



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA



**DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE**

**CORSO DI LAUREA IN INGEGNERIA BIOMEDICA**

**“Gesture recognition per interfacce uomo-macchina: una panoramica dei sistemi e delle applicazioni”**

**Relatore: Prof. / Dott Sarah Tonello**

**Laureando/a: Gumierato Giacomo**

**ANNO ACCADEMICO 2021–2022**

**Data di laurea 21/07/2022**

# Abstract

Nel mondo moderno la tecnologia sta diventando sempre più preponderante nella vita dell'uomo, con dispositivi aventi un numero crescente di funzionalità e che quindi necessitano di un numero altrettanto crescente di comandi, che, sul lungo termine, potrebbero risultare scomodi e poco spontanei per l'utente finale. Uno sviluppo di un'interfaccia tra uomo e macchina intuitiva e semplice può portare ad una semplificazione drastica dell'utilizzo dei dispositivi elettronici permettendo alle macchine una comprensione più diretta del linguaggio verbale umano. Lo scopo di questo elaborato è quello di analizzare una possibile soluzione a questo problema mediante un sistema di riconoscimento dei gesti, offrendo una panoramica delle applicazioni (il riconoscimento dei gesti stessi, il comando di attuatori a distanza e la realtà virtuale) e degli aspetti tecnologici principali da considerare durante la progettazione di un sistema di *gesture recognition* come i sensori (in questo elaborato verranno trattati accelerometri, giroscopi ed estensimetri) , il sistema di acquisizione o modulo DAQ, il trasferimento dei dati (mediante Bluetooth, Wi-Fi ed NFC) e la loro successiva elaborazione con algoritmi deterministici, correlazioni tra dati e apprendimento automatico. Nell'elaborato viene presentata anche una dimostrazione pratica di un sistema di tracking della mano che, utilizzando l'app Phyphox che sfrutta i sensori presenti nello smartphone per effettuare misure dell'accelerazione del telefono collegato alla mano per poi elaborare i dati con Matlab, è in grado di rappresentare il movimento in uno spazio tridimensionale.

# Indice

<b>1. Introduzione ed applicazioni</b> .....	1
1.1. I gesti e le loro caratteristiche.....	1
1.2. Storia dell'interazione uomo-macchina.....	3
1.3. Applicazioni.....	6
1.3.1. Riconoscimento dei gesti.....	7
1.3.2. Comando di attuatori esterni.....	10
1.3.3. Realtà Virtuale.....	12
<b>2. Struttura Hardware</b> .....	15
2.1. Sensori.....	15
2.1.1. Accelerometri.....	16
2.1.2. Giroscopi.....	20
2.1.3. Estensimetri.....	22
2.1.4. Telecamera.....	24
2.2. Sistema di acquisizione.....	25
2.3. Trasmissione dati.....	30
2.3.1. Sistema Bluetooth.....	30
2.3.2. Wi-Fi.....	32
2.3.4. NFC.....	34
<b>3. Analisi ed elaborazione dei dati</b> .....	35
3.1. Estrazione di features.....	36
3.2. Correlazione tra segnali.....	40
3.3. Machine learning.....	42
<b>4. Movement tracking: un'esempio pratico</b> .....	49
4.1. Acquisizione tramite Phyphox.....	50
4.2. Elaborazione dati e risultati.....	52
<b>5. Conclusioni</b> .....	65
<b>6. Ringraziamenti</b> .....	67
<b>7. Riferimenti</b> .....	69

# 1. Introduzione ed applicazioni

La tecnologia sta diventando sempre più preponderante nella vita di tutti i giorni, grazie allo sviluppo dei sistemi di connessione Wi-Fi viviamo costantemente connessi ad internet e possiamo sfruttare computer o smartphone per svariate attività (e.g. acquisti online, effettuare operazioni bancarie, comunicare con altre persone e guardare la nostra serie TV preferita). L'enorme numero di funzionalità che possono offrire questi dispositivi richiede però un altrettanto numero di comandi per poter eseguire tali operazioni, che, sul lungo termine può portare ad avere interfacce particolarmente complesse e poco intuitive.

Lo studio e l'ideazione di interfacce *user-friendly* che facilitino la comunicazione uomo-macchina, abbattendone i confini e rendendola molto simile alla comunicazione tra esseri umani, risulta quindi un'area di particolarmente rilevante, e di interesse nel mondo attuale.

Un possibile mezzo per semplificare l'interazione tra una persona e una macchina è rappresentato dai gesti, che costituiscono una parte del linguaggio del corpo e sono la parte più studiata della comunicazione non verbale, proprio perché sono facilmente osservabili.

## 1.1. I gesti e le loro caratteristiche

I gesti presentano una forma di comunicazione semplice ed intuitiva: i gesti con le mani, in particolare, si possono osservare durante una normale conversazione e possono rivelare molte informazioni riguardo lo stato d'animo, i pensieri e anche le intenzioni di un soggetto. Si può pensare al fatto che parlare con le mani nelle tasche può rappresentare un indice di insicurezza oppure che una persona che si tocca il collo, la bocca o il naso potrebbe mentire riguardo alcune sue affermazioni.



I gesti possono essere classificati in due principali categorie:

- *Statici*: L'utente (in questo caso la mano) assume determinate configurazioni/pose in un dato istante
- *Dinamici*: Successione temporale di più gesti statici

Con l'avvento di tecnologie indossabili è stata facilitata la comunicazione tra esseri umani e computers costituendo una possibile interfaccia uomo-macchina (HMI), permettendo in futuro l'utilizzo di questa tecnologia nell' *internet of things* (IoT) in un modo estremamente semplificato e intuitivo.

I gesti possono infatti essere acquisiti mediante dispositivi indossabili, aventi varie tipologie di sensori (e.g. elettromiografici, magnetici, inerziali, di deformazione, giroscopi ecc.) oppure mediante telecamere e tecniche a computer vision. Ciascuna di queste tecnologie differisce dalle altre secondo accuratezza, risoluzione, raggio di azione, latenza, *user-comfort* e costi.



Figura 1.2: Bepop Forte Data Glove, un guanto utilizzato per la realtà virtuale e aumentata



Figura 1.3 Smart TV Samsung Serie H in grado di riconoscere i gesti delle mani

Nonostante la loro indubbia utilità, i gesti possono presentare criticità e limitazioni nel loro univoco riconoscimento

Generalmente i gesti delle mani possono essere ambigui: per esempio per indicare il concetto di stop si potrebbe alzare la mano con il palmo affacciato in avanti, ma si potrebbero anche agitare in modo forsennato le mani verso l'alto. Quindi sia come nelle parole che nella scrittura i gesti possono variare di significato tra individui, o anche nello stesso individuo ma in situazioni differenti.

Il riconoscimento dei gesti con componenti dinamiche richiede una finestra temporale: deve essere specificato infatti l'inizio e la fine dell'acquisizione. Un gesto inoltre può essere contestualizzato dai gesti precedenti e influenzare i gesti successivi.

Infine, i gesti possono anche essere caratteristici di un linguaggio o di una cultura particolare.

## 1.2. Storia dell'interazione uomo-macchina

I primi computer non erano particolarmente interattivi: l'utente finale, ovvero un tecnico lasciava al centro di calcolo i pacchi(batches) di attività(job) da svolgere e passava dopo qualche ora se non addirittura il giorno dopo a ritirare i risultati. Nel 1961 fu introdotto il *Computer time sharing system* (CTSS) dando la possibilità agli utenti di interagire con il computer mediante un terminale interattivo.

Il primo terminale interattivo fu il terminale scrivente o *teletype* composto da tastiera e stampante integrata: veniva inviato il comando a cui seguiva la risposta del calcolatore e assieme all'eventuale sollecito(prompt) successivo da parte dell'utente, tutto questo veniva poi stampato su un rullo carta. Tale modalità di interazione relegava il computer ad un ruolo passivo, limitandosi semplicemente ad eseguire i comandi dell'utente e fornire poi una risposta.

Vennero sviluppati sistemi in cui il ruolo passivo e occupato dall'utente, che si limita a rispondere alle domande richieste dal software: per esempio il sistema Mycin (Medical Expert System) sviluppato nel 1970 con il linguaggio di programmazione LISP poneva al medico delle domande predeterminate dipendenti dalle risposte precedenti suggerendo infine degli antibiotici adatti per curare infezioni batteriche specifiche.



Figura 1.4 modelli Teletype 33 F usati nel 1970

Nel '71 il teletype venne sostituito da un terminale video e la tastiera viene arricchita con nuovi tasti contenenti funzioni utili agli utenti. Vi fu anche l'introduzione innovativa di un cursore spostabile mediante appositi tasti che permetteva all'utente di indicare una specifica posizione dello schermo.

La presenza di un terminale video stravolse completamente la modalità di comunicazione tradizionale del teletype, permettendo all'utente di scrivere su uno schermo visibile e non più su uno invece immaginario.

I menù permettono di selezionare le attività che il calcolatore dovrà svolgere, semplificando drasticamente la comunicazione tra computer ed essere umano ma relegando il computer ad un ruolo ancora più passivo, molto più simile a un telecomando o ad una lavatrice.

Negli anni '80 viene introdotto sul mercato il personal computer con potenza di calcolo maggiore, capacità di archiviazione su floppy-disc e stampa istantanea, arrivando su tutte le scrivanie di studenti, impiegati e professionisti. Il mouse, inventato da Douglas Engelbart a metà degli anni 60(brevettato il 21 luglio 1967) diventerà poi la dotazione standard di ogni personal computer. Questo dispositivo, mediante lo spostamento sulla scrivania, è in grado di muovere un cursore sullo schermo con grande agilità e precisione, introducendo così una nuova forma di comunicazione con il calcolatore (una comunicazione quasi gestuale)



Figura 1.5 Computer Amiga 500 della Commodore, uno dei primi PC ad avere il mouse in dotazione

Dagli anni '90 iniziano ad affermarsi sul mercato i *palmtop*: piccoli computer in grado di stare in una mano e dotati di *touchpad* basato sul rilevamento capacitivo il dito che, toccando lo schermo diventa un conduttore con una propria capacità, che causa quindi una variazione capacitiva che viene rilevata dai sensori.



Negli stessi anni esplose il mercato della telefonia mobile: i primi telefoni erano modelli molto semplici con come uniche funzionalità quelle di telefonare e di inviare messaggi, fino ad evolversi nei più moderni cellulari, in particolare, l'iPhone della Apple, era completamente sprovvisto di tasti fisici e dotato unicamente di un grande schermo touch. In termini di interazione uomo-macchina l'iPhone non si può più considerare né come un computer, tanto meno come telefono, si tratta di qualcosa di nuovo, ovvero un *mobile*, perché destinato a un uso strettamente personale da parte dell'utente, diventando di fatto una sua estensione: è in grado di raccogliere informazioni sull'utente e sull'ambiente che lo circonda.



Figura 1.6 Steve Jobs mentre presenta il primo modello di iPhone

Nel settore videoludico, la Nintendo rilascia la console Nintendo Wii che rivoluziona il concetto di controller con il suo *Wiimote* dotato di accelerometro per la rilevazione del movimento e un sensore IR posto sulla sua sommità che abbinato alla *sensor bar* in dotazione permetteva il puntamento sullo schermo, mentre Microsoft rilascia *Kinect*, che consentiva all'utente di controllare il gioco senza l'utilizzo di controller mediante il tracciamento dei movimenti del giocatore. Iniziano a comparire i primi dispositivi che sfruttano la realtà virtuale come i visori Oculus Rift, Playstation VR e HTC Vive in grado di creare ambienti virtuali 3d.



Figura 1.7 Modello di Oculus Rift con in dotazione due controller

Il settore videoludico è stato certamente la culla di questa tecnologia, i gesti tuttavia risultano lo strumento di comunicazione per eccellenza tra uomo e macchina e quindi tale tecnologia può essere utilizzata in vari ambiti.



## 1.3. Applicazioni

Il riconoscimento dei gesti (definito comunemente anche con l'espressione *gesture recognition*) rappresenta un utile strumento che trova applicazione in settori molto vari.

Interessanti esempi possono essere infatti individuati per apportare un miglioramento in:

- ambito lavorativo, nelle tecnologie informatiche impiegate, (e.g. comunicazione durante video-conferenze)
- ambito educativo, consentendo l'interazione sia in presenza che a distanza (e.g. facilitare l'interazione dei tra bambini e computers, apprendimento a distanza o *tele-teaching*,)
- ambito ricreativo, consentendo maggiore realismo (e.g. motion capture, videogiochi in realtà virtuale)
- ambito forense, progettando sistemi biometrici e di riconoscimento (e.g. tecniche di identificazione, *lie detection*);
- ambito di sicurezza, per migliorare i sistemi di allerta durante attività a rischio (e.g. valutazione dei livelli di attenzione/stanchezza di guidatori)
- ambito medico, per migliorare le tecniche di riabilitazione e la qualità di vita di pazienti con disabilità (e.g. navigazione e manipolazione di ambienti virtuali in ambito riabilitativo, controllo medico di stati emozionali e di stress di pazienti, controllo di attuatori esterni, progettazione di strumenti di ausilio per persone con incapacità uditive o vocali).

Un metodo conveniente per classificare applicazioni tanto variegata può essere quello di raggrupparle sulla base dell'utilizzo che viene fatto del gesto. Le principali categorie sono quindi: applicazioni in cui l'obiettivo stesso è il riconoscimento stesso dei gesti e la loro eventuale replica; applicazioni in cui il riconoscimento dei gesti è convertito a comandi per controllare attuatori esterni; applicazioni in cui il riconoscimento dei gesti viene sfruttato per creare una realtà virtuale e consentire l'interazione dell'utente con essa.

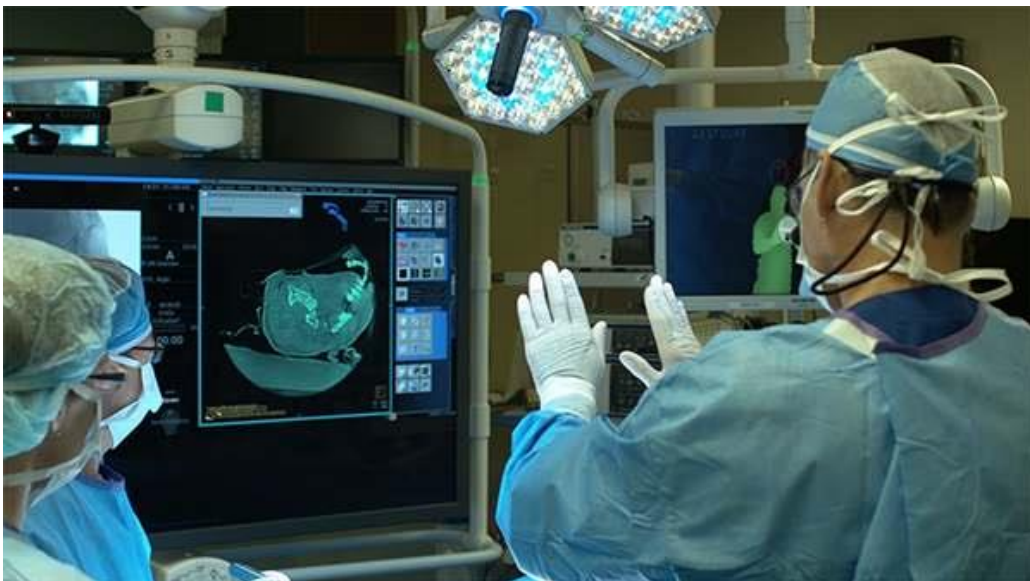


Figura 1.8: Esempio di operazione chirurgica touchless

### 1.3.1. Riconoscimento dei gesti

Nel linguaggio dei segni ogni gesto, espressione facciale e movimento del torso ha un proprio significato con specifiche regole grammaticali e di contestualizzazione per facilitare la comprensione della conversazione. Tale linguaggio, per esempio *l'American sign language* (ASL) è anche dotato di uno spelling effettuato col movimento delle dita per esprimere parole particolari e nomi propri. Mediante l'utilizzo di un guanto contenente opportuni sensori o mediante una telecamera a colori è possibile tenere traccia dei movimenti della mano e poter associare tali movimenti a una frase codificata nell'ASL. Tale dispositivo sarebbe in grado di facilitare drasticamente la comunicazione tra persone sorde o mute con altre persone non informate sul linguaggio dei segni.

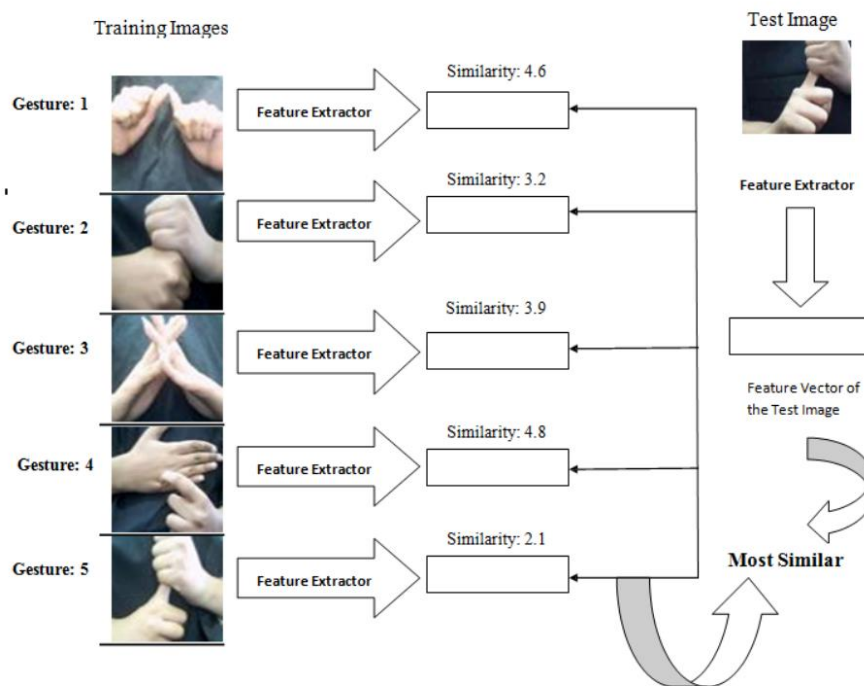


Figura 1.9: Possibile schema di organizzazione del riconoscimento del linguaggio dei segni

Il riconoscimento dei gesti è anche la base del *motion capture* ovvero la registrazione dei movimenti di persone, animali e oggetti. L'esecutore del movimento doveva inizialmente indossare marcatori in prossimità delle proprie articolazioni affinché durante il movimento le distanze e gli angoli tra i vari marker fossero misurabili. I marcatori potevano essere acustici, magnetici, LED, inerziali o una combinazione di questi. Con l'avvento delle tecnologie moderne del XXI secolo molti sistemi di riconoscimento sono in grado di estrarre la silhouette dell'esecutore separandola dallo sfondo (background) calcolando la posizione degli arti e delle articolazioni mediante modelli matematici partendo dalla sagoma estratta.

Il motion capture trova impiego in vari settori: sportivo se si vuole acquisire e valutare un gesto tecnico per poi individuarne raffinementi e migliorare di conseguenza la performance di un soggetto o eventualmente poterne prevedere infortuni , industria cinematografica e videoludica per la realizzazioni di controparti digitali realistiche(celebre esempio Gollum del Signore degli anelli) ,militare per l'analisi delle prestazioni fisiche di un soggetto , cliniche per la ricerca di disfunzioni nell'attività motoria, monitoraggio dei progressi della riabilitazione di un paziente o per eventuali interventi mirati. I vantaggi di questa tecnologia consistono nel fatto di poter avere un'animazione iperrealistica in tempi relativamente brevi e a costi contenuti e a un carico di lavoro non dipendente dalla complessità della performance o della lunghezza del movimento da acquisire.



Figura 1.10: Gollum, personaggio de "Il signore degli anelli", celebre esempio di controparte digitale realizzata mediante Motion Capture

Altro settore di interesse per il riconoscimento dei gesti è rappresentato dalla telemedicina, ovvero la digitalizzazione delle informazioni mediche e il loro scambio attraverso un'infrastruttura medica connessa alla rete. Un'interessante tecnologia utilizzabile nel settore della telemedicina è quella del sistema *telepointer*, in grado di potenziare la comunicazione a grandi distanze e aumentando la qualità delle prestazioni del sistema sanitario. Questo sistema funziona inviando informazioni aggiuntive sotto forma di gesti che possono contenere (e quindi trasmettere) istruzioni o informazioni più accurate. Porta a una comunicazione più efficace, una diagnosi precisa e una decisione migliore attraverso la discussione e la consultazione tra l'esperto e l'eventuale medico "junior".

La migliore modalità di comunicazione è certamente quella "faccia a faccia", grazie alla presenza di interazioni, gesti, linguaggio del corpo ,espressioni facciali e varie tonalità di voce , la comunicazione a distanza, pur essendo necessaria in molte situazioni, manca di tutte queste componenti, portando spesso gli interlocutori a malintesi, tale tecnologia potrebbe rafforzare la conversazione a distanza mediante i gesti per permettere all'interlocutore di esprimersi correttamente e dando anche un senso di presenza alla conversazione.

Il telepointer può svolgere due funzioni:

- Puntare un oggetto o una regione di interesse.
- Creare un *pattern* che, in base al tipo di movimento, ha un proprio significato.

Un sistema di telepointer offre molti vantaggi come:

- Fornire informazioni sulle coordinate dell'oggetto target.
- Creare il senso di presenza.
- Aumentare l'attenzione del pubblico utilizzando la forma multipla del telepointer come colore, dimensione e immagine.

La telemedicina può quindi essere utilizzata come strumento di supporto per facilitare la condivisione delle conoscenze e la tecnologia del telepointer ha un grande potenziale in campo medico, poiché:

- Migliora le prestazioni nell'esecuzione di un compito.
- Aumenta l'accuratezza della diagnosi
- Evita la tragedia dovuta alla mancanza di uno specialista durante momenti cruciali in operazioni o manovre mediche.

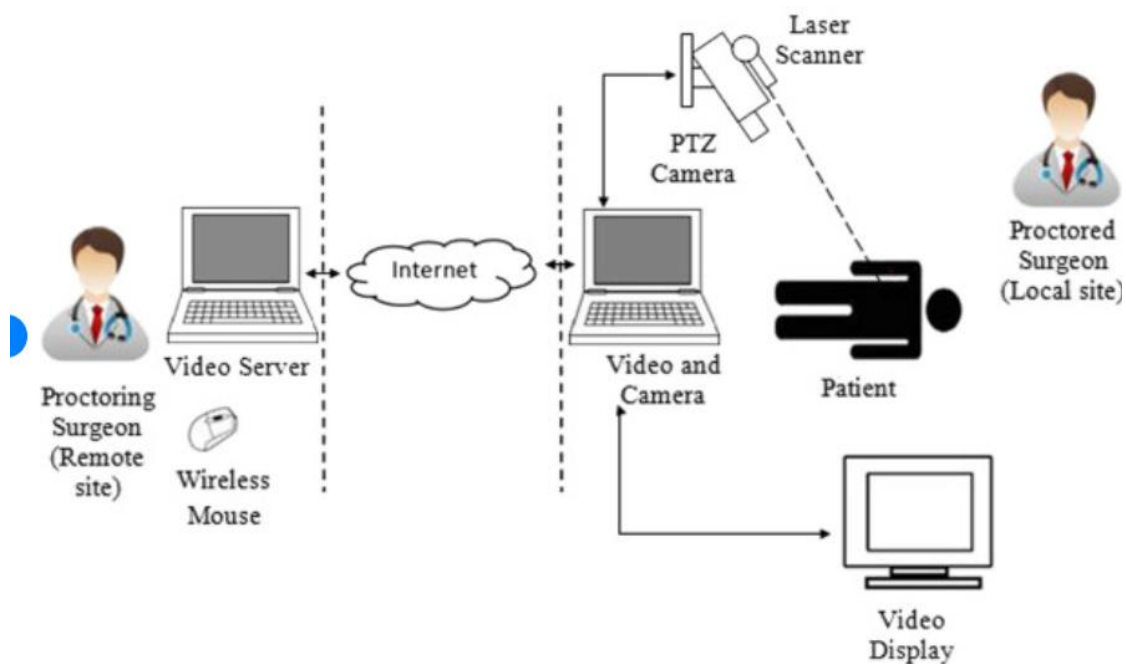


Figura 1.11 Rappresentazione schematica di un sistema telepointer nell'ambito della telemedicina

## 1.3.2. Comando di attuatori esterni

Con lo sviluppo delle moderne tecnologie e il naturale uso dell'uomo ad usare gesti per comunicare, il riconoscimento dei gesti permette alle macchine di eseguire istruzioni specifiche dell'utente. Tutto questo avviene mediante dei meccanismi, definiti come attuatori, con cui un uomo, animale o sistema artificiale intelligente è in grado di interagire con l'ambiente come, per esempio, mettere in moto un processo.

Gli attuatori trovano applicazione nel settore biomedico, dell'automazione, robotico e aerospaziale e possono essere classificati in base al fenomeno che sfruttano per il loro funzionamento, che può essere:

- Pneumatico
- Idraulico
- Elettrico
- Meccanico

In una possibile situazione in cui un soggetto non possa essere in grado di attivare in modo sicuro, comodo o semplicemente pratico un determinato dispositivo, vengono utilizzati gli attuatori.

Le auto moderne stanno offrendo un numero sempre maggiore di funzionalità, richiedendo però un altrettanto numero di comandi, un ipotetico guidatore non sarebbe in grado di prestare attenzione sia alla strada che a un'ipotetica interfaccia grafica con tutti i comandi in questione, un'interfaccia basata sul riconoscimento dei gesti potrebbe rivelarsi molto importante quindi per ridurre il livello di attenzione necessario per controllare le varie funzionalità del veicolo. I gesti, che includono movimenti di mano e di dita, possono essere utilizzati per controllare vari dispositivi elettronici nel veicolo.

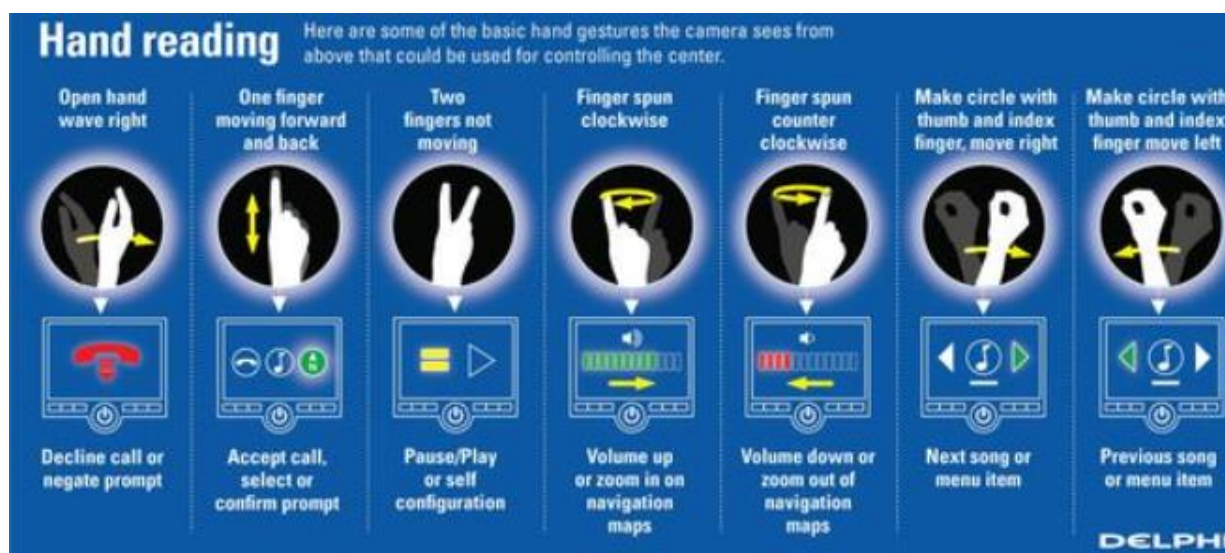


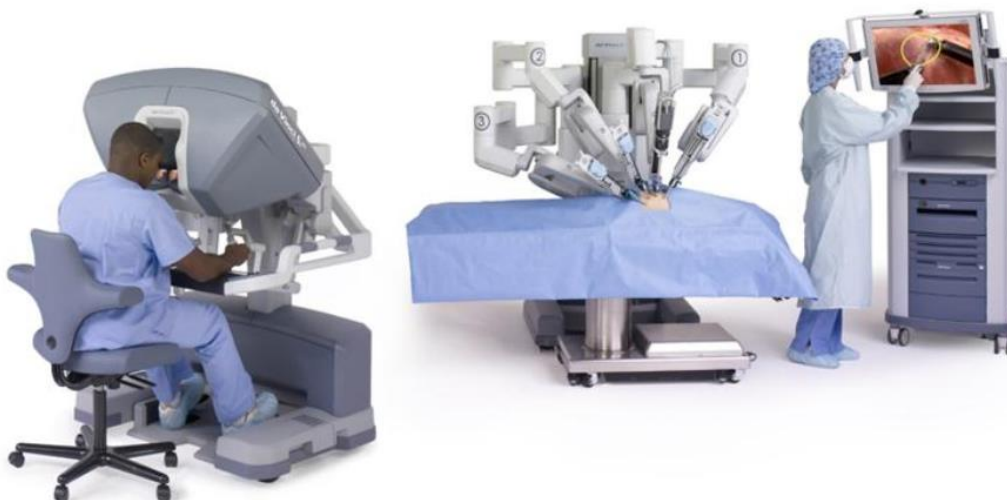
Figura 1.12 Possibile schema di comandi a gesti per l'auto



Il riconoscimento dei gesti può essere utilizzato anche per controllare droni da remoto, semplificando di molto i comandi di un drone non rendendo più necessaria la conoscenza professionale di pilotaggio di un drone. Oltre ai droni possono essere controllati robot, in grado di assistere sia persone con problemi di mobilità, in condizioni particolarmente scomode, sia operai in ambiente industriale nel caso di compiti particolarmente difficili per un semplice essere umano (per esempio alzare carichi pesanti, lavorare in ambienti poveri di ossigeno o ad alta pressione ecc.), sia in ambito militare nell'assistenza ai soldati di fanteria che saranno in grado di comunicare in modo efficace con i robot mediante gesti, rendendoli molto simili a ei compagni di squadra.

La popolazione anziana sta crescendo sempre di più negli ultimi anni, uno dei principali problemi dell'invecchiamento è certamente quello della riduzione della mobilità, mediante il riconoscimento dei gesti è possibile implementare un sistema di navigazione che, se inserito in una sedia a rotelle, porterebbe alla possibilità di movimento domestico, senza la necessità dell'assistenza di un'altra persona.

A livello sanitario, la possibilità di poter controllare da remoto dispositivi medici mediante l'utilizzo di semplici gesti per poter far svolgere alla macchina compiti più complicati, permettendo al medico di mantenere la concentrazione e di avere un minor tempo di risposta, il tutto in un ambiente sterile, migliorando significativamente l'efficacia di interventi e operazioni.



*Figura 1.13 Esempio di operazione chirurgica effettuata da remoto*

E' da molto tempo presente la tecnologia smart home, derivata dall'internet of things, che permette un collegamento dei dispositivi domestici tra loro, rendendo possibile il loro controllo da remoto mediante smartphone o altro dispositivo di controllo che, nel nostro caso, potrebbe trattarsi di un dispositivo in grado di rilevare i movimenti della mano aumentando il comfort dell'abitazione. Si sarà in grado mediante gesti si potrà controllare il climatizzatore/riscaldamento, controllo di porte, tende, rubinetti, elettrodomestici.



### 1.3.3. Realtà virtuale

Si definisce realtà virtuale la capacità di simulare situazioni realistiche attraverso l'uso di computer e di appositi dispositivi di interfaccia. Una persona che utilizza tali interfacce è in grado infatti di vedere, visitare uno scenario artificiale che possiede caratteristiche e oggetti interattivi per l'utente. Vengono utilizzati giroscopi e sensori di movimento per tracciare la posizione della testa, del corpo e delle mani. Uno schermo HD permette la visione stereoscopica delle immagini, il tutto supportato da processori leggeri e veloci.

La realtà virtuale è principalmente utilizzata nel settore dell'intrattenimento ovvero videogiochi (Oculus Rift e playstation VR per citarne alcuni), cinema e ambienti virtuali social (basti pensare al recentissimo Meta). Nella psicologia può invece essere utile per poter creare ambienti controllati per poter trattare l'ansia dovuta a particolari fobie e disturbi da stress post traumatico(PTSD).E' anche molto utile per trattare la malattia di Alzheimer e altre malattie neurodegenerative ,infatti questa tecnologia è in grado di far rivivere situazioni stimolanti ai pazienti che nel loro attuale stato mentale non sarebbero in grado di rivivere nel mondo reale(in particolare per le persone anziane che di norma presentano grosse limitazioni in termini di mobilità). In medicina permette di simulare scenari clinici (come interventi chirurgici per esempio) a un costo ridotto, permettendo al personale sanitario (in particolare la componente "alle prime armi") di fare pratica su tali scenari senza rischi per i pazienti. Inoltre, l'utilizzo della realtà virtuale può tornare utile nel trattamento di disabilità funzionali e cognitive dovute a danni cerebrali, disturbi dell'apprendimento e disordini neurologici.



*Figura 1.14 Soggetto affetto da demenza mentre indossa un visore VR*

La realtà virtuale potrebbe inoltre essere utilizzata come interfaccia per e-commerce permettendo la visione di un articolo che si vorrebbe acquistare in una situazione di confort da casa in una situazione verosimile.

Nel settore dell'apprendimento la realtà virtuale è in grado di simulare ambienti di lavoro in totale sicurezza permettendo all'utente di affinare una propria abilità senza possibili conseguenze di fallimento che potrebbero capitare in ambienti di lavoro virtuali come scenari clinici, militari, aerospaziali ecc.



*Figura 1.15 Addestramento di un astronauta della NASA mediante l'utilizzo della realtà virtuale*



## 2. Struttura hardware

In questo capitolo verranno discusse le tipologie di componenti hardware che possono essere utilizzate in un sistema di tracciamento dei movimenti della mano.

### 2.1. Sensori

Il sistema può acquisire i dati attraverso due possibili configurazioni di sensori:

- Mediante tecnologie indossabili: che dovranno per forza essere indossate dall'utente, inoltre il dispositivo sarà collegato attraverso una serie di cavi a un computer, limitando potenzialmente la mobilità dell'utente.
- Mediante riconoscimento visivo: Non necessitano il collegamento dell'utente ad un computer, risolvendo quindi la limitazione delle tecnologie indossabili, tuttavia presentano la problematica (non presente nei sensori indossabili) dell'occlusione di alcune parti del corpo dell'utente dalla visuale della telecamera dovuto ad oggetti esterni o altre parti del corpo dell'utente spesso (una telecamera non sarebbe in grado di rilevare i movimenti dell'utente se questo fosse nascosto dietro un muro per esempio)

In termini di prestazioni, invece, i sensori indossabili risultano più veloci ed efficienti dei sensori visivi: si rivelano in grado, infatti, di rilevare movimenti della mano e delle dita in modo preciso e veloce a differenza dei sensori visivi che sono in grado solamente di dare indicazioni qualitative (e non quantitative) del movimento della mano. A favore dei sensori visivi va però attribuito il fatto di essere in grado di analizzare particolari proprietà come il colore e la texture (anche se per l'analisi di gesti questo vantaggio risulta trascurabile).

Le tecniche di acquisizione visive si basano inoltre su alcuni parametri come:

- Il numero di telecamere utilizzate
- La loro velocità e latenza
- Condizioni ambientali come ostacoli visivi, luci particolari ecc.
- Requisiti richiesti all'utente come dover indossare qualcosa di speciale per facilitare il riconoscimento
- La scelta di rappresentazione 3D o 2D
- Se è presente una rappresentazione temporale

In seguito, si discuteranno i vari tipi di sensori ritrovabili sia nelle tecnologie indossabili che visive.

## 2.1.1. Accelerometri

L'accelerazione risulta essere una proprietà dinamica di un corpo, in quanto, secondo la seconda legge della dinamica (o legge di Newton), la presenza di tale proprietà implica la presenza di una forza:

$$F = m \times a$$

Posizione, velocità e accelerazione sono tutte correlate tra loro, infatti la velocità risulta essere la derivata dello spostamento rispetto al tempo, l'accelerazione risulta invece essere la derivata della velocità rispetto al tempo (risultando quindi come la derivata seconda della posizione). In ambienti rumorosi (specialmente nel caso in cui siano presenti rumori ad alta frequenza non trascurabili) l'utilizzo di filtri derivatori potrebbe non risultare un'idea ottimale a causa di possibili uscite completamente differenti dal risultato posto come obiettivo. Sono quindi necessari sensori in grado di misurare l'accelerazione in ambienti rumorosi che non utilizzino possibilmente operatori derivata, dovendo lavorare anche in ambienti potenzialmente rumorosi.

La vibrazione è un fenomeno meccanico che consiste nel movimento oscillatorio di un corpo attorno a una posizione di riferimento: tale fenomeno può essere presente o meno nei casi di interesse.

L'accelerometro è un tipo di sensore in grado di rilevare variazioni di accelerazione (*proper acceleration*) nell'unità 'g' (ovvero la misura viene effettuata in relazione all'accelerazione gravitazionale) su uno, due o tre (caso più comune) assi. A livello puramente concettuale, l'accelerometro risulta composto da una massa di prova (che risulta nota) collegata ad una molla: nel momento in cui la massa subisce una forza (e quindi un'accelerazione), questa si sposterà fino al punto in cui la molla tirerà alla stessa forza di quella applicata: la compressione/estensione della molla ci permetterà quindi di risalire all'accelerazione incognita.

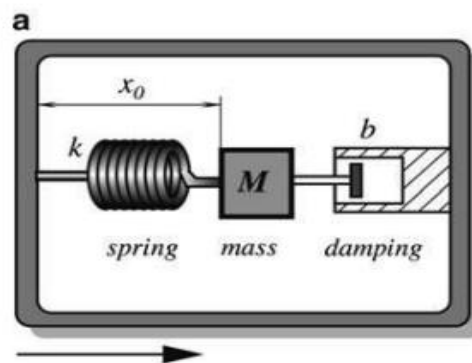


Figura 2.1: Modello schematico di un accelerometro

Dalla seconda legge di Newton otteniamo:

$$Ma = -kx - b \frac{dx}{dt} \quad (2.1)$$

Con M massa di prova, a accelerazione incognita relativa alla terra (l'accelerazione di gravità non viene contata), k costante elastica della molla, b coefficiente di smorzamento e x spostamento del sistema. L'accelerazione a, in uno spazio bidimensionale è data da:

$$a = \frac{d^2x}{dt^2} - \frac{d^2y}{dt^2} \quad (2.2)$$

Sostituendo l'equazione 2.2 nell'equazione 2.1:

$$M \frac{d^2x}{dt^2} + b \frac{dx}{dt} + kx = M \frac{d^2y}{dt^2} \quad (2.3)$$

Tale equazione differenziale risulta essere del secondo ordine, implicando la possibile presenza di moti oscillatori.

Mediante l'utilizzo della trasformata di Laplace è possibile risolvere l'equazione 2.3:

$$Ms^2X(s) + bsX(s) + kX(s) = -MA(s) \quad (2.4)$$

Con X trasformata di Laplace di x e A trasformata di  $\frac{d^2y}{dt^2}$ .

Risolvendo l'equazione 2.4 per X:

$$X(s) = -\frac{MA(s)}{Ms^2 + bs + k} \quad (2.5)$$

Sostituendo  $\omega_0 = \sqrt{\frac{k}{M}}$  e  $2\zeta\omega_0 = \frac{b}{M}$ :

Otteniamo:

$$X(s) = -\frac{A(s)}{s^2 + 2\zeta\omega_0s + \omega_0^2} \quad (2.6)$$

$\omega_0$  rappresenta la frequenza fondamentale dell'accelerometro,  $\zeta$  rappresenta invece il fattore di smorzamento normalizzato.

Portando A(s) al primo membro riscriviamo l'equazione 2.6 per  $G(s) = \frac{X(s)}{A(s)}$ :

$$G(s) = -\frac{1}{s^2 + 2\zeta\omega_0s + \omega_0^2} \quad (2.7)$$

X(s) risulta uguale a G(s)A(s), anti trasformando e ricordando che il prodotto delle trasformate corrisponde alla convoluzione nel dominio del tempo:

$$x(t) = \int_0^t g(t - \tau)\alpha(t)d\tau \quad (2.8)$$



Con  $\alpha(t)$  l'impulso tempo variante del corpo, e  $g(t)$  la trasformata inversa di  $G(s)$ , ponendo

$\omega = \omega_0 \sqrt{1 - \zeta^2}$  l'equazione 2.8 può avere due soluzioni:

- $x(t) = \int_0^t -\frac{1}{\omega} e^{-\zeta\omega_0(t-\tau)} \sin \omega(t - \tau) \alpha(t) d\tau$
- $x(t) = \int_0^t -\frac{1}{\omega} e^{-\zeta\omega_0(t-\tau)} \sinh \omega(t - \tau) \alpha(t) d\tau$  (caso di sovra smorzamento  $\zeta > 1$ )

Con  $\omega = \sqrt{1 - \zeta^2}$

Queste soluzioni possono essere ricavate da qualsiasi accelerazione di input applicata all'accelerometro. Un accelerometro di buona qualità deve possedere una frequenza di risonanza unica e distinta e una banda piatta nel range di frequenze misurabili: in tale regione l'accelerometro sarà in grado di misurare la grandezza di input senza distorcere il segnale di ingresso.

Per valutare quindi lo spostamento della molla sarà necessario impiegare un sensore di posizione, che può essere:

- Estensimetrico: l'accelerometro sfrutta la variazione di resistenza di un estensimetro
- Piezoresistivo: al posto degli estensimetri vengono utilizzati sensori piezoresistivi
- LVDT: che sfrutta i sensori LVDT (ovvero *linear variable differential transformer*) o trasduttori a spostamento induttivo che sfruttano la variazione di induttanza di una bobina
- Capacitivo: sfrutta la variazione della capacità di un condensatore al variare della distanza delle due armature
- Piezoelettrico: sfrutta il segnale elettrico generato da un cristallo piezoelettrico quando sottoposto a sollecitazioni
- Laser: sfrutta un interferometro laser che misura, istante per istante, lo spostamento della massa di prova mandando i dati a un computer che, effettuando la derivata seconda dello spostamento, ottiene direttamente l'accelerazione del corpo
- Gravimetro: il principio alla base è simile all'accelerometro laser e viene utilizzato per misurare l'accelerazione di gravità

La scelta di un accelerometro può dipendere da diversi fattori:

- Larghezza di banda: Indica il range di vibrazioni che il sensore è in grado di misurare in modo ottimale, nel caso di movimenti umani, questi non sono in grado di superare i 10 Hz quindi una frequenza di taglio a 50-60 Hz risulta più che sufficiente, il comportamento del sensore in funzione della frequenza viene definito "risposta in frequenza".
- Sensibilità: ovvero il quantitativo minimo di accelerazione misurabile che può quindi modificare l'output del sensore, in termini matematici risulta essere il rapporto tra output di tensione e variazione meccanica in input, la sensibilità viene determinata in Europa mediante delle misure alla frequenza di riferimento di un'onda meccanica a 160 Hz.
- Resistenza al rumore: Inversamente proporzionale alla larghezza della banda, più la frequenza in cui si lavora aumenta, più il sensore sarà soggetto a rumore influenzando significativamente le prestazioni dell'accelerometro.
- Voltaggio indipendente dall'accelerazione di gravità  $g$ : Il sensore non è affetto dall'accelerazione di gravità (costituirebbe un offset della misura).
- Range dinamico: descrive l'intervallo della grandezza misurabile del sensore ovvero dal valore minimo misurabile (sensibilità) al valore massimo (fondoscala).
- Linearità: porzione del range dinamico in cui il sensore ha comportamento lineare.

## Funzionamento di un accelerometro capacitivo

L'obiettivo principale di qualsiasi accelerometro è quello di misurare lo spostamento della massa di prova, uno dei possibili metodi per rilevare tale spostamento è quello di rilevare la variazione della capacità di un condensatore, un accelerometro che sfrutta questo fenomeno viene definito capacitivo.

Gli accelerometri capacitivi sono largamente usati grazie al loro basso consumo di energia, struttura semplice e stabile oltre alla loro facilità di integrazione in sistemi più avanzati. Tale sensore è generalmente costituito da due elementi: un'armatura del condensatore fissata (stazionaria) e l'altra armatura solidale alla massa di prova che si muoverà con essa. Per evitare interferenze e *drift* alle basse frequenze viene adottata una configurazione differenziale in cui viene aggiunto un secondo condensatore nelle stesse condizioni (un'armatura fissa e una mobile) il secondo condensatore dovrebbe avere una capacità molto simile al primo e un cambio di fase di  $180^\circ$ : l'accelerazione misurata verrà rappresentata dalla differenza di capacità tra i due condensatori.

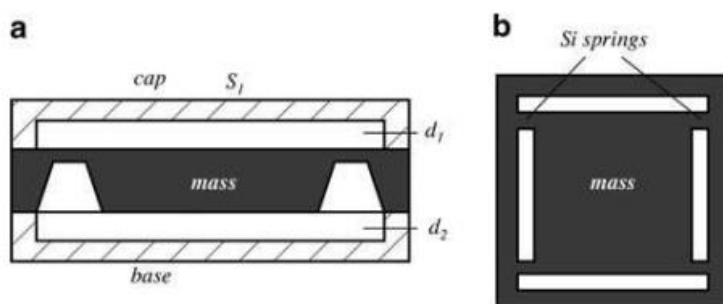


Figura 2.2: (a) Visone sul piano frontale dell'accelerometro capacitivo differenziale, (b) visione trasversale della massa di prova supportata da molle in silicio

Guardando la figura 2.2a possiamo notare la configurazione dell'accelerometro differenziale con la massa di prova inserita nell'intercapedine delle due armature (*base* e *cap*), tale massa è supportata da molle in silicio (figura 2.2b). La grandezza  $d_1$  rappresenta la distanza della piastra superiore dalla massa, mentre  $d_2$  la distanza della inferiore.  $S_1$  è invece la superficie dell'armatura del condensatore.

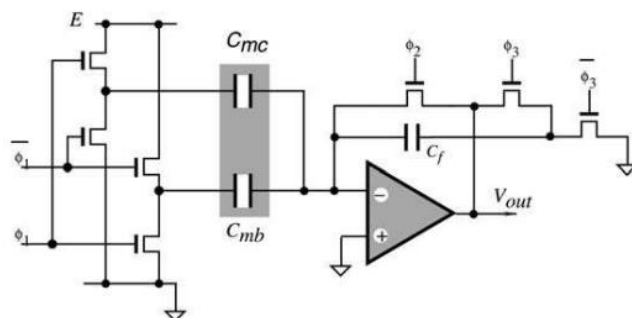


Figura 2.3 Schema circuitale per convertire la variazione di capacità in tensione

Nella figura 2.3 si può osservare uno schema circuitale di conversione capacità-tensione, con  $C_{mc}$  e  $C_{mb}$  che rappresentano i due condensatori del sensore, quando il circuito viene sollecitato, l'uscita  $V_{out}$  dell'amplificatore dipende dalle due capacità, e quindi anche dalla forza applicata alla massa di prova (e quindi infine dall'accelerazione). Tale schema risulta stabile solo nel caso in cui il sensore non venga sollecitato da forze elettrostatiche, inoltre la temperatura potrebbe influenzarne le prestazioni.

## 2.1.2. Giroscopi

Il giroscopio è un dispositivo rotante che se agganciato su un corpo permette di misurare sia la direzione che la velocità angolare se tale corpo iniziasse a ruotare. Oltre a poter misurare tali grandezze, il giroscopio è anche in grado di mantenerle, il funzionamento del giroscopio è basato sulla conservazione del momento angolare: nel caso di assenza di forze esterne, in un sistema di particelle il momento angolare rispetto a un punto fissato nello spazio rimane costante.

Il giroscopio tradizionale meccanico consiste in un disco (toroide per precisione) libero di ruotare attorno al proprio asse, che è fissato ad un supporto avente a sua volta libertà di rotazione attorno a uno o due assi di rotazione. Tale configurazione permette al giroscopio di svolgere due incarichi fondamentali:

- La direzione dell'asse di rotazione del giroscopio rimarrà costante nello spazio (previa assenza di forze esterne).
- La capacità di fornire un segnale di output proporzionale alla velocità angolare di un asse perpendicolare a quello di rotazione.

Quando l'elemento rotante è libero di muoversi, questo tende a preservare la direzione del proprio asse di rotazione, se il supporto del giroscopio ruotasse, questo svilupperebbe un momento torcente attorno all'asse perpendicolare di output, ruotando il proprio asse attorno a questo. Secondo la legge di Newton:

$$T = \omega I \Omega$$

Con  $\omega$  velocità angolare rispetto all'asse di rotazione (input),  $I$  momento di inerzia del corpo rotante,  $\Omega$  velocità angolare rispetto all'asse di output.

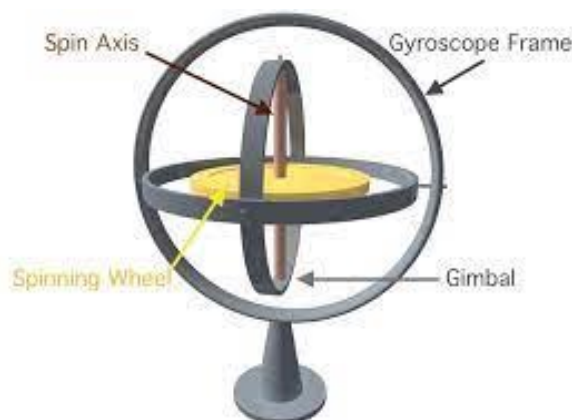


Figura 2.4 Modello di giroscopio meccanico

Il giroscopio meccanico, pur essendo stato utilizzato per moltissimo tempo, contiene sostegni, giunti cardanici, rotori e motori che richiederebbero una fabbricazione e una calibrazione eccessivamente accurate, e quindi costose. Risulta quindi fondamentale rilevare meccanismi diversi per rilevare variazioni angolari. L'utilizzo della tecnologia *Micro Electro-Mechanical System* (MEMS) ha permesso la realizzazione di microscopi in miniatura *low-cost*.

I giroscopi MEMS sono giroscopi confezionati in microchip utilizzati nei circuiti elettronici, utilizzano masse vibranti (che sostituiscono gli elementi rotanti tradizionali) per determinare la velocità angolare, non sono quindi dotati di elementi rotanti come i giroscopi tradizionali (questo permette la loro miniaturizzazione). Questi giroscopi sono inoltre dei sistemi più robusti della loro controparte tradizionale.

Il concetto di forza di Coriolis è largamente utilizzato nei giroscopi: Fissato un sistema di riferimento, se un corpo ruota a velocità angolare  $\Omega = \Omega x_i$  (con  $x_i$  vettore unitario lungo l'asse  $x$ ) attorno all'asse  $x$ , un osservatore, solidamente ancorato all'asse  $z$ , vede la particella muoversi (pur non subendo una forza reale) lungo l'asse  $z$  con un'accelerazione di Coriolis pari a:

$$a_c = 2v \times \Omega$$

Il giroscopio MEMS sostituisce la rotazione con la vibrazione e l'accelerazione risultante risulta correlata allo spostamento angolare del corpo.

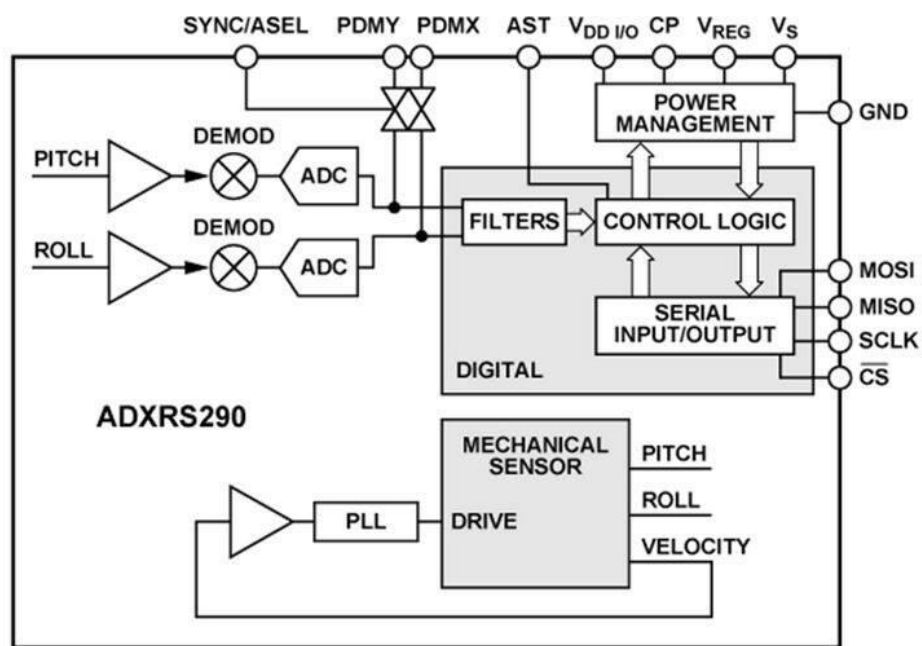


Figura 2.5: Possibile schema circuitale di un giroscopio MEMS

Tra le caratteristiche più importanti da prendere in considerazione per l'acquisto di un giroscopio troviamo:

- *Angle random walk*: l'output del sensore è sempre soggetto a rumore, questo può essere causato nel caso dei MEMS dalle masse vibranti.

- *Offset*: nel caso di input nullo anche l'uscita dovrebbe risultare zero; tuttavia, nei casi reali è sempre presente un ingresso diverso da zero pur non dando input, tendenzialmente tale offset viene misurato alla temperatura di 25°C dal produttore che lo inserisce nella scheda dati del sensore, rendendolo quindi compensabile.
- Sensibilità alla temperatura: Le prestazioni del giroscopio cambiano con la temperatura. Per verificare che le prestazioni del giroscopio soddisfino gli obiettivi del sistema è necessaria una caratterizzazione di parametri come rumore, offset di polarizzazione e fattore di scala sulla temperatura.

## 2.1.3. Estensimetri

La deformazione consiste nella variazione di una o più dimensioni di un volume solido.

L'estensimetro è una tipologia di sensore in grado di rilevare piccole variazioni dimensionali di un'oggetto sottoposto a tensioni meccaniche, conoscendo a priori le caratteristiche meccanico/fisiche del materiale (come modulo di Young e coefficiente di Poisson), misurando le deformazioni si possono facilmente ricavare gli sforzi meccanici a cui il materiale è sottoposto.

L'estensimetro elettrico a resistenza è costituito da una griglia di sottilissimo filo metallico (solitamente costantana) rigidamente applicata su un supporto di materiale plastico.

L'estensimetro viene utilizzato incollandolo sulla superficie del corpo di cui si vogliono misurare le deformazioni generalmente utilizzando collanti istantanei come le colle acriliche. Il filo dell'estensimetro segue le deformazioni della superficie a cui è incollato, allungandosi ed accorciandosi insieme ad essa; queste variazioni dimensionali causano una variazione della resistenza elettrica del filo. Misurando, tramite un ponte di Wheatstone o altro sistema, tali variazioni, si può risalire all'entità della deformazione che le ha causate.

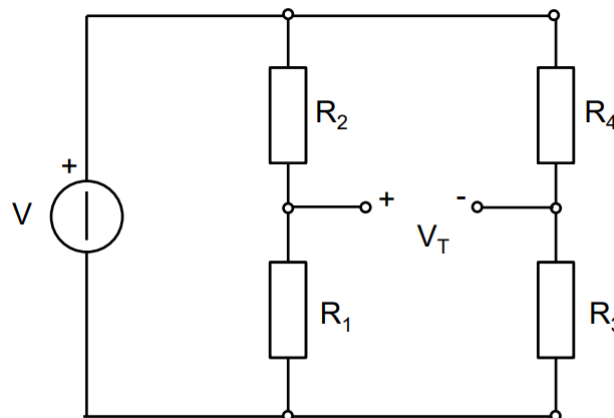


Figura 2. 6 Ponte di Wheatstone che permette di associare a una variazione di resistenza una tensione

Le dimensioni di un estensimetro possono variare da pochi micron ad alcuni centimetri. Quelli di dimensioni minori sono utilizzati per misure puntuali, cioè per conoscere il valore delle deformazioni in un punto preciso, mentre quelli di maggior lunghezza sono utili per rilevare la deformazione media relativa ad un'area più estesa.

La resistenza R è data da:

$$R = \rho \frac{S}{l} \quad (2.9)$$

Con  $\rho$  densità del materiale, S superficie della sezione trasversale del volume, l lunghezza del volume. La variazione infinitesima della resistenza dR sarà data dalla somma delle derivate parziali:

$$dR = \frac{dR(\rho,l,S)}{d\rho} d\rho + \frac{dR(\rho,l,S)}{dl} dl + \frac{dR(\rho,l,S)}{dS} dS \quad (2.10)$$

Sostituendo l'equazione 2.9 nell'equazione 2.10 e dividendo il tutto per R:

$$\frac{dR}{R} = \frac{d\rho}{\rho} + \frac{dl}{l} - \frac{dS}{S} \quad (2.11)$$

Ipotizzando una sezione circolare:

$$\frac{dS}{S} = \frac{d(\pi r^2)}{\pi r^2} = \frac{2dr}{r} \quad (2.12)$$

Ricordando inoltre che il coefficiente di Poisson è dato dal rapporto tra deformazione trasversale e longitudinale:

$$\nu = \frac{\frac{dr}{r}}{\frac{dl}{l}} \quad (2.13)$$

Introducendo l'equazione 2.13 nell'equazione 2.12:

$$\frac{dS}{S} = -2\nu \frac{dl}{l} \quad (2.14)$$

Sostituendo la 2.14 nella 2.11:

$$\frac{dR}{R} = \frac{dl}{l} (1 + 2\nu) + \frac{d\rho}{\rho} \quad (2.15)$$

La variazione relativa della resistenza viene definita sensibilità o fattore di guadagno (Gauge factor) G:

$$G = (1 + 2\nu) + E\Pi \quad (2.16)$$

Con E modulo di Young e  $\Pi$  costante piezoresistiva. La prima componente della 2.16 viene definito effetto geometrico mentre la seconda componente viene definita come effetto piezoresistivo. Nel caso di materiali metallici (nel nostro caso la costantana è una lega di nickel e rame) l'effetto piezoresistivo risulta trascurabile, la resistenza dipende quindi linearmente dalla variazione longitudinale:

$$R = R_0(1 + x) \quad (2.17)$$

Con  $x = (1 + 2\nu)\frac{dl}{l}$



Gli estensimetri possono sperimentare cali di prestazioni dovuti a temperatura, umidità, isteresi del materiale.

## 2.1.4. Telecamere

Il riconoscimento dei gesti *vision-based* prevede l'utilizzo di una o più telecamere con lo scopo di analizzare la sequenza video da più prospettive.

Analogamente ai dispositivi indossabili, esistono varie tipologie di telecamere, alcune di queste in particolare:

-Telecamere ad infrarossi: utilizzate principalmente per le immagini notturne, producono immagini sfuocate e disturbate; tuttavia, esse sono in grado di ottenere una mappa di profondità tridimensionale che colloca un corpo in uno spazio tridimensionale. Queste telecamere sono dotate di un diodo a luce infrarossa (fondamentalmente si tratta di un led) spesso rilevabile in un anello posizionato attorno all'obiettivo.



Figura 2.7: Telecamera ad infrarossi

-Telecamere stereo: tipo di fotocamere dotte di due obiettivi distanziati tra loro circa 6.35 cm (distanza simile a quella tra i due occhi umani) che, mediante un processo di triangolazione delle immagini, sono in grado di fornire direttamente una mappa tridimensionale

-Telecamere monoculari: le più comuni a causa del loro prezzo più abbordabile, sono spesso associate ad algoritmi di riconoscimento dei gesti (spesso *machine learning*).



Figura 2.8 Telecamera per VR a 8 lenti Insta360

-Telecamere PTZ: telecamere che sono in grado di orientarsi e utilizzare lo zoom autonomamente, ovvero quando nota dei cambiamenti dei pixel tra un fotogramma e l'altro.



Figura 2.9: Telecamere PZT Sony

## 2.2. Sistema di acquisizione

Per sistema d'acquisizione dati (abbreviabile con DAQ) si intende un sistema elettronico realizzato per monitorare, registrare ed eventualmente post-elaborare una o più grandezze fisiche misurate sotto forma di segnale. Trova utilizzo nella strumentazione da laboratorio/elettromedicale, automazione, nei sistemi di rilevamento ambientale ecc. Il sistema di acquisizione è infatti in grado di misurare temperature, tensioni, correnti, pressioni o sforzi meccanici, shock e vibrazioni, spostamenti e distanze, angoli ecc.



Figura 2.10 Modulo di acquisizione DAQ National Instrument (Ni)

Le funzioni principali che un DAQ deve possedere sono:

- Acquisizione segnali
- Registro dati
- Visualizzazione dati in real-time
- Analisi dei dati dopo la registrazione
- Analisi dei dati utilizzando funzioni statistiche
- Generazione di report

I segnali di grandezze fisiche vengono acquisiti mediante l'utilizzo di sensori e convertiti in grandezze elettriche da trasduttori.

### Struttura di un modulo DAQ

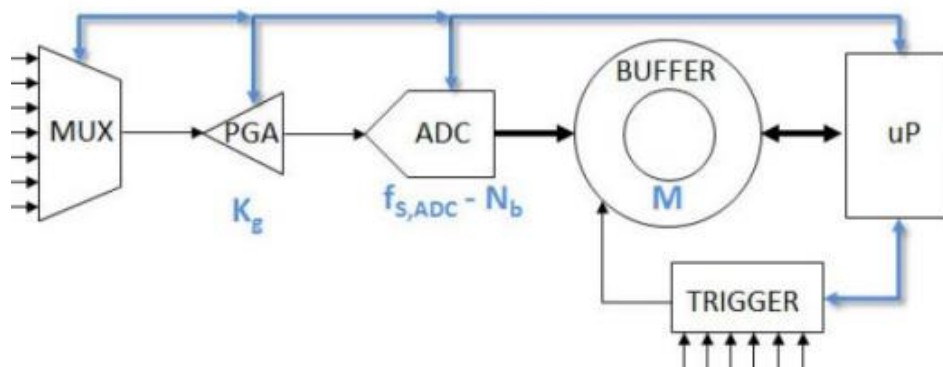


Figura 2.11 Possibile architettura di un DAQ

La struttura di un modulo di acquisizione può essere schematizzata nella figura 2.11, le componenti principali sono:

- Un multiplexer (MUX): Componente che seleziona i segnali in ingresso.
- Un amplificatore a guadagno programmabile (PGA): Componente avente la funzione di adattare l'ampiezza del segnale di ingresso al range dinamico del modulo DAQ.
- Un convertitore analogico digitale: Componente che si occupa di discretizzare nei tempi e nelle ampiezze il segnale in esame.
- Un buffer di memoria: O memoria di acquisizione, è composta da un numero M di celle di memoria, e serve per la memorizzazione temporanea del segnale digitale, tendenzialmente la sua architettura è di tipo FIFO (*First In First Out*), ovvero rimane costantemente piena, eliminando i campioni acquisiti meno recenti.
- Un circuito di trigger: circuito che invia dei segnali di inizio e fine acquisizione al buffer di memoria, regolando di fatto l'inizio e la fine della memorizzazione effettiva del segnale di interesse.
- Un microprocessore: riceve i dati utili dal buffer, inviando in seguito tali dati ad altri dispositivi connessi al modulo mediante interfacce di comunicazione.

Prima dell'ingresso nel modulo, il segnale elettrico in uscita dal trasduttore può subire una serie di operazioni che lo predispongono meglio all'acquisizione del DAQ, questo insieme di operazioni viene definito condizionamento del segnale.

Una possibile operazione di condizionamento del segnale è la rimozione del rumore: il rumore è un segnale indesiderato che si sovrappone al segnale di interesse, abbassando quindi la qualità della misura, il rumore può essere originato da svariati fenomeni, generalmente elettromagnetici (per esempio l'interferenza di rete).

L'influenza del rumore sul segnale misurato può essere ridotta mediante l'utilizzo di filtri, che possono essere classificati in base alla banda di frequenza in cui operano:

- Filtri passa-basso: eliminano le componenti del segnale aventi frequenza maggiore della frequenza di taglio, risultando quindi utili per la rimozione di disturbi alle alte frequenze con un segnale più *smooth*.
- Filtri passa-alto: eliminano le componenti del segnale aventi frequenza minore della frequenza di taglio, risultando quindi utili per la rimozione di drift alle basse frequenze con un segnale filtrato risulta più propenso a variazioni repentine.
- Filtri passa-banda: accentuano o inibiscono le componenti del segnale aventi frequenza contenuta nel range definito dalla banda.

Altre operazioni di condizionamento sono: amplificazione e traslazione di livello, isolamento (per proteggere il dispositivo da correnti di dispersione dannose) e la linearizzazione.

Un punto certamente importante del modulo DAQ è certamente la conversione del segnale da analogico a digitale. Tale conversione consiste nella discretizzazione dell'asse dei tempi e delle ampiezze.

La discretizzazione dell'asse dei tempi viene definito campionamento e consiste nella convoluzione del segnale di interesse con il treno di impulsi di periodo  $T_s$  (periodo di campionamento). Tale operazione porta alla discretizzazione del segnale nel dominio del tempo e ad una ripetizione periodica di periodo  $\frac{1}{T_s}$  nel dominio della frequenza (figura 2.12).

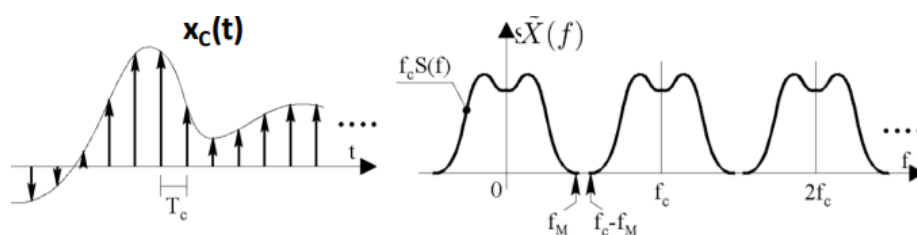


Figura 2.12 Effetti del campionamento nei domini del tempo e della frequenza

Il teorema di Shannon afferma che il campionamento di un segnale a banda limitata è avvenuto correttamente se e solo se la frequenza di campionamento (ovvero il reciproco del periodo di campionamento) è pari almeno al doppio della massima frequenza del segnale (frequenza di Nyquist). Nel caso di non osservanza di tale teorema si verifica una perdita di informazione detta aliasing (ovvero le alte frequenze si nascondono dietro un alias) provocando una distorsione del segnale originale.

La quantizzazione consiste invece nella discretizzazione dell'asse delle ampiezze, dividendolo in B intervalli: l'insieme di valori che ricadono in un intervallo viene approssimato ad un numero valore unico (valore digitale), la distanza tra due livelli digitali viene definita passo di quantizzazione.

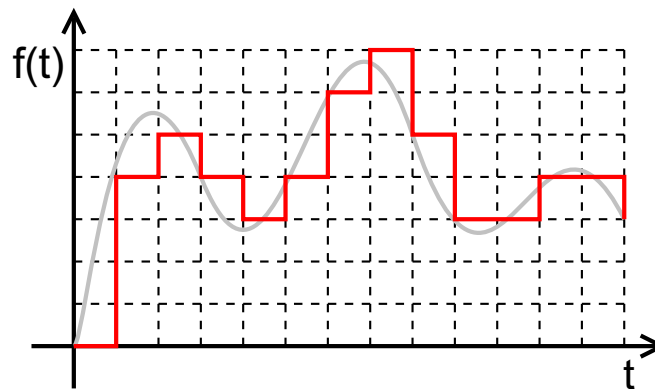


Figura 2.13 Esempio di quantizzazione di un segnale  $f(t)$

Affinché la quantizzazione avvenga in modo ottimale, la variazione dell'ingresso deve essere minore della metà del passo di quantizzazione: oltre alla limitazione del teorema di Shannon viene aggiunta l'ulteriore condizione sulla frequenza massima del segnale:

$$f_0 = \frac{1}{\pi T_{ADC} 2^b}$$

Con  $T_{ADC}$  massima cadenza di campionamento del convertitore analogico-digitale e b i livelli di quantizzazione possibili.

Se tale condizione non verrà rispettata sarà necessario l'utilizzo di un *sample and hold amplifier* (SHA), tale dispositivo è infatti in grado di congelare il segnale permettendo una conversione corretta da parte dell'ADC.

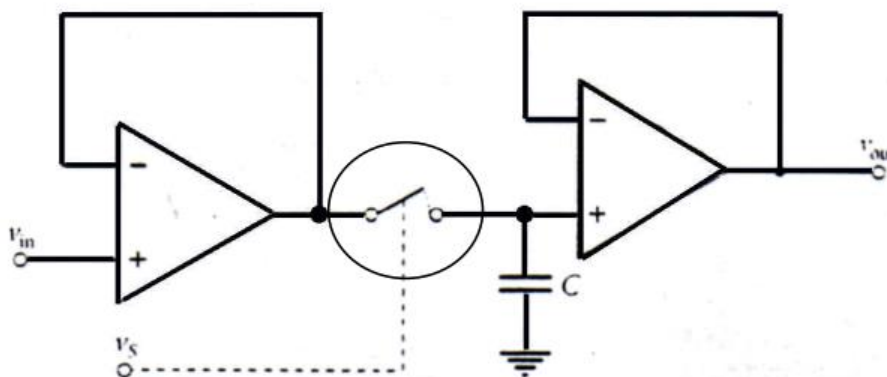


Figura 2.14: Schema circuitale di uno SHA

Lo SHA deve quindi acquisire velocemente il segnale e tenerlo poi bloccato per molto tempo, l'amplificatore all'ingresso ha una resistenza di uscita molto bassa, permettendo la carica veloce di un condensatore, l'amplificatore in uscita avendo resistenza di ingresso molto alta provoca una scarica del condensatore molto lenta.

Per quanto riguarda gli ADC, sono presenti varie architetture, ognuna con vantaggi e svantaggi da tenere in considerazione:

- *Flash converter*
- Ad approssimazioni successive
- Ad integrazione
- A conteggio

La scelta migliore è certamente l'ADC ad integrazione:

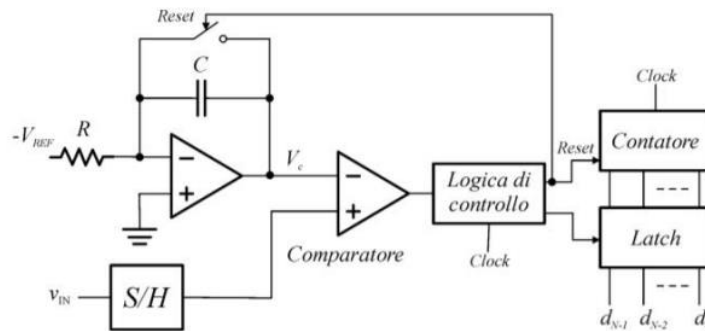


Figura 2.15: Schema a blocchi di un ADC ad integrazione

Questa tipologia di convertitore consiste di una porta logica che darà un valore in uscita alto o basso a seconda dei valori di entrata.

Il circuito integratore genererà un segnale rampa (fungendo anche come filtro passa basso): nel momento in cui tale rampa raggiunge l'ampiezza del segnale da convertire, il comparatore e la porta logica daranno un'uscita negativa, il contatore fornirà il valore digitale, un altro segnale avvierà poi il reset del circuito.

Questo tipo di convertitore si rivela particolarmente adatto per segnali a bassa frequenza ( i movimenti della mano non vanno oltre i 10 Hz) permettendo anche una reiezione automatica del rumore.

## 2.3. Trasmissione dei dati

L'hardware ha il compito di acquisire i dati, la loro elaborazione però potrebbe richiedere l'utilizzo di algoritmi particolarmente complessi eseguiti da calcolatori. La trasmissione dei dati, aspetto altrettanto importante, può avvenire mediante cavi o wireless. Per flessibilità e costi minori è preferibile la seconda tipologia, le principali tecnologie di comunicazione wireless sono:

- Bluetooth.
- Wi-Fi.
- NFC.

### 2.3.1. Sistema Bluetooth

Da dove deriva il nome Bluetooth? Il nome proviene da Aroldo I "Dente Azzurro" (Bluetooth appunto) Gormsson di Danimarca, famoso storicamente per aver unito la popolazione danese in un'unica nazione nel X secolo: la funzione del sistema Bluetooth, infatti, è quella di unire i dispositivi elettronici come fece questo personaggio storico con il suo popolo. Il simbolo del sistema Bluetooth non è altro che l'unione di due rune nordiche che rappresentano le iniziali del re danese (figura 2.16).

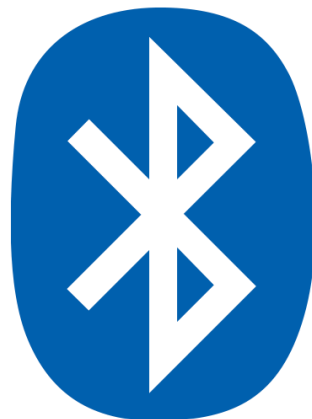


Figura 2.16: Simbolo del sistema Bluetooth

Il sistema Bluetooth è un sistema di trasmissione dati wireless a corto raggio che sfrutta le onde radio per scambiare informazioni tra due o più dispositivi ad un consumo di energia ridotto. Tali dispositivi possono essere computer, telefoni o, nel nostro caso, strumenti di misura.

Il Bluetooth lavora alla frequenza radio di 2.402 e i 2.480 GHz con delle bande di guardia (con lo scopo di prevenire l'interferenza con dispositivi appartenenti ad un altro utente) di 2 MHz (banda inferiore) e 3.5 MHz (banda superiore). Il protocollo per evitare ulteriori interferenze divide la banda in cui lavora in 79 canali commutando tra questi circa 1600 volte al secondo.

Sono presenti tre classi di dispositivo:

- Classe 1: avente raggio di azione di cento metri, questa classe ha impiego industriale
- Classe 2: avente raggio di azione di dieci metri
- Classe 3: avente raggio di azione di un metro

L'effettiva distanza di propagazione dipenderà dalla qualità delle componenti elettroniche, condizioni di propagazione, occlusione di materiali, configurazione dell'emettitore di onde e condizioni della batteria del dispositivo.

Ogni dispositivo Bluetooth è caratterizzato da due parametri fondamentali:

- Un indirizzo a 48 bit assegnato al dispositivo durante la fabbricazione che è unico per ciascun dispositivo
- Un *clock* da 28 bit che "ticchetta" ogni 312.25 microsecondi, questa tempistica corrisponde alla metà del tempo di commutazione quando il dispositivo commuta 1600 volte al secondo.

I dispositivi Bluetooth possono comunicare tra loro mediante l'utilizzo di queste due informazioni.

La *piconet* è un network di connessione di dispositivi Bluetooth comunicanti tra loro, tale rete viene formata ad hoc senza l'utilizzo di infrastrutture di supporto e perdura fin quando l'utente lo desidera. Questa rete contiene almeno un dispositivo definito *master* e altri dispositivi, il clock di questo dispositivo sarà il clock della *piconet*, gli altri dispositivi, detti *slave*, dovranno sincronizzare il loro *clock* (mediante l'aggiunta di un *offset*) con quello della *piconet*. Per identificare ciascuno *slave* (un *master* può comunicare con massimo sette dispositivi), il *master* assegna un indirizzo locale attivo a ciascuno di essi. I dispositivi connessi alla rete ma non attivi vengono invece definiti *parked*. La *piconet* può essere allargata ad una *scatternet*, ovvero una rete formata da due o più *piconet* in cui un dispositivo svolge il ruolo di *master* in una rete e il ruolo di *slave* in un'altra (all'interno di una *scatternet* sono presenti più *master*).

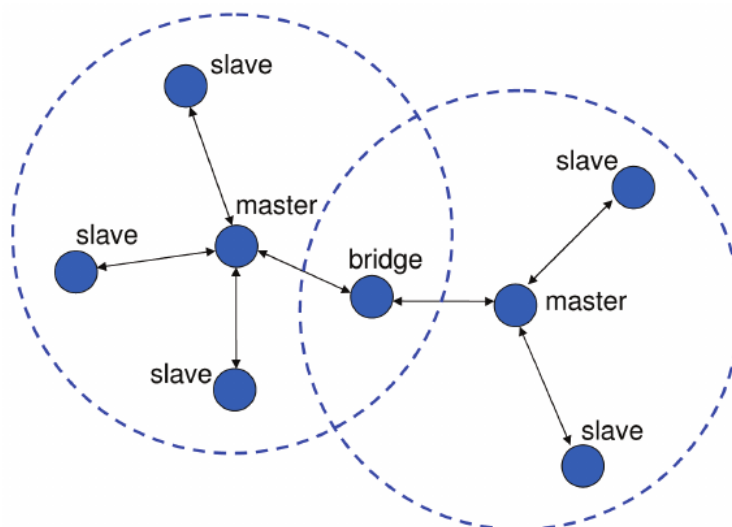


Figura 2.17: Esempio di scatternet



## 2.3.2. Wi-Fi

Il Wi-Fi (acronimo di *wireless fidelity*) o IEEE 802.11 è un insieme di protocolli che permette a un dispositivo elettronico il collegamento ad una rete locale di *devices* o ad internet mediante l'utilizzo di onde radio (frequenza compresa tra i 2.4 GHz e i 6 GHz, è quindi possibile l'interferenza con dispositivi Bluetooth), la banda in cui opera viene divisa in più canali, ognuno di questi viene utilizzato da più reti. Il collegamento ad Internet avviene mediante punti particolari detti punti di accesso wireless (o più comunemente noti come *hotspot*) che collegano l'utente alla rete elettronica e la portata della connessione risulta variabile: dai venti metri all'interno di un edificio fino a chilometri quadrati mediante la sovrapposizione di più *hotspot* collegati tra loro mediante il cablaggio nella rete locale, l'architettura verrà descritta in seguito in modo più dettagliato.

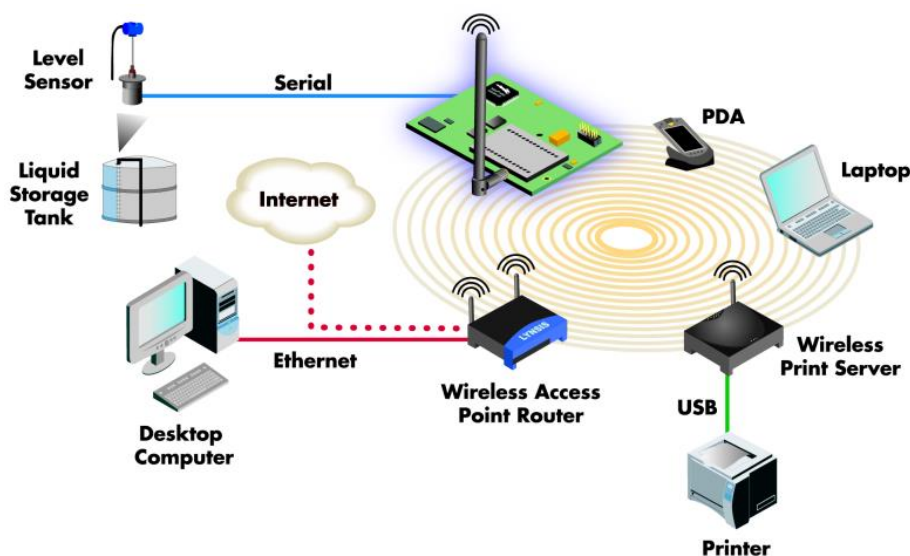


Figura 2.18: Rete locale wireless connessa ad internet

I vantaggi offerti dalla rete Wi-Fi sono molteplici, prime fra tutti l'aumento della flessibilità dei sistemi di monitoraggio, permettendo l'acquisizione di informazioni da posizioni anche remote a costi significativamente inferiori. I vantaggi non sono però solo questi e possono essere riassunti nel seguente elenco:

- Ethernet reso wireless: Il Wi-Fi non è altro che un rimpiazzo della tecnologia Ethernet, entrambi possiedono elementi fondamentali in comune.
- Accesso esteso: La mancanza di cablaggio permette la connessione in zone in cui non questo non sarebbe fisicamente o economicamente realizzabile.
- Riduzione dei costi: Come menzionato prima, il collegamento wireless abbassa i costi di realizzazione, non sono più necessari i cavi di connessione e nemmeno tutti i lavori necessari per consentire la loro presenza (e.g. scavare, trivellare).
- Mobilità: è possibile cambiare posizione con il proprio PC/cellulare senza preoccuparsi di perdite di connessione (rimanendo chiaramente nell'area di ricezione).
- Flessibilità: I vantaggi precedentemente elencati permettono la realizzazione di nuove tecnologie e nuove soluzioni alle applicazioni già note.

Tutti i dispositivi capaci di connessione Wi-Fi vengono definiti sono chiamati stazioni wireless (STAs), la connessione di due o più STAs formano un *basic service set* (BSS) che è controllato da una funzione di coordinazione (CF) che stabilisce quando una STA riceve o trasmette segnale

(non è detto che le STA appartenenti ad un BSS comunichino tra loro). I BSS possono essere indipendenti (costituiscono la forma più basilare di Network) e pur non essendo connessi ad altre reti o network, possono essere utilizzati in situazioni in cui non è presente un'infrastruttura, o questa risulta danneggiata, che permette la connessione al web (e.g. scenari di guerra, siti di emergenza). Per poter connettere più BSS è quindi necessaria una particolare infrastruttura che richiede punti di accesso (o *hotspot*), che non sono altro che STA aventi una funzionalità aggiuntiva: permettere l'accesso alla rete cablata che collega altre BSS mediante altri *hotspot*, tale rete cablata viene definita come sistema di distribuzione (DS).

La rete Wi-Fi può essere infatti quindi suddivisa in due porzioni principali:

- Rete di accesso, che collega gli utenti finali al primo commutatore della rete di trasporto (rappresentata da BSS contenenti STAs).
- Rete di trasporto, che provvede al trasporto dei dati degli utenti su scala geografica (DS).

La rete di accesso risulta essere wireless, quindi affidata al trasporto di dati mediante le onde radio mentre la rete di trasporto si serve di dorsali cablate: sono possibili architetture completamente wireless, apportando sicuramente costi inferiori ma con prestazioni altrettanto inferiori dovute alla diminuzione dell'efficienza spettrale.

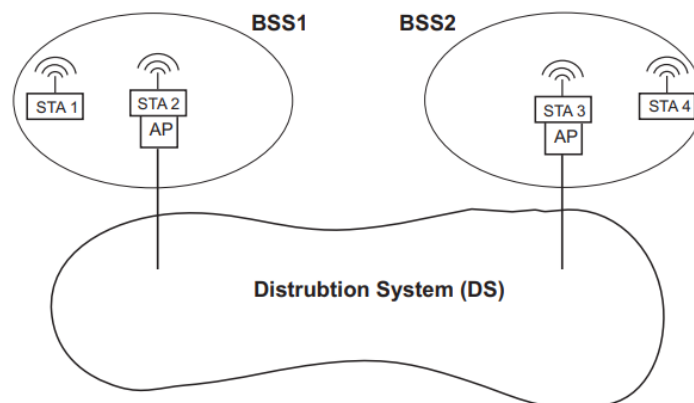


Figura 2.19: Collegamento concettuale di più BSS

Un DS e due o più BSS costituiscono invece un *extended service set* (ESS). Un ESS è una rete Wi-Fi di complessità variabile.

Nelle reti Wi-Fi tendenzialmente vengono utilizzate due frequenze principali:

- 2.4 GHz: Comprende 14 canali, ciascuno con banda compresa tra i 20 e i 22 MHz, buona velocità di traffico dati, penalizzata però dall'interferenza di altri dispositivi (Bluetooth, microonde, radio) essendo una banda particolarmente affollata.
- 5 GHz: Comprende 13 canali, ciascuno con banda di circa 20 MHz, minore velocità di traffico dati, col vantaggio di essere meno affollata (in futuro probabilmente lo sarà di più).

I segnali ad alta frequenza risultano attenuati da ostacoli fisici (e.g. muri) a causa del trasferimento di energia del campo magnetico all'oggetto fisico, portando quindi ad una riduzione del segnale.

La forza di un segnale è funzione della potenza di uscita del trasmettitore, la frequenza utilizzata e le perdite di energia dovute al percorso dalla sorgente al punto di arrivo del segnale. La forza del segnale subisce variazioni dinamiche e imprevedibili: piccoli cambiamenti dell'ambiente di propagazione possono portare a drastici cambiamenti della qualità del segnale. Oltre alla forza del segnale è importante controllare il rapporto segnale-rumore (SNR), il rumore è infatti una componente indesiderata del segnale trasmesso ed è funzione della potenza della sorgente di interferenza, della larghezza di banda e della vicinanza alla fonte di trasmissione. Alcuni accorgimenti per aumentare l'SNR sono:

- Posizionamento ottimale dei punti di accesso, distante il più possibile dalle fonti di interferenza.
- Antenne ad alto guadagno posizionate in prossimità o dell'*hotspot* o delle STAs con SNR alto (di per sé l'interferenza non viene rimossa, ma resa trascurabile).

### 2.3.3. NFC

La comunicazione di prossimità o *near-field communication* (NFC), fondata e inizializzata da Sony e Philips, è un'attività di ricetrasmisione che fornisce connettività wireless a corto raggio (massimo con un massimo di 10 cm) bidirezionale, permettendo quindi una connessione peer-to-peer (i due nodi, in questo caso i due dispositivi, ricevono e trasmettono contemporaneamente informazioni, comportandosi contemporaneamente sia da client che da server). Questo sistema di comunicazione ha permesso ai telefoni di sostituire le carte di credito per effettuare pagamenti (basti pensare alle varie app come Satispay e Yapp). Il sistema NFC consente quindi una connessione rapida, semplice e non necessita di infrastrutture di supporto (a differenza del Wi-Fi), inoltre una particolarità della connessione NFC è la sua sicurezza: essendo un tipo di comunicazione a corto raggio è difficile intercettare i dati. NFC può essere implementata nei dispositivi mediante chip integrato o schede SD dotate di chip NFC.

La frequenza operativa di NFC risulta essere di 13.56 MHz (non è presente quindi il pericolo di interferenza con Wi-Fi e Bluetooth) con una velocità di trasmissione che varia da 106 a 424 kbit/sec. La diffusione dei dati avviene mediante l'utilizzo di un campo magnetico alternato, in questo modo non abbiamo dispersione di energia sotto forma di onde radio.



Figura 2.20: Possibili utilizzi della connessione NFC

### 3. Analisi ed elaborazione dei dati

Negli ultimi decenni l'elaborazione dei segnali ha subito una forte evoluzione, passando dalla semplice analisi del segnale nel dominio del tempo ad analisi più avanzate nel dominio della frequenza ad analisi complesse in domini a dimensione superiore. Il segnale estratto dal modulo DAQ sarà una funzione in una o più variabili indipendenti tra loro. Nella maggior parte dei casi, l'analisi di questo segnale consisterà nell'identificare il comportamento del segnale, estrarre proprietà lineari e non lineari, compressione oppure espansioni a dimensioni minori o maggiori, e il riconoscimento di patterns.

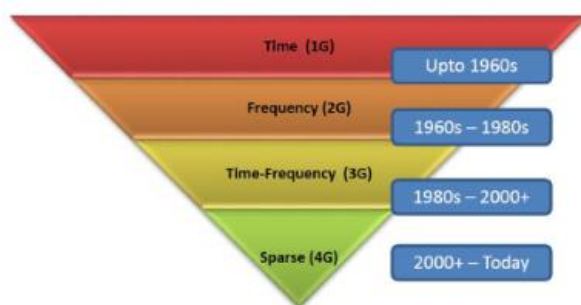


Figura 3.1 Evoluzione delle tecniche di analisi dei segnali

Il primo obiettivo dell'elaborazione dei segnali è quello di aumentare il rapporto segnale-rumore (SNR), definito come il rapporto del contributo del segnale utile rispetto al contributo del rumore (ovvero la componente non interessante del campione acquisito). Questo passaggio prevede l'utilizzo di tecniche di *denoising*: tali tecniche variano dall'utilizzo dei semplici filtri a media mobile o dai filtri aventi frequenze di taglio selezionate in base alle componenti armoniche del segnale da attenuare, fino alle più avanzate tecniche basate su algoritmi di scomposizione del segnale (e.g. wavelet functions) oppure su algoritmi che separano i contributi delle varie sorgenti in gioco, noti come Blind Source Separation (e.g. Principal Components Analysis, Independent Components Analysis).

Eseguito il *denoising*, la successiva elaborazione dei dati (e quindi dei segnali) coinvolgerà vari passaggi, significativamente dipendenti dalle applicazioni desiderate.

In generale gli obiettivi principali comprendono:

- Estrazione di caratteristiche utili (features)
- Estrazione di informazioni tramite confronto e correlazione con segnali noti
- Derivazione di segnali di controllo per il comando di dispositivi tramite la classificazione delle features estratte, con algoritmi ad hoc.

# 3.1. Estrazione di features

L'estrazione delle features di un segnale si pone i seguenti obiettivi:

- Identificare i parametri che caratterizzano al meglio il segnale di interesse e il suo comportamento
- Riconoscere eventi

Siamo quindi in grado di classificare il segnale di interesse da un insieme di features caratteristiche di un determinato comportamento o pattern del segnale. L'estrazione di features consiste spesso nella riduzione delle variabili indipendenti (dimensioni) del segnale oppure alla compressione/riduzione dei dati con lo scopo di ridurre il numero di sorgenti necessarie per l'analisi, in modo da rendere i processi di estrazione robusti ed efficienti.

Le features possono essere estratte:

- Nel dominio del tempo: utilizzando tecniche autoregressive, codifica predittiva lineare (LPC), riconoscimento di forme, analisi Cepstrum e *kernel-based modelling*
- Nel dominio della frequenza: utilizza tecniche che visualizzano il segnale nel dominio della frequenza come la *Fast Fourier Transform* (FFT), Discrete Cosine Transform (DCT), la stima spettrale e la trasformata di Hilbert.
- In entrambi i domini: ovvero analizza il segnale sia nel tempo che nella frequenza, per esempio la Short Time Fourier Transform (STFT), CWT e DWT (Continuous and Discrete Wavelet Transform), Wavelet Packet Transform (WPT), ambiguity function e Non-Negative Matrix Factorization(NMF)
- Decomponendo ed estraendo caratteristiche di dominio sparse: come l'Empirical mode Decomposition (EMD), Intrinsic Time-Scale Decomposition (ITD), singular value decomposition (SVD) e analisi tensoriale

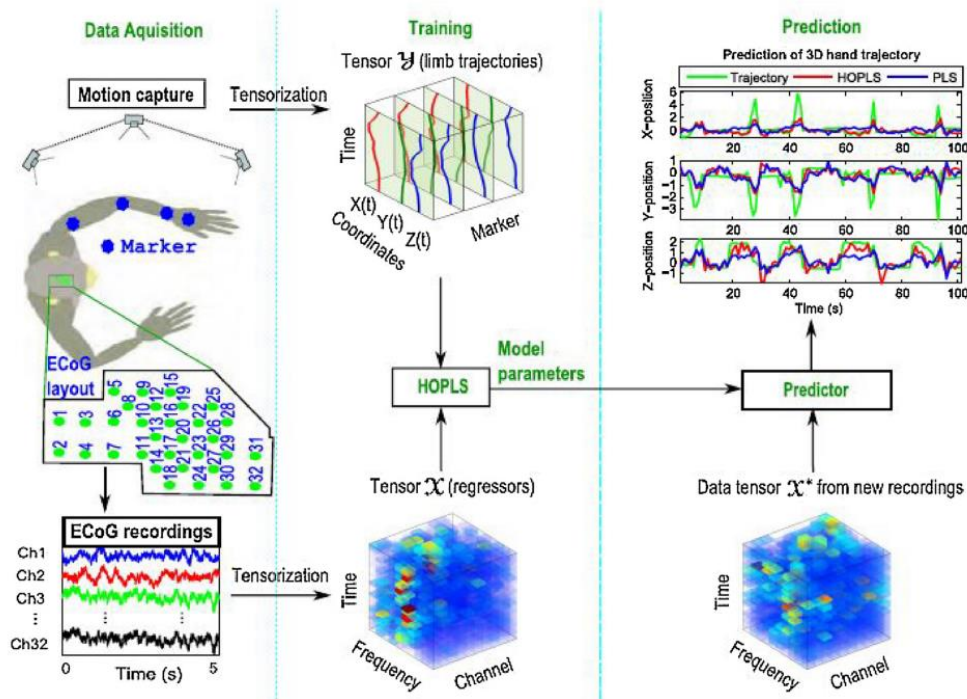


Figura 3.2: Esempio di estrazione di features nel motion capture mediante l'analisi tensoriale

Nel caso di un'immagine (motion capture per esempio) i metodi di estrazione possono essere divisi in due tipologie:

- Pixel-based: famiglia di metodi più tradizionali, che analizza i pixel in modo individuale, spesso utilizzando anche l'analisi spettrale dell'immagine.
- Object-based: famiglia di metodi più moderni che studia invece le correlazioni tra pixel contigui, orientandosi quindi all'analisi degli oggetti, ovvero insieme di pixel con caratteristiche comuni.

Il secondo metodo permette di estrarre molte più informazioni (quindi più features) in quanto analogo all'interpretazione visiva umana di un'immagine e non dovendo ricorrere ad analisi nel dominio della frequenza risulta, inoltre, più efficiente.

Un esempio interessante ed utile per descrivere in modo completo l'estrazione di quasi tutte le features utili, per poi sfruttarle nel tracciamento dei movimenti della mano, risulta essere l'analisi di un tracciato elettromiografico EMG. Per EMG si intende la registrazione dei potenziali elettrici che si formano in un muscolo durante la sua contrazione volontaria, tali potenziali sono causati dalla depolarizzazione elettrica delle fibre muscolari in risposta all'arrivo di un impulso elettrico alla sinapsi neuromuscolare.

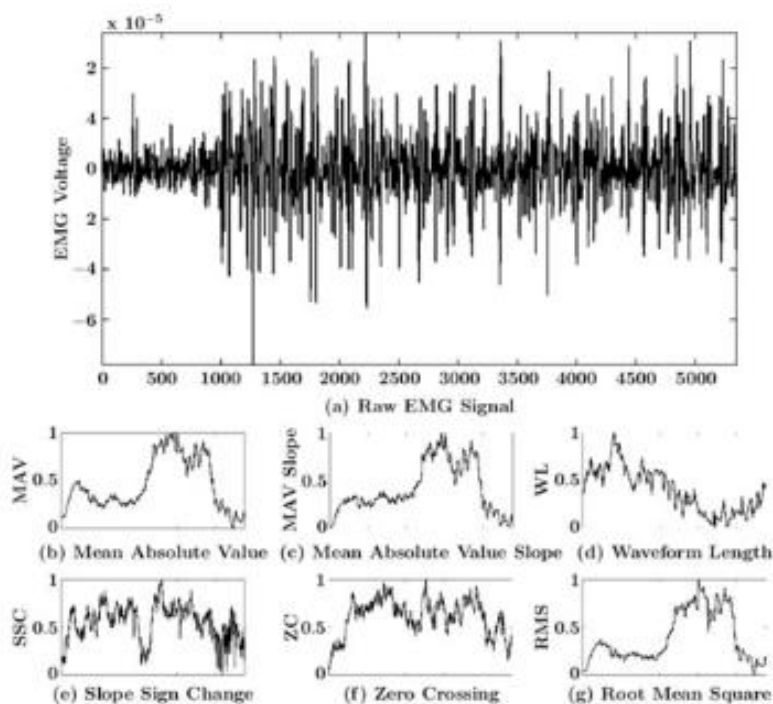


Figura 3.3: Esempio di estrazione di features da un tracciato EMG

Tali features raggruppano in tutto 15 elementi sia dal dominio del tempo (13) che della frequenza (2):

- Integrale dell'EMG (IEMG): utilizzato come valore di onset per rilevare l'attività dei muscoli (un movimento provocherà un aumento dell'integrale) ed è dato dalla somma dei valori assoluti dei campioni acquisiti

$$IEMG = \sum_{n=1}^N |x_n|$$

- Media dei valori assoluti (MAV): Avente funzione simile a quella dell'integrale

$$MAV = \frac{1}{N} \sum_{n=1}^N |x_n|$$

- Valore assoluto modificato 1 (MAV1): Versione modificata del MAV, utilizzando una finestra di pesatura per aumentare la robustezza della feature estratta

$$MAV1 = \frac{1}{N} \sum_{n=1}^N w_n |x_n|$$

$$\text{Con } w_n = \begin{cases} 1 & \text{se } 0.25N \leq n \leq 0.75N \\ 0.5 & \text{altrimenti} \end{cases}$$

- MAV2: feature correlata a MAV1 ma con una finestra più complessa:

$$w_n = \begin{cases} 1 & 0.25N \leq n \leq 0.75N \\ \frac{4n}{N} & 0.25N > n \\ \frac{4(n-N)}{N} & 0.75N < n \end{cases}$$

- Pendenza media del valore assoluto (MAVS): versione alternativa del MAV

$$MAVS_n = MAV_{i+1} - MAV_i$$

- Integrale quadratico (SSI): feature che esprime l'energia del segnale

$$SSI = \sum_{n=1}^N |x_n|^2$$

- Varianza (VAR): feature che esprime la potenza del segnale

$$VAR = \frac{1}{N-1} \sum_{n=1}^N x_n^2$$

- Scarto quadratico medio (RMS): associato a contrazione costante e senza fatica

$$RMS = \sqrt{\frac{1}{N} \sum_{n=1}^N x_n^2}$$

- Lunghezza d'onda (WL): lunghezza cumulativa dell'onda nel tempo di misura

$$WL = \sum_{n=1}^{N-1} |x_{n+1} - x_n|$$

- Zero crossing (ZC): il numero di volte che il segnale si interseca con l'asse x (ovvero quando assume valore nullo), viene utilizzata una soglia per prevedere l'effetto del rumore, può fornire alcune informazioni sul dominio della frequenza

$$ZC = \sum_{n=1}^{N-1} [\text{sgn}(x_n \times x_{n+1}) \cap |x_n - x_{n+1}| \geq \text{threshold}]$$

$$\text{sgn}(x) = \begin{cases} 1 & x > \text{threshold} \\ 0 & \text{altrimenti} \end{cases}$$



- Cambio di pendenza (SSC): ovvero il numero di cambi di pendenza (o il numero di cambi di segno della derivata del segnale)

$$SSC = \sum_{n=2}^{N-1} [f[(x_n - x_{n-1}) \times (x_n - x_{n+1})]]$$

$$\text{Con } f(x) = \begin{cases} 1 & x > \text{threshold} \\ 0 & \text{altrimenti} \end{cases}$$

- Ampiezza di Willison (WAMP): il numero di volte in cui la differenza tra due campioni adiacenti supera una certa soglia, riduce il rumore e permette di calcolare in modo ottimale lo zero crossing, viene associato all'inizio della contrazione dell'unità motoria

$$WAMP = \sum_{n=1}^{N-1} f(|x_n - x_{n+1}|)$$

$$\text{Con } f(x) = \begin{cases} 1 & x > \text{threshold} \\ 0 & \text{altrimenti} \end{cases}$$

Per ZC, SSC, WAMP il valore di *threshold* risulta essere compreso tra i 10 e i 100 mV, e dipende dalle impostazioni di guadagno della strumentazione utilizzata.

- Coefficienti auto regressivi (AR): utilizzati per descrivere un campione di segnale EMG come combinazione lineare dei campioni di EMG precedenti e di rumore bianco

$$x_n = - \sum_{i=1}^p a_i x_{n-1} + w_n$$

- Frequenza mediana (MDF): frequenza in cui lo spettro del segnale è diviso in due porzioni aventi la stessa ampiezza totale

$$\sum_{n=1}^{MDF} P_n = \sum_{n=MDF}^M P_n = \frac{1}{2} \sum_{n=1}^M P_n$$

- Frequenza media: calcolata come sommatoria del prodotto tra frequenza e sua ampiezza, tutto diviso per l'intensità spettrale totale

$$MNF = \frac{\sum_{i=1}^M f_i P_i}{\sum_{i=1}^M P_i}$$

Le formule descritte per ricavare queste features possono essere generalizzate anche per l'analisi di altri segnali (e.g. accelerazioni, posizione) utili per la valutazione dei gesti.

Le features estratte verranno poi compattate in un vettore di input inserito in una macchina di supporto vettoriale (SVM) descritta nel paragrafo 3.3.



## 3.2. Correlazione tra segnali

In alcune specifiche applicazioni, a causa di specifiche caratteristiche del segnale di interesse (e.g. segnali stocastici, tempo variabili), può non risultare possibile o conveniente estrarre le features importanti direttamente dal segnale. In questi casi può risultare possibile e utile invece comparare tale segnale con un template di un segnale invece conosciuto, e, se questi coincidono, estrarne le caratteristiche utili.

La correlazione tra due segnali ( $x(t)$  e  $y(t)$  per esempio) invece consiste nella misura della similitudine tra segnali, matematicamente viene definita come:

$$r_{x,y}(\tau) = \int_{-\infty}^{+\infty} x^*(\tau)y(t + \tau)d\tau$$

Con  $x^*(t)$  complesso coniugato del segnale  $x(t)$ , serve infatti per garantire che le funzioni con valori complessi contribuiscano positivamente all'integrale.

Nella realtà si ha però sempre a che fare con segnali discreti  $x(n)$  e  $y(n)$ , la correlazione diventa quindi:

$$r_{x,y}(\tau) = \sum_{m=-\infty}^{+\infty} x^*(m)y(m + n)$$

Il significato fisico della correlazione non è altro che una misura quantitativa della somiglianza tra due segnali, dando un valore molto alto nel caso in cui i picchi coincidono o sono di valore opposto (il loro contributo sarà molto rilevante nell'integrale/sommatoria).

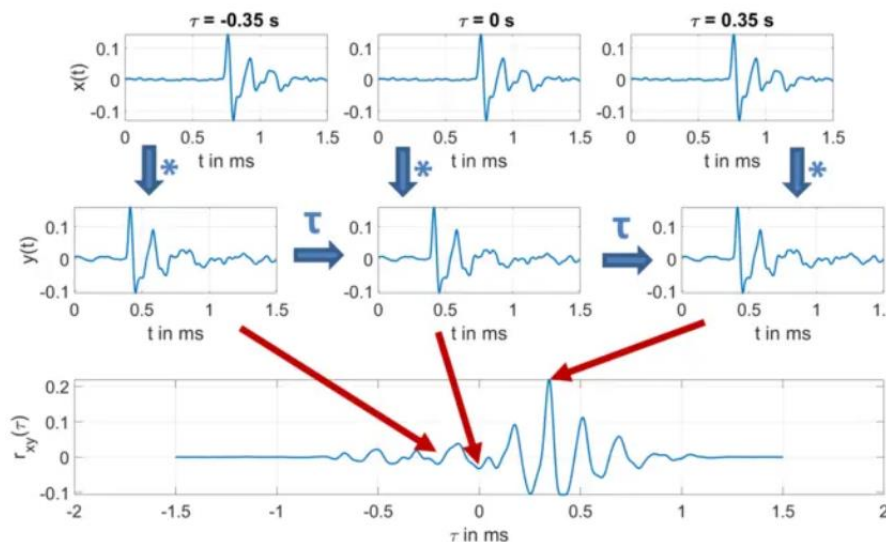


Figura 3.4: Esempio di correlazione tra due segnali continui  $x(t)$  e  $y(t)$

Nel caso di segnali stocastici, ovvero non descrivibili da un'espressione matematica, l'espressione  $r_{x,y}$  è data dalla correlazione dei valori dei due segnali in tempi differenti, rendendo l'operazione dipendente da due variabili.

Ipotizzando che  $X_{t1}$  e  $Y_{t2}$  siano i valori rispettivamente del segnale X al tempo t1 e del segnale Y al tempo t2, con media  $\mu_x(t)$  e  $\mu_y(t)$  e varianza  $\sigma_x^2(t)$  e  $\sigma_y^2(t)$  ad un dato tempo t, la correlazione sarà data da:

$$r_{x,y}(t1, t2) = E[X_{t1}Y_{t2}^*]$$

Ovvero il valore atteso del prodotto dei due segnali (Y complesso coniugato).

La correlazione può essere applicata anche ad immagini, utilizzando una maschera (definita anche kernel): risultati alti implicheranno una somiglianza alta della regione analizzata con la maschera applicata. Ipotizzando che la maschera applicata sia composta da un numero dispari (nel caso di parità verrà effettuato lo zero padding) di pixel (quindi della forma  $(2N+1) \times (2N+1)$ , con N numero naturale), rappresentabile come una matrice della stessa dimensione. In pratica allineando un pixel dell'immagine con il centro del filtro, moltiplicare tutti i fattori sovrapposti per poi sommarli, matematicamente:

$$F \bullet I(x, y) = \sum_{j=-N}^N \sum_{i=-N}^N F(i, j)I(x + i, y + j)$$

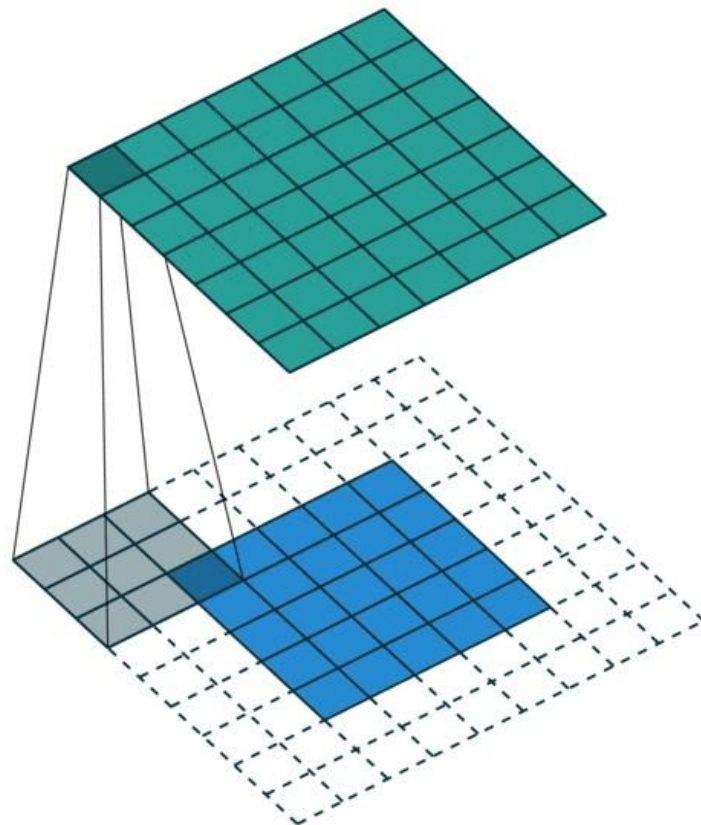


Figura 3.5: Esempio di correlazione in un'immagine

### 3.3. Machine Learning

Il machine learning, o apprendimento automatico, è una branca dell'intelligenza artificiale che si occupa di migliorare la performance di un algoritmo nel rilevare particolari pattern di dati mediante l'utilizzo di metodi statistici, attraverso tali metodi si potranno poi fare previsioni sui dati analizzati.

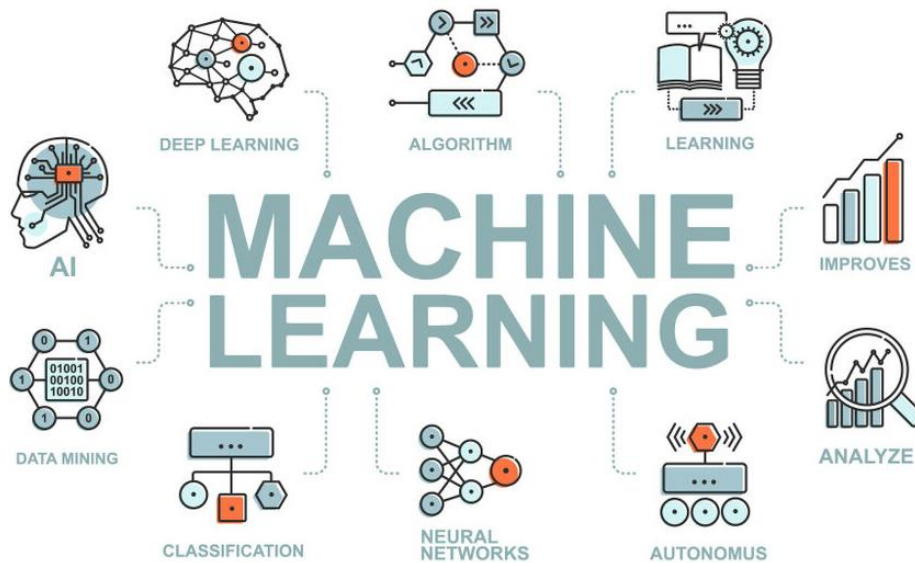


Figura 3.6: possibili applicazioni del machine learning

Il machine learning trova applicazione in ambiti in cui l'utilizzo di algoritmi espliciti risulta impraticabile come nel nostro caso il riconoscimento dei gesti.

Per parlare di apprendimento automatico è però sufficiente e necessaria una condizione: il programma deve migliorare le proprie prestazioni dopo aver svolto un determinato compito.

L'obiettivo cardine del machine learning è quindi quello di permettere ad un programma di fare ragionamenti di tipo induttivo, in pratica la macchina deve essere in grado di portare a termine compiti che non ha mai eseguito precedentemente utilizzando l'esperienza acquisita da un particolare tipo di dati, detto dati di addestramento (*training examples*), considerati rappresentativi del fenomeno da apprendere, e costruendo quindi un modello probabilistico da cui fare previsioni.

La realizzazione dell'ipotesi induttiva risulta quindi di particolare importanza per determinare le prestazioni dell'algoritmo, in particolare la complessità dell'ipotesi sarà strettamente legata alla complessità del *task* da svolgere, nel caso di violazione di tale vincolo possono svilupparsi due fenomeni negativi:

- *Overfitting*: la realizzazione dell'ipotesi è eccessivamente complicata rispetto al problema da svolgere, il modello elaborato non è quindi una generalizzazione dei training examples forniti per risolvere altri pattern di dati, ma una struttura costruita ad hoc per i

dati di addestramento: la performance del caso di *overfitting* è infatti ottima nel caso di *training examples* ma pessima nel caso di dati generici in input.

- *Underfitting*: la realizzazione dell'ipotesi è eccessivamente semplice rispetto al problema da risolvere, portando alla realizzazione di un modello troppo semplificato che non è in grado di fare previsioni corrette su un pattern di dati fornito in input: la performance in questo caso risulta pessima sia per i *training examples* sia per i dati in input.

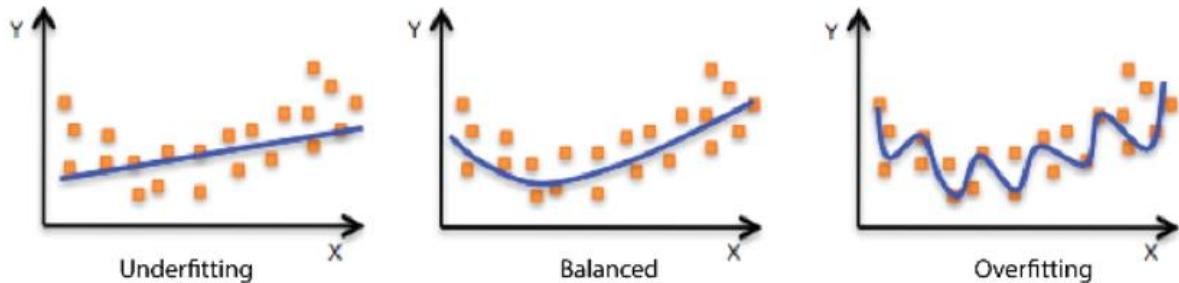


Figura 3.7: Esempio di overfitting e underfitting

Oltre alla complessità del modello, bisogna tenere conto della fattibilità dell'apprendimento, che deve avvenire in tempi polinomiali, ovvero tempistiche che sono limitate superiormente da un'espressione polinomiale che sarà O-grande della funzione che associa il tempo di apprendimento al numero di campioni forniti alla macchina.

Per poter realizzare l'ipotesi induttiva è necessario l'apprendimento partendo dai dati di addestramento, questo può essere di tre possibili tipologie:

- Supervisionato: vengono forniti degli input a cui vengono associati gli output desiderati con lo scopo di estrarre una regola di associazione tra entrata e uscita.
- Non supervisionato: il modello deve trovare una struttura negli input forniti, senza l'associazione ad output desiderati.
- Di rinforzo: il modello interagisce con un ambiente dinamico, venendo migliorato in maniera iterativa fino a quando un operatore (o insegnante) conferma il raggiungimento dell'obiettivo desiderato.

I dati in output finali potranno poi essere forniti in tre possibili modi:

- Classificati: ovvero l'output viene diviso in due o più classi, per esempio un programma che filtra la posta elettronica dallo spam fornisce due classi di output, ovvero le mail spam e le mail non spam.
- Regrediti: l'output viene fornito sotto forma di funzione matematica, in pratica viene fornito il modello matematico.
- In *clustering*: l'input viene diviso in gruppi (*cluster*) dall'algorithm, non inizialmente noti, rendendo tali output tipici di addestramenti non revisionati.

I principali algoritmi di apprendimento automatico sono:

- Programmazione logica induttiva
- Albero di decisione
- Regole di associazione
- Reti neurali artificiali
- Programmazione genetica
- Reti bayesiane
- *Support Vector Machine* (SVM)

- Apprendimento profondo
- *Clustering*

Il machine learning trova utilizzo nell'apprendimento dei gesti umani. Recentemente, infatti, gli algoritmi di machine learning, in particolare le reti neurali artificiali (ANN) e le *support vector machines* (SVM), hanno acquisito molta popolarità grazie alla loro versatilità e tendenza ad automigliorarsi. Ogni algoritmo ha inoltre i propri vantaggi e svantaggi dovuti non solo al costo computazionale, ma anche dalle prestazioni dipendenti anche dai training examples forniti.

La segmentazione della mano e l'estrazione di features (quali media e varianza dei livelli di grigio della mano, il suo perimetro e le aree di convessità, l'orientazione e la firma radiale) risulta un passo cruciale dello stadio di pre-processione dei movimenti della mano. Sarà inoltre necessario rimuovere le componenti di rumore mediante l'utilizzo di filtri come filtri Kalman, gaussiani, mediani ecc.

L'estrazione di features avviene mediante l'utilizzo combinato di tecniche di sottrazione di background con le tecniche di riconoscimento cromatico della pelle della mano. La pelle viene individuata mediante il modello HSV (*Hue-Saturation-Value*): in pratica l'immagine di input è data dal format RGB, ovvero il colore di un pixel sarà dato dalla combinazione lineare dell'intensità di rosso (R), verde (G) e blu (B) e verrà convertita in HSV mediante le seguenti equazioni:

$$H = \begin{cases} 0 & \text{se } G \geq B \\ 2\pi - \theta & \text{se } G < B \end{cases}$$

$$S = \frac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B)}$$

La soglia in cui solitamente ricade la pelle risulta trovarsi nelle condizioni in cui  $0^\circ < H < 25^\circ$  e  $70^\circ < S < 180^\circ$ . La segmentazione della mano avviene riconoscendo delle regioni adiacenti aventi lo stesso valore di H ed S. Il riconoscimento dei gesti avverrà analizzando le strutture di connessione tra le varie regioni della pelle, si assume infatti che all'interno del fotogramma sarà solo la mano ad avere le proprietà cromatiche della cute.

Vengono poi operate delle trasformazioni morfologiche in cui le zone corrispondenti alla mano vengono identificate con il numero binario "1" e i pixel che risultano avere H ed S non inclusi nel range di appartenenza vengono identificati come "0"; vengono poi applicate tecniche di erosione e dilatazione per dissolvere le connessioni più deboli delle regioni della mano e per rimuovere il rumore dello sfondo che potrebbe portare ad identificare pixel di background come regioni della mano.

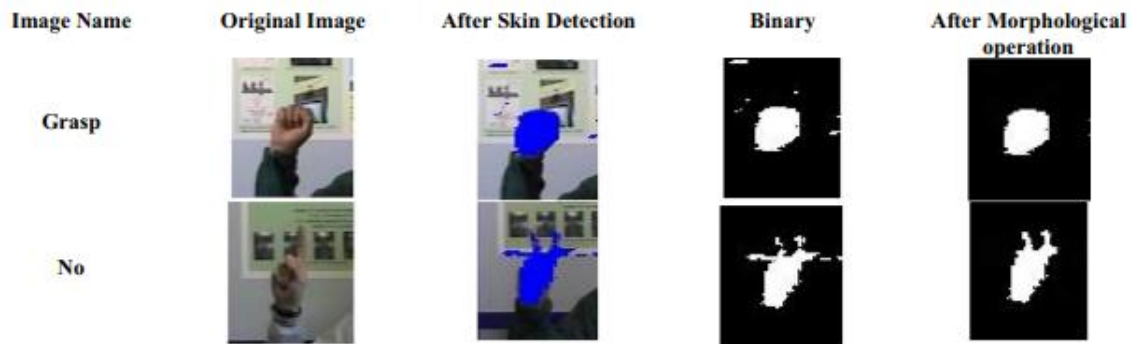


Figura 3.8: Isolamento della mano dal background

Le features estratte sono tendenzialmente:

- L'integrale dell'immagine:  $\sum_{m=1}^{m<x} \sum_{n=1}^{n<y} I(m, n)$ , con x e y dimensioni della matrice di pixel e I il valore del livello di grigio del pixel alle coordinate m ed n.
- I punti chiave (*keypoints*), ovvero regioni spaziali del fotogramma particolarmente ricchi di informazioni locali dell'immagine. I metodi di estrazione di *keypoints* sono molteplici, i principali sono SIFT (*Scale-Invariant Feature Transform*), SURF (*Speed Up Robust Feature*) e PCA (*Principal Component Analysis*); ciascuno con i propri vantaggi e svantaggi.
- Orientamento della mano.
- Parametri (descriptors) estratti dai punti chiave.

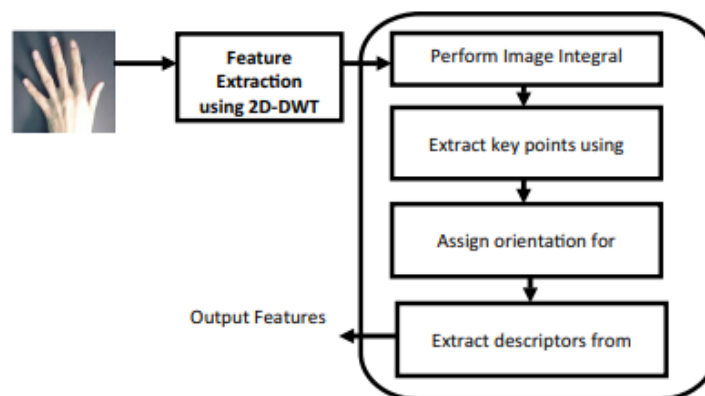


Figura 3.9: Estrazione delle features

Le features di ogni immagine verranno poi raccolte in un vettore composto da 60 elementi (in alcuni casi questo numero può variare dal numero di *keypoint* estratti).

La parte finale del processo consisterà invece nella classificazione dei gesti riconosciuti, che consiste nella classificazione del vettore di input contenente le features estratte.

La SVM utilizzata costruisce inizialmente un iperpiano in uno spazio multidimensionale, che nel nostro caso sarà costituito da 60 dimensioni, il vettore di input verrà rappresentato come un punto in questo spazio, l'iperpiano ha la funzione di separare due classi tra loro per identificare le regioni associate alla classe: se un vettore di input ricade in questa regione verrà classificato come il gesto associato; nel caso di una classificazione multipla serviranno più iperpiani, ciascuno infatti separerà una classe dall'altra, questi poi verranno riadattati per permettere una divisione coerente delle classi.

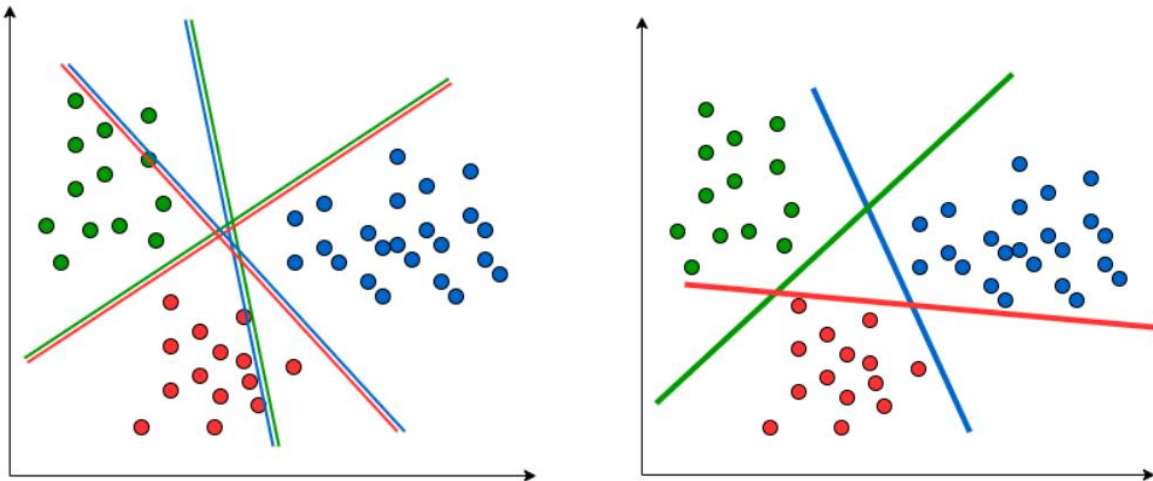
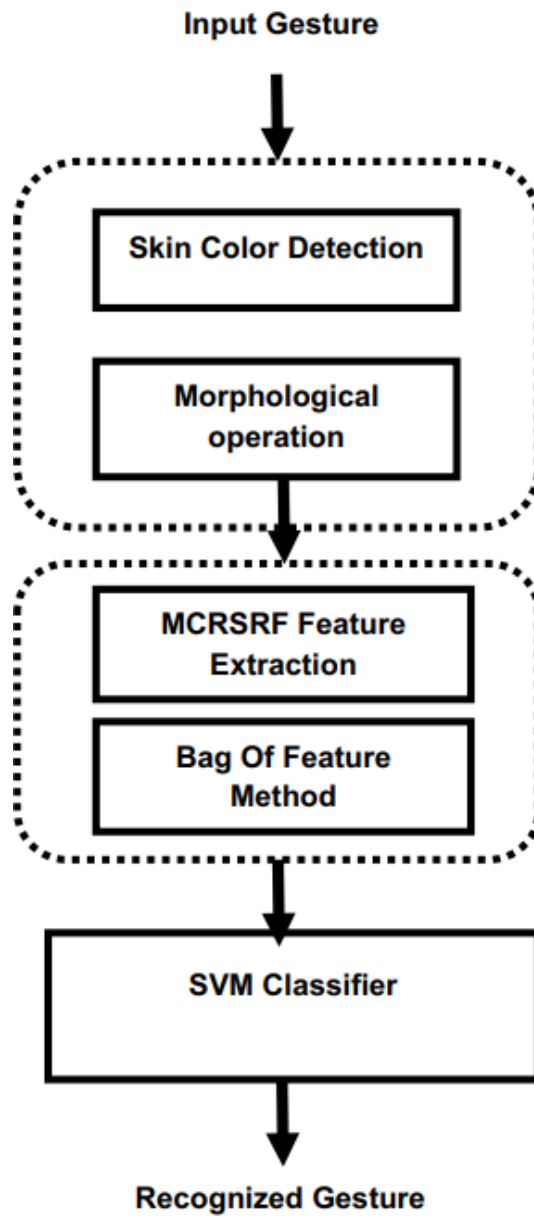


Figura 3.10: Divisione delle categorie di campioni mediante l'utilizzo di iperpiani da parte della SVM

Inizialmente i vettori da classificare verranno separati dai vettori derivati dai training examples. I vettori di addestramento daranno come uscita dei punti nell'iperpiano che saranno etichettati come la classi dei gesti predefiniti, associate alle features estratte. Dopo aver analizzato i training examples la SVM realizzerà un modello matematico in grado di prevedere la classe di appartenenza del vettore di input analizzando le sue features.

Questo sistema di riconoscimento è stato applicato su 600 immagini, dimostrando un'accuratezza dei risultati di circa il 96.5%, rivelandosi quindi uno strumento utile per la rivoluzione delle interfacce uomo-macchina.



*Figura 3.10: Procedimento riassunto*





## 4. Movement tracking nella pratica

Nella pratica è possibile realizzare un sistema di tracking degli spostamenti della mano mediante l'utilizzo di un cellulare, un sistema Wi-Fi e un computer: sfruttando gli accelerometri presenti nel cellulare, mediante l'app Phyphox, è possibile misurare l'accelerazione presente nell'intervallo di tempo del movimento, e sapendo che questa è la derivata seconda della posizione rispetto al tempo, integrando due volte l'accelerazione misurata sarebbe quindi possibile tracciare i movimenti del telefono (quindi anche i movimenti della mano che lo tiene).

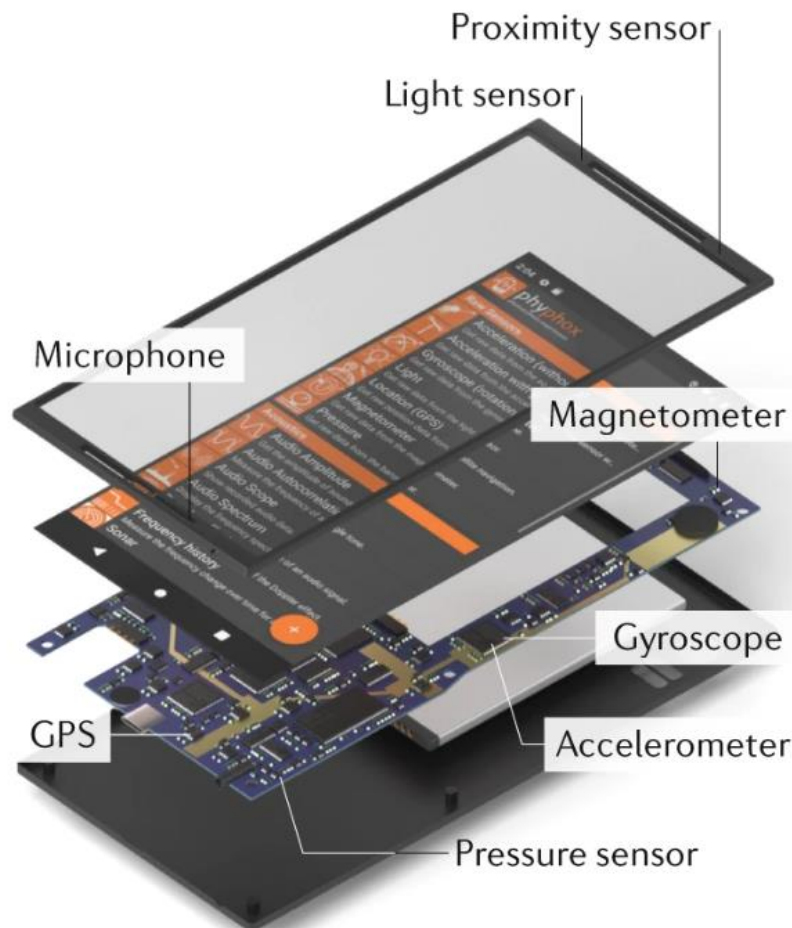


Figura 4.1.1 alcuni sensori presenti in uno smartphone utilizzabili mediante l'app Phyphox

Nella pratica però, sorgono numerose problematiche da tale procedura, in particolare il fatto che l'errore, causato da dei drift alle basse frequenze, del segnale misurato aumenta nel tempo, per misure effettuate in lunghi periodi il sistema fornirebbe errori intollerabile.

A livello teorico partendo da un oggetto fermo all'istante iniziale, la velocità all'istante T è ottenuta integrando l'accelerazione  $a$  in  $dt$ , tale operazione però deve essere applicata su un dominio discreto portando alla sostituzione dell'integrale con una sommatoria:

$$v(T) = \int_0^T a(t) dt = \lim_{n \rightarrow \infty} \frac{T}{n} \sum_{s=1}^n a(t_s)$$

La velocità può essere calcolata come la somma delle aree sottostanti la curva di accelerazione nel piano Cartesiano.

Per evitare l'amplificazione di errori dovuti al drift dell'accelerometro sarà però prima necessario applicare un filtro passa-alto (trend removal), riducendo quindi le componenti di errore ed offset non desiderate. Tale operazione andrà ripetuta sul segnale della velocità ottenuto dall'integrazione.

In modo simile, la posizione del corpo da tracciare è calcolata mediante l'integrazione (nel nostro caso sommatoria) della velocità:

$$d(T) = \int_0^T v(t) dt = \lim_{n \rightarrow \infty} \frac{T}{n} \sum_{s=1}^n v(t_s)$$

## 4.1. Phyphox

Phyphox, acronimo di PHYSical PHOne eXperiment, è un'applicazione che permette di sfruttare i sensori presenti negli smartphone come accelerometri, giroscopi, microfoni, magnetometri e barometri e rendendolo di fatto un validissimo strumento di misura "tascabile".



Figura 4.2.1 logo dell'applicazione Phyphox

La funzionalità utilizzata nell'esperienza è la misura dell'accelerazione (con l'accelerazione di gravità  $g$  esclusa), il telefono viene fissato alla mano (in modo tale da essere solidale con essa) e vengono compiuti 8 tipi di movimenti: movimento parallelo all'asse delle  $x$ , poi delle  $y$  e delle  $z$ , un movimento in diagonale sul piano  $xy$ , un movimento descrivente una traiettoria circolare, un'adduzione seguita da un'abduzione e un movimento in diagonale tridimensionale.

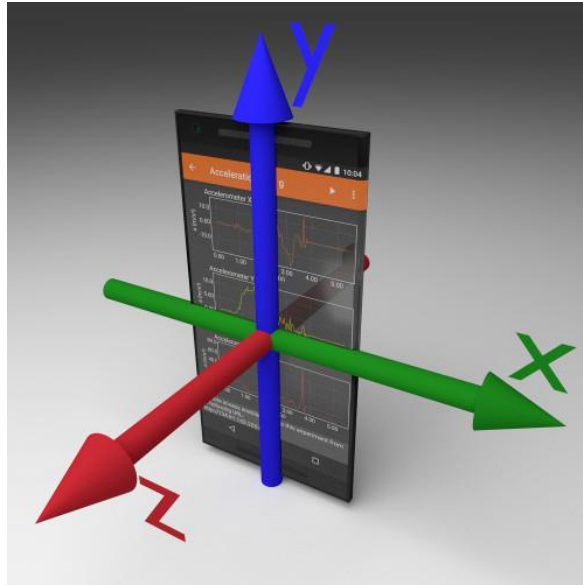


Figura 4.2.2 Sistema di riferimento triassiale adottato dall'applicazione

L'apparato sperimentale è costituito da uno smartphone, un pc, e una rete Wi-Fi che permette lo scambio di dati tra le due componenti, per poi essere successivamente elaborati al computer mediante l'utilizzo di Matlab.

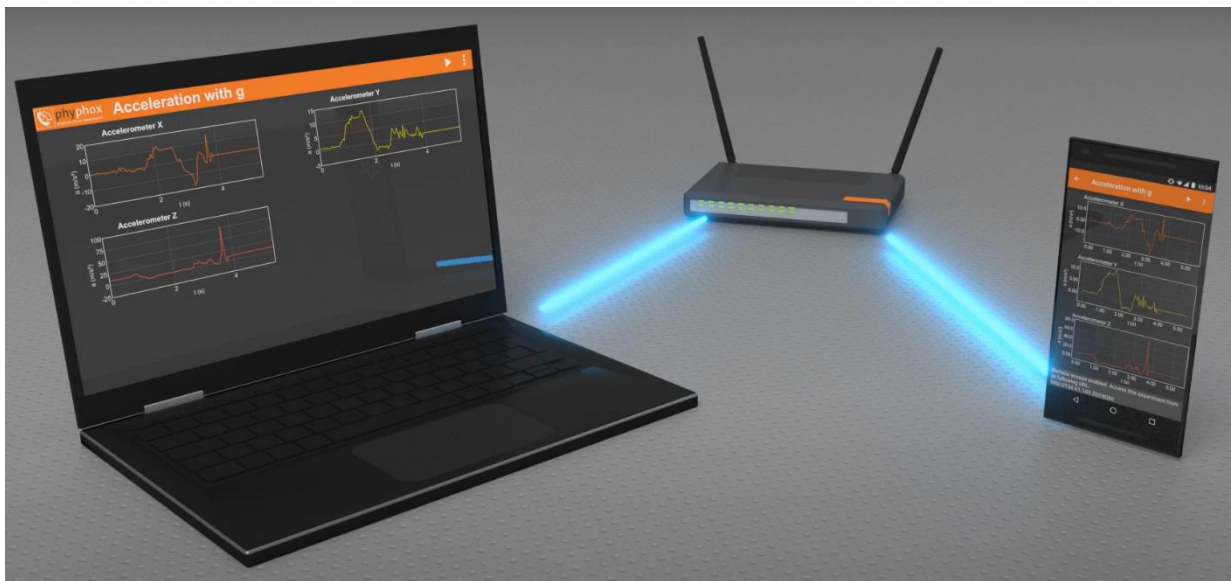


Figura 4.2.3 Schema dell'apparato sperimentale

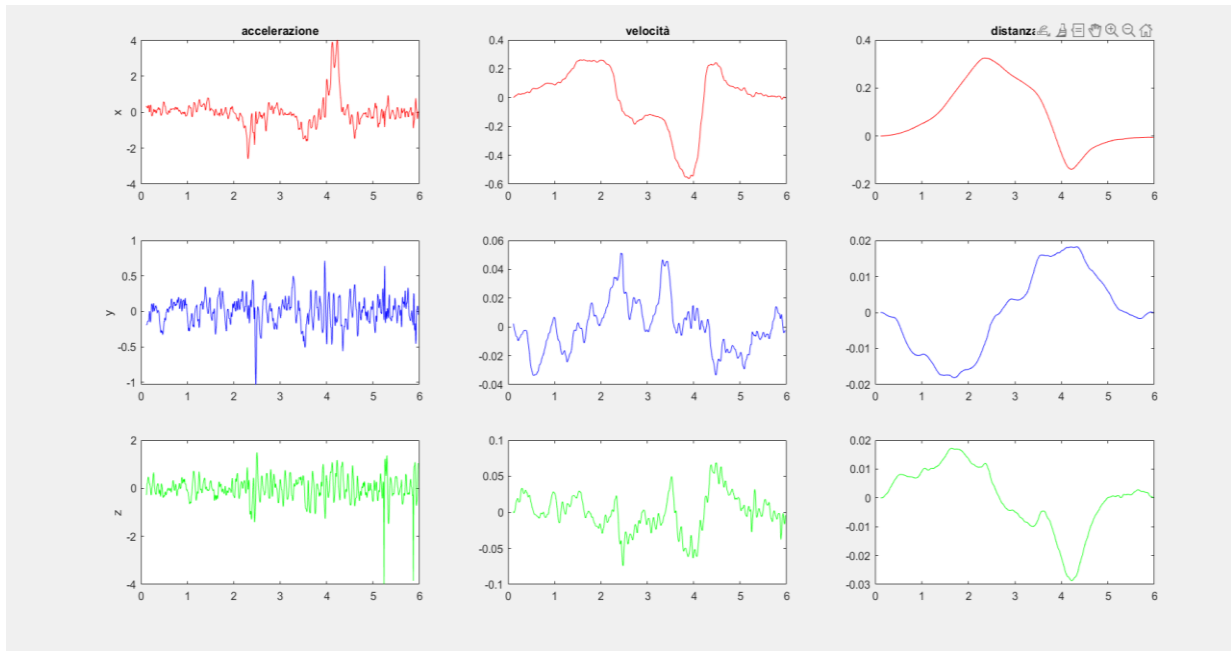
## 4.2. Elaborazione dei dati e risultati

I dati vengono scaricati come file .csv che verrà poi elaborato in Matlab con il seguente script:

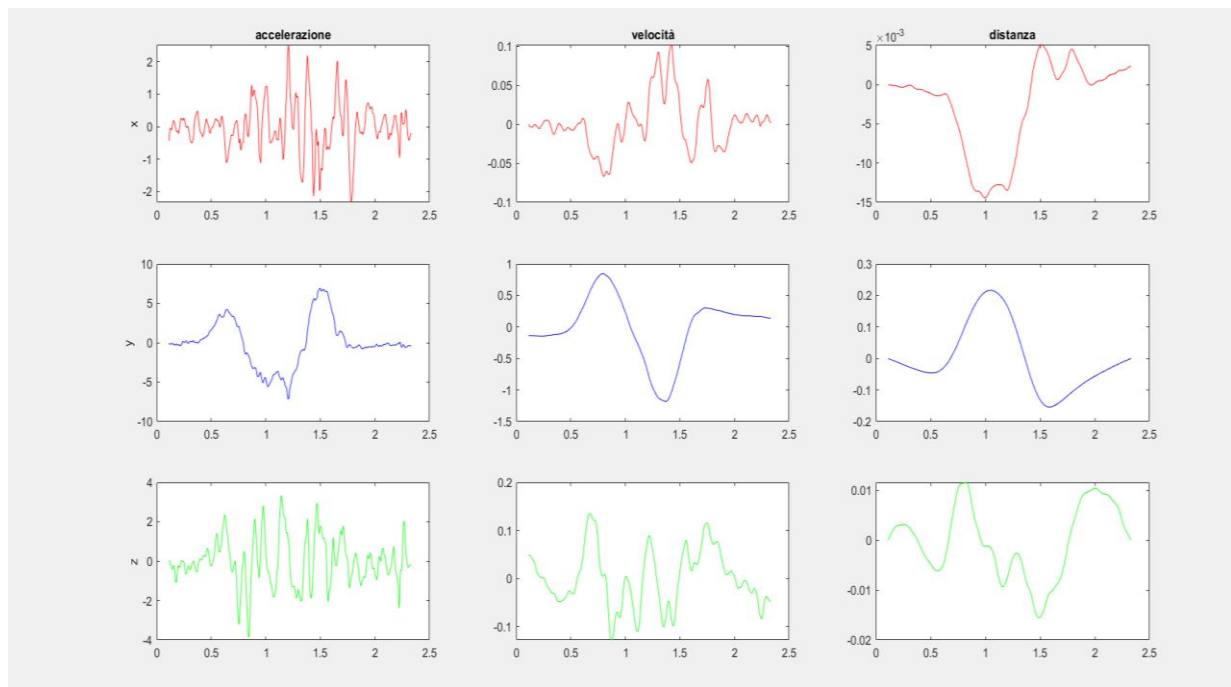
```
1  %carico i dati
2  data=readtable('RitornoY.csv');
3  %separo i dati dalla matrice principale e li converto in vettori
4  t=table2array(data(:,1));
5  accelerationX=table2array(data(:,2));
6  accelerationY=table2array(data(:,3));
7  accelerationZ=table2array(data(:,4));
8  absValue=table2array(data(:,5));
9  %calcolo dt, ovvero la distanza temporale tra due periodi di campionamento
10 dt=t(2)-t(1);
11 %applico il comando detrend che permette la rimozione del drift dello
12 %strumento
13 accelerationX=detrend(accelerationX);
14 accelerationY=detrend(accelerationY);
15 accelerationZ=detrend(accelerationZ);
16 %integro le accelerazioni lungo i tre assi rispetto a t, utilizzando la
17 %funzione cumtrapz, utile quando si conoscono i valori ma non la
18 %funzione matematica
19 Vx=cumtrapz(dt,accelerationX);
20 Vy=cumtrapz(dt,accelerationY);
21 Vz=cumtrapz(dt,accelerationZ);
22 %riapplico il tend-removal
23 VX=detrend(Vx);
24 Vy=detrend(Vy);
25 Vz=detrend(Vz);
26 %Posso applicare nuovamente l'integrale
27 X=cumtrapz(dt,Vx);
28 Y=cumtrapz(dt,Vy);
29 Z=cumtrapz(dt,Vz);
30 minC=min([min(X),min(Y),min(Z)]);
31 maxC=max([max(X),max(Y),max(Z)]);
32 %plotto lo spostamento nel tempo
33 figure
34 subplot(3,3,1)
35 plot(t,accelerationX,'r')
36 ylabel('x')
37 title('accelerazione')
38 subplot(3,3,2)
39 plot(t,Vx,'r')
40 title('velocità')
41 subplot(3,3,3)
42 plot(t,X,'r')
43 title('distanza')
44 subplot(3,3,4)
45 plot(t,accelerationY,'b')
46 ylabel('y')
47 subplot(3,3,5)
48 plot(t,Vy,'b')
49 subplot(3,3,6)
50 plot(t,Y,'b')
51 subplot(3,3,7)
52 plot(t,accelerationZ,'g')
53 ylabel('z')
54 subplot(3,3,8)
55 plot(t,Vz,'g')
56 subplot(3,3,9)
57 plot(t,Z,'g')
```

Accelerazione, velocità e spostamento vengono rappresentate su grafici (con distanza espressa in metri, velocità in metri al secondo, e accelerazione in metri su secondo quadro):

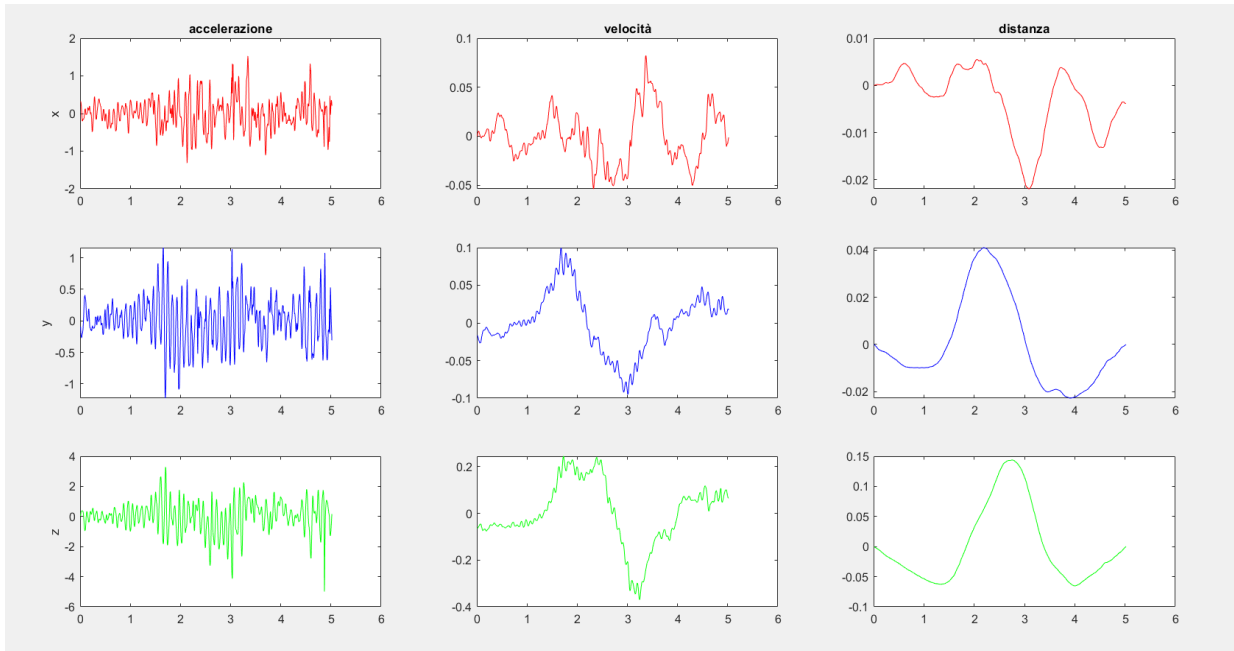
### Movimento lungo l'asse delle X



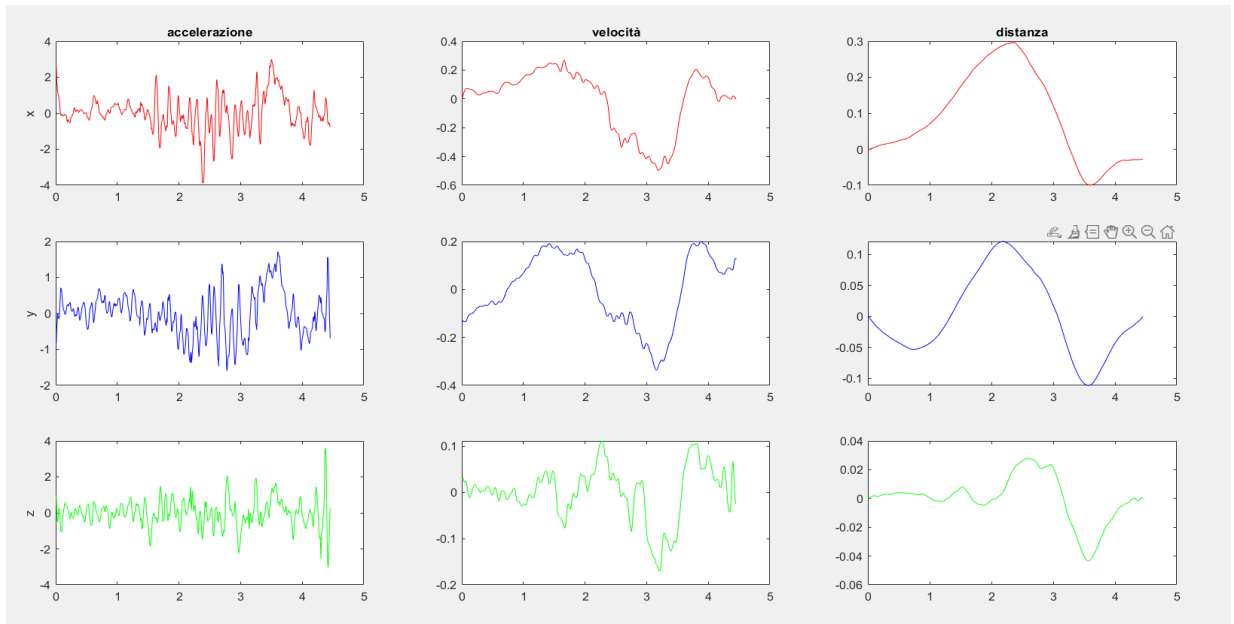
### Movimento lungo l'asse delle Y



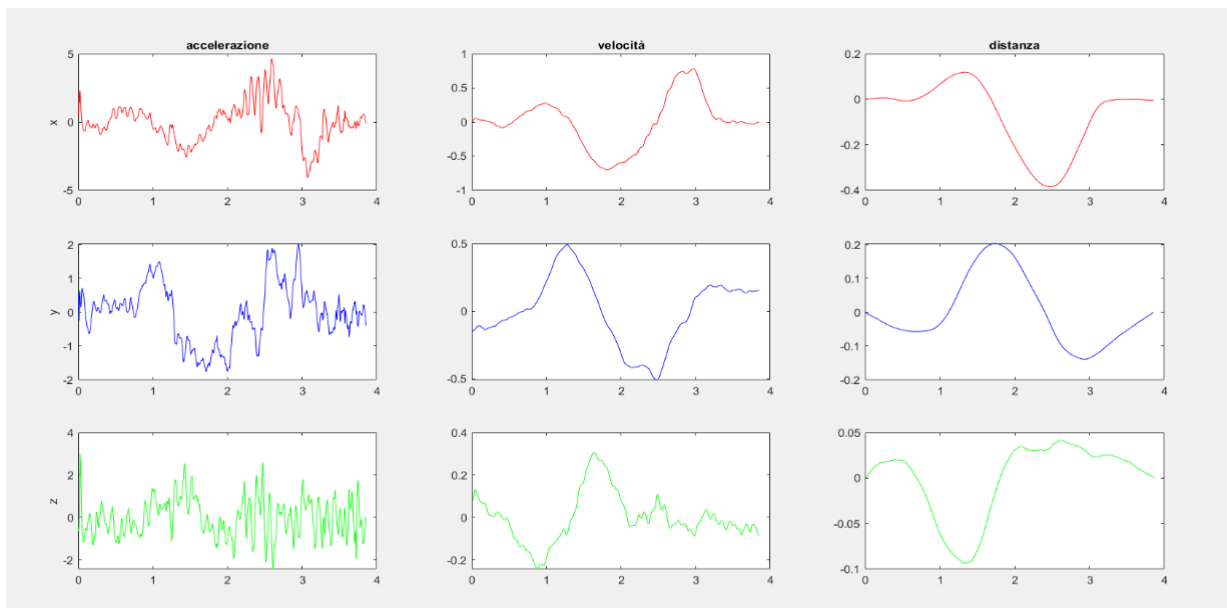
### Movimento lungo l'asse delle Z



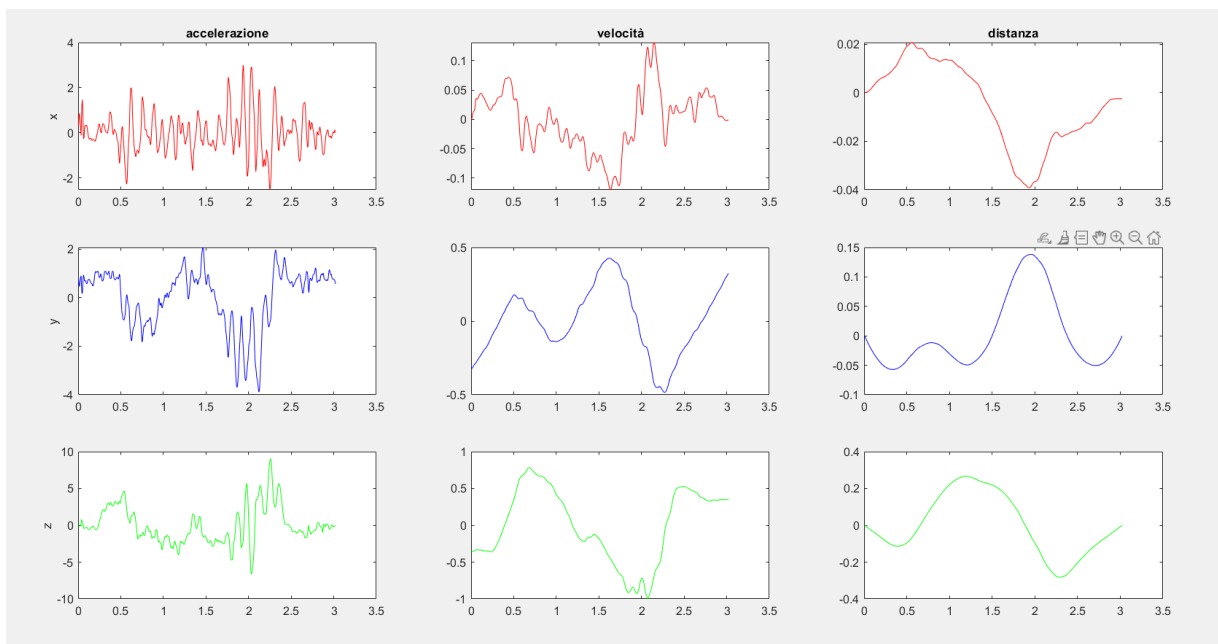
### Movimento in diagonale sul piano xy



## Movimento descrivente una circonferenza lungo il piano xy

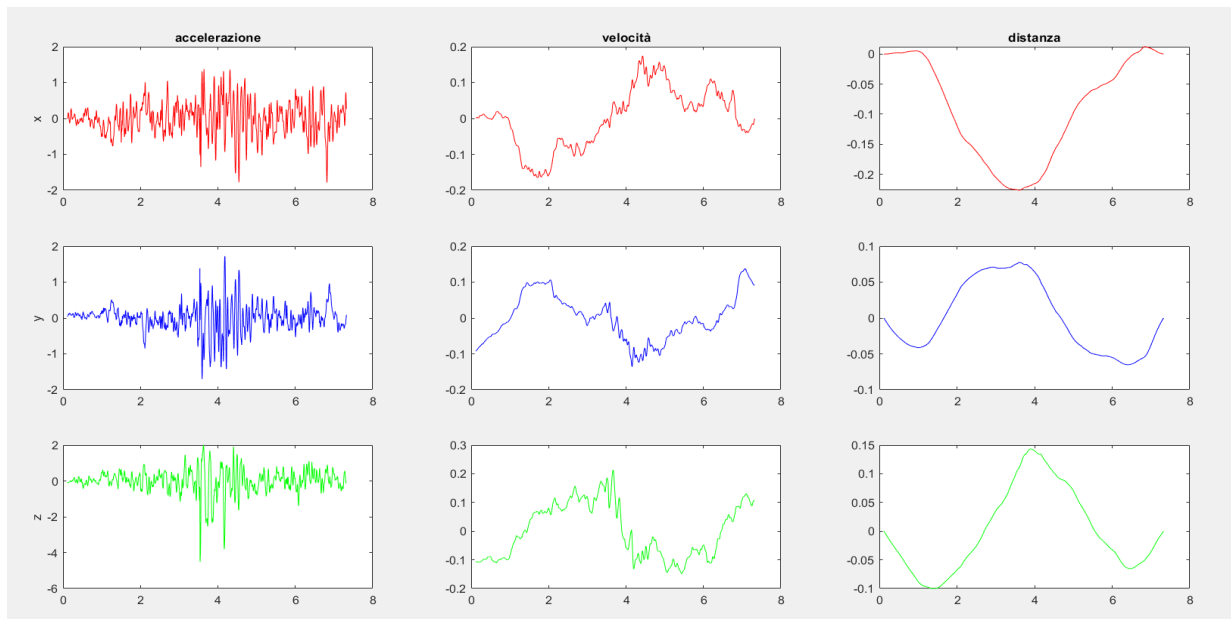


## Flessione dell'avambraccio





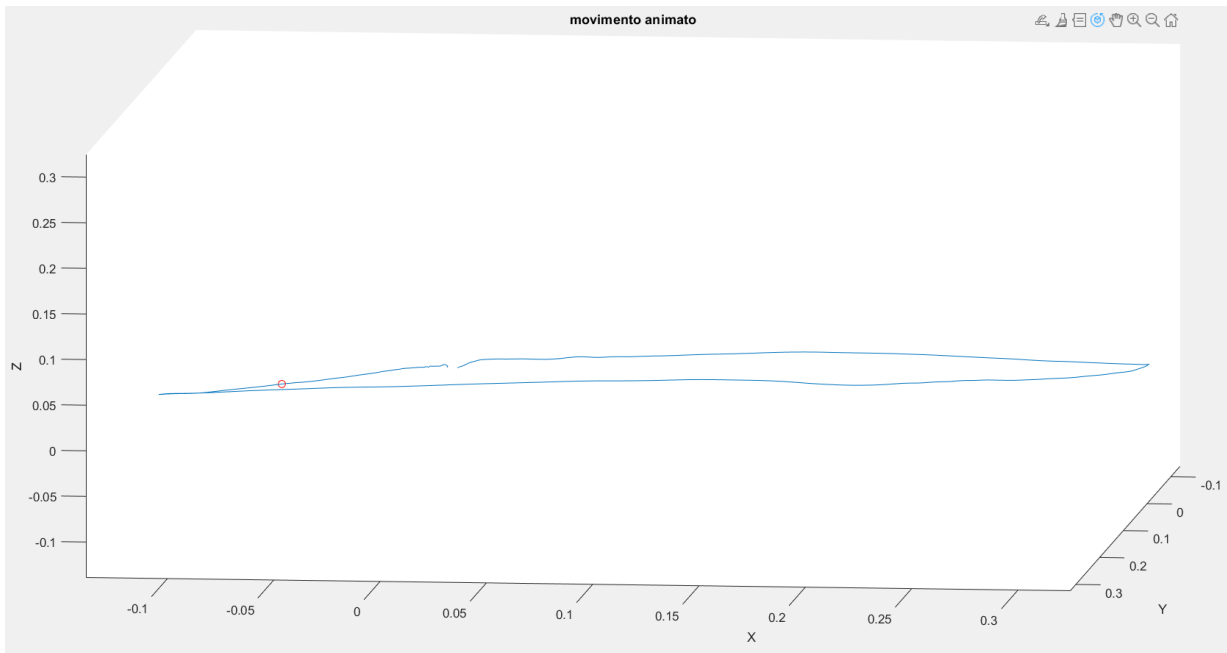
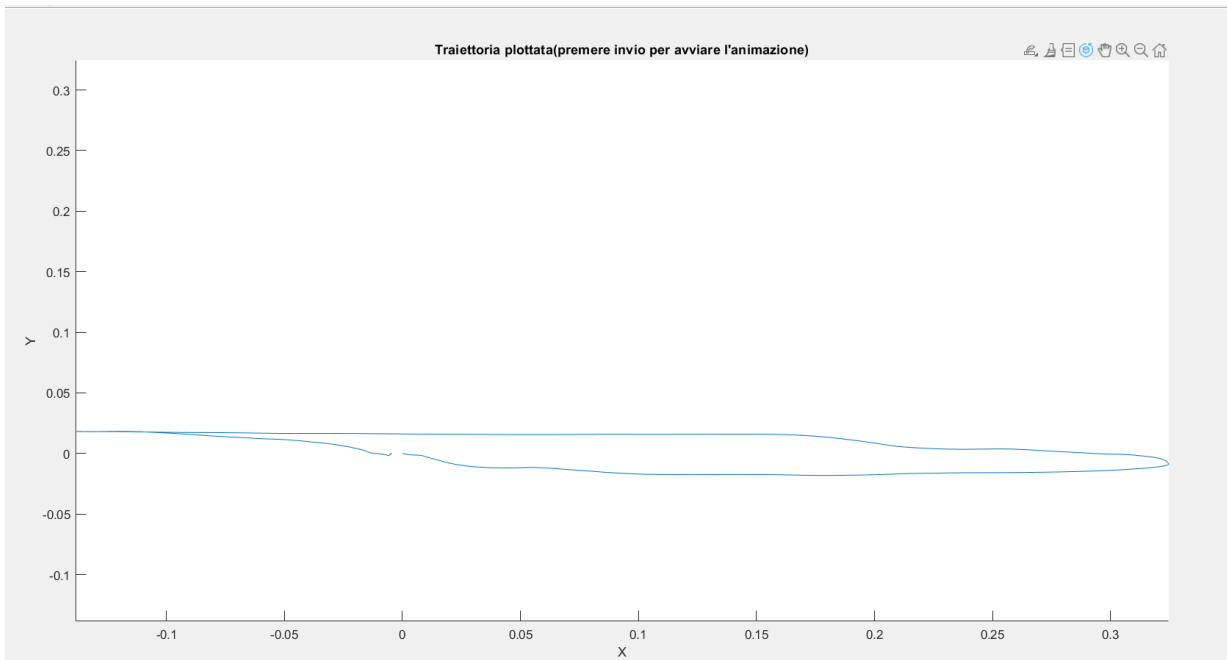
## Diagonale tridimensionale



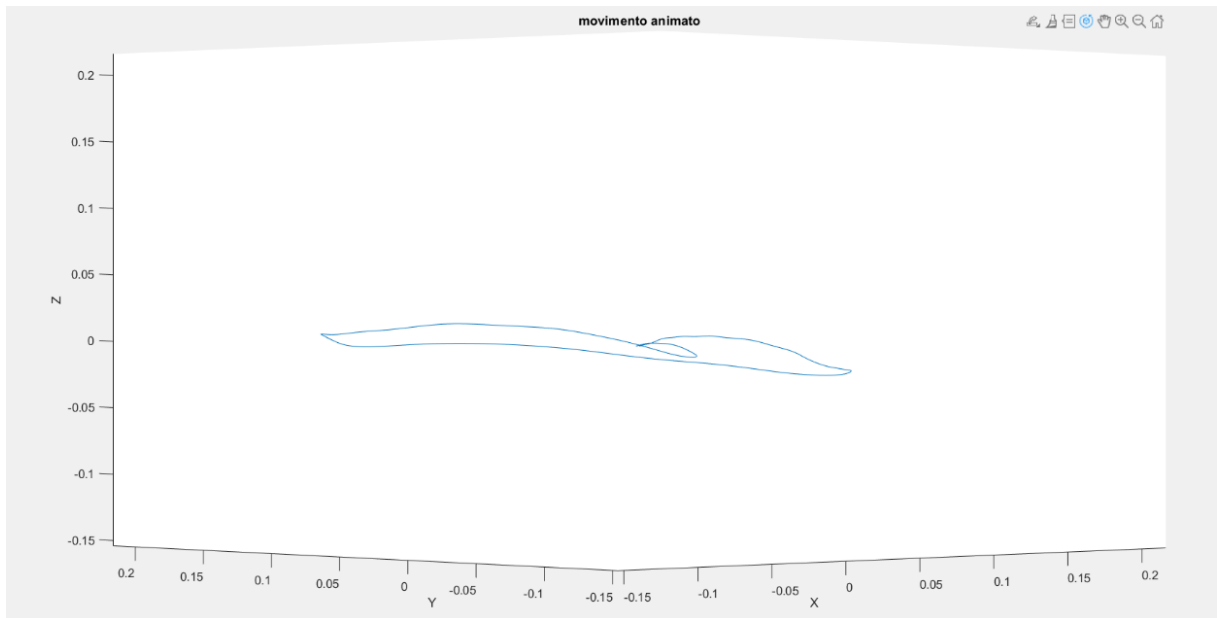
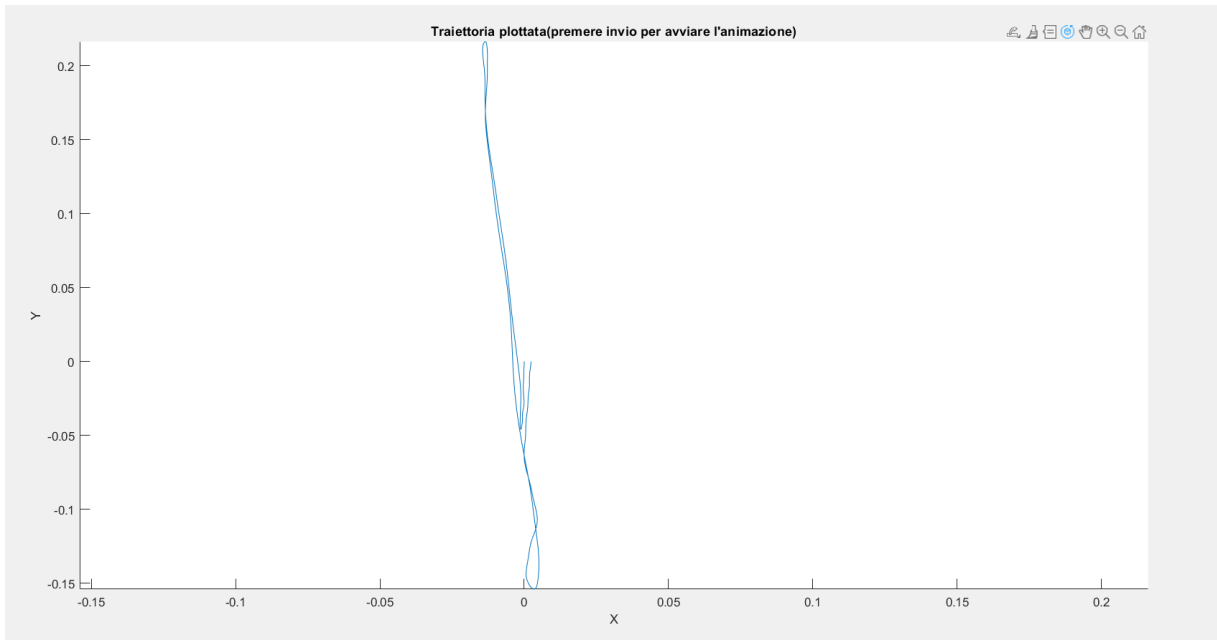
Vengono quindi plottate le componenti di accelerazione, velocità e posizione percorsa. Questi dati risultano però difficilmente interpretabili per alcuni movimenti (come la flessione e il cerchio) e potrebbe quindi essere più utile un plot in 3D. In seguito, viene quindi anche riprodotto il movimento acquisito in un grafico 3D, con un punto rosso che rappresenta lo smartphone in movimento.

```
63 %%plot in 3d
64 plot3(X,Y,Z)
65 xlabel('X')
66 ylabel('Y')
67 zlabel('Z')
68 xlim([minC,maxC])
69 ylim([minC,maxC])
70 zlim([minC,maxC])
71 title('Traiettoria plottata(premere invio per avviare l''animazione)')
72 pause
73 hold on
74 for i = 1:1:length(t)
75     h = plot3(X(i),Y(i),Z(i), 'or');
76     xlim([minC,maxC])
77     ylim([minC,maxC])
78     zlim([minC,maxC])
79     title('movimento animato')
80     pause(0.03)
81     delete(h)
82 end
```

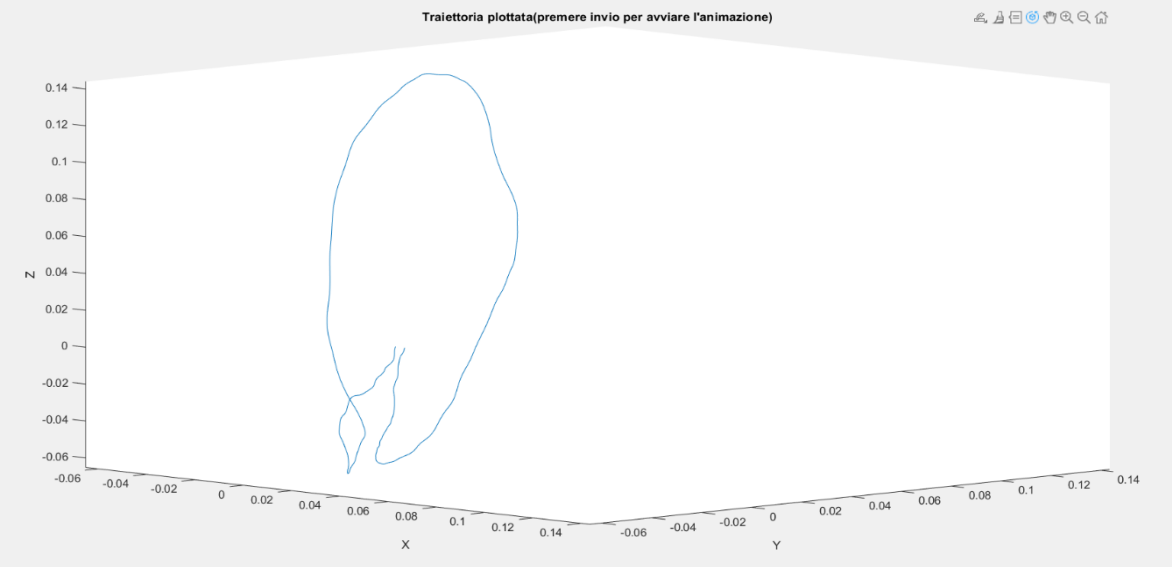
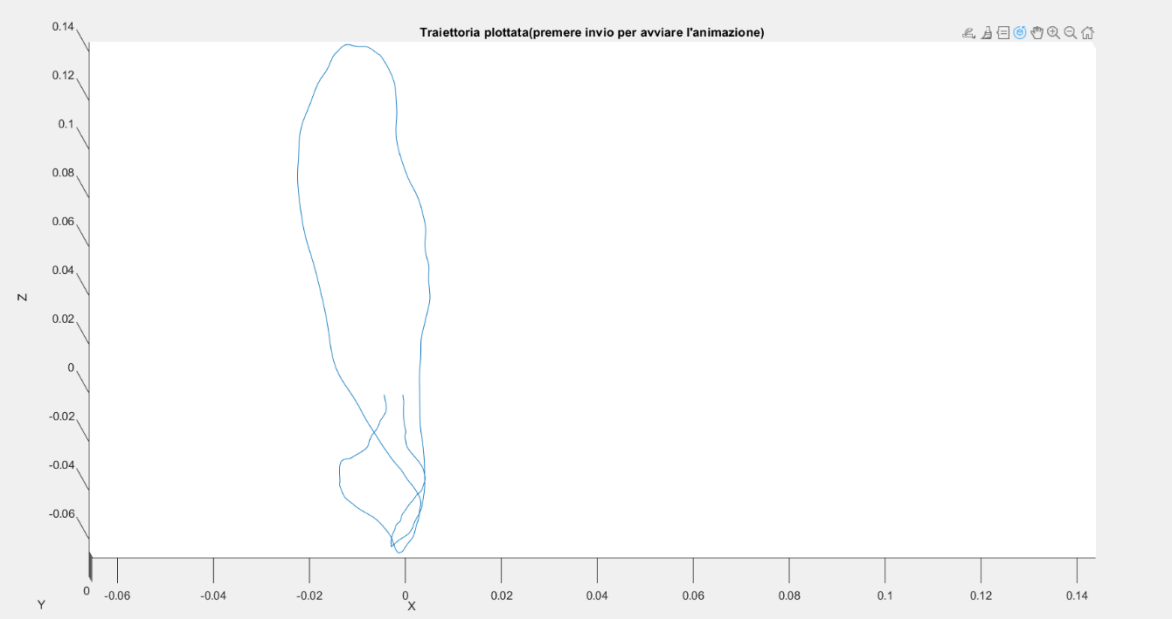
# Movimento lungo l'asse delle X (prospettiva sia del piano xy che tridimensionale)



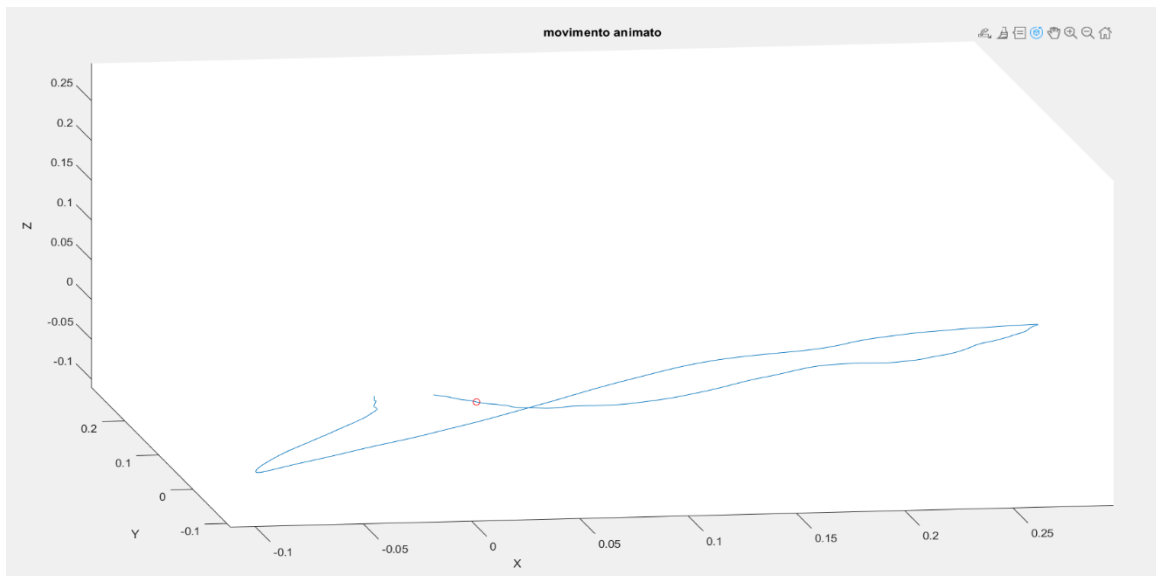
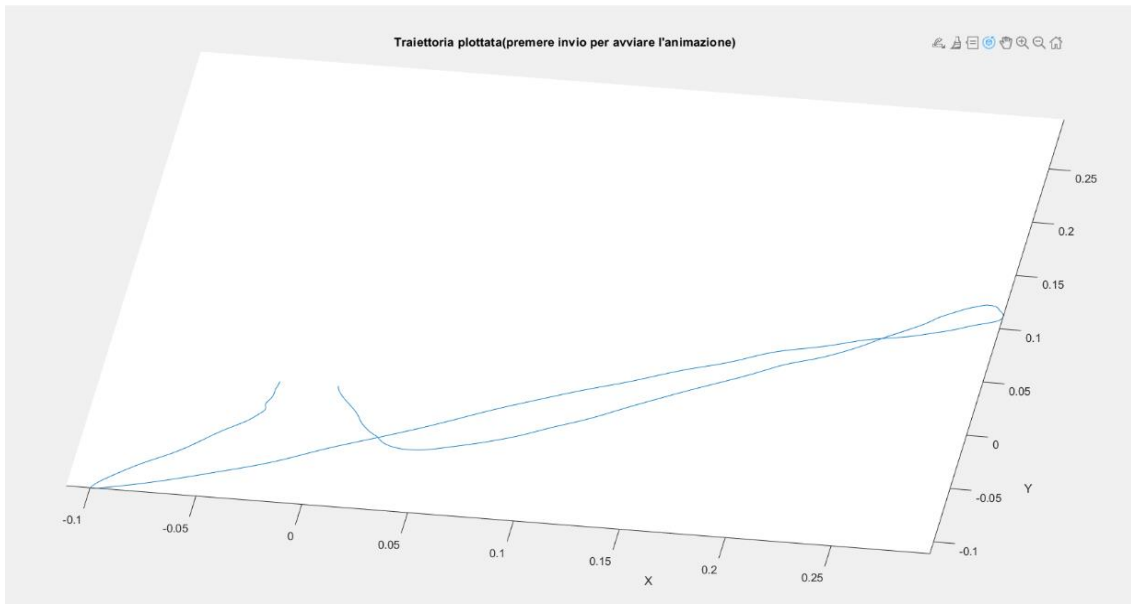
### Movimento lungo l'asse delle Y (prospettiva sia del piano xy che tridimensionale)



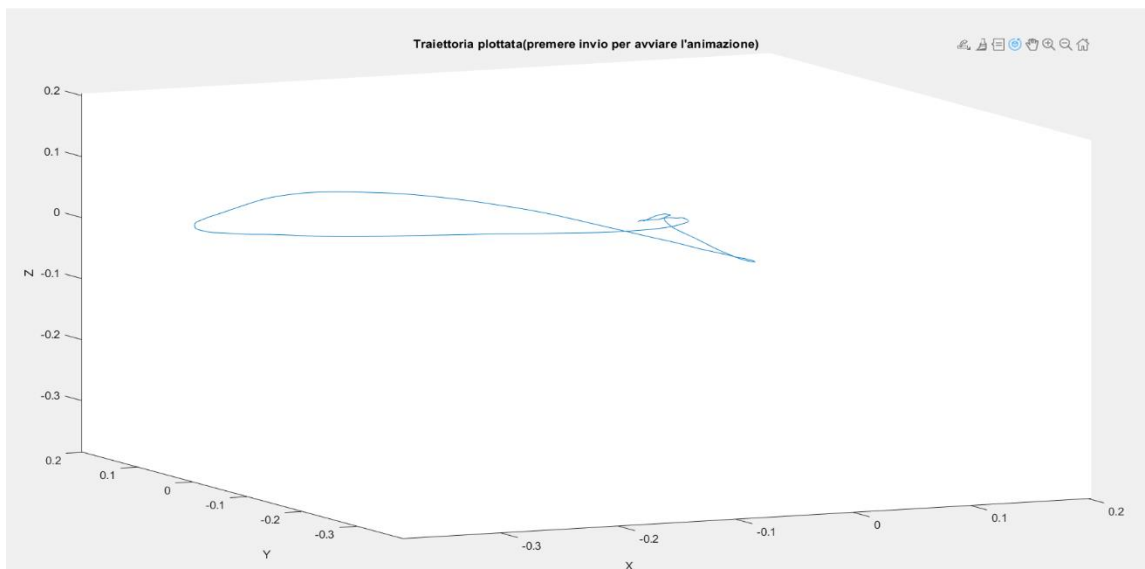
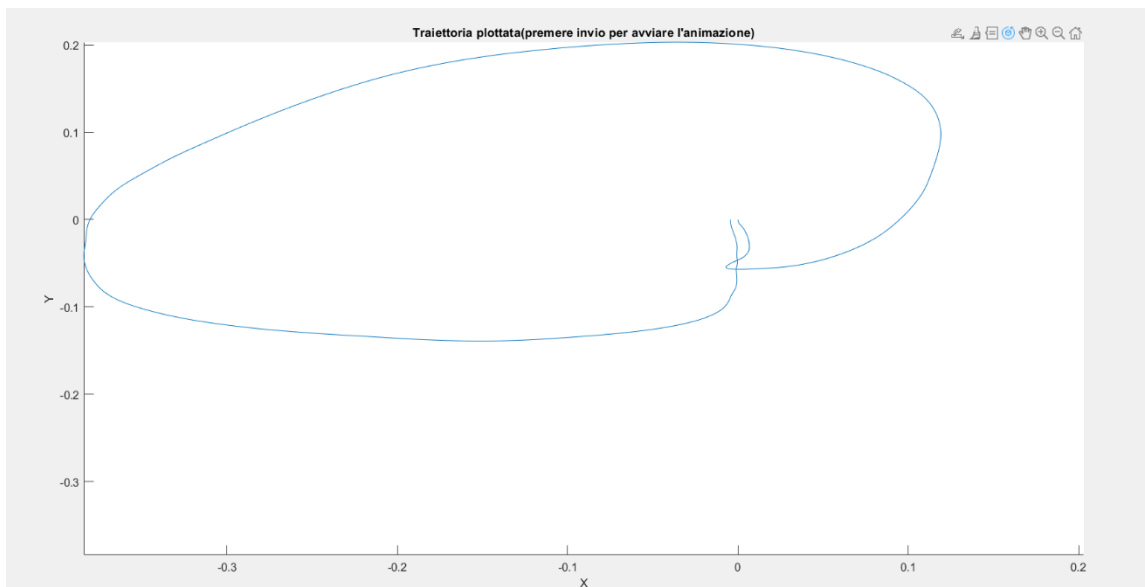
Movimento lungo l'asse delle Z (prospettiva sia del piano xz che tridimensionale)



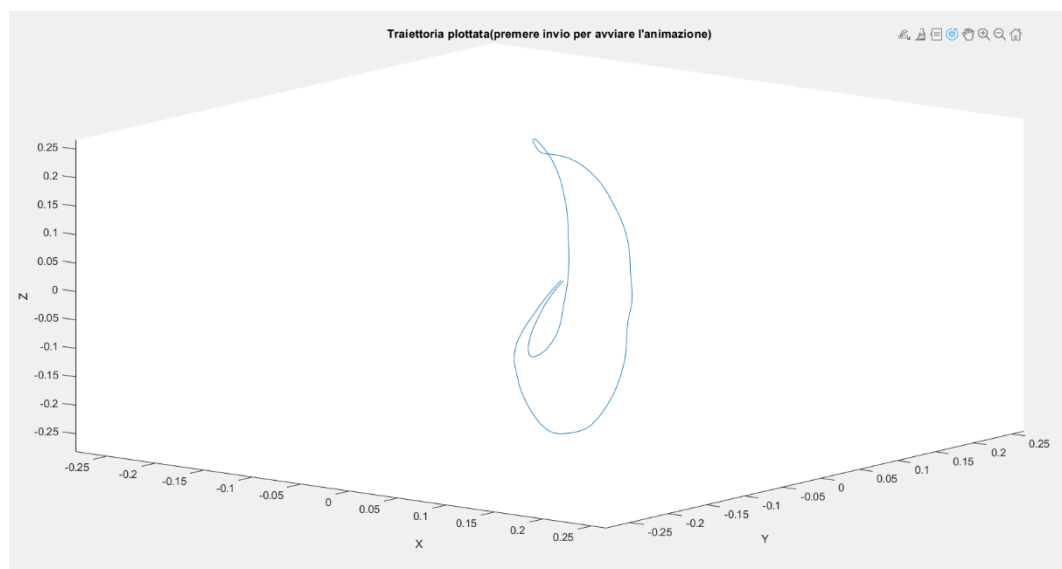
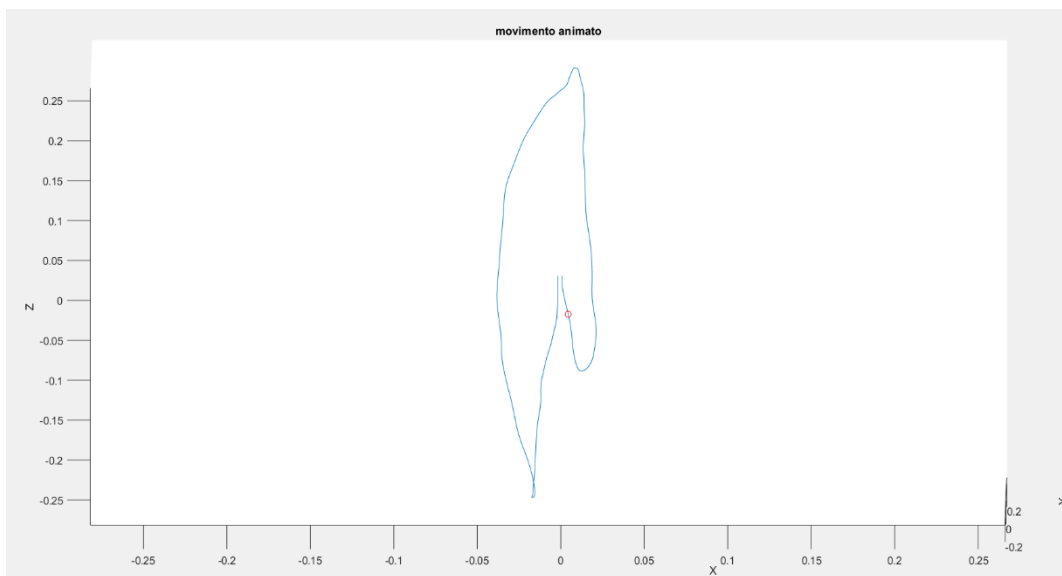
# Movimento in diagonale sul piano xy (piano xy e spazio tridimensionale)



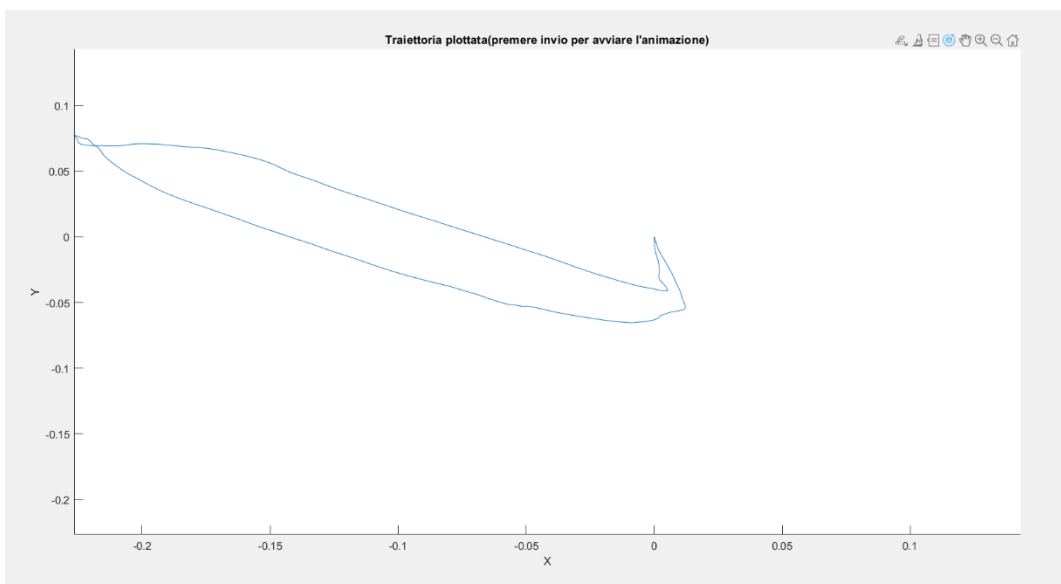
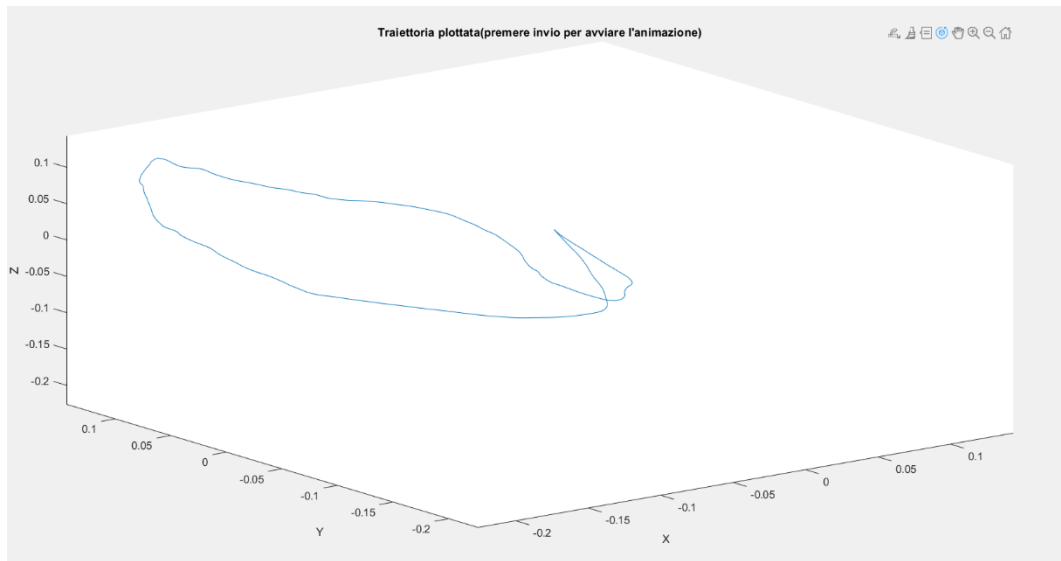
# Movimento descrivente una circonferenza lungo il piano xy (piano xy e 3D)



# Flessione dell'avambraccio (piano xz e spazio 3D)



## Diagonale tridimensionale



I movimenti ottenuti si discostano significativamente dalle traiettorie ideali, per esempio la traiettoria della flessione idealmente potrebbe essere un arco di circonferenza a 90 gradi, tuttavia considerando la biomeccanica articolare e il fatto che l'essere umano non esegue movimenti meccanici precisi le traiettorie tracciate risultano verosimili.





## 5. Conclusioni

L'analisi delle componenti principali di un sistema di acquisizione dei gesti eseguita in questo elaborato rappresenta una panoramica dei vari aspetti tecnologici da prendere in considerazione durante la progettazione di un sistema di *gesture recognition* e quindi un possibile punto di partenza, per poi procedere con la considerazione di altri aspetti più dettagliati e raffinati per arrivare quindi alla concreta realizzazione del sistema desiderato. Il primo aspetto generale, e indubbiamente il più importante, è l'ambito di applicazione del sistema da progettare: nel caso di formazione di personale mediante l'utilizzo della realtà virtuale potrebbe essere più utile l'utilizzo di sensori indossabili per permettere il riconoscimento di gesti veloci e non facilmente rilevabili con una telecamera (come per esempio i movimenti delle dita) per poter permettere una maggiore esperienza formativa ai dipendenti grazie alla maggiore sensibilità del sistema di rilevazione; nel settore videoludico, invece, l'utilizzo di sensori indossabili connessi mediante cavi alla console potrebbe limitare la libertà di movimento, rovinando potenzialmente l'esperienza di gioco. Altro aspetto da considerare è il tipo di sensore/telecamera che può portare a problematiche dovute alle non idealità delle condizioni di acquisizione, come il rumore (basti pensare al drift alle basse frequenze presente nell'accelerometro nel capitolo 5, ma potrebbe anche essere l'elettrocardiogramma che si somma all'elettromiogramma che stiamo raccogliendo), per quanto riguarda la trasmissione dei dati acquisiti dal modulo DAQ, sicuramente il protocollo NFC non risulta adatto a causa della sua connessione a cortissimo raggio (2 centimetri), la connessione Bluetooth risulta affidabile in termini di resistenza alle interferenze, tuttavia la connessione Wi-Fi, pur essendo esposta al potenziale rischio di interferenza di banda, possiede un raggio di comunicazione estremamente maggiore, rendendola adatta in particolare per le applicazioni di telemedicina o per il comando di attuatori a grande distanza. Per quanto riguarda l'elaborazione dei dati, pur essendo presenti diversi metodi deterministici, il *machine learning* risulta essere il più interessante, flessibile e promettente tra gli algoritmi analizzati, in particolare mediante l'uso di SVM e reti neurali artificiali. Concludendo con questo riassunto degli aspetti citati nell'elaborato, le possibili applicazioni dei sistemi di riconoscimento dei gesti sono molte, tutte con le potenzialità di rivoluzionare le modalità di interazione tra umani e macchine, consentendo la possibilità di sblocco di dispositivi in tempi sensibilmente minori, attivare menù, ridimensionare e ruotare immagini. Alcune aziende si stanno inoltre occupando di portare nel mercato dispositivi basati su questa tecnologia come la suite software Elliptic Labs, in grado di offrire funzioni di gesture e di prossimità riutilizzando l'auricolare e il microfono esistenti, precedentemente utilizzati solo per l'audio, per emettere e ricevere segnali ad ultrasuoni inviati attraverso l'aria dagli altoparlanti integrati in smartphone e tablet che, rimbalzando contro una mano / oggetto / testa, sono registrati dai microfoni, anch'essi integrati in questi dispositivi. In questo modo, la tecnologia di Elliptic Labs riconosce i gesti delle mani e li utilizza per spostare oggetti su uno schermo, in modo simile al modo in cui i pipistrelli utilizzano l'ecolocalizzazione per navigare; Airwriting è una tecnologia che permette di scrivere testi in aria; Eyesight permette di navigare su un monitor senza toccarlo fisicamente.



## 6. Ringraziamenti

E' stato un lungo viaggio, a tratti emozionante e divertente e a tratti difficile e stressante, ma la soddisfazione di essere arrivato a questo punto è impagabile, e per questo devo ringraziare molte persone.

Ringrazio tutti i miei parenti, ma prima di tutti i miei genitori, che hanno sempre creduto in me, che mi hanno sostenuto anche nei momenti peggiori della mia vita e, soprattutto, mi hanno insegnato che non importa quante volte la vita ti butterà giù, che sia una brutta malattia, un licenziamento o un semplice 23 ad un esame, l'importante è rialzarsi e che (lo so che me lo ripeterai ancora un'infinità di volte papà) volere è potere.

Ringrazio i miei amici, in particolare Gianluca (anche se ogni tanto potevi presentarti in orario quando uscivamo in compagnia eh), Marco, Riccardo, Chiara, Natalia, Martina, ed Emanuele per avermi aiutato a superare questi 3 anni con il cuore più leggero grazie alle risate e le pazzie passate assieme, alle notti al Pavan o dalla Dina e vi chiedo anche scusa per tutte le volte che non sono uscito con voi perché ero troppo fissato con lo studio o che vi ho risposto male per l'ansia e il nervosismo causati dalla sessione, non ve l'ho mai detto e mi scuso per non averlo mai fatto, ma non saprei cosa fare senza di voi.

Ringrazio tutto lo staff di Pizzeria Alla Torre, per avermi accolto nel loro gruppo e avermi insegnato la lealtà, la disciplina e l'umiltà. Ringrazio in particolare Romario, il titolare, potevi sfruttarmi, potevi darmi una paga misera come fanno la maggior parte delle pizzerie, ma non l'hai fatto e sono felice di considerarti più come un amico che il mio capo anche se non mi hai ancora detto se il forno è a legna o a gas.

Ringrazio i compagni dell'università, ne abbiamo passate tante assieme dandoci forza tra noi prima di ogni esame, abbiamo condiviso paure, insicurezze, ansia ma anche gioia, felicità e tante risate; spero di vedervi presto con la corona d'alloro perché tutti voi lo meritate e sappiate che mi mancherete.

Ringrazio la professoressa Tonello per la disponibilità concessami come relatrice di questo elaborato e per aver reso la stesura del documento divertente e per niente stressante, oltre agli utilissimi consigli forniti per migliorare l'elaborato.



## 7. Riferimenti

Choi, E., Kwon, S., Lee, D., Lee, H., and Chung, M. K. (2012), Design of Hand Gestures for Smart Home Appliances based on a User Centered Approach, *Journal of the Korean Institute of Industrial Engineers*, 38(3), 182-190. <https://doi.org/10.7232/JKIIIE.2012.38.3.182>

Khan, Rafiqul Zaman, and Noor Adnan Ibraheem. "Hand gesture recognition: a literature review." *International journal of artificial Intelligence & Applications* 3.4 (2012): 161.

Freeman, William T., and Michal Roth. "Orientation histograms for hand gesture recognition." *International workshop on automatic face and gesture recognition*. Vol. 12. 1995.

Y. Fang, K. Wang, J. Cheng and H. Lu, "A Real-Time Hand Gesture Recognition Method," *2007 IEEE International Conference on Multimedia and Expo*, 2007, pp. 995-998, doi: 10.1109/ICME.2007.4284820.

Garg, Pragati, Naveen Aggarwal, and Sanjeev Sofat. "Vision based hand gesture recognition." *World academy of science, engineering and technology* 49.1 (2009): 972-977.

Rautaray, Siddharth S., and Anupam Agrawal. "Vision based hand gesture recognition for human computer interaction: a survey." *Artificial intelligence review* 43.1 (2015): 1-54.

Galván-Ruiz J, Travieso-González CM, Tejera-Fettmilch A, Pinan-Roescher A, Esteban-Hernández L, Domínguez-Quintana L. Persecutive and Evolution of Gesture Recognition for Sign Language: A Review. *Sensors (Basel)*. 2020 Jun 24;20(12):3571. doi: 10.3390/s20123571. PMID: 32599793; PMCID: PMC7349703.

Han J, Shao L, Xu D, Shotton J. Enhanced computer vision with Microsoft Kinect sensor: a review. *IEEE Trans Cybern*. 2013 Oct;43(5):1318-34. doi: 10.1109/TCYB.2013.2265378. Epub 2013 Jun 25. PMID: 23807480.

Oudah M, Al-Naji A, Chahl J. Hand Gesture Recognition Based on Computer Vision: A Review of Techniques. *J Imaging*. 2020;6(8):73. Published 2020 Jul 23. doi:10.3390/jimaging6080073

G. Incertis, J. G. Garcia-Bermejo and E. Z. Casanova, "Hand Gesture Recognition for Deaf People Interfacing," 18th International Conference on Pattern Recognition (ICPR'06), 2006, pp. 100-103, doi: 10.1109/ICPR.2006.619.

S. Vanaja, R. Preetha and S. Sudha, "Hand Gesture Recognition for Deaf and Dumb Using CNN Technique," *2021 6th International Conference on Communication and Electronics Systems (ICCES)*, 2021, pp. 1-4, doi: 10.1109/ICCES51350.2021.9489209.

Jagdish L., Raheja, A., Sadab "Android based Portable Hand Sign Recognition System" Machine Vision Lab CSIR-CEERI, Pilani India

Kang, Sung-Kwan, Kyungyong Chung, and Jung-Hyun Lee. "Real-time tracking and recognition systems for interactive telemedicine health services." *Wireless personal communications* 79.4 (2014): 2611-2626.

Karim, Rohana Abdul, et al. "Telepointer technology in telemedicine: a review." *Biomedical engineering online* 12.1 (2013): 1-19.

Owens GE, Connor CW. Controlling Anesthesia Hardware With Simple Hand Gestures: Thumbs Up or Thumbs Down? *Anesth Analg*. 2021 Jul 1;133(1):243-250. doi: 10.1213/ANE.0000000000005071. PMID: 32701544; PMCID: PMC7855226.

M. Zobl, M. Geiger, B. Schuller, M. Lang and G. Rigoll, "A real-time system for hand gesture controlled operation of in-car devices," 2003 International Conference on Multimedia and Expo. ICME '03. Proceedings (Cat. No.03TH8698), 2003, pp. III-541, doi: 10.1109/ICME.2003.1221368

Zhao Z, Luo H, Song GH, Chen Z, Lu ZM, Wu X. Web-based interactive drone control using hand gesture. *Rev Sci Instrum*. 2018 Jan;89(1):014707. doi: 10.1063/1.5004004. PMID: 29390725.

Zhang J, Wang B, Zhang C, Xiao Y, Wang MY. An EEG/EMG/EOG-Based Multimodal Human-Machine Interface to Real-Time Control of a Soft Robot Hand. *Front Neurobot*. 2019 Mar 29;13:7. doi: 10.3389/fnbot.2019.00007. PMID: 30983986; PMCID: PMC6449448.

Sathiyarayanan, Mithileysh, et al. "Gesture controlled robot for military Purpose." *International Journal for Technological Research in Engineering* 1.11 (2014): 1300-1303.

Prasad, N. H., A. Mariyan Richard, and B. N. Lakshmi Narayan. "Hand gesture controlled robot." *International Conference on Communication, Computing and Electronics Systems*. Springer, Singapore, 2021.

Eadie, L. H., A. M. Seifalian, and aBR Davidson. "Telemedicine in surgery." *Journal of British Surgery* 90.6 (2003): 647-658.

Mahmood, Mufrath, et al. "Design of a low-cost hand gesture controlled automated wheelchair." *2020 IEEE Region 10 Symposium (TENSYP)*. IEEE, 2020.

Goyal, Diksha, and S. P. S. Saini. "Accelerometer based hand gesture controlled wheelchair." *International Journal on Emerging Technologies* 4.2 (2013): 15-20.

Han, Hobeom, and Sang W. Yoon. 2019. "Gyroscope-Based Continuous Human Hand Gesture Recognition for Multi-Modal Wearable Input Device for Human Machine Interaction" *Sensors* 19, no. 11: 2562.

Byun, Sung-Woo, and Seok-Pil Lee. "Hand gesture recognition suitable for wearable devices using flexible epidermal tactile sensor array." *Journal of Electrical Engineering and Technology* 13.4 (2018): 1732-1739.

Salchow-Hömmen, Christina, Leonie Callies, Daniel Laidig, Markus Valtin, Thomas Schauer, and Thomas Seel. 2019. "A Tangible Solution for Hand Motion Tracking in Clinical Applications" *Sensors* 19, no. 1: 208

Ferri, Josue et al. "A Wearable Textile 3D Gesture Recognition Sensor Based on Screen-Printing Technology." *Sensors (Basel, Switzerland)* vol. 19,23 5068. 20 Nov. 2019, doi:10.3390/s19235068

Okada, Kazuhiro. "Tri-Axial Piezoresisive Accelerometer." *Technical digest of the 11th sensor symposium*. 1992.

Li, X.; Hu, J.; Liu, X. A High-Performance Digital Interface Circuit for a High-Q Micro-Electromechanical System Accelerometer. *Micromachines* **2018**, *9*, 675.

Faisal, I. Arun, T. Waluyo Purboyo, and A. Siswo Raharjo Ansori. "A Review of accelerometer sensor and gyroscope sensor in IMU sensors on motion capture." *J. Eng. Appl. Sci* 15.3 (2019): 826-829.

Passaro, Vittorio M. N., Antonello Cuccovillo, Lorenzo Vaiani, Martino De Carlo, and Carlo Edoardo Campanella. 2017. "Gyroscope Technology and Applications: A Review in the Industrial Perspective" *Sensors* 17, no. 10: 2284. <https://doi.org/10.3390/s17102284>

da Silva, Josivaldo Godoy, Aparecido Augusto de Carvalho, and Doriedson Dutra da Silva. "A strain gauge tactile sensor for finger-mounted applications." *IEEE Transactions on Instrumentation and measurement* 51.1 (2002): 18-22.

Miron, Casian, et al. "Hand gesture detection using a stereo camera system and simulation of movement." *2017 10th International Symposium on Advanced Topics in Electrical Engineering (ATEE)*. IEEE, 2017.

C. Bisdikian, "An overview of the Bluetooth wireless technology," in *IEEE Communications Magazine*, vol. 39, no. 12, pp. 86-94, Dec. 2001, doi: 10.1109/35.968817.

Darroudi, S.M.; Gomez, C. Bluetooth Low Energy Mesh Networks: A Survey. *Sensors* **2017**, *17*, 1467. <https://doi.org/10.3390/s17071467>

Bikonis, Krzysztof & Demkowicz, Jerzy. (2013). Data Integration from GPS and Inertial Navigation Systems for Pedestrians in Urban Area. *TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation*. 7. 401-406. 10.12716/1001.07.03.12.

Handbook of modern sensors

Introduction to blueetooth

An introduction to Wi-Fi

Beggining NFC

Sridhar Krishnan, Yashodhan Athavale, Trends in biomedical signal feature extraction, *Biomedical Signal Processing and Control*, Volume 43, 2018, Pages 41-63, ISSN 1746-8094, <https://doi.org/10.1016/j.bspc.2018.02.008>.

P. Trigueiros, F. Ribeiro and L. P. Reis, "A comparison of machine learning algorithms applied to hand gesture recognition," *7th Iberian Conference on Information Systems and Technologies (CISTI 2012)*, 2012, pp. 1-6.

Frank, J. (1980). The Role of Correlation Techniques in Computer Image Processing. In: Hawkes, P.W. (eds) *Computer Processing of Electron Microscope Images*. Topics in Current Physics, vol 13. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-81381-8\\_5](https://doi.org/10.1007/978-3-642-81381-8_5)

A. Phinyomark, S. Hirunviriya, C. Limsakul and P. Phukpattaranont, "Evaluation of EMG feature extraction for hand movement recognition based on Euclidean distance and standard deviation," *ECTI-CON2010: The 2010 ECTI International Confernce on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, 2010, pp. 856-860.



Ullah, Asad & Ali, Sarwan & Khan, Imdadullah & Khan, Muhammad & Faizullah, Safiullah. (2020). Effect of Analysis Window and Feature Selection on Classification of Hand Movements Using EMG Signal. 10.1007/978-3-030-55190-2\_30.

appunti e slide del corso "Misure e Acquisizione di Dati Biomedici", corso di laurea triennale in ingegneria biomedica, A.A: 2021-2022, docente: Sarah Tonello

appunti e slide del corso "Tecnologia e strumentazione biomedica", corso di laurea triennale in ingegneria biomedica, A.A: 2021-2022, docenti: Simone Del Favero, Michele Schiavon, Porcaro Camillo

appunti e slide del corso "Elaborazione di segnali biomedici", corso di laurea triennale in ingegneria biomedica, A.A: 2021-2022, docenti: Mariapia Saccomani, Francesca Bettini

R. Ferrero, F. Gandino, M. Hemmatpour, B. Montrucchio and M. Rebaudengo, "Exploiting accelerometers to estimate displacement," *2016 5th Mediterranean Conference on Embedded Computing (MECO)*, 2016, pp. 206-210, doi: 10.1109/MECO.2016.7525741