada la varianza.

Finalmente, con el objetivo de reducir la dimensionalidad de los datos, no todos los autovectores deberán formar la matriz $\mathbf{P}$, sino solo aquellos con mayores autovalores (denominados *componentes principales*, PC), pues serán los que contengan la mayor cantidad de información (varianza) posible de los datos originales.

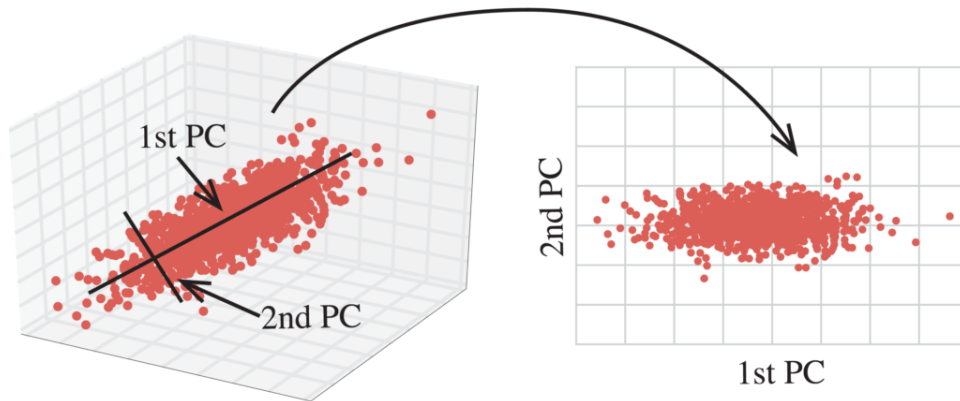


Figura 4-8. Ejemplo de reducción de dimensionalidad aplicando PCA

Hay que tener algo en cuenta en el anterior algoritmo, y es que si el problema inicial era la gran dimensionalidad de los datos $\mathbf{X}$, el cálculo de la matriz de covarianza (4.13) será muy costoso. Es por ello que una de las soluciones más recurrentes es aplicar una descomposición en valores simples (SVD, Singular Value Decomposition) [38] a dichos datos (menos costoso), con lo que se consiguen calcular los autovectores de su matriz de covarianza de forma indirecta. El método consiste en descomponer $\mathbf{X}$ en un producto de tres matrices (4.14). La propiedad que nos es útil de la SVD es que la matriz $\mathbf{V}$ contendrá los autovectores buscados de la matriz $\mathbf{X}\mathbf{X}'$.

$$\mathbf{X} = \mathbf{U} \cdot \mathbf{S} \cdot \mathbf{V} \quad (4.14)$$

5 IMPLEMENTACIÓN DEL SISTEMA

5.1 Entorno de desarrollo

5.1.1 Lenguaje de programación

El lenguaje de programación elegido para la implementación es Python [39], ya que es un lenguaje de programación orientado a objetos muy utilizado en el desarrollo de algoritmos de inteligencia artificial, por lo que existen diferentes librerías y herramientas que pueden ser muy útiles para conseguir los objetivos de una manera más sencilla. La versión utilizada es la 3.6, a pesar de que la última versión es la 3.7. Se ha escogido esta versión al ser la más estable para trabajar actualmente.

5.1.2 Entorno de trabajo

El entorno de trabajo utilizado es Spyder, un entorno de programación para Python que incluye la distribución multiplataforma Anaconda desarrollada por Continuum Analytics [40]. Esta distribución se caracteriza por ser libre y por incluir una gran colección de paquetes y librerías para análisis de datos, computación científica e ingeniería. La versión utilizada es Anaconda 5.1.0, la última versión estable en la fecha del comienzo del trabajo.

5.1.3 Librerías y extensiones

Para el desarrollo de la aplicación se hace uso de diferentes librerías que facilitarán la implementación de los algoritmos descritos. Las principales son las siguientes:

- **OpenCV** [41] (versión 3.4.1) es una librería Open Source originalmente desarrollada por Intel para procesamiento de imágenes y visión por computador. Dispone de más de 500 funcionalidades, entre las que destacan la detección y el reconocimiento de objetos y algoritmos del cálculo del flujo óptico como el de Lucas-Kanade.
- **Numpy** [42] (versión 1.14) librería también Open Source que ofrece un gran conjunto de funciones matemáticas. Destaca también por su utilidad a la hora de operar con vectores y matrices.
- **Scikit-learn** [43] (versión 0.19.1) librería muy utilizada en aprendizaje automático (Machine Learning) que incluye varios algoritmos de clasificación como por ejemplo el análisis de componentes principales (PCA) que será utilizado en este trabajo.
- **SciPy** [44] (versión 1.0.0) librería compuesta de herramientas y algoritmos matemáticos tales como optimización, álgebra lineal, interpolación, FFT y procesamiento de señales e imágenes.

5.2 Algoritmo propuesto

El algoritmo propuesto es una reproducción lo más fiel posible del algoritmo desarrollado por Balakrishnan en el MIT [14], aunque puede haber partes que no estén implementadas de igual forma por no describir el artículo suficientes detalles de cómo se implementaban.

Como se ha visto en el estado del arte, este algoritmo se basa en la obtención del pulso mediante el movimiento inducido por el flujo sanguíneo en la cabeza (BCG). Para ello, se ha seguido el esquema de la siguiente figura:

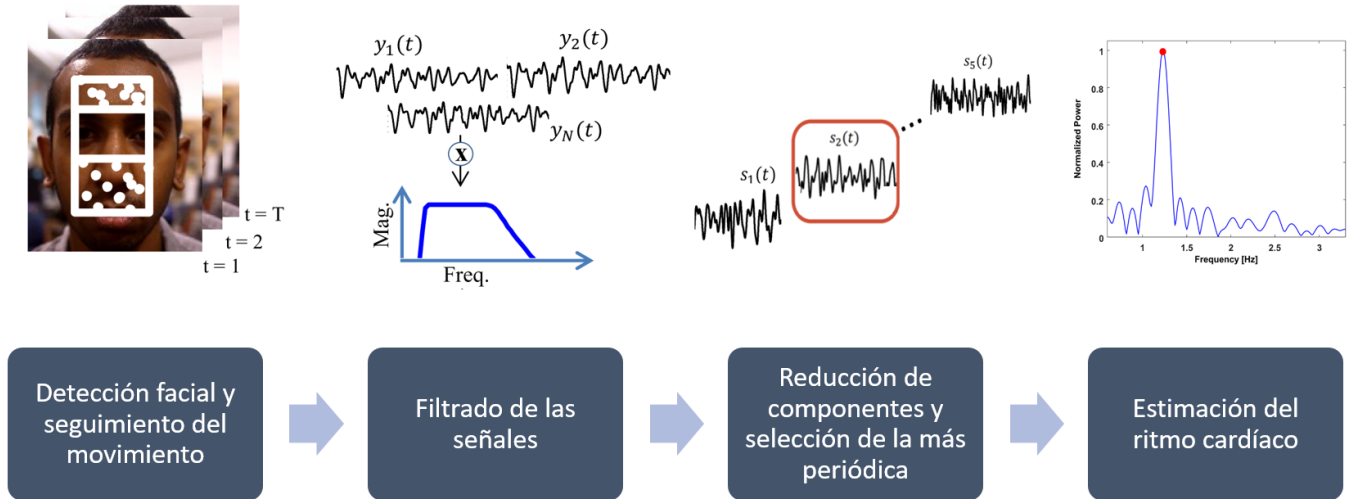


Figura 5-1. Diagrama de las fases del algoritmo propuesto

A continuación se irá detallando cómo se han implementado cada una de las fases del esquema anterior.

5.2.1 Detección facial y selección de ROI

Para la detección de rostros y ojos se han utilizado los clasificadores *Haar-like features* bajo el algoritmo de Viola-Jones [15], cuyo funcionamiento se explicó en el capítulo 4. La ventaja de este método es su robustez y bajo coste computacional, lo que permite la detección de rostros en tiempo real.

La implementación se ha hecho con la librería OpenCV [45], que incluye funciones tanto para entrenar clasificadores en cascada como para detectar objetos. Esta librería facilita además una serie de clasificadores ya entrenados, entre los que se incluyen clasificadores entrenados con rostros y ojos, por lo que se han aprovechado dichos clasificadores evitando un previo entrenamiento.

El procedimiento consiste primero en cargar de la librería el clasificador *Haar* del objeto que se quiere detectar, en nuestro caso el de detección frontal de rostros `haarcascade_frontalface`. Posteriormente, utilizando la función `detectMultiScale`, es posible variar el tamaño de la ventana mínima de detección (`minNeighbors`) e imponer un escalado a la propia imagen (`scaleFactor`) para reducir los tiempos de cómputo. En la siguiente figura se muestra un ejemplo de detección facial usando esta cascada.

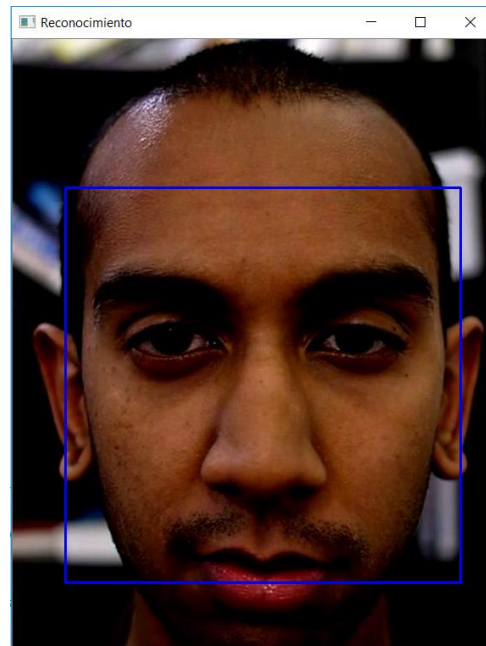


Figura 5-2. Detección facial inicial

Sin embargo, como ya se explicó en el estado del arte, a partir de esta detección se extrae un par de regiones de interés (ROIs) sobre la frente y la zona de las mejillas y el labio superior. La justificación para estas ROIs es que las áreas seleccionadas contienen gran parte de los capilares que se ramifican desde las arterias carótidas, excluyendo así los artefactos de movimiento que añadirían el parpadeo de los ojos a nuestra estimación del pulso cardíaco. En la siguiente figura se muestra la selección de las ROIs:

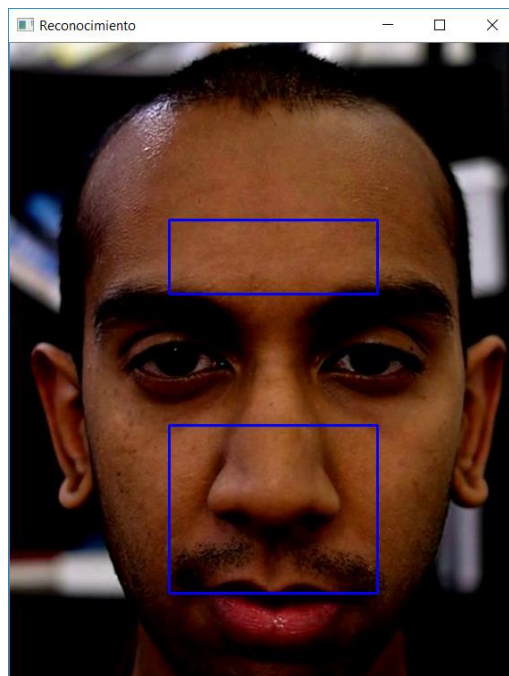


Figura 5-3. Selección final de las ROIs

5.2.2 Selección y seguimiento de puntos de interés

Para medir el movimiento inducido por el flujo sanguíneo de la cabeza a lo largo de una secuencia de imágenes se seleccionan y rastrean puntos característicos dentro de las regiones de interés del paso anterior. Para la selección de los puntos característicos se utiliza el criterio de Shi y Tomasi [46], que encuentra regiones que contienen bordes y esquinas dentro de la imagen, las cuales resultan fáciles de seguir.

La implementación del criterio de Shi y Tomasi se realiza por medio del detector de Harris [47] con la función `GoodFeaturesToTrack` de OpenCV. Esta función calcula primero el valor propio mínimo para cada píxel de la imagen de origen, usando la función `CornerMinEigenVal` y lo almacena en la variable `eig_image`. Finalmente, la función asegura que todas las esquinas detectadas se encuentran separadas unas de otras por una distancia mayor a `min_distance`, eliminando las que se encuentran muy cerca y dándole prioridad a las esquinas más fuertes.

La ejecución de la función `GoodFeaturesToTrack`, da como resultado las coordenadas de un conjunto de puntos denominados puntos de interés, los cuales en su mayoría, pertenecen a esquinas y contornos de los objetos presentes en la escena como se aprecia en la Figura 5-4. La cantidad de puntos seleccionados por la función no puede superar a `maxCorners`.

Como se mencionó anteriormente, el flujo óptico de los puntos encontrados se estima mediante el criterio de Shi y Tomasi, pero utilizando a su vez la versión piramidal de Lucas y Kanade que se vio en capítulo 4. Este nuevo método, como ya se comentó en el estado del arte, recibe el nombre KLT [16]. Para su implementación, se utiliza la función `CalcOpticalFlowPyrLK` de OpenCV, que ejecuta de forma iterativa la versión piramidal de LK. Esta función estima el flujo óptico sobre los puntos de interés, encontrados por medio de `GoodFeaturesToTrack` para un frame $n-1$, y da como resultado las coordenadas finales del flujo óptico para dichos puntos en el frame n del video. El algoritmo utiliza pirámides de máximo 3 niveles, ventanas fijas de 15×15 píxeles y máximo 10 iteraciones.

Debido a que solamente nos interesa el movimiento vertical de la cabeza, se desecha la componente $x_n(t)$ quedándonos con la componente $y_n(t)$. Además, como los ECG modernos operan alrededor de los 250 Hz y los vídeos fueron grabados a 30 Hz, es necesario aplicar una interpolación para aumentar la frecuencia de muestreo de cada señal a 250 Hz. Por último, muchos de estos puntos pueden tener trayectorias erróneas, por eso, se descartan aquellos puntos que sobrepasan una distancia máxima de un frame a otro.

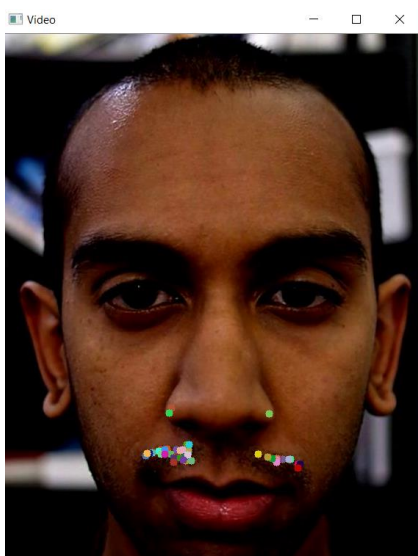


Figura 5-4. Representación de los puntos de interés

5.2.3 Filtrado temporal

El ritmo cardíaco de una persona adulta normal está entre 45 y 120 bpm o lo que es lo mismo, entre 0.75 y 2 Hz. Como no todas las frecuencias de las trayectorias son útiles para la detección del pulso cardíaco, es necesario filtrar estas señales. Se conoce que frecuencias menores a 0.75 Hz corresponden a movimientos de la respiración, de la postura o de rasgos faciales, afectando a nuestro propósito. Por otro lado, las mayores a 2 Hz aportan bastante precisión a la hora de obtener los picos. Teniendo esto en cuenta, se filtra cada señal con un filtro paso-banda Butterworth de quinto orden entre 0.75 y 5 Hz.

Para implementar el filtro se hace uso del método `signal` de la librería SciPy, y una vez en este método la función `butter` para crear el filtro con los parámetros deseados. Una vez creado el filtro, este se aplica a nuestras señales interpoladas obtenidas en el paso anterior haciendo uso de la función `lfilter`.

5.2.4 Reducción de componentes

Como se ha visto en el capítulo del marco teórico, los métodos de reducción de dimensionalidad (MDR) son algoritmos que mapean el conjunto de los datos a subespacios de menor dimensión, permitiendo hacer una descripción de los datos a un menor coste conservando la mayor información intrínseca posible. En nuestro caso, al tener N puntos en la cara, la posición de la cabeza en cada frame también será N -dimensional.

Aplicando PCA de la misma forma que se vio en el capítulo 4, se obtienen N señales o formas de onda 1-D resultado de proyectar el movimiento de los puntos (N -dimensional) en cada uno de los N autovectores. Esto se hace para descomponer el movimiento de los puntos característicos en subseñales, pudiendo aislar el pulso de una manera más sencilla como se explicará a continuación.

La cuestión ahora es ver con que autovector quedarse para la extracción de la señal del pulso. En el artículo del MIT [14] se comprueba que basta con seleccionar en primer lugar los primeros cinco autovectores con mayor autovalor, ya que en ellos se concentra la mayor información para la estimación del pulso cardíaco. Con esto se consigue descomponer el movimiento de los puntos en cinco subseñales, al proyectar dicho movimiento en estas cinco componentes principales. De entre estas cinco componentes, en dicho artículo se menciona que no siempre aquella con mayor autovalor será la mejor opción para extraer la señal del pulso, sino que es preferible quedarse con aquella señal que presente una mayor periodicidad.

Para implementar esto, se ha contado con la ayuda de la librería Scikit-learn, la cual tiene varios métodos para el cálculo de la PCA. Uno de estos métodos es `decomposition.PCA`, utilizado para configurar el número de componentes a conservar. Después de esto, se usa el método `fit` para aplicar dicha descomposición PCA a los datos. Finalmente dichos datos son proyectados en el nuevo subespacio con el método `transform`.

5.2.5 Obtención de la frecuencia cardíaca

Para saber cuál es la señal que presenta una mayor periodicidad, se calcula para cada una de las cinco señales la DFT mediante el método `fft.rfft` que nos proporciona SciPy. A partir de las DFTs es fácil obtener la potencia, y por tanto, la densidad espectral de potencia o PSD de cada una de ellas. Según el artículo del MIT, la señal más periódica será aquella que presente una mayor potencia en su frecuencia dominante.

Una vez seleccionada la señal 1-D que mejor caracteriza el movimiento de la cabeza, para la estimación de la frecuencia cardíaca en bpm se multiplica la frecuencia dominante en hercios por 60 segundos/minutos, como se hace mención en el artículo del MIT.

En la siguiente figura se muestra un ejemplo de PSD, donde se puede observar que la frecuencia dominante es aproximadamente 1.3 Hz, que equivaldría a 78 bpm.

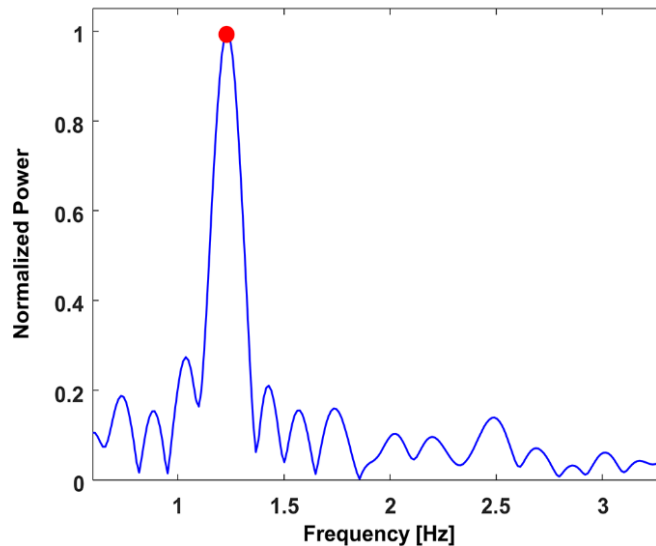


Figura 5-5. Ejemplo de estimación del pulso a partir de la PSD

Como resultado final de la implementación, el sistema muestra la estimación del pulso cardíaco sobre el vídeo como se muestra en la siguiente figura:

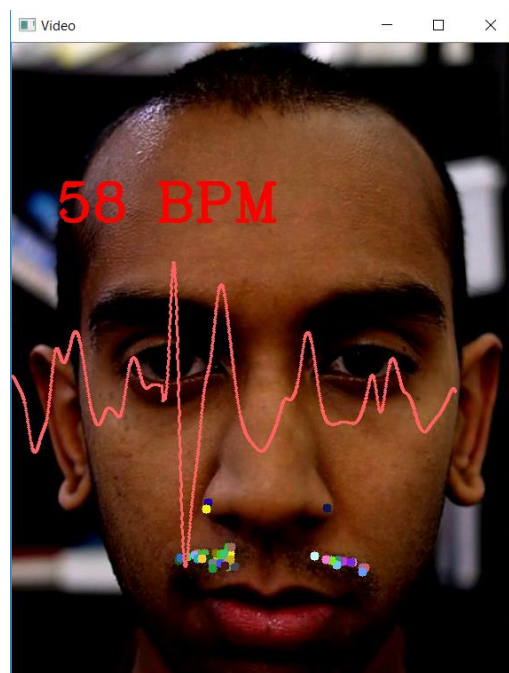


Figura 5-6. Estimación final del ritmo cardíaco

6 EXPERIMENTOS REALIZADOS Y RESULTADOS

El objetivo de este capítulo es evaluar el rendimiento que ofrece el sistema implementado. Los experimentos que se llevarán a cabo consistirán en una serie de pruebas para verificar la detección del pulso cardíaco de diferentes personas tanto en vídeos previamente grabados como en vídeos en tiempo real. Para poder comprobar la veracidad del sistema, las medidas estimadas del pulso cardíaco en cada uno de los experimentos serán contrastados con las medidas obtenidas por un ECG en el caso de los vídeos grabados y por un pulsómetro o medidor de tensión en los vídeos en tiempo real.

6.1 Evaluación de la detección

La detección es una de las fases más importante de la aplicación, ya que una mala detección conllevaría un error en el resto de las fases. El sistema implementa para la detección el algoritmo de Viola-Jones [15], el cual necesita que los individuos aparezcan en las imágenes de frente para que la detección sea lo más correcta posible. En la Figura 6-1 se puede comprobar el funcionamiento del algoritmo de la detección.

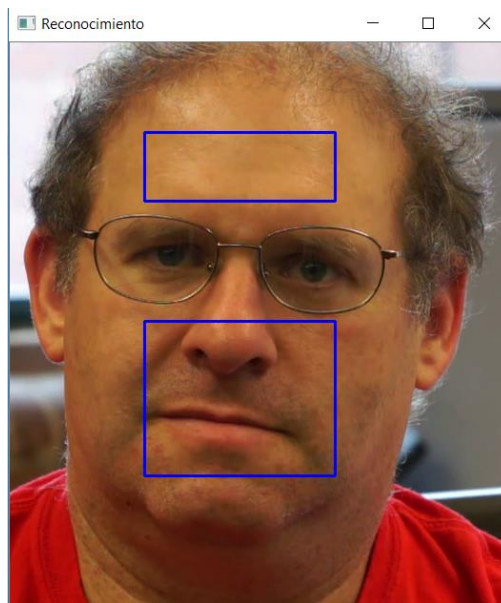


Figura 6-1. Detección correcta

Existen casos particulares en los que el algoritmo no funciona correctamente. A continuación se muestran algunos ejemplos de imágenes que no servirían, ya que no es posible la detección del rostro.

- **Individuos de perfil:** en este caso no se consigue detectar el rostro
- **Condiciones de iluminación:** la detección, además de estar condicionada por la posición y orientación de la cara del sujeto con respecto a la cámara, también depende fuertemente de las condiciones de iluminación. En situaciones donde la imagen es muy oscura por condiciones adversas no es posible realizar la detección.

6.2 Evaluación de la estimación en vídeos grabados

El objetivo de este primer conjunto de pruebas con los vídeos utilizados por el artículo original del MIT [14], es corroborar que con nuestro algoritmo base se obtienen los mismos resultados que con el algoritmo desarrollado por el MIT. En estas pruebas se han comparado los resultados publicados en [14] con los obtenidos con nuestra implementación.

Para la evaluación de los algoritmos que se han desarrollado se han seleccionado las siguientes métricas usadas en [14]:

- Media del pulso medido: Es la media del pulso medido en la duración de la secuencia, en pulsaciones por minuto.

$$\text{Media pulso medido (ppm)} = \text{media}(P, \text{duracion})$$

- Error: Es el error entre la media de la pulsación grabada del pulsómetro en la duración de la secuencia y la media del pulso mostrado por el algoritmo en la duración de la secuencia.

$$\text{Error}(\%) = \left| \frac{\text{Media pulso}_{ECG}(P, \text{duración}) - \text{Media pulso}_{medido}(P, \text{duración})}{\text{Media pulso}_{ECG}(P, \text{duración})} \right| \times 100$$

Para estas pruebas solamente hemos podido acceder a 3 de los vídeos usados por el MIT, disponibles en [48], ver Figura 6-2. Dos de ellos son secuencias colaborativas del sujeto en posición frontal a una distancia reducida respecto de la cámara, con una duración de 10 segundos. El tercero de los vídeos es una grabación de un bebé a una mayor distancia y con una posición frontolateral de la región facial y una duración de 30 segundos.

Como se puede ver en la Tabla 6-1, los resultados que se han obtenido con el algoritmo propuesto son muy próximos a los resultados presentados en el artículo [14]. Se puede observar que el error entre los resultados del algoritmo del MIT y nuestra implementación del algoritmo base es inferior al 3.7% en los tres casos bajo estudio.



a)



b)



c)

Figura 6-2. Vídeos del MIT. a) 54 bpm b) 55 bpm c) 152 bpm

Vídeo	Media del pulso (ppm)			Error (%) MIT-propuesto
	ECG	Algoritmo MIT (% error)	Algoritmo propuesto (% error)	
Cara 54 bpm	54	55 (1.85)	57 (5.55)	3.7
Cara 55 bpm	55	56 (1.78)	58 (5.45)	3.67
Bebé 152 bpm	152	152 (0)	153 (0.65)	0.65

Tabla 6-1. Resultados de los algoritmos para cada cara

Teniendo en cuenta que en el artículo usan vídeos de 90 segundos, de los cuales solo hemos tenido disponibles 10 segundos, estos resultados se pueden interpretar como que nuestro algoritmo propuesto funciona de forma muy similar al algoritmo propuesto por el MIT a pesar de tener menos información para la estimación del pulso cardíaco.

Para profundizar un poco más, se va a detallar uno de los casos anteriores mostrando la representación que realiza el sistema sobre el vídeo y se van a sacar una serie de gráficas explicativas de diferentes partes del proceso.

En la Figura 6-3 se puede observar que el sistema va mostrando para cada frame del vídeo la estimación del ritmo cardíaco, además de ir mostrando los puntos característicos de la cara a través de los cuales lo está estimando.

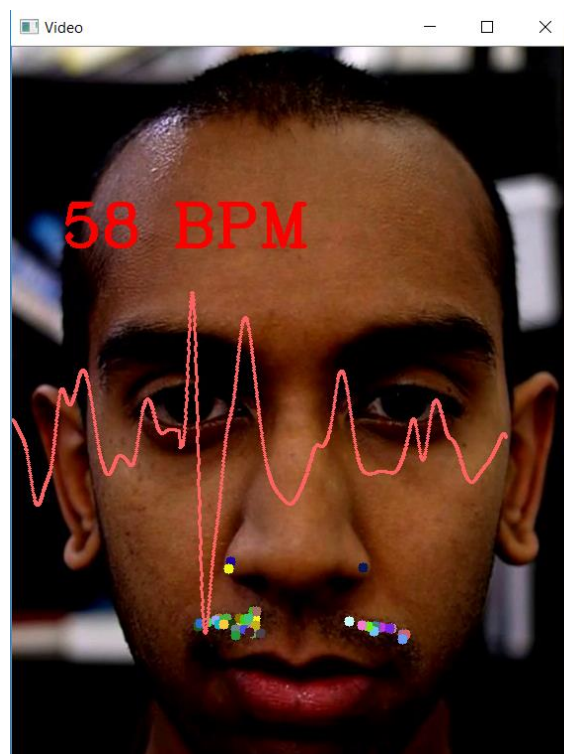


Figura 6-3. Estimación final del ritmo cardíaco del caso 55bpm

También se ha diseñado el sistema para que muestre las siguientes gráficas. En la Figura 6-4 se representa la variación de la posición de los puntos seleccionados por el método de Lucas-Kanade para realizar el seguimiento del movimiento de la cabeza de la persona.

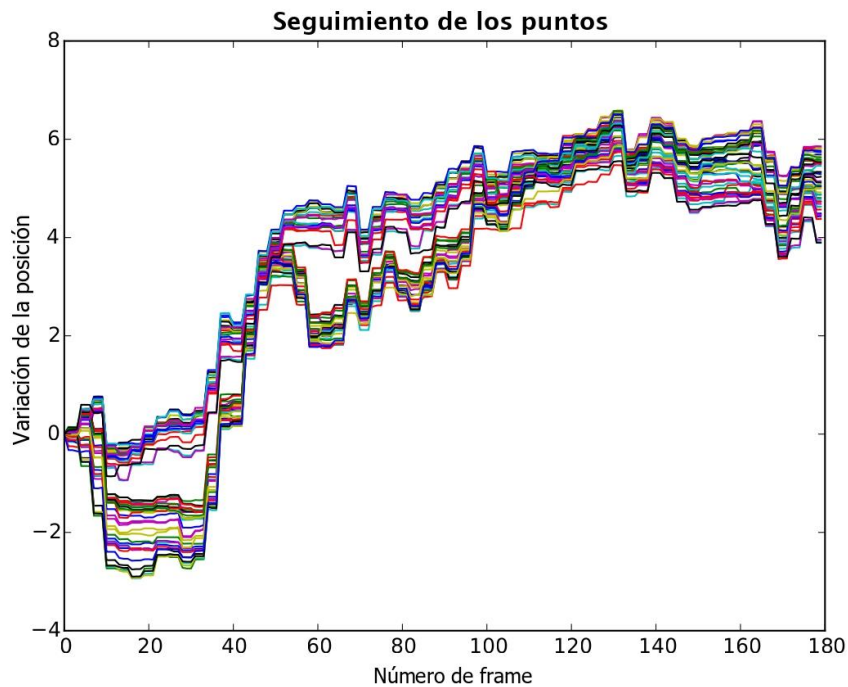


Figura 6-4. Variación de la posición de los puntos seleccionados en el vídeo

En la gráfica de la Figura 6-4 se muestran todas las componentes de los movimientos, destacando en color verde la componente cuyo movimiento es más periódico, que será la seleccionada para la estimación del pulso.

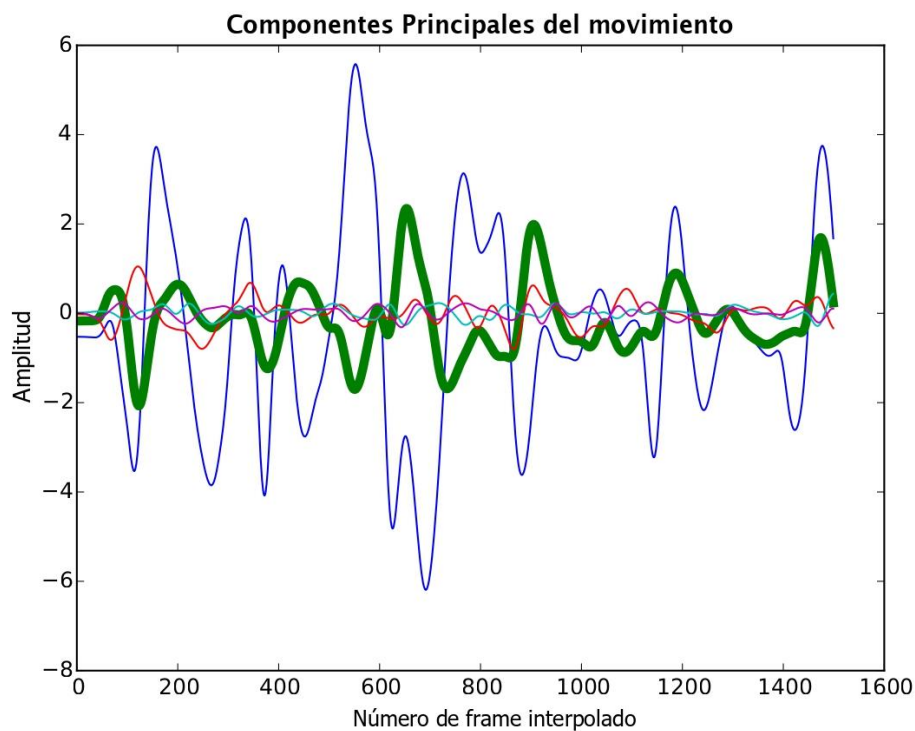


Figura 6-5. Componentes principales, destacando la más periódica

Por último también se representa en la Figura 6-6 la densidad espectral de la componente principal seleccionada, donde a partir de la máxima potencia se puede estimar la frecuencia cardíaca, siendo en este caso 58 bpm.

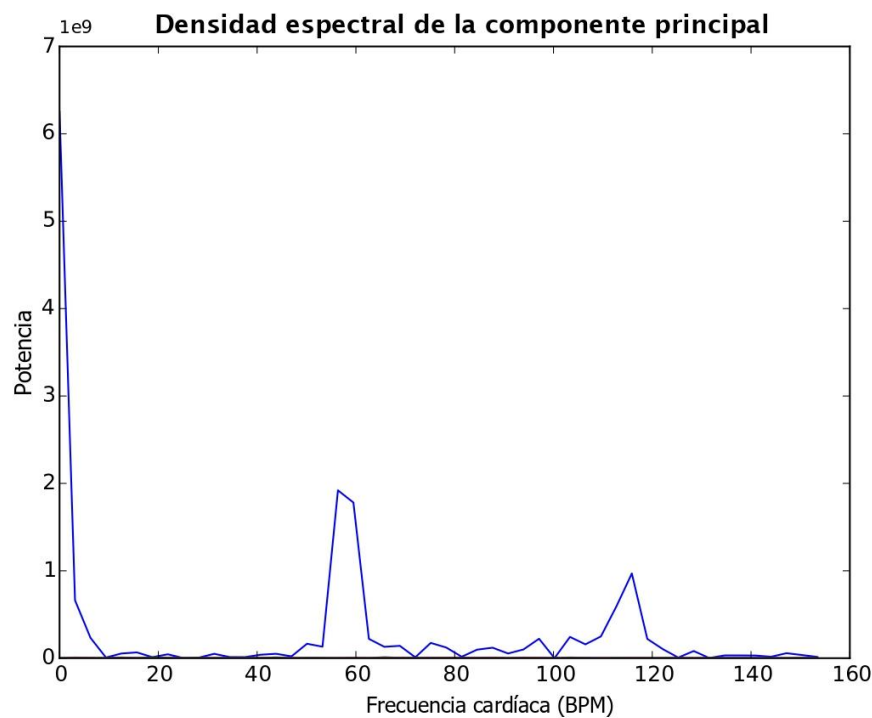


Figura 6-6. Densidad espectral de la componente principal

6.3 Evaluación de la estimación en vídeos en tiempo real

El sistema implementado no solo ha sido preparado para trabajar con vídeos de personas previamente grabados, sino también para poder calcular el ritmo cardíaco de una persona mediante la grabación de su cara en tiempo real.

Para poder evaluar esta funcionalidad, me he grabado a mí mismo con la cámara web del ordenador, mandando en tiempo real cada frame al sistema implementado. La representación sería la misma que en el caso anterior como se muestra en la Figura 6-7.

Durante la grabación también me he ido tomando el pulso con dos pulsómetros de la misma marca pero de modelos diferentes, mostrando las mediciones a la cámara. De esta forma se puede comprobar la precisión en la estimación frente a dos mediciones reales, donde se ve que se han conseguido unos resultados muy buenos.



Figura 6-7. Estimación del ritmo cardíaco en reposo (pulsómetro 1)



Figura 6-8. Estimación del ritmo cardíaco en reposo (pulsómetro 2)

Ya que hasta ahora todos los experimentos han sido de personas en reposo, para finalizar este capítulo se ha decidido realizar un último experimento sobre una persona con el pulso acelerado. Para ello he realizado una serie de ejercicios físicos y he vuelto a grabarme junto con los dos pulsómetros con la cámara web. Los resultados se muestran en las siguientes figuras:

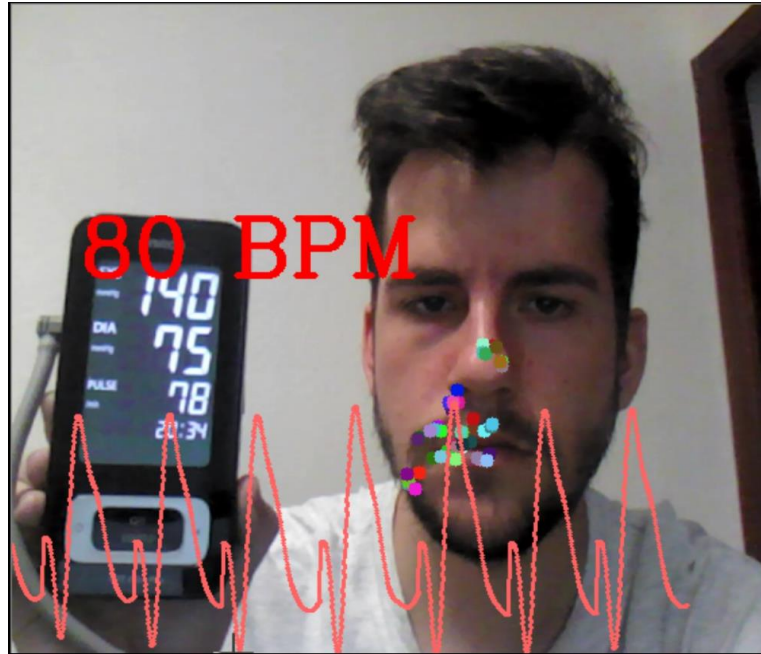


Figura 6-9. Estimación del ritmo cardíaco acelerado (pulsómetro 1)



Figura 6-10. Estimación del ritmo cardíaco acelerado (pulsómetro 2)

Se vuelve a observar en estos casos que los resultados de la estimación también son muy próximos a los reales. Por último se muestran en la siguiente tabla un resumen de los resultados así como el error de las estimaciones.

Vídeo	Media del pulso (ppm)		Error (%)
	Pulsómetro	Algoritmo propuesto	
Reposo 1	64	62	3.12
Reposo 2	61	63	3.27
Acelerado 1	78	80	2.56
Acelerado 2	76	75	1.31

Tabla 6-2. Resultados del algoritmo para cada caso

7 CONCLUSIONES Y LÍNEAS FUTURAS

A lo largo de este trabajo se ha podido ver un estudio de las técnicas de obtención del ritmo cardíaco por análisis de vídeo así como una clasificación de las mismas. En nuestro caso se ha optado por implementar el método de estimación por movimientos. También se ha visto el funcionamiento de los sistemas que intervienen en el proceso, como son el corazón y el electrocardiograma, al ser el corazón el que produce el movimiento que se mide y el electrocardiograma el aparato con el que se contrasta el pulso cardíaco estimado. Se han descritos los pasos teóricos y técnicos para el diseño y desarrollo de una aplicación de estas características, para finalmente evaluar los resultados del sistema implementado en una fase de pruebas. En este capítulo final se estudia el resultado obtenido para después proponer posibles mejoras y líneas de investigación futuras.

El objetivo principal de este trabajo era desarrollar un algoritmo capaz de medir el ritmo cardíaco de una persona mediante los movimientos imperceptibles que son inducidos por el flujo sanguíneo en la cabeza tanto en vídeos grabados como a tiempo real a través de la cámara web. Después del análisis realizado en este documento de las distintas implementaciones realizadas, podemos concluir que el objetivo se ha cumplido de forma razonable.

7.1 Conclusiones

Una de las conclusiones más importantes que se puede obtener es la gran importancia que se lleva el paso de la detección. A pesar de ser un trabajo sobre estimación del ritmo cardíaco, como ya se ha comentado en varias ocasiones siempre hay un paso previo que es la detección y que es crucial para el proceso. Para llevarla a cabo se ha utilizado uno de los algoritmos más usados en la actualidad y se ha verificado su correcto funcionamiento. No obstante, si no se trata de imágenes con caras frontales, su rendimiento no está totalmente optimizado y el rendimiento obtenido puede ser inferior.

Otra de las conclusiones que se puede obtener de los experimentos realizados es que se dan unas bajas tasas de error, de unos 2 o 3 bpm en el caso de los vídeos grabados y de 1 o 2 bpm para los de tiempo real. Esto es debido a que en la estimación en tiempo real no tenemos limitación en la duración de los vídeos, por lo que al tener más datos del movimiento de los puntos de interés, se puede ver que la estimación del pulso cardíaco es mejor.

Por último, destacar las ventajas que el sistema implementado presenta ante las poblaciones como los recién nacidos prematuros y personas de edad avanzada cuya piel es frágil y vulnerable a los sensores tradicionales, pues se trata de un método de estimación sin contacto que presenta unos muy buenos resultados, como se vio en el experimento del bebé.

7.2 Líneas futuras

Con el objetivo de continuar y mejorar lo estudiado en este proyecto, se proponen posibles líneas futuras de trabajo. No obstante, además de lo aquí mencionado, al tratarse de un sistema completamente modular siempre se pueden cambiar fácilmente las técnicas aquí empleadas en cada una de las fases por otras de las mencionadas en el estado del arte, ya sea con el propósito de comparar resultados o de buscar alguna mejoría. Además, en cuanto a la base de datos utilizada en los experimentos de este trabajo, se podría ampliar con secuencias de otros sujetos con el fin de poder evaluar si el error se mantiene en torno al mismo porcentaje.

Uno de los puntos más buscados y también más costosos consiste en que no se requiera que el individuo aparezca de frente. Para ello se podría intentar mejorar tanto la detección como la estimación del flujo óptico del sistema para su correcto funcionamiento en dichas situaciones, pues como se ha comentado anteriormente tiene esta limitación. De esta manera se ampliarían las posibilidades de realizar la estimación en otras poses o incluso en movimiento.

El movimiento de la cabeza puede ofrecer más información sobre el ciclo cardíaco. La dirección del movimiento puede revelar asimetrías en el flujo sanguíneo hacia dentro o hacia afuera de la cabeza. Esto podría ser útil para el diagnóstico de una estenosis, un bloqueo parcial, de las arterias carótidas.

Finalmente, el método propuesto puede ser útil para otras aplicaciones además de extraer el pulso del movimiento de la cabeza. Esto abriría futuras investigaciones sobre la combinación de señales en todo el cuerpo para obtener información fisiológica más profunda. Por ejemplo, al medir la diferencia de tiempo entre los movimientos en el pecho, la carótida y las manos, podemos obtener una buena estimación de la velocidad de la sangre. También se podría medir la respiración desde el movimiento de la cabeza o el pecho simplemente cambiando la banda de frecuencia de interés.

REFERENCIAS

- [1] J. L. I. Pavlidis, "Thermal image analysis for polygraph testing", IEEE Engineering in Medicine and Biology Magazine, 2002.
- [2] C. Takano y Y. Ohta, "Heart rate measurement based on a time-lapse image", Medical engineering & physics, 2007.
- [3] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand y W. Freeman, "Eulerian video magnification for revealing subtle changes in the world", ACM Transactions on Graphics, 2012.
- [4] Heart Rate Monitor, http://sport.com/heart_rate_monitor/.
- [5] Cardiio app, <https://www.cardiio.com/>, 2013.
- [6] A. Lam y Y. Kuno, "Robust heart rate measurement from video using select random patches", Proceedings of the IEEE International Conference on Computer Vision, 2015.
- [7] M. Poh, D. McDuff y R. Picard, "Non-contact automated cardiac pulse measurements using video imaging and blind source separation", Opt. Express 18, 2010.
- [8] M. Poh, D. McDuff y R. Picard, "Advancements in noncontact, multiparameter physiological measurements using a webcam", IEEE Trans.Biomed, 2011.
- [9] M. Lewandowska, J. Ruminski y T. Kocejko, "Measuring pulse rate with a webcam a non-contact method for evaluating cardiac activity", Federated Conference on Computer Science and Information Systems, 2011.
- [10] L. Aarts, V. Jeanne, J. Cleary, C. Lieber y J. Nelson, "Non-contact heart rate monitoring utilizing camera photoplethysmography in the neonatal intensive care unit", Early Hum. Dev, 2012.
- [11] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge y D. Clifton, "Non-contact video-based vital sign monitoring using ambient light and auto-regressive models", Physiological measurement, 2014.
- [12] L. Carvalho, M. Virani y M. Kuty, "Analysis of Heart Rate Monitoring Using a Webcam", International Journal of Advanced Research in Computer and Communication Engineering, 2014.

- [13] O. T. Inan, "Recent advances in cardiovascular monitoring using ballistocardiography", Engineering in Medicine and Biology Society, 2012.
- [14] G. Balakrishnan, F. Durand y J. Guttag, "Detecting pulse from head motions in video", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013.
- [15] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", In Computer Vision and Pattern Recognition. Proceedings of the 2001 IEEE Computer Society Conference, 2001.
- [16] C. Tomasi y T. Kanade, "Detection and Tracking of Point Features", School of Computer Science, 1991.
- [17] L. Shan y M. Yu, "Video-based heart rate measurement using head motion tracking and ICA", 6th International Congress on Image and SignalProcessing (CISP), 2013.
- [18] R. Irani, K. Nasrollahi y T. B. Moeslund, "Improved Pulse Detection from Head Motions Using DCT", International Conference on Computer Vision Theory and Applications, 2014.
- [19] «Electrocardiograma,» [En línea]. Available: <http://es.wikipedia.org/wiki/Electrocardiograma>.
- [20] A. Lopez y C. Macaya, "Salud cardiovascular. Libro de Salud Cardiovascular", Fundacion BBVA, 2009.
- [21] A. C. Guyton y J. E. Hall, "Músculo cardíaco: el corazon como bomba y la función de las válvulas cardiacas", Tratado de Fisiología Médica, 2011.
- [22] M. Espinosa Sánchez del Villar, "Proyecto Fin de Carrera: Técnicas de procesamiento y análisis para la señal electrocardiográfica", Universidad de Sevilla, 2002.
- [23] J. P. Amigo, «¿Cuántas son las pulsaciones normales por edad y estado?,» [En línea]. Available: <https://revistadigital.inesem.es/biosanitario/pulsaciones/>.
- [24] L. D. Estany, «¿Tienes una frecuencia cardíaca normal?,» [En línea]. Available: <http://entrenamientoysaludcoruna.com/tienes-una-frecuencia-cardiaca-normal/>.
- [25] "Enfermedades y trastornos cardíacos", Heart Rhythm Society, 2013.
- [26] H. R. Society, Enfermedades y trastornos cardíacos, [www.hrsonline.org/content/download/.../Enfermedades y trastornos cardiacos.pdf](http://www.hrsonline.org/content/download/.../Enfermedades+y+trastornos+cardiacos.pdf), 2013.
- [27] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation", IEEE Transactions on Pattern Analysis and Machine Intelligence, 1989.
- [28] Y. Freund and R. Schapire, "A short Introduction to Boosting", Journal of Japanese Society for Artificial Intelligence, 1999.

- [29] J. Gonzalo Pajares, "Visión por Computador: Imágenes digitales y aplicaciones", Editorial Ra-Ma, 2ª Edición.
- [30] F. Calderón Solorio y J. Marroquín Zaleta, "Un Nuevo Algoritmo para el Cálculo de Flujo Óptico y su Aplicación al Registro de Imágenes", Computación y Sistemas Vol. 6, 2003.
- [31] A. Bruhn y J. Weickert, "Lucas/Kanade Meets Horn/Schunck: Combining Local and Global Optic", International Journal of Computer Vision, 2005.
- [32] L. Yee Siong, S. Salasiah Mokri y A. Aini Hussain, "Motion detection using Lucas Kanade algorithm and application enhancement", International Conference on Electrical Engineering and Informatics, 2009.
- [33] J. Kanawathi, S. Salasiah Mokri y N. Ibrahim, "Motion detection using Horn Schunck algorithm and implementation", International Conference on Electrical Engineering and Informatics, 2009.
- [34] B. Lucas y Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision", DARPA Image Understanding Workshop, 1981.
- [35] B. Horn y G. Schunck, "Determining Optical Flow", Artificial Intelligence, 1981.
- [36] F. Tsai y K. Luk Chan, "Dimensionality reduction techniques for data exploration", 6th International Conference on Information, Communications & Signal Processing, 2007.
- [37] J. Shlens, "A Tutorial on Principal Component Analysis", Systems Neurobiology Laboratory, Salk Institute for Biological Studies, 2005.
- [38] K. Baker, "Singular Value Decomposition Tutorial", 2005.
- [39] Python, <https://www.python.org/>.
- [40] Continuum Analytics Anaconda, <https://www.continuum.io/>.
- [41] OpenCV, <http://opencv.org/>.
- [42] Numpy, www.numpy.org.
- [43] Scikit-learn, <https://scikit-learn.org/stable/documentation.html>.
- [44] SciPy, <https://docs.scipy.org/doc/scipy/reference/>.
- [45] Face Detection using Haar Cascades, https://docs.opencv.org/3.4.1/d7/d8b/tutorial_py_face_detection.html.

- [46] S. J y C. Tomasi, "Good Features to Track", 9th IEEE Conference on Computer Vision and Pattern Recognition, 1994.
- [47] Harris Corner Detection, https://docs.opencv.org/trunk/dc/d0d/tutorial_py_features_harris.html.
- [48] Eulerian Video Magnification for Revealing Subtle Changes in the World, <http://people.csail.mit.edu/mrub/evm/>, 2012.

