



UNIVERSITÀ POLITECNICA DELLE MARCHE
FACOLTÀ DI INGEGNERIA

Corso di Laurea Triennale in Ingegneria Biomedica

**Elaborazione di immagini RGB dal diametro
della pupilla e valutazione della dipendenza
della variabilità cardiaca.**

**RGB pupil diameter images processing and
evaluation of the dependent on heart rate
variability**

Relatore:
Prof. Ennio Gambi

Tesi di Laurea di:
Fabio Bartolucci

Correlatore:
Ing. Manola Ricciuti

Indice

Introduzione	3
1 Stato dell'arte	5
2 Setup e dispositivi.....	11
2.1 GoPro Hero 6	12
2.2 Pulsossimetro	13
3 Descrizione metodi	14
3.1 Metodi per l'estrazione del battito cardiaco	14
3.1.1 Fotopletismografia	15
3.1.2 Videopletismografia	16
3.1.3 EVM	16
3.1.4 Face detection	18
3.1.5 Estrazione del rapporto LF/HF	19
3.2 Metodi per l'estrazione del diametro della pupilla.....	21
3.2.1 Eye detection.....	21
3.2.2 Canny edge detection	21
3.2.3 Trasformata di Hough.....	23
4 Algoritmo di estrazione del battito cardiaco	25
4.1 Decomposizione del video in sequenza di immagini	25
4.2 Variabili fondamentali	25
4.3 Rilevazione volto.....	26
4.4 Decomposizione piramidale e filtraggio temporale	26
4.5 Creazione delle ROI	27
4.6 Estrazione del segnale VPG.....	28
4.7 Estrazione HRV e analisi dei picchi	28
5 Algoritmo di estrazione del diametro della pupilla	31
5.1 Decompressione del video in immagini	31

5.2	Variabili fondamentali	32
5.3	Eye detection e taglio della pupilla	32
5.4	Binarizzazione immagine e inizializzazione del ciclo	33
5.5	Segmentazione pupilla.....	35
5.6	Blob ed equazione ellisse.....	35
5.6.1	Circle Blob.....	36
5.6.2	Metodo dell'ellisse	36
5.7	PSD.....	36
5.8	Plot dei grafici	37
6	Risultati	39
	Conclusioni	44
	Bibliografia	46

Introduzione

L'elaborato di tesi è stato svolto basandosi su due studi condotti separatamente e poi messi a confronto: il primo riguarda l'analisi delle variazioni del diametro pupillare modificando le condizioni di luminosità; la seconda parte tratta invece lo studio delle componenti a bassa e alta frequenza della densità spettrale di potenza con particolare attenzione al rapporto LF/HF mediante valutazioni sull'Heart Rate Variability (HRV).

I due oggetti di studio in questo elaborato di tesi verranno infine confrontati perché correlati al sistema nervoso centrale; entrambi ci forniscono una panoramica del livello di stress del soggetto in esame. Ad oggi è possibile effettuare queste misure solo con il pupillografo a livello clinico per cui è nostro interesse valutare se è possibile raggiungere buone prestazioni anche utilizzando solo una fotocamera digitale commerciale RGB.

La trattazione svolta è avvenuta mediante l'acquisizione video utilizzando GoPro Hero6 seguita da un successivo trasferimento in ambiente di programmazione Matlab, sfruttando due tool differenti per i relativi aspetti analizzati.

L'acquisizione dei video fatta su nove soggetti è avvenuta per una durata di quaranta secondi: i primi venti sono stati dedicati allo studio della fluttuazione del diametro pupillare, tra il decimo e il quindicesimo secondo è stata incrementata l'intensità luminosa per evidenziare la contrazione della pupilla così che il sistema così sviluppato potesse effettivamente riconoscerla e visualizzarla; i successivi venti secondi sono invece stati attribuiti alle microvariazioni, (impercettibili ad occhio nudo) del colore della pelle che consentono l'estrazione, attraverso la tecnica EVM (Eulerian Video Magnification), dei dati relativi all'HRV (Heart Rate Variability); i risultati ottenuti dalle due differenti tecniche sono stati infine confrontati.

Gli esperimenti sono stati condotti nel laboratorio di telecomunicazioni dell'Università utilizzando la luce presente all'interno mediante illuminazione artificiale, mentre come fonte di luce per la prima trattazione è stata utilizzata la torcia del telefono cellulare.

Comune alle due elaborazioni è il lavoro di detection presente all'interno dei due codici, relativo all'occhio per il primo (eye detection) e al viso per il secondo (face detection). Gli algoritmi focalizzati sull'identificazione e sul tracciamento di oggetti in movimento a partire da analisi video sono materia di ricerca e di particolare interesse negli ultimi anni.

Nel capitolo 1 è stata effettuata una ricerca in letteratura per ottenere un quadro complessivo dei lavori sviluppati finora e per ottenere un punto di partenza per il lavoro svolto. Nel capitolo 2 viene data una panoramica generale degli strumenti utilizzati e dei soggetti analizzati. Nel capitolo 3 è fornita un'attenta descrizione dei due metodi impiegati nel lavoro di tesi. Nel capitolo 4 e 5 viene illustrato l'algoritmo di estrazione del battito cardiaco e il codice per la rilevazione delle fluttuazioni del diametro della pupilla. Di seguito nel capitolo 6 vengono riportati i risultati ottenuti e nel capitolo ultimo le conclusioni.

1 Stato dell'arte

In questo capitolo vengono riportati ed illustrati gli studi precedenti relativi agli argomenti utilizzati ed approfonditi nel mio lavoro di tesi per lo studio della variazione del diametro pupillare e per le valutazioni dell'HRV (Heart Rate Variability), messi in correlazione dallo studio del rapporto ad alta e bassa frequenza (LF/HF).

La misurazione della variazione della pupilla e la direzione dell'occhio umano sono connessi ad eventuali problemi di particolare rilievo ed è materia di ricerca nel campo della medicina, psicologia e ingegneria biologica. In particolare in medicina ci sono studi sulla relazione tra movimento dell'occhio e malattie caratterizzate da disturbi di equilibrio e vertigini [1].

Comune alle due tecniche messe a confronto in questo elaborato è il lavoro di detection sviluppato in precedenza da Paul Viola e Micheal J. Jones [2] e poi standardizzato: il cui obiettivo è il rilevamento generale di oggetti. Questo approccio è stato ampiamente utilizzato per rilevare diversi target nell'immagine, ricevendo grande attenzione a quelle relative al volto umano e ai suoi elementi [3]. Il metodo proposto da Viola e Jones, basato su AdaBoost, ha grande popolarità grazie ad un alto tasso di rilevamento, bassa complessità e solide basi teoriche. La velocità elevata del metodo AdaBoost è principalmente dovuta all'uso di semplici funzionalità simili a Haar e una struttura di classificatori a cascata, che esclude rapidamente la maggior parte delle ipotesi della finestra dell'immagine [4]. Una volta ottenuto un solido ed efficace tracciamento oculare è possibile applicare un dispositivo in grado di rilevare condizioni psicologiche e cognitive di un soggetto; un campo sicuramente interessato a questo aspetto, per cui si cerca di predisporre un dispositivo, è il settore automobilistico. Si vuol ottenere un'analisi della fatica del conducente fino a poter consigliare una pausa dalla guida attraverso un sistema di allarme opportunamente impiantato nella vettura. Il rilevamento della pupilla implica una localizzazione di essa nell'immagine e consiste in due fasi: rimozione dell'interferenza di illuminazione e rilevamento della pupilla [5]. Il dispositivo prende il nome di eye tracker il quale consente la misurazione della posizione degli occhi e i relativi movimenti annessi.

La stanchezza del conducente influisce gravemente sulla prontezza e la vigilanza per rispondere a una situazione pericolosa ed è anche spesso sottovalutato. Le statistiche dimostrano che tra il 10% e il 20% degli incidenti stradali sono dovuti a stanchezza o livello di vigilanza ridotto. In Cina la stanchezza ha provocato 925 morti su un totale di 3056 registrati nel 2004, un

equivalente al 14,8%. Questo problema ha aumentato la necessità di sviluppare sistemi di sicurezza attiva che possono ridurre gli incidenti stradali avvertendo, come detto in precedenza, i conducenti della loro cattiva condizione di guida. La stanchezza del conducente può essere sostanzialmente suddivisa in tre principali categorie:

- affaticamento basato sui parametri fisiologici del conducente come onde cerebrali, frequenza cardiaca e respirazione, ma queste tecniche sono invadenti per il driver;
- rilevamento della fatica in base alle prestazioni di guida come: cambiamenti del volante, posizione laterale del veicolo, forza di presa del conducente sul volante. Ma in questo caso lungo il tempo di guida comporterebbe la traspirazione sui sensori, diminuendo la loro capacità di monitorare accuratamente;
- rilevamento della fatica da segnali visivi: monitoraggio del movimento delle palpebre e lo sguardo. L'utilizzo di videocamere per monitorare il conducente è un metodo non invasivo e diventa sempre più pratico, ma allo stesso tempo possono essere sensibili a fattori esterni come la luminanza, aspetto del conducente o vibrazioni durante la guida.

Nel complesso i comportamenti degli occhi forniscono un'alternativa migliore all'analisi che si intende fare sulla stanchezza del conducente, prevedendo così lo stato di sonnolenza o vigilanza. Nel documento [6] viene quindi proposto un sistema di rilevamento della fatica del conducente in tempo reale e il suo algoritmo.

L'approccio utilizzato per lo studio della pupilla consiste nella trasformazione dell'immagine RGB (a colori) in una in scala di grigi per ridurre il tempo di elaborazione e la complessità. Utilizzando due rilevatori dei bordi circolari (CED) riusciamo ad ottenere i confini interni ed esterni delle regioni dell'iride. Gli errori di rilevamento dovuti a fattori di rumore, come le occlusioni dell'occhio dovute agli occhiali e i capelli, sono comuni. Pertanto, le immagini rilevate vengono divise in due casi, vale a dire "casi di rilevamento corretto" e "rilevamento errato", basati sull'esistenza della riflessione speculare corneale (SR). Questa distinzione viene basata sul livello di grigio presente, infatti se è superiore di 250 viene classificato come SR corneale; se il numero di punti SR è maggiore a 1 nel file regioni della pupilla dell'iride allora viene studiata e catalogata come "caso di buona rilevazione", in caso opposto viene considerato come "caso di rilevamento errato". In caso negativo viene applicato AdaBoost per rilevare nuovamente la regione dell'iride. AdaBoost è un algoritmo che costituisce un classificatore forte accoppiando classificatori deboli, richiede molto tempo per rilevare deboli classificatori ma presenta vantaggi come la velocità di rilevamento e delle buone prestazioni di classificazione. In sostanza se l'immagine non effettua una rilevazione del riflesso allora

conclude che l'occhio è chiuso, in caso contrario eseguendo una segmentazione del colore lo elimina [7].

Nel documento [8] viene mostrato lo sviluppo di un metodo di rilevamento dello sguardo degli occhi utilizzando un sistema di supporto per persone fisicamente svantaggiate come pazienti con sclerosi multipla amiotropica. In precedenza è stata proposta la tecnica di rilevamento della pupilla utilizzando due sorgenti luminose e il metodo della differenza di immagine per rilevare facilmente l'area della pupilla. In questo studio viene costruito un sistema hardware e un algoritmo per il rilevamento degli occhi.

La variazione nel diametro della pupilla e quindi anche il suo essere reattivo agli stimoli dipende da diverse cause. Essa aumenta di dimensione in condizioni di scarsa luminosità e viceversa ne diminuisce la sua grandezza in presenza di condizioni di luce elevata. Altri indicatori sono legati alla sfera emotiva ed allo stato clinico del paziente. Il sistema simpatico influenza l'aumento della dimensione mentre il sistema parasimpatico, al contrario, regola la restrizione di essa. Uno studio di Bilman ha confermato la relazione tra l'attività del sistema nervoso e le bande di frequenza, quindi è possibile la correlazione tra la categorizzazione dello stress e il rapporto LF/HF. Viene quindi consentita la rigorosità nell'esame dell'HRV riducendo le ambiguità, attraverso la stima dell'attività in LF e HF [9]. Le attività dei sistemi descritti in precedenza influiscono e si riflettono nelle bande a bassa (0,04-0,15 Hz) e alta frequenza (0,15-0,4 Hz) dell'HRV. Il rapporto tra le due bande (LF/HF) viene utilizzato per la valutazione dello stato psicologico e fisiologico del soggetto e viene analizzato da un punto di vista bidimensionale vincolando i gradi di libertà e fornendo una solida valutazione dello stress. Questo rapporto sarà tanto maggiore quanto elevato è il grado di stress del soggetto e in termini numerici un valore desiderabile si aggira attorno a 0,5. Altra correlazione proposta nel documento [9] è quella con le abitudini del soggetto da un punto di vista prettamente di sedentarietà.

L'Heart Rate Variability è possibile misurarla attraverso le fluttuazioni spontanee del diametro della pupilla, nel documento [10] viene presentato un metodo per poter estrarlo. Attraverso un eye tracker, che sfrutta una sorgente nell'infrarosso, viene localizzata la regione dell'occhio; in seguito l'immagine viene elaborata eliminando i riflessi corneali e infine la pupilla viene distinta dall'iride con un operatore integrodifferenziale, calcolandone il diametro. Il codice infine analizza lo spettro di potenza, in frequenza, generato dalle fluttuazioni pupillari alle basse e alle alte frequenze per estrarre l'indice HRV.

Un ulteriore motivo per cui è funzionale questo approccio non invasivo e non contactless è il disagio che può far sentire i soggetti nervosi se applicati sensori su di essi. Questo potrebbe fornire una variabilità nei risultati che renderebbe lo studio non attendibile. I dispositivi senza contatto consentono la misurazione priva di avviso per il soggetto monitorandolo senza suscitare in lui emozioni che potrebbero interferire. Tramite tecnica VPG è possibile stimare l'HRV partendo da un semplice video in cui viene inquadrato il volto. [11]

Attraverso un segnale fotopletismografico (PPG) è possibile monitorare ed estrarre dati relativi all'HRV ed inoltre è un metodo non invasivo. Nel documento [12] vengono confrontati HRV e variabilità del segnale fotopletismografico (PPGV) calcolando nel tempo e nella frequenza parametri relativi alle misurazioni HRV. I risultati ottenuti nel lavoro svolto in questo documento rilevano una forte correlazione tra HRV e PPGV, quindi permette di considerare il segnale in questione un'ottima alternativa per misurare l'HRV.

Nel documento [13] viene rimarcato, come già fatto negli articoli in precedenza, come la variabilità della frequenza cardiaca sia diventata un indicatore per varie condizioni di salute. I sensori di pletismografia (PPG) integrati in dispositivi indossabili come smartwatch e telefoni sono ampiamente utilizzati per misurare l'attività cardiaca. L'HRV richiede una stima accurata dell'intervallo di tempo tra i picchi consecutivi nel segnale PPG. Tuttavia, il segnale PPG è molto sensibile all'artefatto da movimento che può portare a una scarsa stima dell'HRV se vengono rilevati falsi picchi. In questo documento viene proposto un approccio probabilistico basato sull'apprendimento bayesiano per stimare meglio l'HRV dal segnale PPG.

Secondo quanto riportato dal documento [14] al giorno d'oggi l'analisi della frequenza cardiaca e della variabilità della frequenza cardiaca è materia di grande interesse. I recenti progressi nelle tecniche di misurazione HR e HRV senza contatto aumenteranno ulteriormente il loro interesse per la stima emotiva tramite HRV. Tali metodi di misurazione comportano l'estrazione del segnale PPG dal volto umano attraverso una videocamera. Uno dei metodi più utilizzati per questa elaborazione è l'Eulerian Video Magnification (EVM) per rilevare sottili cambiamenti nel colore della pelle associato al segnale PPG. La tecnica EVM ha dimostrato di essere efficace quando si tratta di estrazione HR. Si basa sul principio euleriano che afferma che la pressione e la velocità si sviluppano nel tempo. Quando si usa la prospettiva euleriana sui video, ogni pixel viene elaborato indipendentemente e trattato in funzione del tempo. L'EVM utilizza un pool spaziale localizzato e un filtraggio temporale per estrarre il segnale del polso cardiaco.

Questa tecnica consente l'ingrandimento dei sottili cambiamenti nel colore della pelle provocati dal flusso sanguigno nei vasi facciali.

Una volta confermata la correlazione tra cambiamenti impercettibili di colore e HRV, l'aspetto pratico da analizzare è l'algoritmo annesso al rilevamento del volto. Nel documento [15] viene presentata la combinazione di un approccio invariante dall'illuminazione del viso con un meccanismo di tracciamento utilizzato per migliorare velocità e precisione del sistema. Vengono introdotte le caratteristiche della struttura locale per il rilevamento di soggetti indipendentemente dall'illuminazione. Il monitoraggio è eseguito mediante rilevamento continuo. Derivano tre condizioni per una sequenza di stati valida nel tempo. Per stimare lo stato ottimale di un volto tracciato dai risultati di rilevamento viene utilizzato un filtro Kalman. Questo porta ad una levigatura istantanea della traiettoria del viso. Si può dimostrare sperimentalmente che levigare le traiettorie del viso porta a una significativa riduzione dei falsi rilevamenti rispetto al rilevatore statico senza l'estensione di tracciamento. Viene mostrato inoltre come sfruttare le informazioni ridondanti in una sequenza video naturale per velocizzare l'esecuzione del rilevatore statico mediante una procedura di scansione temporale che chiamiamo "slicing". Le prime tre fasi di questo processo vengono svolte con un AdaBoost.

L'algoritmo AdaBoost è una tecnica tra gli algoritmi di maggior successo in rilevamento facciale, ma impiega una quantità di tempo rilevante. Nel documento [16] viene proposto un metodo di rilevamento basato sulla segmentazione del colore della pelle che ha una velocità di funzionamento relativamente elevata rispetto AdaBoost, ma perde di precisione rispetto quest'ultimo. Quindi viene presentato un algoritmo che combina il colore YcbCr con la rilevazione della pelle e l'algoritmo AdaBoost per il rilevamento preciso.

Per ovviare il problema della bassa precisione e della bassa velocità di elaborazione per il rilevamento e il tracciamento dei volti in tempo reale nei sistemi viene proposto un approccio alla regione di interesse basato sul margine con concetti di margine fisso e dinamico, che consente di accelerare il tempo di elaborazione. Inoltre viene sviluppato un sistema ibrido per aumentare la precisione e superare la carenza del principale algoritmo di rilevamento. Questo approccio consiste in due routine: una routine principale e una di fuga. Tre algoritmi vengono utilizzati indipendentemente come routine principale per valutare l'efficacia dell'approccio ibrido proposto. Questi algoritmi sono Haar cascade, Joint cascade e reti neurali convoluzionali multitask. La routine fuga basata sull'algoritmo di corrispondenza dei modelli è progettata per

valutare l'efficacia dell'approccio ibrido proposto e migliorare l'accuratezza del rilevamento. I risultati dell'esperimento del documento [17] confermano che l'approccio ibrido è in grado di rilevare e tracciare i volti con orientamento frontale con una migliore precisione e una maggiore velocità di elaborazione, ovvero quattro volte più veloce rispetto alle tecniche di scansione full frame convenzionali.

2 Setup e dispositivi

In questo lavoro di tesi sono stati presi in esame nove soggetti che in maniera individuale sono stati posti frontalmente alla videocamera e ripresi per un tempo di 40 secondi. La videocamera utilizzata per questo studio è GoPro Hero6, fondamentale è l'impegno del soggetto nel cercare di muovere il meno possibile la testa e per i primi 20 secondi gli occhi.

I 40 secondi presi in considerazione sono stati così suddivisi:

- dal secondo 1 al 20 utilizzati per lo studio delle fluttuazioni della pupilla variando le condizioni di luminosità
- dal secondo 21 al 40 sono stati processati tramite tecnica EVM (Eulerian Video Magnification) per l'estrazione di dati relativi all'Heart Rate Variability (HRV)

Precedentemente all'acquisizione abbiamo rilevato in maniera manuale la dimensione della pupilla attraverso un righello e la frequenza cardiaca attraverso un pulsossimetro. Il soggetto per la registrazione veniva posto frontalmente alla camera ad una distanza di circa 40 cm in posizione seduta su una sedia e GoPro Hero6 posizionata sul treppiede (*Figura 2.1*). Nei capitoli successivi verrà fornita una panoramica relativa agli strumenti utilizzati e sopraccitati.



Figura 2.1 Postazione soggetto in esame

2.1 GoPro Hero 6

GoPro Hero 6 (*Figura 2.2*) è una videocamera messa in commercio nel 2017, di dimensione 4,44 x 6,20 x 3,20 centimetri e peso 120 grammi con un display touchscreen di 2 pollici in diagonale posizionato nella parte posteriore dello strumento. È stata scelta per questo lavoro di tesi, oltre che per la risoluzione, anche per l'ampio intervallo dinamico che permette di ottenere colori migliori rispetto alle versioni precedenti con contrasti evidenti, colori fedeli alla realtà e immagini vivide. Presente nella parte anteriore troviamo l'elemento più sporgente ovvero la lente all'interno del quadrato che ospita anche il display secondario monocromatico il quale mostra in concomitanza con lo schermo principale il tempo di registrazione del video. Sulla parte superiore si trova il tasto scatto/registra, sul bordo destro il tasto Modalità/Alimentazione, nella parte inferiore si trova il vano batterie e sul bordo sinistro la porta USB-Type-c e le porte Mini-HDMI. Questo strumento è resistente all'acqua e in particolare a immersioni fino a 10 metri di profondità. Permette di registrare video con qualità 4K acquisendo 30, 60 fino a 240 frame per secondo (fps) a 12 megapixel. Per effettuare lo studio sono stati girati video della durata di 40 secondi con la videocamera in modalità 30 fps e con una risoluzione pari a 1920x1080. I soggetti sono stati posizionati frontalmente allo strumento seduti ad una distanza di circa 40 cm e l'apparecchio sistemato su un treppiede. Di fondamentale rilievo è l'impegno da parte del soggetto di minimizzare al massimo i movimenti sia degli occhi che della testa, per evitare segnali rumorosi o indurre errori al video tracker del volto. Un'ulteriore accortezza è la condizione di illuminazione del volto del soggetto.



Figura 2.2 GoPro Hero6

2.2 Pulsossimetro

Il pulsossimetro o saturimetro (*Figura 2.3*) è uno strumento pletismografico nel quale viene inserito un dito e lavora attraverso un diodo che emette della luce (nel rosso e nell'infrarosso) che attraversa l'arto arrivando ad un secondo diodo di ricezione ottenendo un segnale PPG. La fotopletismografia (PPG) è la registrazione della variazione dell'intensità luminosa ottenuta dalla variazione del volume dell'arto in esame. La frequenza cardiaca è un parametro vitale che varia in funzione di età, patologie, stato del soggetto (riposo, stress, fatica) e fattori ambientali. Prima di effettuare rilevazioni con GoPro Hero6 abbiamo utilizzato il pulsossimetro (pulse oxymeter) per ottenere dei dati relativi all'HR da poter poi riportare su codice matlab per la tecnica EVM. Ottenendo così nello schermo dello strumento il valore bpm che ha in quel dato momento. Altro dato che fornisce il pulsossimetro è la percentuale di emoglobina satura di ossigeno, che non verrà tenuto in considerazione per questo lavoro di tesi.



Figura 2.3 Pulsossimetro

3 Descrizione metodi

In questo capitolo verranno descritti in maniera prettamente teorica i metodi utilizzati per questo lavoro di tesi, nel 3.1 si espongono tutti i metodi di estrazione del battito cardiaco e nel 3.2 quelli di estrazione del diametro della pupilla.

3.1 Metodi per l'estrazione del battito cardiaco

Nei seguenti capitoli verranno illustrati i metodi presi in considerazione in questo elaborato per l'estrazione del battito cardiaco da acquisizioni video. La frequenza cardiaca è da sempre argomento di grande interesse per i sistemi sanitari in quanto possono aiutare la diagnosi di diverse patologie, tra cui le malattie cardiache. La frequenza cardiaca è il numero di battito del cuore al minuto (bpm). L'HR varia in funzione di alcune condizioni fisiche del soggetto in esame tra cui età, allenamento, riposo, ansia e stress. La frequenza cardiaca di un soggetto adulto a riposo è di circa 70/75 bpm, più un soggetto è allenato e minore sarà l'HR e varia in maniera opposta in relazione alla diminuzione dell'età. Il lavoro di pompaggio del cuore dipende dalla propagazione dell'impulso che genera fenomeni di depolarizzazione e ripolarizzazione nelle diverse parti di esso.

Un segnale ECG ha un tracciato che si ripete ad ogni ciclo cardiaco composto da:

- onda P: depolarizzazione degli atri
- complesso QRS: depolarizzazione dei ventricoli
- onda T: ripolarizzazione dei ventricoli

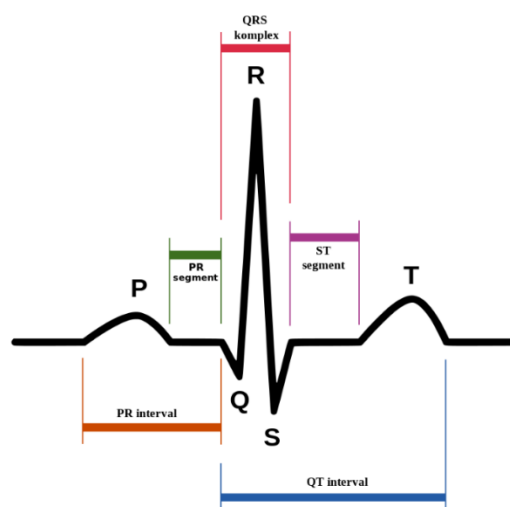


Figura 3.1 Tracciato ECG

Il tracciato ECG può variare in presenza di problemi o anomalie come aritmie, le quali vanno anche a ricadere nei range dei valori di bpm. Le tecniche illustrate per l'estrazione dell'HR sono fotopleletismografia, videopleletismografia, Eulerian Video Magnification (EVM), Face detection e metodo relativo al rapporto LF/HF.

3.1.1 Fotopleletismografia

La fotopleletismografia (Photopleletismography, PPG) è la tecnica per determinare la frequenza cardiaca fondata sullo studio dell'irrorazione dei tessuti (orecchio, polsi, dita delle mani e dei piedi) mediante la diffusione dei raggi e ne misura dati legati all'HR. Il pulsossimetro precedentemente descritto sfrutta questa tecnica e fornisce informazioni legate alla trasmissione (*Figura 3.2*). In base al posizionamento dei sensori otteniamo diverse acquisizioni del segnale. Vado a determinare la trasmissione se il tessuto è posto tra la sorgente e il rilevatore (dita) e determino la riflessione se sono posti adiacentemente (polso).

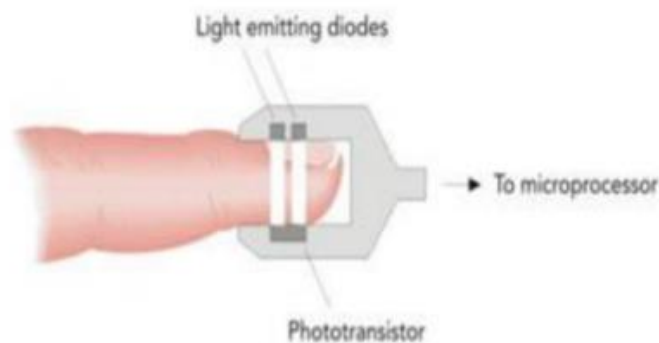


Figura 3.2 Fotopleletismografia pulsossimetro

I segnali PPG risultano precisi e a volte preferibili data la loro non invasività e la facile applicazione anche in ambiti esterni a quello sanitario. Il segnale PPG a differenza del segnale ECG rileva solo le variazioni di pressione sanguigna e non un tracciato con i vari complessi PQRS (*Figura 3.3*). Otteniamo inoltre un ritardo riguardo il picco R in quanto abbiamo un lasso di tempo in cui il sangue deve raggiungere il sito di misurazione.



Figura 3.3 Comparazione tra segnale PPG e segnale ECG

3.1.2 Videoplethysmografia

La tecnica di Videoplethysmografia (Video Plethysmography, VPG) è una tecnica completamente non invasiva e contactless, utilizza un foto-rilevatore come una fotocamera andando a rilevare le variazioni di colore impercettibili ad occhio nudo della pelle causate dal pompaggio del sangue nei vasi sanguigni. Le zone prese in considerazione da questa tecnica sono gli zigomi e la fronte in cui è rilevante in maniera significativa questa variazione.

3.1.3 EVM

La tecnica EVM è utilizzata per estrarre l'HRV e la PSD (densità spettrale di potenza) attraverso l'elaborazione delle immagini ottenute dalla conversione del video registrato. La frequenza cardiaca viene misurata in Hertz (unità di misura): moltiplicando il valore in hertz per 60 (secondi) si ottengono i battiti per minuto (bpm). I video sono stati registrati con la videocamera GoPro Hero6 in modalità 30 frame per secondo. Tramite l'EVM viene amplificato il colore della pelle nelle immagini in corrispondenza di un aumento volumetrico del flusso sanguigno nei vasi dovuto al pompaggio cardiaco (Figura 3.4). L' algoritmo in questione studia la sequenza di fotogrammi eseguendo una decomposizione spaziale ed un filtraggio temporale per captare i suddetti cambiamenti in intervalli temporali specifici. I movimenti e le variazioni di colore vengono in seguito amplificati per riuscire a renderli riconoscibili nel video di output. Nel lavoro di tesi svolto viene sfruttata la correlazione tra cambiamenti di colore della tonalità della pelle nel viso del soggetto e la circolazione del sangue in vene e arterie. Le variazioni sopracitate possono essere utilizzate per determinare la frequenza cardiaca di un individuo e per poter diagnosticare irregolarità nel flusso sanguigno.

L'algoritmo riconosce la variazione periodica nella sequenza di fotogrammi selezionando quelle che ricadono nell'intervallo di frequenze selezionate in input. La variazione di colore è quindi moltiplicata per un fattore di magnificazione α : variabile di ingresso che viene fornita.



Figura 3.4 Amplificazione delle sottili variazioni di colore attraverso EVM

I filtri spaziale e temporale vengono utilizzati per estrarre il segnale del polso cardiaco sfruttando il principio euleriano che asserisce che la pressione e la velocità, in questo caso sanguigna, si sviluppano nel tempo. Attraverso il filtraggio ricavo un'amplificazione del rapporto segnale/rumore ottenendo la scomposizione in diverse bande di frequenza spaziale amplificate della sequenza di fotogrammi. In prima battuta applica un filtro passa-basso utilizzando un sistema di piramidizzazione Laplaciana o Gaussiana per ricevere una sottocampionatura dell'immagine, per ricavare una riduzione della risoluzione e del rumore e allo stesso tempo per avere l'efficacia computazionale desiderata. In secondo luogo applica l'elaborazione temporale per tutte le bande spaziali ed infine applica un filtro passa banda per ottenere i parametri spettrali di interesse. In base al rapporto segnale rumore la sottocampionatura varia tra un pixel e l'altro, mentre l'elaborazione temporale rimane uniforme ai pixel di uno stesso livello spaziale. Il segnale filtrato viene poi moltiplicato per il coefficiente di amplificazione α e sovrapposto al segnale di origine. L'obiettivo dell'analisi è la rilevazione degli impercettibili cambiamenti di colore sulla base della video magnificazione euleriana (EVM) estraendo dal video, quindi dalle immagini, il segnale fotopleletismografico (VPG). Il codice MATLAB utilizzato per questo studio è formato dai vari passaggi che verranno spiegati e analizzati nel capitolo 4.

3.1.4 Face detection

La problematica legata alla rilevazione facciale ed al suo riconoscimento è argomento di grande interesse, resa complessa da diversi fattori ambientali i quali potenzialmente ne ostacolano la riuscita. Nel 2001 Paul Viola e Micheal Jones proposero un algoritmo in grado di classificare all'interno di un'immagine la presenza di uno o più volti. Esso utilizza una singola immagine o una sequenza video facendo scorrere una finestra di ricerca (le cui dimensioni possono variare) estraendone le feature presenti e classificandole come volto o non volto. Vengono determinate e rilevate delle bande luminose che corrispondono a dettagli legati alla fisionomia del volto che quindi se vengono identificate all'interno di un riquadro allora l'algoritmo identifica la presenza di un viso. La rilevazione avviene attraverso quattro passi: funzione di base Haar, creazione di un'immagine integrale, l'algoritmo AdaBoost e l'uso dei rilevatori a cascata. Attraverso il calcolo delle feature di Haar determiniamo delle zone in cui è presente una differenza tra grigi medi e regioni rettangolari per andare a stabilire la presenza di un volto mediante la fisionomia umana indipendentemente dalla posizione, l'espressione e l'orientamento. Quindi la determinazione è resa possibile grazie alla presenza di zone più scure come sopracciglia, narici e zone più chiare come le guance. Quindi col calcolo di feature Haar otteniamo una serie di filtri formati come sommatoria di sotto-parti rettangolari soprannominate "Haar-Like-feature" dell'immagine (*Figura 3.5*).

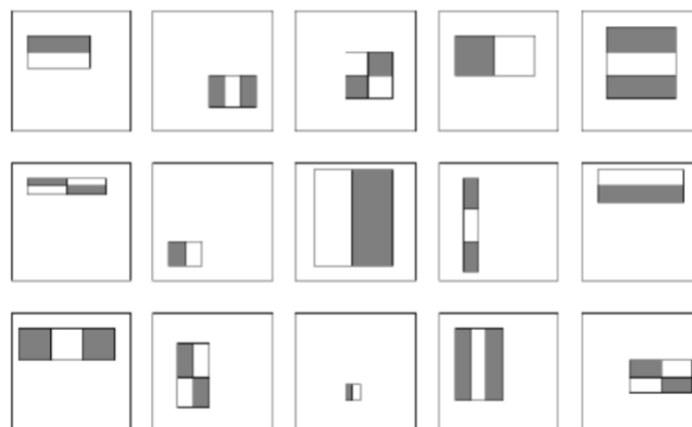


Figura 3.5 Funzioni di rettangoli relative a un certo elemento

Ad esempio in un volto umano abbiamo la presenza di una parte più scura in corrispondenza della regione degli occhi e una parte più chiara in corrispondenza delle guance, perciò una Haar-Like-feature considera due rettangoli adiacenti identificandone uno scuro in corrispondenza dell'occhio e uno chiaro in concomitanza della guancia (*Figura 3.6*).

Il vantaggio di questa tecnica è la sua velocità di calcolo grazie all'utilizzo dell'algoritmo di immagini integrali. Essendo le peculiarità da analizzare un numero elevato, facendo aumentare in maniera esagerata il tempo di computazione, ne segue la necessità di identificare le zone più discriminanti da distinguere.



Figura 3.6 Individuazione dei rettangoli relativi a parti del volto attraverso la funzione di Haar

Questo problema viene ovviato dall'utilizzo dell'algoritmo AdaBoost (acronimo di Adaptive Boosting) il quale è un classificatore "forte" risultante da un insieme di classificatori "deboli". L'Adaboost utilizza alberi decisionali imputati alla determinazione di classificatori per ogni peculiarità del modello: ad ogni iterazione va introdotto un nuovo classificatore debole che ne va a compensare le carenze per indurlo a diventare un classificatore forte. L'ultima parte del Viola-Jones è costituita dai sopracitati classificatori messi a cascata in fasi di cui ogni fase è costituita da uno di essi forte. La fase ha l'obiettivo di concludere la presenza di un viso in una sottofinestra la quale viene scartata in caso contrario.

3.1.5 Estrazione del rapporto LF/HF

Il parametro analizzato nello studio svolto è il rapporto LF/HF ovvero la correlazione tra le basse frequenze e le alte frequenze della densità spettrale di potenza dei picchi RR calcolati nel tempo ottenibile dal segnale VPG e HRV. Il segnale HRV è indicativo della regolazione autonoma della frequenza cardiaca e in alcuni casi potrebbe essere utilizzato come marcatore non invasivo per determinare lo stato fisiologico di un individuo. Il metodo più comune per derivare il segnale HRV è l'acquisizione dell'elettrocardiogramma (ECG), ottenendo una rilevazione del complesso QRS attraverso appropriati algoritmi, localizzando poi i picchi R così da ottenere in seguito gli intervalli RR ed infine eseguire un ricampionamento per andare a costruire il tacogramma. È possibile che a causa di interferenze biologiche o elettromagnetiche

e alla morfologia del segnale ECG il processo può rilevare errori nel segnale HRV. Questo problema è possibile ovviarlo tramite l'utilizzo del segnale VPG (videopleletismografia). La tecnica VPG è una tecnica non invasiva che utilizza una videocamera come fotorilevatore per evidenziare la variazione di colore nelle regioni di interesse per lo studio da effettuare.

In *Figura 3.7* è mostrata la possibilità di dedurre l'attività di stress del soggetto dai valori di HF e LF; vengono prese in considerazione singolarmente i valori di HF e LF per utilizzare tutte le informazioni.

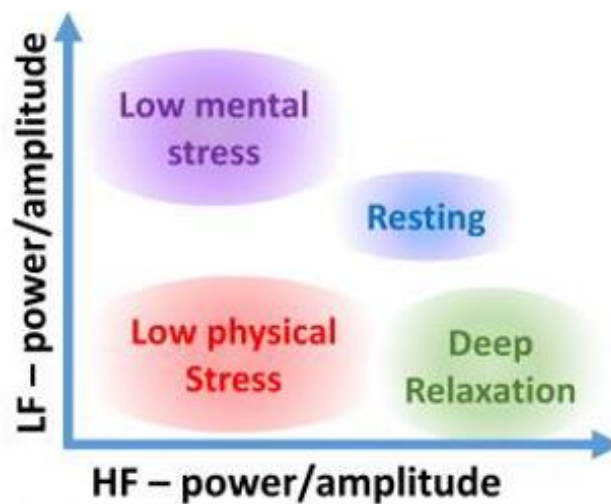


Figura 3.7 Classificazione del livello di stress

Se il valore di LF risulta molto alto rispetto al valore di HF allora il soggetto analizzato mostra uno stato di stress fisico alto; mentre viceversa se il valore di LF è molto basso rispetto HF allora il soggetto esprime uno stato di stress fisico basso fino a raggiungere, per valori elevati di HF, uno stato di profondo rilassamento. Nel caso in cui il rapporto raggiunga il valore 2, ovvero LF è il doppio di HF questo è indice di eventuali patologie o alto livello di stress. Se i due valori non differiscono di molto il dato sta a significare un soggetto in fase di riposo. In questo studio per l'analisi spettrale è stato utilizzato il segnale in uscita dalla Fast Fourier Transformation (FFT) scegliendo l'utilizzo delle sole componenti HF ed LF del segnale HRV.

3.2 Metodi per l'estrazione del diametro della pupilla

Nei seguenti capitoli verranno illustrati i metodi presi in considerazione in questo elaborato per l'estrazione del diametro della pupilla da acquisizioni video.

3.2.1 Eye detection

Per l'estrazione del diametro della pupilla, nel codice utilizzato è inserita una funzione imputata all'individuazione di ROI (zone di interesse). Questa funzione permette di rilevare diversi box per poter selezionare quello da esaminare per tutti frame (*Figura 3.8*). La selezione del frame da analizzare va fatta solo per il primo frame e verrà poi iterato per tutti i successivi. Questa funzione utilizza l'algoritmo di Viola Jones (già descritto nel capitolo 3.1.4).

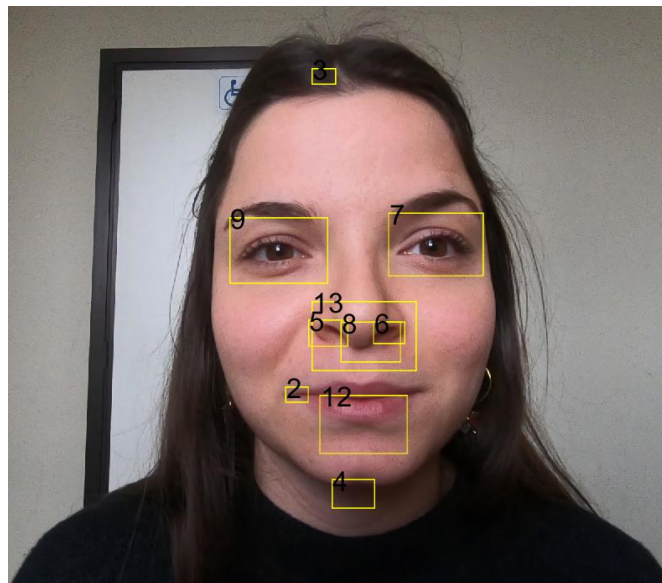


Figura 3.8 Individuazione ROI

3.2.2 Canny edge detection

Dopo aver determinato con Viola-Jones il posizionamento del volto attraverso Canny andiamo ad individuare la regione dell'occhio nell'immagine digitale e più nello specifico il contorno della pupilla. John F. Canny ideò questo algoritmo nel 1986 il quale è un operatore in grado di riconoscere i contorni (edge detection). Questo passaggio costituisce lo stadio preliminare per l'estrazione di caratteristiche utili per l'analisi che verrà fatta, vengono generate delle immagini dei bordi attraverso operatori derivati. Il suddetto prevede quattro passaggi base: smoothing

gaussiano dell'immagine, calcolo del gradiente, ricerca di massimi locali e selezione degli edge significativi mediante isteresi.

Le specifiche che devono essere soddisfatte dall'algoritmo sono:

- a) Capacità nell'individuazione del bordo analizzato: deve marcare nel migliore dei modi e con un elevato grado di precisione i contorni presenti nell'immagine massimizzando il rapporto segnale/rumore avendo un'elevata sensibilità.
- b) Precisa localizzazione: minimizzazione della distanza tra bordo reale e localizzato dall'algoritmo.
- c) Risposta minima: ottimizzazione del rapporto segnale utile/rumore in modo che l'operatore fornisca una sola risposta in concomitanza con un bordo reale.

All'interno dell'algoritmo è stato applicato un filtro a causa della presenza del rumore che si sovrappone all'immagine il quale inficia il riconoscimento del bordo. Il filtro in questione è di tipologia gaussiana (derivata prima di una funzione gaussiana). Eliminando il rumore otteniamo un'immagine con leggera sfocatura però ottenendo così la privazione di rumore per ogni pixel presente nell'immagine. Nel secondo e terzo passaggio (calcolo del gradiente e ricerca dei massimi locali) si va ad utilizzare quattro filtri per delimitare i contorni orizzontali, verticali e diagonali. Nell'ultimo passaggio vado a confrontare ogni pixel con due valori di soglia (uno basso e uno alto) per poter stabilire se il pixel in questione fa parte del contorno oppure no, ottenendo così un'immagine binaria. La difficoltà di questo passaggio sta nell'inserire i due valori che mi classificano il pixel appartenente oppure no al contorno (livello di binarizzazione).



Figura 3.9 Applicazione dell'algoritmo di Canny

3.2.3 Trasformata di Hough

La trasformata di Hough è una tecnica utilizzata per effettuare la rilevazione di linee estratte da immagini digitali. Attraverso questa tecnica nel mio lavoro di tesi ho estratto e successivamente calcolato il diametro pupillare dei soggetti presi in esame. In prima battuta nel 1962 questa tecnica veniva utilizzata per determinare la presenza e il riconoscimento di segmenti rettilinei, in seguito venne ampliato anche ad altre forme geometriche descritte dalle loro equazioni analitiche fino ad arrivare in fase finale ad ogni tipo di poligono. Principalmente il lavoro svolto è quello di passaggio dal piano dell'immagine allo "Spazio di Hough" nel quale vengono considerati dei parametri che ci identificano l'istanza della forma cercata.

Le proprietà della quale la trasformata gode sono:

- ad ogni punto nello spazio dell'immagine reale (SI) corrisponde un'isosuperficie nello spazio di Hough (SP)
- dagli n punti nello spazio dell'immagine appartenenti a una curva si creano lo stesso numero di superfici che si intersecano in un determinato punto nello spazio SP
- ad ogni punto in SP corrisponde una singola istanza della curva in SI
- n punti dello stesso SP appartenenti alla stessa isosuperficie corrispondono ad n punti che si intersecano in un punto in SI.

La trattazione effettuata tramite questa funzione viene utilizzata per rilevare in maniera chiara il cerchio che identifica la pupilla in esame. Considerando l'equazione costitutiva della circonferenza definita:

$$(y - y_c)^2 + (x - x_c)^2 = r^2$$

vengono definiti $C = (x_c; y_c)$ il centro e con r il raggio. Analizzando il caso in cui $C = (0,0)$ e $r = 1$ quindi la circonferenza di equazione $x^2 + y^2 = 1$ otteniamo tre punti come nell'esempio in *Figura 3.10* (parte sinistra) di coordinate $A = (0,1)$ $B = (1,0)$ $C = (0,-1)$. In *Figura 3.10* (parte destra) vengono riportate le tre circonferenze relative ai tre punti scelti in SI nello spazio di Hough.

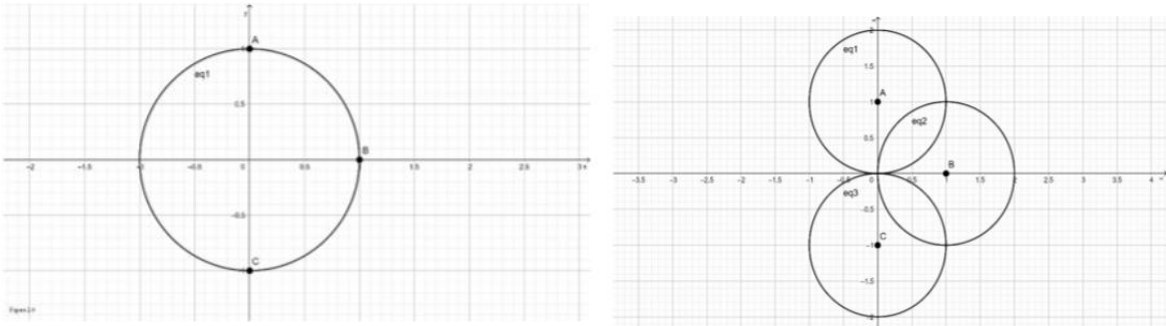


Figura 3.10 Circonferenza con centro origine e raggio unitario in SI e corrispondenti circonferenze legate ai tre punti in SP

L'obiettivo rilevante della trasformata di Hough è quello di identificare il raggio della circonferenza che circoscrive la pupilla in esame. In questo caso lo SP è tridimensionale generando così un cono di equazione:

$$\rho^2 = (C_c - C)^2 + (r - k)^2$$

dove k e C_c sono le coordinate del centro del vertice e ρ è l'apotema del cono. Ottenendo così un cono come in *Figura 3.11*.

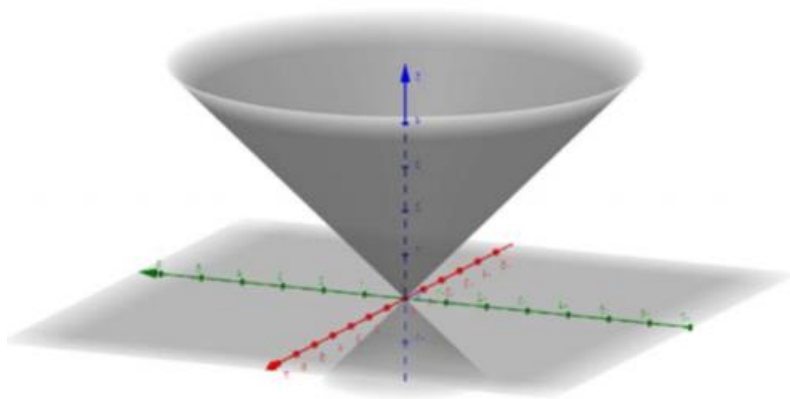


Figura 3.11

In conclusione ogni punto del SI crea una superficie in SP ed il punto di intersezione di più superfici è un indizio sulla presenza di un'istanza della curva che si vuole cercare. Generalmente è necessario un numero di punti pari alla quantità di parametri.

4 Algoritmo di estrazione del battito cardiaco

In questo capitolo verrà fornita un'analisi passo per passo dell'implementazione di un algoritmo e l'utilizzo della tecnica EVM (Eulerian Video Magnification) per poter estrarre la frequenza cardiaca di un soggetto tramite acquisizione video con GoPro Hero6.

4.1 Decomposizione del video in sequenza di immagini

La prima parte del codice MATLAB è imputata alla scomposizione del video in fotogrammi e più precisamente 30 immagini per secondo per un totale di 600 immagini dato che l'analisi viene fatta per un tempo di 20 secondi. Lo script che sviluppa questa funzione è

`“ffmpeg_decompres_script”`.

4.2 Variabili fondamentali

I parametri definiti in questa tecnica che non dipendono dal soggetto esaminato sono:

- “alpha”: coefficiente di amplificazione per il quale viene moltiplicata la saturazione
- “level”: livello di piramidazione della piramide Gaussiana
- “chromAttenuation” rimasto invariato a 1.

Le frequenze di taglio del filtro passa-banda prese per la banda di interesse sono:

- “fh”: frequenza di taglio passa-basso
- “fL”: frequenza di taglio passa-alto.

Le precedenti determinano la banda passante per il battito cardiaco per ottenere così la frequenza di taglio ottimale evitando l'eliminazione di bande del segnale utili. Nel caso in cui il pulsossimetro ci rileva alte frequenze la frequenza di taglio passa-alto f_L si porta a 1.16 Hz (70bpm/60bpm) altrimenti rimane invariata a 1 Hz (60 bpm/60 bpm). La frequenza di taglio passa-basso f_h utilizzata è sempre stata pari a 2,5 Hz.

4.3 Rilevazione volto

In questa parte di codice viene utilizzato l'algoritmo di Viola-Jones che ci determina se nel box identificato dal codice è presente il viso del soggetto in esame: questa rilevazione viene chiamata Face detection (*Figura 4.1*).

In caso affermativo l'algoritmo crea una ROI che avviene solo la prima volta che si esegue il codice per quel determinato soggetto. Attraverso l'algoritmo viene anche richiesto se la Face detection ha rilevato ho meno un volto attraverso una finestra.



Figura 4.1 Face detection e ROI

4.4 Decomposizione piramidale e filtraggio temporale

Nel metodo EVM è importante la combinazione tra l'elaborazione spaziale e quella temporale col fine di evidenziare i cambiamenti temporali invisibili ad occhio nudo all'interno del video. Questa tecnica effettua prima la decomposizione delle immagini in diverse bande di frequenza, le quali potrebbero essere amplificate in maniera diversa in relazione al rapporto segnale-rumore presente; in seguito avviene un filtraggio passa-basso.

Di maggiore utilizzo è il campionamento dall'alto verso il basso delle immagini che prende il nome di piramidazione Gaussiana. Per piramide di un'immagine si intende la stessa in diverse scale ricampionate in seguito fino al punto desiderato.

Esistono due tipi di piramidazione:

- Gaussiana: vengono mantenute sempre le informazioni a basse frequenze ed eliminate sempre più informazione ad alte frequenze. Ad ogni strato si applica un filtro Gaussiano e si sottocampiona il risultato per il successivo strato. Ogni strato della piramide diventa una predizione di quello precedente e la differenza genera lo strato effettivo e quello predetto produce un'immagine errore. La collettività delle immagini errore crea una piramide Laplaciana.

- Laplaciana: risultato di un filtraggio passa banda in cui il primo strato contiene alte frequenze e l'ultimo le frequenze basse, mentre ad ogni strato intermedio viene rispecchiata una certa banda. Attraverso la piramide Laplaciana, formata dalla totalità delle immagini errore, è possibile ricostruire inversamente la piramide Gaussiana in cui l'ultimo strato è l'immagine originale.

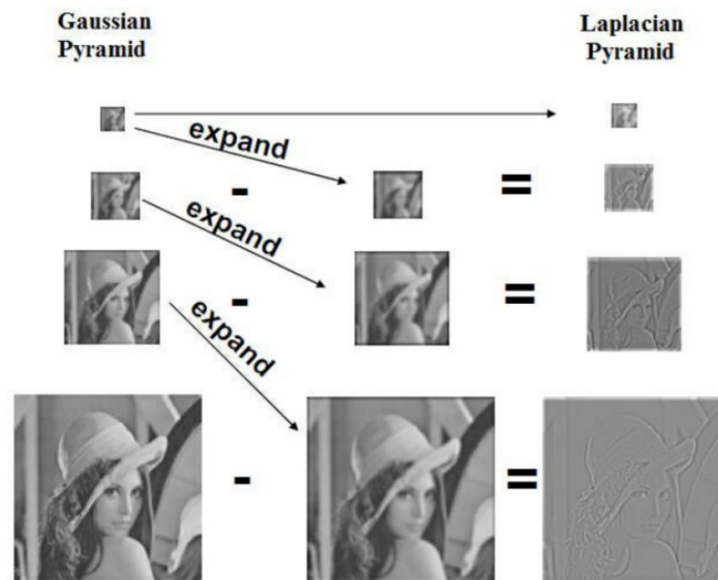


Figura 4.2 Piramidazione Laplaciana e Gaussiana

4.5 Creazione delle ROI

Dopo aver identificato il volto attraverso la Face detection è necessario identificare le zone da cui estrarre il segnale desiderato. Ovviamente non tutte le zone del viso possono essere imputate all'estrazione del ritmo cardiaco, per esempio le sopracciglia non forniscono alcun dato riguardo l'HRV.

Le zone anatomiche che considero nel mio lavoro di tesi che prendono il nome di ROI (Figura 4.3) sono:

- Fronte: "ROI_forehead"
- Guancia destra (zigomo): "ROI_cheecks"
- Guancia sinistra (zigomo): "ROI_cheecks".

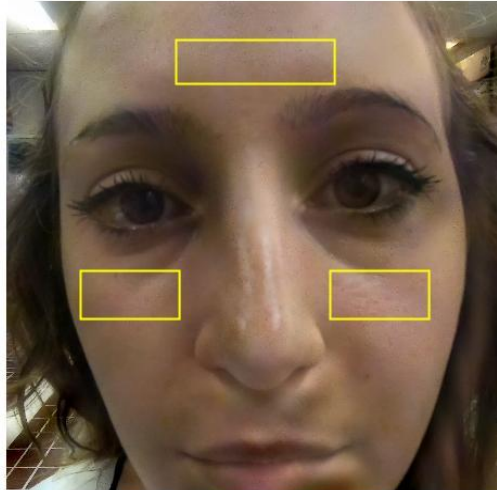


Figura 4.3 ROI: zone anatomiche analizzate

Ogni ROI è soggettiva ad ogni individuo sia in correlazione alla fisionomia del volto e sia alla distanza alla quale si pone il soggetto dalla videocamera. Quindi in maniera del tutto manuale è stata analizzata l'immagine e facendo tentativi si sono modificati i parametri relativi alle coordinate della ROI da analizzare.

4.6 Estrazione del segnale VPG

Al fine di estrarre il segnale VPG si calcola la media dei valori da analizzare dei pixel contenuti nelle ROI, così facendo è possibile ridurre il rumore sul segnale in quanto non vado a considerare i dati di un solo pixel, innalzando così il rapporto segnale-rumore. Il segnale VPG consiste in una rappresentazione nel dominio del tempo delle variazioni di luminanza nelle ROI. I contributi delle ROI sono considerati tutti uguali e pesati in base l'area ricoperta ed ogni punto è ritenuto ugualmente informativo. Dopo aver estratto i valori RGB dei pixel di tutti i frame si effettua una conversione in luminanza e cromaticanza.

4.7 Estrazione HRV e analisi dei picchi

Si individuano i picchi (R) di massimo locale e quindi il battito cardiaco del soggetto in analisi. Con "picco" andiamo ad indicare il massimo dell'onda R del complesso PQRS dell'elettrocardiogramma. I picchi cerchiati di rosso nella *Figura 4.4* sono quelli rilevati dalla funzione di peak detection. Il numero di picchi ottenuti andrà diviso per la durata della

registrazione (rapporto tra numero di frame e frequenza dei fotogrammi) e poi moltiplicata per 60 ottenendo così i bpm.

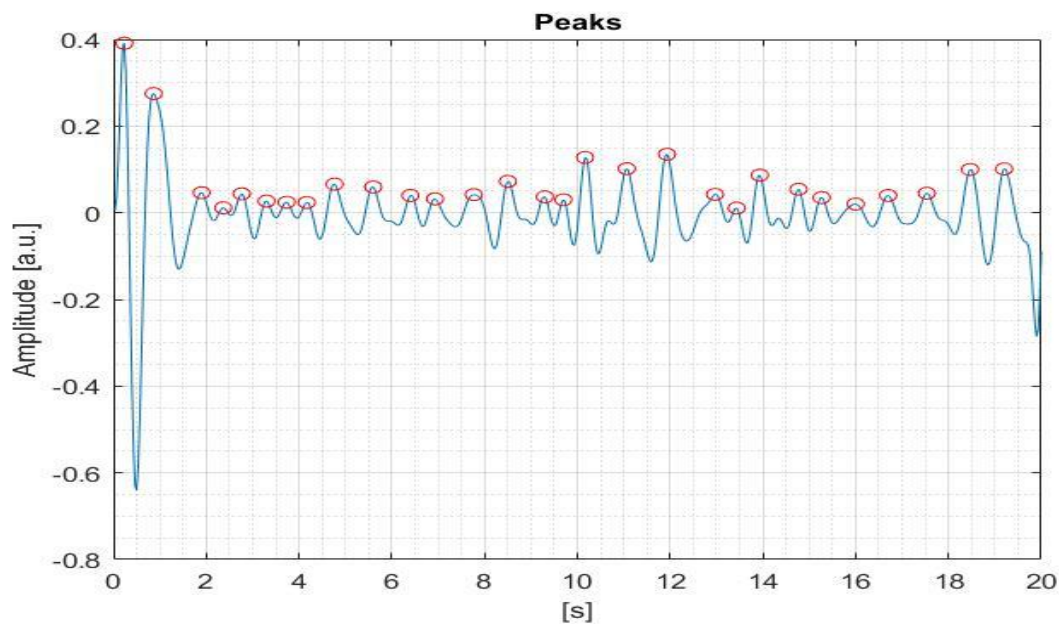


Figura 4.4 Applicazione funzione "findpeaks"

Viene applicato un filtro passa banda di Butterworth così da eliminare il rumore di sottofondo e selezionare la banda di frequenze di interesse utili per calcolare il numero di battiti al minuto. Con la finalità di evitare errori son state imposte delle prerogative per essere presi in considerazione come dei picchi di massimo locale. Prima di tutto il picco per essere considerato deve essere maggiore di una percentuale stabilita dall'utente (ampiezza minima) e la seconda caratteristica è una distanza minima che due picchi consecutivi devono avere al fine di evitare di considerare picchi troppo vicini come se fossero picchi distinti (distanza temporale minima). Per estrarre il picco in frequenza dal quale verrà calcolato il battito cardiaco medio nella sequenza dei 20 s si applica la FFT (Figura 4.5). Un dato di particolare interesse nel lavoro di tesi è il rapporto LF/HF cioè il rapporto tra le bande a bassa e alta frequenza che viene utilizzato per determinare lo stato psicologico del soggetto. Questo dato viene restituito dalla densità spettrale di potenza calcolata con il metodo Welch (Welch's power spectral density).

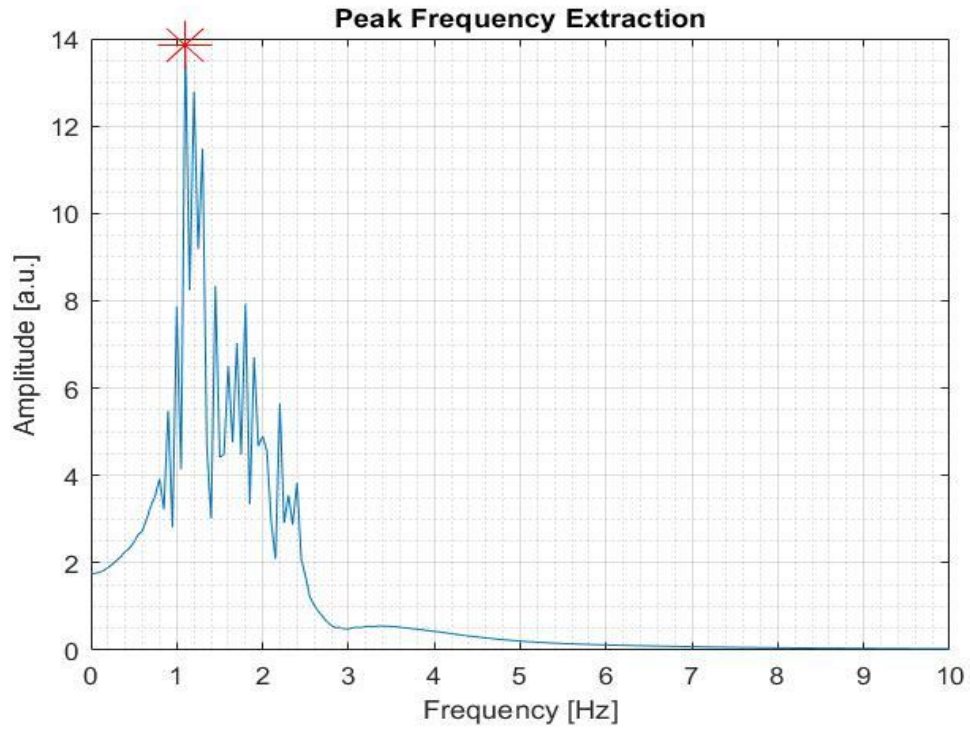


Figura 4.5 Picco massimo

Come mostrato in *Figura 4.5* per estrarre il battito cardiaco viene preso in considerazione il picco di massimo assoluto presente nello spettro moltiplicandola per 60 ottengo il valore in bpm.

5 Algoritmo di estrazione del diametro della pupilla

In questa parte del lavoro di tesi vengono utilizzati i primi 20 secondi del video girato con GoPro Hero6 in cui nei primi 10 secondi viene ripreso il soggetto, con attenzione all'occhio, in condizioni di luce normale di laboratorio, dal secondo 10 al 15 viene accesa una torcia del telefono cellulare posta frontalmente all'individuo e poi ripristinata la luce ambientale nel tempo rimanente. La variazione della pupilla viene affrontata in termini dinamici nei 600 frame generati nei 20 secondi di video in quanto l'acquisizione è sempre in modalità 30 frame per secondo in full HD con risoluzione 1920 x 1080 pixel. Il lavoro svolto avviene attraverso un codice MATLAB alla quale vengono modificati i controlli sperimentalmente considerando i due livelli di binarizzazione nella situazione di luce ambientale e luce della torcia. In concomitanza alla trasformata di Hough sono stati utilizzati altri due metodi di analisi della suddetta variazione: Areabloob che va a valutare l'area relativa al livello di binarizzazione utilizzato e la trasformata dell'ellisse che approssima la forma della pupilla a un'ellisse calcolandone l'area. Le due tecniche verranno poi analizzate e spiegate nei capitoli successivi. Di fondamentale rilievo è la considerazione del fatto che entrambi le tecniche hanno riportato gli stessi risultati durante l'analisi dei diversi soggetti in esame. Il programma individua in primis la regione dell'occhio che corrisponde alla pupilla e attraverso un processo di segmentazione prosegue per i frame successivi, ottenendo una curva per ogni metodo sopracitato che esprime la variazione del diametro della pupilla. Dalla curva è possibile osservare oscillazioni con picchi positivi e negativi in relazione all'aspetto da noi preso in esame, ma allo stesso tempo possono anche essere dovuti a movimenti arbitrari e involontari da parte del soggetto. Nei successivi capitoli verranno illustrati i vari passaggi, le relative righe di codice e i grafici ottenuti dalle tecniche descritte.

5.1 Decompressione del video in immagini

In questa prima fase viene decompresso il video in immagini raccolte in una cartella directory di output, nel nostro caso "bpmdata". Dato in ingresso un video di 20 secondi e impostando il frame rate a 30 otteniamo 600 frame.

5.2 Variabili fondamentali

Prima di lavorare con il codice MATLAB vanno identificate e definite alcune variabili e parametri fondamentali. Il codice richiede il frame dal quale partire l'analisi "startIndex" e il frame finale "stopIndex", dato che il periodo del video analizzato attraverso questo codice sono i primi 20 secondi l'intervallo di frame da analizzare è [1:600]. I frame RGB vengono inizialmente convertiti in scala di grigi per ottenere una minimizzazione del tempo di elaborazione e in seguito binarizzate secondo due livelli di binarizzazione: uno in caso di torcia accesa "reduced" e uno in presenza della sola luce del laboratorio "normal". Vengono determinati effettuando dei tentativi e valutando la corretta delimitazione della pupilla durante lo scorrimento del codice per tutti i frame portando una maggiore attenzione in corrispondenza del passaggio tra luce normale e torcia accesa. Questi due livelli di binarizzazione fanno corrispondere al box selezionato una matrice composta da 0 (nero) e 1 (bianco) in questo modo passiamo da un'immagine in scala di grigi ad una composta solo da bianco e nero. Attraverso il parametro "radii" viene stabilito il range di lavoro della trasformata di Hough, questo dato è stato stimato empiricamente in maniera manuale andando a misurare manualmente e grossolanamente la pupilla del soggetto studiato. "calibrationFactor" è il fattore di passaggio da pixel a millimetri.

5.3 Eye detection e taglio della pupilla

In questa parte di codice avviene la ricerca e la selezione del box relativo all'occhio che vogliamo analizzare, il codice identifica diversi box all'interno del primo frame e l'operatore dovrà fornire in input il numero del box per il quale il codice dovrà effettuare l'analisi (*Figura 5.1*). Questo lavoro di identificazione va fatto solo per il primo frame e poi il codice itererà per i seguenti.

La funzione imputata a questo lavoro è "vision.CascadeObjectDetector".

È possibile modificare le dimensioni del box manualmente per poter escludere parti che disturbano l'identificazione della pupilla, durante il lavoro le parti più interferenti con l'identificazione sono il trucco e le sopracciglia, in quanto entrambi possono confondere il codice.

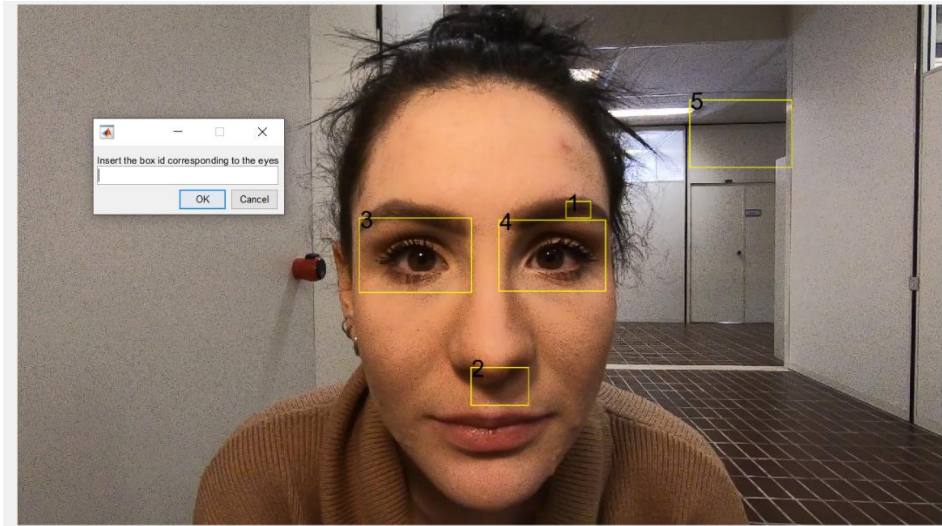


Figura 5.1 Identificazione Roi e finestra per inserire il box da analizzare

5.4 Binarizzazione immagine e inizializzazione del ciclo

Come già accennato nel capitolo 4.2 e dopo aver selezionato il box che il codice deve considerare durante il lavoro di analisi della variazione pupillare vanno inseriti manualmente due livelli di binarizzazione distinti che il codice dovrà considerare in concomitanza dei frame con illuminazione del laboratorio (“normal”) e quelli con l’illuminazione del telefono cellulare (“reduced”). Fornendo in ingresso diversi livelli di binarizzazione per uno stesso frame può causare una esagerata rilevazione di nero nell’immagine che non sempre corrisponde alla pupilla, in *Figura 5.2* è possibile notare come due diversi livelli di binarizzazione per una stessa pupilla variano consistentemente la dimensione rilevata; il giusto livello di binarizzazione è quello considerato nella parte “a”.

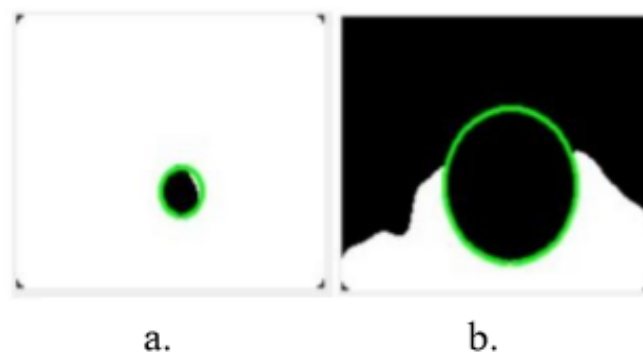


Figura 5.2 Diversi livelli di binarizzazione per una stessa pupilla

L'individuazione dei perfetti livelli da utilizzare dipende della fisionomia del soggetto in quanto potrebbe presentare delle ciglia/sopracciglia molto scure (a volte dovute anche ai cosmetici utilizzata) che quindi vanno a interferire con la distinzione della pupilla dal resto (*Figura 5.3*).



Figura 5.3 Binarizzazione di una pupilla con presenza di sopracciglia che potrebbe disturbare la rilevazione

La funzione imputata alla binarizzazione dell'immagine è "rgb2gray" che permette il passaggio dal formato RGB ad un'immagine in scala di grigi. Successivamente viene applicato un filtro che consente l'eliminazione del riflesso presente in essa. Il modello RGB (Red-Blue-Green) consiste nell'addizione dei tre colori primari, alla loro massima saturazione, per ottenere il bianco. Viene basato sui tre colori primari (da quali prende nome ovvero rosso, verde e blu) in cui un'immagine viene scomposta attraverso l'applicazione di filtri e grazie alla loro miscelazione consentono l'ottenimento di tutto lo spettro dei colori visibili. La binarizzazione utilizza la funzione "imbinarize" la quale prende in ingresso la variabile "level" che assume in base ai comandi inseriti nel codice i valori "normal" o "reduced". I comandi nominati in precedenza indirizzano il codice ad utilizzare dal frame 301 al 450 il livello imputato alla situazione di illuminazione della torcia del telefono cellulare e nei restanti frame quello con la luce del laboratorio. Per motivi tecnici dal punto di vista del metodo sperimentale per cui la torcia veniva applicata dall'operatore e quindi non ottenevamo un perfetto tempismo in cui fornivamo la fonte luminosa nel frame 301 esatto, ci siamo trovati a dover variare i controlli precedenti andando ad inserire il numero del frame preciso in corrispondenza dell'applicazione di essa. A causa di questo problema capitava che l'algoritmo in cui non veniva inserito con precisione il frame in questione, e che quindi rimaneva a priori a 301, non riusciva a rilevare per qualche frame il contorno della pupilla e andava fuori scala (*Figura 5.4*).

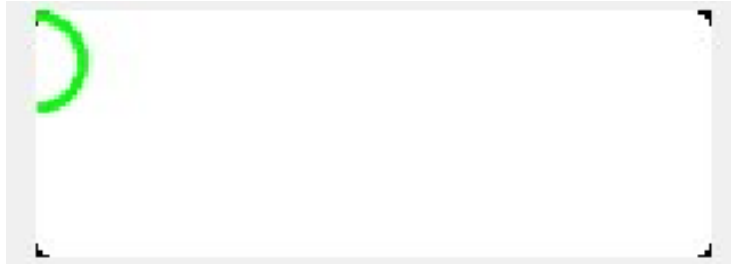


Figura 5.4 Rilevazione fuori dal box al momento di cambio illuminazione

5.5 Segmentazione pupilla

In prima battuta in questo passaggio vengono definiti attraverso l’algoritmo di Canny (3.2.2) e quindi la funzione “edge” i contorni dell’immagine procedendo al rilevamento del diametro della pupilla. Si parte con la creazione della variabile

```
circle=edge(ap, 'canny');
```

che appunto sfrutta la canny edge detection, successivamente viene utilizzata la trasformata di Hough che frame dopo frame riconosce il cerchio contenente la pupilla con il raggio compreso nel range dalla variabile “radii”. Il vettore che risulta da Hough è denominato “peaks” (Figura 5.5) ed esprime le coordinate del centro del cerchio e la dimensione del raggio in pixel.

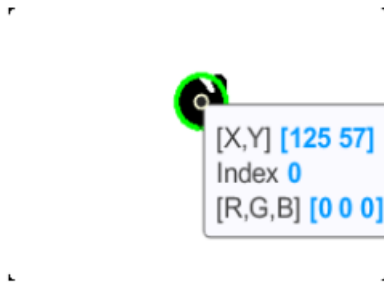


Figura 5.5 Coordinate del centro del cerchio

5.6 Blob ed equazione ellisse

In questo capitolo verranno illustrate le due metodologie usate in aggiunta alla trasformata di Hough per la determinazione del diametro della pupilla. Le tre trattazioni verranno poi confrontate nelle conclusioni.

5.6.1 Circle Blob

Il metodo del diametro equivalente o Circle Blob è un metodo alternativo per determinare il diametro della pupilla, in sostanza viene estratto da una stima dell'area del blob. Parte da un ciclo for che consente l'individuazione di tutte le macchie presenti nel box, a questo punto attraverso le funzioni BlobArea, BlobPerimeter e BlobCentroid individua area perimetro e centro della macchia in questione. Una volta resi noti questi parametri viene implementata la funzione blobECD (Equivalent Circular Diametre) il quale restituisce il diametro della pupilla.

5.6.2 Metodo dell'ellisse

Un'altra tecnica che porta alla determinazione del diametro della pupilla è la tecnica dell'ellisse che verte sull'assunzione per la quale approssima la pupilla ad un'ellisse calcolandone i parametri di interesse.

Le funzioni imputate sono "fit_ellipse" e "calculateEllipse". Per il calcolo del raggio l'algoritmo effettua una media dei due assi e poi attraverso un'equazione lo definisce utilizzando il fattore di calibrazione precedentemente indicato.

5.7 PSD

La Densità Spettrale di Potenza è l'energia fornita dal segnale in frequenza; in questo lavoro di tesi viene utilizzata per osservare il rapporto tra LF (basse frequenze) e HF (alte frequenze). I dati relativi al rapporto sopraindicato ci forniscono dati relativi all'HRV che in seguito verranno confrontati con l'altro tipo di lavoro svolto. La stima della PSD viene fornita dal codice tramite la funzione "pwelch". Vengono calcolate due stime in base alla situazione in cui il soggetto si trova: una quando il soggetto non è stimolato dalla torcia del telefono cellulare (*Figura 5.6*) e una quando viene fornita luce (*Figura 5.7*).

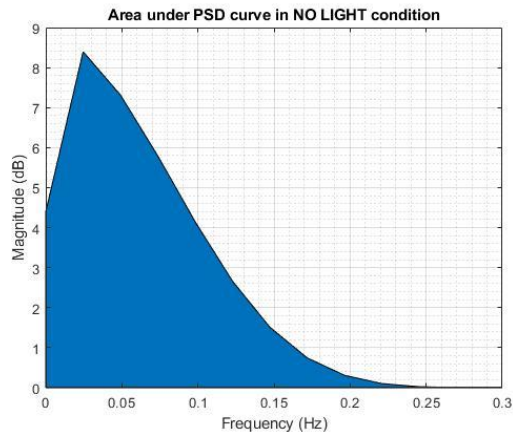


Figura 5.6 Area PSD In condizioni di non luce

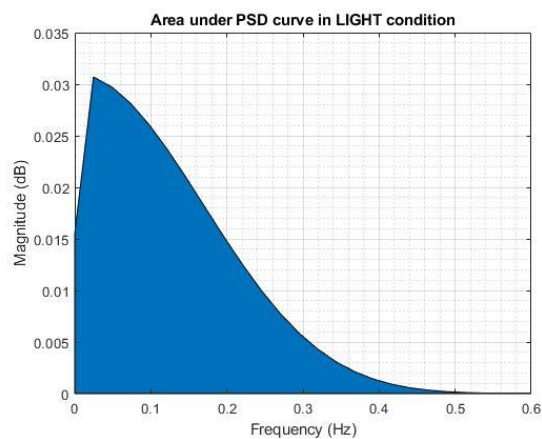


Figura 5.7 Area PSD in condizioni di luce

5.8 Plot dei grafici

Il passaggio finale per ottenere un'illustrazione della variazione del diametro pupillare nel tempo è la costruzione del grafico. Nel grafico vengono mostrate in tre diversi colori le tecniche utilizzate così da poter attuare un confronto visivo (Figura 5.8).

Le tre diverse tecniche utilizzate sono:

- Trasformata di Hough (magenta)
- Metodo dell'Area del Blob (nero)
- Metodo dell'ellisse (verde)

È possibile notare come le prime due tecniche utilizzate si sovrappongano in ogni punto mentre la terza mostra un leggero scostamento. È possibile osservare dalla Figura 5.8 che al secondo 10 il diametro della pupilla mostra una notevole diminuzione della dimensione dovuta, come

precedentemente affermato, all'introduzione della torcia del telefono cellulare; tornando poi a dimensioni originali dal secondo 15 fino a 20.

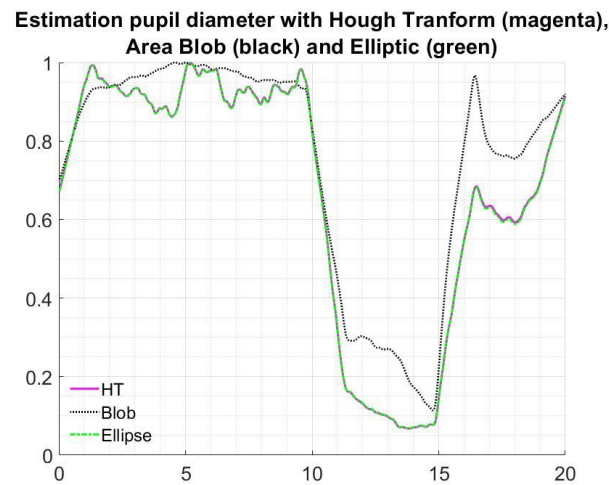


Figura 5.8 Variazione diametro pupillare in relazione al tempo in secondi

Il grafico risultante a volte può mostrare delle imprecisioni dovute ad artefatti da movimento del soggetto spesso causati da una chiusura degli occhi involontaria durante il processo. Per questo motivo a volte vengono applicati due filtri chiamati “midfilt” e “smooth” per smussare picchi sia negativi che positivi non legati alla variazione studiata.

6 Risultati

L'elaborato di tesi è stato incentrato sul calcolo del rapporto tra le componenti ad alta e bassa frequenza dell'HRV utilizzando due diversi tool. Nel metodo della variazione pupillare sono state utilizzate a loro volta tre diverse tecniche per ottenere una rappresentazione grafica della variazione della dimensione della pupilla in funzione del tempo. Argomento di discussione iniziale è stato l'ambiente in cui effettuare l'acquisizione video. Inizialmente si è partiti utilizzando la luce artificiale del laboratorio ma essendo posta nel soffitto non ci forniva una perfetta illuminazione frontale del soggetto e quindi anche in funzione della fisionomia dello stesso non sempre la zona di interesse veniva chiaramente visualizzata: a volte accadeva che l'occhio veniva parzialmente oscurato dalle ciglia. Successivamente si è provato ad effettuare rilevazioni ponendo il soggetto frontalmente a una finestra ottenendo così una perfetta illuminazione del viso; il problema che è scaturito in questo caso è stata la diminuzione consistente della variazione del diametro della pupilla nei due casi che dovevamo considerare fino ad ottenerne una quasi irrilevante nel caso di soggetti con occhi chiari in quanto la loro pupilla risultava già ridotta senza applicare la torcia del telefono cellulare (*Figura 6.1*).

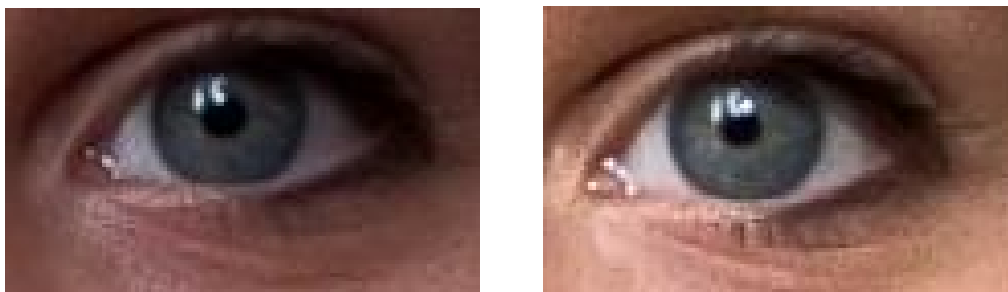


Figura 6.1 Soggetto con occhi chiari posto frontalmente alla finestra in condizioni senza (sx) e con torcia (dx)

Di seguito verrà mostrato in *Figura 6.2* il grafico che mostra l'andamento del diametro pupillare rilevato dal codice il quale non presenta una consistente diminuzione in corrispondenza del secondo decimo; inoltre sono presenti numerose oscillazioni dovute alla difficoltà che il soggetto presentava a tenere gli occhi aperti con un'esposizione luminosa così elevata.

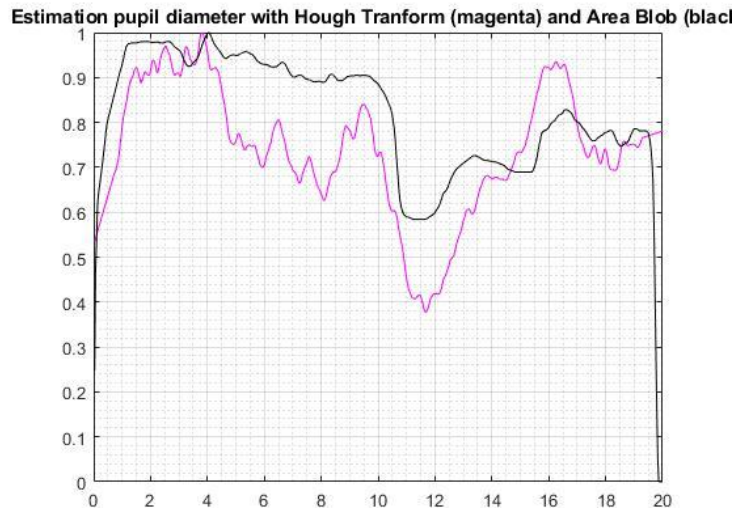


Figura 6.2 Andamento nel tempo della dimensione pupillare di un soggetto con occhi chiari

A questo punto si è deciso di procedere con le acquisizioni all'interno del laboratorio cercando di posizionare tutti i soggetti in maniera tale che il viso non presentava zone d'ombra in corrispondenza della cavità oculare. In questo caso i soggetti che meglio esprimevano la diminuzione in maniera considerevole erano quelli con gli occhi di tonalità chiare. In *Figura 6.3* verrà mostrato come sia percettibile anche ad occhio nudo la variazione nelle due condizioni.

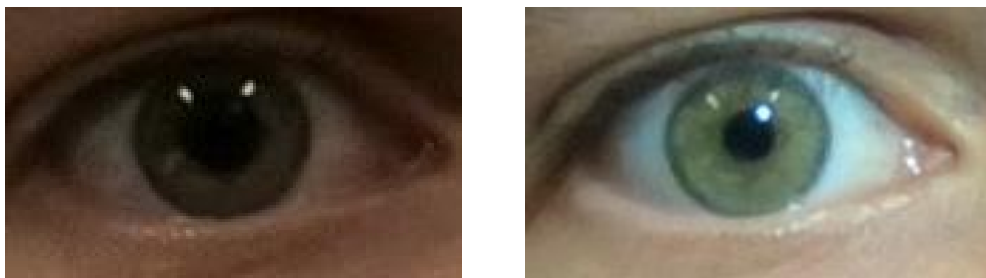


Figura 6.3 Soggetto con occhi chiari in laboratorio in assenza (sx) e in presenza (dx) di torcia telefono cellulare

Di seguito verrà riportato il grafico di uno dei casi migliori, in cui viene mostrata la variazione del diametro della pupilla in funzione del tempo nelle tre diverse tecniche utilizzate: di verde viene mostrata in funzione del metodo dell'ellisse, in nero per l'Area Blob e in magenta attraverso la trasformata di Hough. È importante notare che la curva relativa alla trasformata di Hough è identica punto per punto a quella del metodo dell'ellisse. Il metodo con l'AreaBlob risulta più robusto a interferenze e quindi fornisce valori più approssimati.

Estimation pupil diameter with Hough Tranform (magenta), Area Blob (black) and Elliptic (green)



Figura 6.4 Andamento temporale della dimensione della pupilla nel tempo

Come spiegato in precedenza il repentino calo a 10 secondi è il risultato dell'applicazione della torcia del telefono cellulare fino a 15 secondi. Le oscillazioni che possiamo trovare nelle curve sono dovute ad eventuali chiusure dell'occhio durante la registrazione da parte del soggetto. Nella tabella di seguito verranno forniti i valori ottenuti dal lavoro di tesi mediante questa tecnica, nella prima colonna sono presenti i due livelli di binarizzazione (normale e reduced) utilizzati soggetto per soggetto e determinati sperimentalmente, nella seconda le pulsazioni che il soggetto presentava e nella terza il rapporto LF/HF che il codice ha rilevato.

<i>SOGGETTO</i>	<i>LIVELLI DI BINARIZZAZIONE</i>	<i>BPM</i>	<i>LF/HF</i>
S01	0,052-0,07	90	1,041
S02	0,062-0,1	84	1,006
S03	0,04-0,09	66	0,861
S04	0,04-0,085	78	0,993
S05	0,05-0,1	62	1,040
S06	0,045-0,075	85	1,0098
S07	0,07-0,25	54	1,037
S08	0,06-0,11	95	1,046
S09	0,07-0,12	96	1,038

Tabella 1 Livelli di binarizzazione, pulsazioni e LF/HF dallo studio della variazione del diametro pupillare

Per quanto riguarda lo studio attraverso la tecnica EVM l'acquisizione dei video poteva essere ostacolata dalla presenza di capelli, che copriva o creava ombra sulla fronte del soggetto, ma era sufficiente spostarli o fissarli con una molletta. Altro problema che poteva interferire con il segnale era la previa applicazione di cosmetici coprenti sia nella zona frontale ma principalmente sulle guance, in tal caso si richiedeva la rimozione. Un esempio di soggetto in cui sono ben visibili le zone di interesse è quello in *Figura 6.5*.



Figura 6.5 Soggetto per tecnica EVM

Nella *Tabella 2* vengono riportati nella prima colonna i battiti per minuti rilevati dal pulsossimetro, nella seconda quelli ricavati attraverso la tecnica EVM e nella terza il rapporto LF/HF che la tecnica EVM ha rilevato.

<i>SOGGETTO</i>	<i>BPM</i> <i>PULSOSSIMETRO</i>	<i>BPM EVM</i>	<i>LF/HF</i>
S01	89	90	0,994
S02	84	84	0,8977
S03	66	66	0,9656
S04	78	78	0,9625
S05	63	63	0,468
S06	75	75	0,996
S07	54	54	0,936
S08	90	90	0,9169
S09	96	96	0,967

Tabella 2 Calcolo della frequenza cardiaca e LF7HF mediante tecnica EVM

Infine verrà impostata una tabella dove soggetto per soggetto sono inseriti i dati del rapporto ad alta e bassa frequenza sia del codice relativo alla variazione del diametro pupillare e sia attraverso la tecnica EVM per poter poi sviluppare delle considerazioni nel capitolo delle conclusioni.

<i>SOGGETTO</i>	<i>LF/HF PUPILLA</i>	<i>LF/HF EVM</i>
S01	1,041	0,994
S02	1,006	0,8977
S03	0,861	0,9656
S04	0,993	0,9656
S05	1,040	0,468
S06	1,0098	00,996
S07	1,037	0,936
S08	1,046	0,9169
S09	1,038	0,967

Tabella 3 Confronto LF/HF tra i due codici

Conclusioni

La prima considerazione in ordine di importanza per lo studio svolto è la rilevazione dei valori di LF/HF maggiori per il codice relativo alla pupilla rispetto la tecnica EVM, in quanto il primo codice veniva sottoposto ad una condizione di stress legato all'applicazione della torcia del telefono cellulare.

Dal lavoro svolto è evidente la correlazione tra diametro della pupilla e la quantità di luce presente, in quanto all'aumentare dell'intensità luminosa diminuisce il diametro pupillare.

Le maggiori fonti di errore verificate durante tutto lo svolgimento, in particolare durante le acquisizioni video, sono dovute principalmente a movimenti da parte del soggetto, sia a livello della posizione del viso, per quanto riguarda la tecnica EVM, sia per la chiusura involontaria degli occhi per la tecnica di rilevazione del diametro pupillare.

Per ovviare a questi problemi, in primo luogo, si potrebbe utilizzare un sostegno che ci permette di mantenere stabile e immobile il viso cosicché il codice non deve continuamente ricalcolare la posizione delle ROI. In secondo luogo va richiesta una maggiore attenzione da parte del soggetto in esame nel fissare un punto a favore di camera per poter mantenere inalterata la direzione della pupilla così da avere sempre uno stesso livello di illuminazione focalizzato dall'occhio. Un'ulteriore attenzione che dovrebbe avere il soggetto è quella di riuscire a chiudere il meno possibile le palpebre perché in questo modo la pupilla, nel momento di riapertura, si trova con un diametro maggiore dovuto all'assenza di luminosità contratta nel momento della chiusura dell'occhio.

Un'accortezza aggiuntiva da parte dell'utente per svolgere questo studio è quella di poter oggettivare la luce presente nell'ambiente di rilevazione; questo è reso possibile tramite applicazioni che ci forniscono la misura in LUX rendendo l'esperimento ripetibile e i risultati paragonabili.

Per quanto riguarda la tecnica EVM abbiamo riscontrato un problema sperimentale legato all'acquisizione continua di quaranta secondi per entrambi gli esperimenti in cui è stato necessario una ripresa ad una determinata distanza videocamera-soggetto legata alla perfetta messa a fuoco e rilevazione della pupilla che però allo stesso tempo ha costretto l'eliminazione di una ROI della tecnica EVM.

La distanza di quaranta centimetri è risultata sufficiente per poter mettere a fuoco la pupilla e allo stesso tempo consentire alla tecnica EVM di rilevare il viso; ma nella tecnica EVM le ROI da considerare dovrebbero essere la fronte, gli zigomi e il collo. Non è stato possibile l'utilizzo della ROI del collo per il motivo sopracitato, la quale sarebbe stata di grande rilievo in quanto

è presente la carotide che possiede un importante flusso di sangue rispetto ai capillari nelle altre ROI.

Bibliografía

- [1] Takegami T., Gotoh T. & Ohyama G., “An algorithm for model-based stable pupil detection for eye tracking system “, 2004.
- [2] Micheal J. Jones, Paul Viola, “Robust Real-Time Face Detection”, 2004.
- [3] Modesto Castrillón, Oscar Déniz, Daniel Hernández, Javier Lorenzo, “A comparison of face and facial feature detectors based on the Viola–Jones general object detection framework”, 2010.
- [4] Yang M., Crenshaw J., Augustine B., Mareachen R., Wu Y., “AdaBoost-based face detection for embedded systems”, 2010.
- [5] Qiang Ji and Xiaojie Yang, “Real-Time Eye, Gaze, and Face Pose Tracking for Monitoring Driver Vigilance”, 2002.
- [6] Weiwei Liu, Haixin Sun, & Weijie Shen, “Driver fatigue detection through pupil detection and yawing analysis” 2010.
- [7] Jeong D. S., Hwang J. W., Kang B. J., Park K. R., Won C. S., Park D. K., & Kim J, “A new iris segmentation method for non-ideal iris images.”, 2010.
- [8] Masao Ohtani and Yoshiobu Ebisawa “Eye-gaze detection based on the pupil detection technique using two light sources and the image difference method”.
- [9] Von Rosenberg W., Chanwimalueang T., Adjei T., Jaffer U., Goverdovsky V. & Mandic D. P., “Resolving Ambiguities in the LF/HF Ratio: LF-HF Scatter Plots for the Categorization of Mental and Physical Stress from HRV”, 2017.
- [10] Avinash Parnandi and Ricardo Gutierrez-Osuna “Contactless Measurement of Heart Rate Variability from Pupillary Fluctuations”, 2013.

- [11] *Hamidur Rahman, Mobyen Uddin Ahmed and Shahina Begum, "Non-Contact Physiological Parameters Extraction Using Facial Video Considering Illumination, Motion, Movement and Vibration", 2020.*
- [12] *Sheng Lu, He Zhao, Kihwan Ju, Kunsoo Shin, Myounggho Lee, Kirk Shelley and Ki H. Chon, "Can photoplethysmography variability serve as an alternative approach to obtain Heart Rate Variability information?" 2008.*
- [13] *Ahmed Alqaraawi , Ahmad Alwosheel, Amr Alasaad, "Heart rate variability estimation in photoplethysmography signals using Bayesian learning approach", 2016.*
- [14] *Karim Alghoul, Saeed Alharthi, Hussein Al Osman, Abdulmotaleb El Saddik, "Heart Rate Variability Extraction From Videos Signals: ICA vs. EVM Comparison", 2017.*
- [15] *Christian Küblbeck, Andreas Ernst, "Face detection and tracking in video sequences using the modified census transformation", 2005.*
- [16] *Kun Wang, Haifeng Zhang, "Research on an Improved Algorithm of Face Detection Based on Skin Color Features and Cascaded Ada Boost".*
- [17] *Bacha Rehman, Wee Hong Ong, Abby Chee Hong Tan, Trung Dung Ngo, "Face detection and tracking using hybrid margin-based ROI techniques", 2019.*