



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

CORSO DI LAUREA IN INGEGNERIA BIOMEDICA

Confronto tra stereofotogrammetria ed analisi video nella ricostruzione della cinematica articolare durante il cammino

Relatore: Prof.ssa Zimi Sawacha

Laureando: Gianluca Barbierato

Correlatore: Dott. Silvia Zampato

ANNO ACCADEMICO 2021 – 2022

Data di laurea 19/09/2022

INDICE

SOMMARIO	3
CAPITOLO 1 – ANALISI DEL MOVIMENTO	5
1.1 INTRODUZIONE	5
1.2 GAIT ANALYSIS	5
1.3 STRUMENTAZIONE UTILIZZATA NELLA GAIT ANALYSIS	6
1.4 PROTOCOLLI PER L’ANALISI DEL MOVIMENTO	7
1.4.1 PROTOCOLLO HELEN HAYES – DAVIS	7
1.4.2 PROTOCOLLO C.A.S.T.	8
1.4.3 PROTOCOLLO IOR – GAIT	10
CAPITOLO 2 – TECNOLOGIA MARKERLESS	13
2.1 INTRODUZIONE	13
2.2 TECNOLOGIE MOTION CAPTURE	14
2.2.1 SISTEMI NON OTTICI	14
2.2.1.1 SISTEMI ELETTRMECCANICI	15
2.2.1.2 SISTEMI ELETTRMAGNETICI	15
2.2.1.3 SISTEMI INERZIALI	15
2.2.2 SISTEMI OTTICI	16
2.3 SISTEMI OTTICI MARKER-BASED	17
2.3.1 VIDEO TRACKING	18
2.3.1.1 ALGORITMO KLT	19
2.3.2 CALIBRAZIONE DELLE TELECAMERE	22
2.3.2.1 CAMERA CALIBRATION TOOLBOX FOR MATLAB	26
2.3.3 TRIANGOLAZIONE	29
2.4 SISTEMI OTTICI MARKERLESS	30
2.4.1 ACQUISIZIONE E SEGMENTAZIONE DELLE IMMAGINI	30
2.4.2 VISUALHULL E MODELLO	32
2.5 MARKERLESS ED INTELLIGENZA ARTIFICIALE	35
CAPITOLO 3 – MATERIALI E METODI	36
3.1 SOGGETTI	36
3.2 STRUMENTAZIONE	37
3.3 PROCEDURA MARKER-BASED	38
3.4 PROCEDURA MARKERLESS	39
3.5 VIDEO PROCESSING	40
3.6 INDICI DI CONFRONTO	43

CAPITOLO 4 – DISCUSSIONE DEI RISULTATI	44
CAPITOLO 5 – CONCLUSIONI	57
BIBLIOGRAFIA	59

SOMMARIO

L'acquisizione e l'analisi del movimento umano sono oggi impiegate in diversi campi, da applicazioni grafiche, animazioni ad applicazioni sportive, in particolar modo in ambito biomeccanico e clinico riabilitativo. I metodi attualmente più comuni prevedono un ambiente di laboratorio controllato e l'utilizzo di marker applicati sulla pelle del paziente. Lo scopo è quello di raccogliere informazioni relative alla meccanica del sistema muscolo-scheletrico durante uno specifico gesto motorio, che nel caso specifico di questo progetto di tesi è il cammino. Nello specifico sono state analizzate due diverse tecnologie per la ricostruzione della cinematica del cammino: il *feature-tracking*, in presenza di marker applicati sul corpo del soggetto in punti di reperi anatomico specifico e in assenza di marker. La comparazione tra le due metodologie è avvenuta attraverso il confronto degli angoli articolari calcolati a partire dai due dataset di punti di traiettorie; inoltre, si è proceduto ad un ulteriore confronto per i medesimi soggetti, tra i dati ricostruiti con le due metodologie di tracking e con gli stessi dati acquisiti simultaneamente tramite stereofotogrammetria presenti nel dataset del laboratorio.

Capitolo 1

ANALISI DEL MOVIMENTO

1.1 - INTRODUZIONE

L'analisi del movimento umano ha come obiettivi la misura e la descrizione delle quantità che caratterizzano l'esecuzione di un atto motorio, ovvero del movimento assoluto e relativo dei segmenti ossei del soggetto durante l'esecuzione di task specifici. Quest'analisi nei suoi vari aspetti comprende un vasto insieme di grandezze rilevabili che possono essere il movimento assoluto del centro di massa, il movimento assoluto dei segmenti corporei, la cinematica articolare, le forze e le coppie di forze scambiate con l'ambiente o trasmesse da strutture interne quali tendini, legamenti e muscoli, il lavoro e la potenza muscolare. Le informazioni che si ricercano riguardano dunque sia la cinematica che la dinamica del movimento. Tali dati vengono ricavati da grandezze che possono essere misurate oppure stimate mediante l'utilizzo di modelli morfo-funzionali dei sistemi coinvolti nell'analisi, permettendo dunque di descrivere quantitativamente le funzioni dell'apparato muscolo-scheletrico in condizioni definite normali oppure 'patologiche'. Questi modelli antropomorfi descrivono il corpo in maniera semplificata, modellizzando i segmenti ossei come corpi rigidi (non prendendo dunque in considerazione la loro deformazione) mentre cute e muscolo come tessuto molle (il quale invece rappresenta una fonte d'errore).

1.2 – GAIT ANALYSIS

Il movimento prodotto durante una camminata è molto complesso, una raffinatissima interazione fra diversi muscoli e articolazioni. La Gait Analysis si traduce nello studio della deambulazione e della nostra attività locomotoria. È in pratica lo studio del moto, l'osservazione e misurazione dei movimenti, della meccanica del corpo e dell'attività dei muscoli. Si valutano le condizioni che influenzano la capacità motoria, si identificano problemi posturali o semplicemente si cercano soluzioni per aumentare l'efficienza sportiva.

Con la Gait Analysis si definisce e si ridisegna il contesto deambulatorio attraverso l'uso di apparecchiature sofisticate e integrate tra di loro.

Questo tipo di analisi è caratterizzata dal fatto di non essere un esame invasivo, può essere ripetuta più volte anche a breve distanza di tempo, fornisce una grande quantità di dati precisi e permette l'elaborazione dei dati raccolti mediante modelli tridimensionali.

1.3 – STRUMENTAZIONE UTILIZZATA NELLA GAIT ANALYSIS

I moderni sistemi di analisi del movimento sono spesso costituiti da più strumenti interfacciati, ognuno con lo scopo di indagare un diverso aspetto del movimento stesso: per esempio si utilizzano laboratori con speciali sistemi optoelettronici che misurano le coordinate tridimensionali di marcatori (markers, figura 1,1), elementi sferici che vengono apposti sul corpo del soggetto in posizioni strategiche per la ricostruzione della scena tridimensionale, pedane di forza e pressione (figura 1.2) per lo studio delle forze che il corpo umano scambia col terreno, ed elettromiografi di superficie per la registrazione dell'attività elettrica associata alla contrazione muscolare.



Figura 1.1: Una tipologia di markers [1]



Figura 1.2: Un esempio di pedana di pressione [1]

In particolar modo consente di definire la posizione e l'orientamento, istante per istante, dei punti di reperi anatomico, dei segmenti anatomici nonché del corpo umano. Negli ultimi anni l'attenzione si è rivolta verso sistemi non invasivi basati sull'acquisizione di sequenze video sincronizzate con comuni telecamere.

1.4 - PROTOCOLLI PER L'ANALISI DEL MOVIMENTO

Determinare l'orientamento di un corpo rigido, o la posizione di un punto nello spazio, rispetto ad un dato sistema di riferimento, è un problema che ammette diverse soluzioni. Questo, ed altri problemi, come possono essere quelli di affidabilità di modello, vengono trattati e, in parte, risolti, con lo sviluppo e l'implementazione di protocolli sperimentali. Un protocollo per l'analisi cinematica del movimento è un insieme di regole che permettono di acquisire, elaborare ed analizzare i dati al fine di ottenere i risultati desiderati. Un protocollo consente di eseguire le istruzioni in maniera uniforme e riproducibile, garantendo dunque la possibilità che le misure siano ripetibili. Nel corso degli anni sono stati sviluppati vari protocolli per l'analisi del movimento a seconda delle diverse specifiche di analisi e utilizzabilità. Esistono infatti protocolli per l'analisi del movimento che possono riguardare l'analisi "full body", cioè di anca, ginocchio e caviglia oppure protocolli adatti ad analizzare un determinato tipo di movimento ed alla patologia dei pazienti da studiare e sono stati validati e riconosciuti dalla comunità scientifica internazionale.

1.4.1 - Protocollo Helen Hayes – Davis

Il protocollo Davis et al. (1991) [2] rappresenta uno standard di riferimento ampiamente diffuso per l'analisi del movimento. Esso prevede la rilevazione dei parametri antropometrici del soggetto (altezza, peso, distanza tra le creste iliache anteriori e spessore del bacino, lunghezza gamba, distanza tra i condili femorali o diametro del ginocchio, distanza tra i malleoli o diametro della caviglia), usati per stimare la posizione dei centri articolari. Successivamente si procede con il posizionamento di 22 marker in specifici punti di repere, quali:

- acromion e settima vertebra cervicale (C7)
- creste iliache e osso sacro
- grande trocantere, epicondilo femorale laterale e un marker posizionato su di una barra posta circa a metà coscia
- testa della fibula, malleolo e un marker posto su di una barra posizionata circa a metà tibia
- quinto metatarso e tallone (solo nella fase di standing)

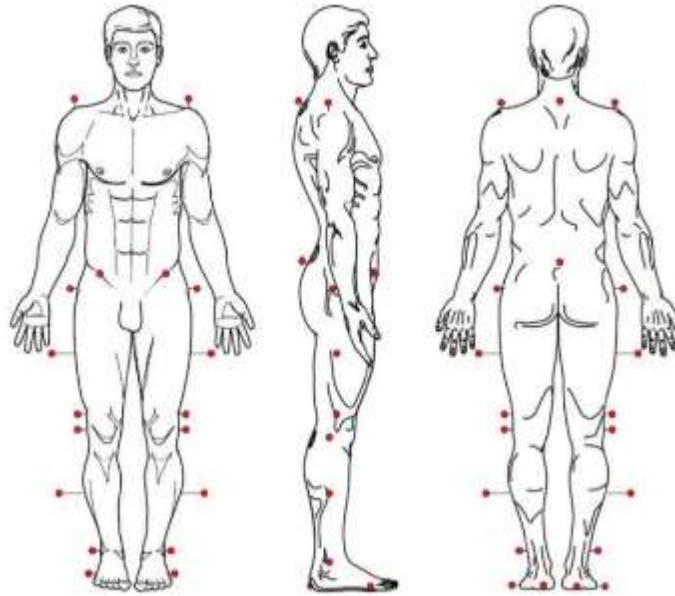


Figura 1.3: posizione markers protocollo Helen Hayes – Davis [2]

Una volta collocati i marcatori si procede con un'acquisizione statica in posizione eretta che consente di registrare le posizioni iniziali e definire i sistemi di riferimento nonché i centri articolari. L'ultimo passaggio consiste nel far camminare il soggetto, con una velocità ritenuta da lui normale, all'interno del volume di acquisizione per un numero di volte necessarie ad ottenere almeno tre camminate buone. Dalle traiettorie dei marcatori infine è possibile ricavare lo spostamento, le velocità ed accelerazioni dei segmenti corporei e gli angoli articolari.

1.4.2 – Protocollo C.A.S.T.

Il protocollo Cast [35] (Calibrated Anatomical System Technique) è stato sviluppato nel 1995 presso l'Istituto di Fisiologia Umana dell'Università La Sapienza, Roma, e il Laboratorio di Analisi del Movimento degli Istituti Ortopedici Rizzoli, Bologna, per rispondere ad un'esigenza di precisione nell'identificazione delle coordinate dei punti di reperi anatomico a fronte dell'artefatto da tessuto molle che ne previene la stima accurata. A tal fine è stata sviluppata una metodologia per stimare le coordinate dei punti di reperi anatomico d'interesse tramite acquisizioni aggiuntive, dette calibrazioni anatomiche, che determinano le coordinate del reperi in riferimento ad un sistema di riferimento tecnico posto sul segmento corporeo, realizzato con un cluster di marcatori, usualmente in numero di 4; è poi necessario procedere per via analitica alla trasformazione delle coordinate anatomiche dal sistema di riferimento locale (tecnico) a quello globale.

La calibrazione anatomica viene effettuata in due modi diversi:

- 1) Se il punto di repere si trova subito al di sotto della cute allora è possibile trascurare l'errore della cute (ATM, artefatto da tessuto molle) collocando il marcatore su tale punto. Si acquisiscono dunque sia la posizione del marker che il cluster (che va a definire il sistema di riferimento tecnico) sul segmento anatomico di interesse. Tale procedura deve essere ripetuta per qualsiasi punto di repere.
- 2) Se il punto di repere si trova in profondità, si impiega una bacchetta con dei marcatori posti a distanza nota, il cui apice viene appoggiato sul punto di repere anatomico di interesse e in questo caso si determina la posizione rispetto al sistema di riferimento tecnico. Anche in questo la procedura deve essere ripetuta per qualsiasi punto di repere.

<i>Hip bone</i>		
(a)	ASIS	anterior superior iliac spine
(a)	PSIS	posterior superior iliac spine
(b)	AC	centre of the acetabulum
<i>Femur</i>		
(a)	FH	centre of the femoral head
(c)	GT	prominence of the greater trochanter external surface
(a)	ME	medial epicondyle
(a)	LE	lateral epicondyle
(b) (c)	LP	anterolateral apex of the patellar surface ridge
(b) (c)	MP	anteromedial apex of the patellar surface ridge
(b) (c)	LC	most distal point of the lateral condyle
(b) (c)	MC	most distal point of the medial condyle
<i>Tibia and fibula</i>		
(c)	IE	intercondylar eminence
(a)	TT	prominence of the tibial tuberosity
(a)	HF	apex of head of the fibula
(a)	MM	distal apex of the medial malleolus
(a)	LM	distal apex of the lateral malleolus
(b) (c)	MMP	most medial point of the ridge of the medial tibial plateau
(b) (c)	NLP	most lateral point of the ridge of the lateral tibial plateau
<i>Foot</i>		
(a)	CA	upper ridge of the calcaneus posterior surface
(a)	FM	dorsal aspect of first metatarsal head
(a)	SM	dorsal aspect of second metatarsal head
(a)	VM	dorsal aspect of fifth metatarsal head

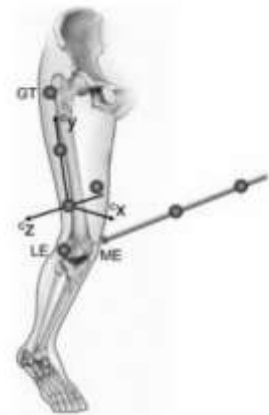


Figura 1.4: a sinistra i punti di repere anatomico di interesse per il protocollo CAST; a destra calibrazione anatomica mediante la wand [3]

1.4.3 - Protocollo IOR-Gait

Il protocollo IOR-gait (Leardini et al. 2007) [4] è stato sviluppato nel 2007 presso l'Istituto Ortopedico Rizzoli, Bologna. Questo protocollo nasce come compromesso tra la riduzione del tempo di preparazione del paziente, grazie all'utilizzo di pochi marker (come nel protocollo Davis), e l'accuratezza e affidabilità dei risultati, grazie alla calibrazione dei punti anatomici (come nel protocollo Cast). Il marker-set e i punti di repere anatomico da considerare sono mostrati in Figura 1.5.

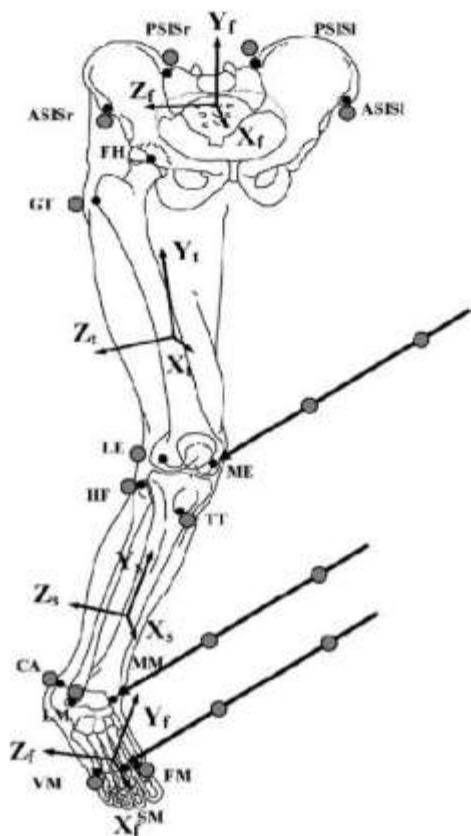


Figura 1.5 – Rappresentazione schematica della posizione dei punti di repere anatomici utilizzati col protocollo IOR-Gait e della disposizione dei marker per la gamba destra. La disposizione è analoga per la gamba sinistra [4].

I punti di riferimento anatomici sono tracciati applicando un marcatore sferico di 10 mm di diametro in specifiche posizioni, quali: i due margini anteriori e posteriori dell'articolazione iliaco – spinale (ASIS, PSIS), le prominenze laterali grande trocantere (GT) e dell'epicondilo laterale femorale (LE), l'estremità prossimale della testa della fibula (HF), il bordo più sporgente della tibia (TT), le sporgenze dei malleoli (LM), l'inserzione del tendine d'Achille in corrispondenza del tallone (CA) e i margini dorsali del piede corrispondenti al primo (FM) e quinto metatarso (VM).

Il centro della testa femorale si assume coincidere con il centro della componente acetabolare, ricostruita mediante un approccio geometrico predittivo basato sul posizionamento dei quattro punti di riferimento del bacino; questo fornisce assieme ai punti (GT) ed (LE) un terzo punto sul femore, che permette la costruzione del sistema di riferimento tecnico. L'epicondilo mediale (ME) è invece calibrato mediante due marker in posizioni note rispetto alla punta. Le calibrazioni vengono eseguite anche per tracciare la prominenza mediale dei malleoli mediali (MM), sfruttando i tre marcatori sulla tibia come sistema di riferimento tecnico, e per tracciare la parte dorsale del secondo metatarso (SM) usando i tre *marker* sul piede (assunto come un unico segmento rigido) come sistema di riferimento tecnico. I centri articolari di anca, ginocchio e la caviglia sono considerati rispettivamente (FH), il punto medio tra (LE) ed (ME), e il punto medio tra (LM) ed (MM). Per ogni articolazione vengono inoltre rilevati, mediante un sistema di riferimento standard, flessione ed estensione attorno all'asse di riferimento Z del segmento prossimale, rotazione interna ed esterna relative all'asse verticale Y del segmento distale, abduzione ed adduzione relative all'asse X perpendicolare ai due precedenti. Questa terminologia è adottata anche per l'articolazione dell'anca e della caviglia; le rotazioni di quest'ultima (Figura 1.6), vengono però indicate come dorsiflessione (flessione) e plantarflexione (estensione), inversione ed eversione, abduzione ed adduzione. Nel caso del bacino (Figura 1.6), vengono valutate con le stesse convenzioni, rotazione, inclinazione ed obliquità.

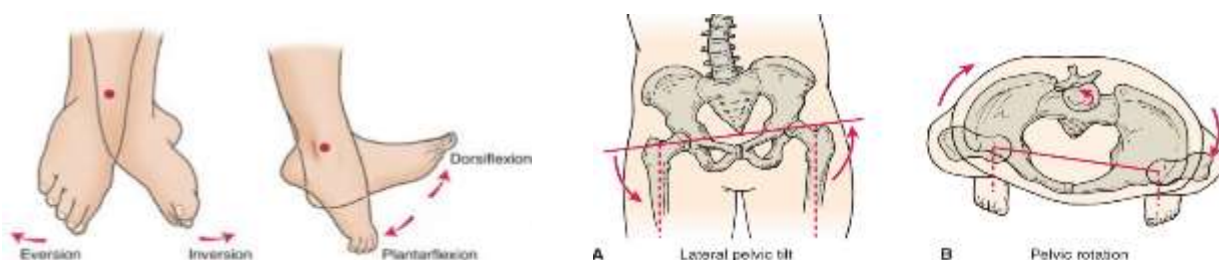


Figura 1.6: esempio movimenti caviglia, a sinistra, e bacino a destra [5], [6]

La scelta di adottare il protocollo IOR-GAIT ricade sulla base delle sue caratteristiche, le quali sono un compromesso dei protocolli precedentemente elencati. Grazie ad un minor numero di marker utilizzati, oltre al fatto di permettere una preparazione abbastanza veloce del paziente, consente una notevole riduzione del tempo necessario alla post elaborazione dei dati; tuttavia, l'accuratezza della ricostruzione dei punti di repere è elevata e comparabile a quella dei protocolli che utilizzano un maggior numero di marker.

Capitolo 2

TECNOLOGIA MARKERLESS

2.1 - INTRODUZIONE

L'applicazione di tecniche di motion capture, ossia del processo di acquisizione del movimento e i meccanismi annessi, richiede grande accuratezza e funzionalità del metodo usato.

Adottando la tecnica marker-based, predisporre il laboratorio, calibrare il sistema e posizionare i *marker* sul soggetto, comporta un impiego di tempo non trascurabile. Inoltre, questa tecnica è affetta da varie sorgenti di errore, strumentali (artefatti dovuti ad una cattiva stima dei parametri di calibrazione, flickering, ecc.) e metodologiche (artefatto da tessuto molle, volume di acquisizione insufficiente per il gesto considerato); infine, l'ingombro dell'attrezzatura stessa può rendere difficoltoso l'utilizzo in particolari condizioni, riducendo così la naturalità del movimento da parte del soggetto sotto esame. Durante l'atto motorio, inoltre, i *marker* sono soggetti a movimenti globali, prodotti dal movimento scheletrico, e a movimenti locali, causati dal movimento della pelle e dei tessuti sottostanti relativamente alle strutture ossee di riferimento (artefatto da tessuto molle); ciò può avere effetti negativi sull'accuratezza e sulla precisione della misurazione.

Per questi motivi, tecniche emergenti e la ricerca nell'ambito della computer vision stanno portando al rapido sviluppo dell'approccio *markerless*, con l'obiettivo di arrivare ad ottenere qualcosa che consenta di facilitare l'intero processo senza perdere affidabilità ed accuratezza.

2.2 - TECNOLOGIE MOTION CAPTURE

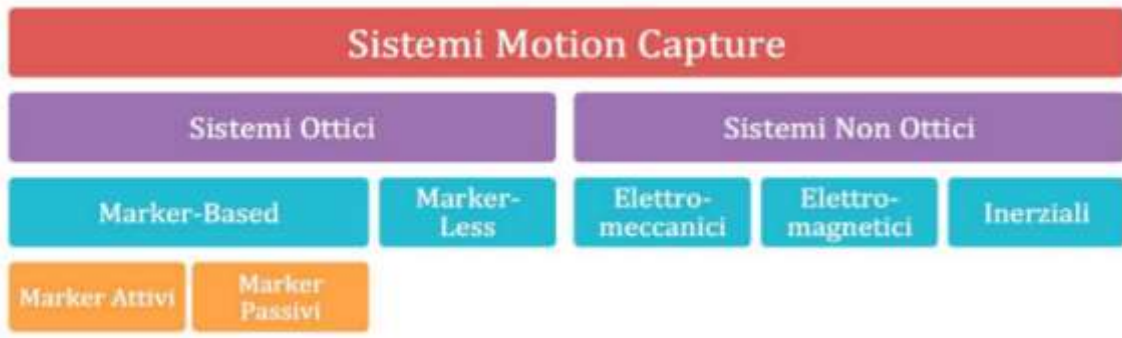


Figura 2.1: Suddivisioni dei sistemi motion capture

Tipicamente con il termine motion capture si identifica un sistema composto da telecamere tra loro interconnesse in grado di rilevare il movimento. Attualmente le tecnologie disponibili si basano su due approcci differenti: sistemi ottici e sistemi non ottici (Figura 2.1). I sistemi non ottici si differenziano ulteriormente in meccanismi elettromagnetici, elettromeccanici ed inerziali, mentre i sistemi ottici si dividono in *marker-based* e in *markerless* [7].

2.2.1 – SISTEMI NON OTTICI

Questo tipo di sistemi è caratterizzato dall'assenza di telecamere, pertanto possono essere usati all'aperto così come all'interno di laboratori senza vincoli sull'illuminazione. Il vantaggio principale è quello di non richiedere la complessa tecnologia di computer vision. Tuttavia, il limite principale è quello di fornire solo dati relativi all'ultima posizione monitorata, inoltre presenta gli stessi problemi dei sistemi ottici basati sui marcatori: è necessario preparare il soggetto installando vari sensori per il rintracciamento del moto. Tale operazione molto spesso comporta un'eccessiva perdita di tempo.

2.2.1.1 – Sistemi elettromeccanici

Il monitoraggio elettromeccanico è generalmente realizzato attraverso barre mobili collegate tra loro con sensori potenziometrici applicati in corrispondenza delle articolazioni tali da formare esoscheletri. È una tecnologia a basso costo in cui il movimento è percepito in tempo reale tramite piccole variazioni dei potenziometri. I movimenti rilevati risultano accurati, senza interferenze e con un volume di acquisizione illimitato; tuttavia, l'ingombro della strumentazione limita i movimenti che possono essere eseguiti. [7, 18]

2.2.1.2 – Sistemi elettromagnetici

L'approccio elettromagnetico sfrutta sensori elettromagnetici posizionati in corrispondenza delle articolazioni in grado di trasmettere la loro posizione e orientazione. Poiché ogni sensore registra le differenze del campo elettromagnetico generato da un trasmettitore principale, è richiesto un ambiente privo di distorsioni del campo magnetico. A differenza dei sistemi elettromeccanici, i volumi di cattura per questo tipo di sistemi sono notevolmente più piccoli, inoltre la risposta del sensore non è una risposta lineare, specialmente verso i bordi della zona di cattura [7, 10].

2.2.1.3 – Sistemi inerziali

I sistemi di tipo inerziale combinano segnali provenienti da giroscopi, da accelerometri e da magnetometri. Per determinare la direzione della verticale si rileva l'accelerazione di gravità attraverso gli accelerometri, mentre i sensori magnetici forniscono la stabilità secondo il piano orizzontale rilevando la direzione del campo magnetico terrestre come una bussola. Per un monitoraggio completo a sei gradi di libertà di ogni singolo segmento corporeo, i dati vengono poi integrati con i dati del giroscopio che fornisce l'orientamento. [8]

2.2.2 – SISTEMI OTTICI

Questi sistemi utilizzano una serie di telecamere calibrate e sincronizzate tra loro, disposte in posizioni differenti che si interfacciano con i *marker* posizionati sul soggetto sfruttando le caratteristiche geometriche (punti, aree) della superficie corporea [9].

La dimensione e la forma dei marker dipendono dalla risoluzione della camera, e possono essere di due tipologie:

- Marker passivi, sono composti da materiale retroriflettente di forma sferica che riflettono la luce che deriva da dei diodi luminescenti (LEDs) equipaggiati dalla camera,
- Marker attivi - in questo caso i marker sono propriamente dei LEDs ad infrarossi che inviano dei segnali luminosi alle camere, possono illuminarsi singolarmente evitando così il bisogno di identificare ogni marker; altri sistemi marker-based attivi si illuminano in contemporanea e serve controllare l'ampiezza e la frequenza del segnale per distinguere i diversi marker.

Grazie ai diversi vantaggi apportati dalle tecniche di elaborazione d'immagine, i sistemi ottici rappresentano la maggior parte dei sistemi utilizzati [9].

A differenza dei sistemi non ottici, essi non forniscono una misura diretta delle variabili cinematiche, bensì una loro stima [11].

2.3 – SISTEMI OTTICI MARKER-BASED

Nei sistemi marker-based l'obiettivo è ricostruire la cinematica dei segmenti ossei attraverso l'individuazione della traiettoria 3D dei marker (piccoli oggetti applicati alla superficie corporea del soggetto in movimento, in zone che rappresentano in modo accurato le azioni delle articolazioni). Il funzionamento delle tecniche marker-based si basa fundamentalmente sul principio di triangolazione: quando almeno due telecamere riprendono contemporaneamente un punto dello spazio 3D allora è possibile, a partire dalle immagini formatesi sui piani principali delle telecamere, ricostruire la posizione di quel punto tramite semplici retroproiezioni (Figura 2.2).

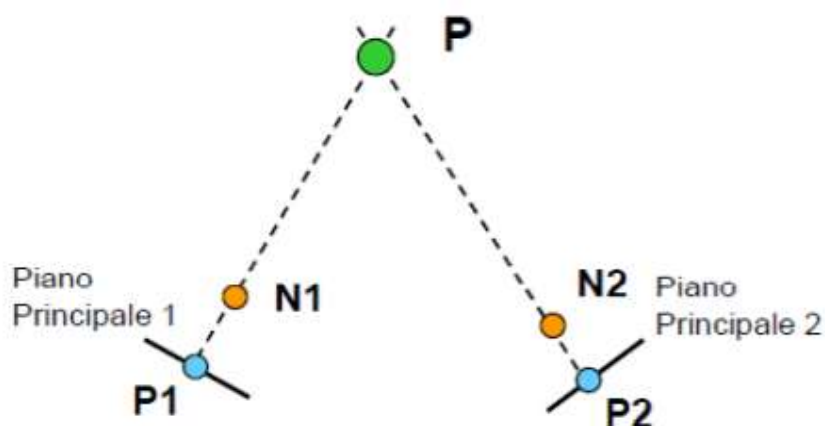


Figura 2.2: Ricostruzione della posizione del punto P, note le sue proiezioni P1 e P2 sui piani principali delle telecamere e le posizioni N1 e N2 dei punti nodali delle telecamere [19].

La tecnica MoCap ottica marker-based di questo tipo è detta anche stereofotogrammetria, dove il prefisso stereo- discende proprio dal principio appena esposto.

Nonostante la popolarità, l'approccio ottico presenta alcune importanti limitazioni che devono essere considerate [9]. Innanzitutto, anche se l'ingombro è limitato, la presenza di marker attaccati alla pelle può influenzare il movimento del soggetto. Essendo inoltre i marker molto piccoli, è possibile che in determinate fasi del movimento uno o più marker vengano oscurati dalla vista della telecamera a causa del movimento delle braccia o delle gambe alle telecamere, facendo così perdere temporaneamente la sua traccia. In corrispondenza di questi eventi capita anche che le traiettorie di marker molto vicini vengano confuse.

Per permettere la ricostruzione delle posizioni dei marker è necessario che ogni indicatore sia visto contemporaneamente da almeno due telecamere.

Questi problemi possono tuttavia essere parzialmente superati con l'utilizzo di un maggior numero di telecamere, le quali però sono dispositivi molto costosi, [11] ed un adeguato posizionamento dei marker. Altri svantaggi riguardano il posizionamento dei marker, il quale è dispendioso in termini di tempo, non completamente ripetibile [12], e la presenza di tessuti molli tra ossa e marker causa un movimento relativo che introduce un artefatto nelle misurazioni, chiamato in gergo ATM, "artefatto da tessuto molle" [13, 14].

2.3.1 – VIDEO TRACKING

La mancanza di strumenti in grado di fornire in modo automatico il monitoraggio dei marker può essere compensata dalla digitalizzazione manuale dei punti caratteristici del movimento. La procedura di *video tracking* consiste nel rilevare, utilizzando delle telecamere, la posizione di markers, collocati nel soggetto in esame nei punti di interesse, da parte di un operatore che analizza le sequenze video grazie all'utilizzo di un *software*, il quale implementa un algoritmo che consente di facilitare l'operatore nelle operazioni di riconoscimento della corretta posizione dei markers nei vari fotogrammi (*frames*).

La maggiore difficoltà legata al *video tracking* è quella di riuscire a identificare la corretta posizione di un marker nella sequenza di fotogrammi, in particolare quando la velocità del gesto motorio è superiore al *frame rate* della telecamera, infatti, può accadere che a causa della complessità dei movimenti e delle variazioni di luminosità, alcuni dei punti vengano persi od oscurati nel tracciamento; in tal caso sarà compito dell'operatore fornire nuovamente la posizione del marker. È importante ricordare che ogni operatore, in quanto essere umano, avrà una percezione visiva differente rispetto ad ogni altro, portando una componente soggettiva, se pur piccola, che indurrà una leggera variabilità nei risultati.

Dato che la posizione dei markers non è casuale, è opportuno definire all'avvio del monitoraggio le posizioni iniziali e il frame di partenza della sequenza video, dopodiché è possibile continuare con l'analisi video di nostro interesse.

2.3.1.1 – *Tracking di features*: l' Algoritmo KLT

Un algoritmo che permette di estrarre caratteristiche di oggetti che non cambiano forma e che presentano una stabilità visiva è l'algoritmo iterativo Kanade-Lucas-Tomasi (KLT). L'algoritmo è stato inizialmente sviluppato da Lucas e Kanade [16], successivamente è stata proposta una variante da Tomasi [17]. Per comprendere meglio il funzionamento dell'algoritmo, si prende in esame una sequenza di immagini $I(x, t)$ definita in funzione di due coordinate spaziali che individuano i punti dell'immagine al tempo t . Se la frequenza di campionamento è sufficientemente elevata, le immagini prese ad istanti vicini sono tra loro correlate, dato che si riferiscono alla stessa scena vista da punti di vista simili. Pertanto, il cambiamento dell'immagine corrente all'istante successivo $t + \tau$ è descrivibile attraverso la seguente relazione:

$$I(u, v, t + \tau) = I(u - \xi(u, v, t, \tau), v - \eta(u, v, t, \tau)) \quad (2.3.1)$$

Le immagini successive, dunque, possono ottenersi spostando ogni singolo punto $\mathbf{x} = (u, v)$ dell'immagine di una quantità $\mathbf{d} = (\xi, \eta)$, detta *displacement*.

La determinazione dello spostamento \mathbf{d} è resa complessa da una serie di problemi derivanti dall'intensità luminosa dei pixel. Per inseguire un pixel è necessario che la sua intensità luminosa sia distinta rispetto ai pixel confinanti, inoltre a causa del rumore è possibile una variazione della luminosità del pixel stesso e pertanto può essere confuso con i pixel adiacenti. Per questo motivo è opportuno considerare una finestra di pixel piuttosto che un solo pixel. Purtroppo, i punti che compongono la finestra molto spesso hanno comportamenti differenti, muovendosi a velocità diverse o, talvolta, scomparendo completamente. A questo punto ci si interroga su come sia possibile essere sicuri di inseguire sempre la stessa finestra, e su come ottenere lo spostamento combinando differenti velocità. Il primo problema è risolvibile monitorando i residui: se l'aspetto della finestra non varia molto si tiene la finestra, altrimenti viene scartata.

Il problema di ricavare lo scostamento combinando velocità differenti può essere risolto introducendo un campo di moto affine $\delta(\mathbf{x})$, *affine motion field*, con il quale è possibile associare diverse velocità a diversi punti della finestra:

$$\delta(\mathbf{x}) = \mathbf{D}\mathbf{x} + \mathbf{d} \quad (2.3.2)$$

dove \mathbf{D} è la matrice di deformazione $\mathbf{D} = \begin{bmatrix} d_{uu} & d_{uv} \\ d_{vu} & d_{vv} \end{bmatrix}$, mentre \mathbf{d} rappresenta il vettore di traslazione del centro della finestra.

Poiché si è supposto che la frequenza di campionamento sia sufficientemente e alta, è possibile trascurare la matrice di deformazione ponendola a zero, approssimando così il movimento con la sola traslazione \mathbf{d} .

Considerando il vettore traslazione \mathbf{d} e la relazione 2.3.1, è possibile definire l'immagine al tempo $t + \tau$ come:

$$J(\mathbf{x}) = I(u, v, t + \tau) \quad (2.3.3)$$

$$I(\mathbf{x} - \mathbf{d}) = I(x - \xi, y - \eta, t) \quad (2.3.4)$$

Poiché il modello di movimento dell'immagine non è perfetto a causa del rumore $r(\mathbf{x})$, si ha:

$$J(\mathbf{x}) = I(\mathbf{x} - \mathbf{d}) + r(\mathbf{x}) \quad (2.3.5)$$

Dato che lo scopo dell'algoritmo è quello di individuare in modo automatico la posizione del punto selezionato per ogni coppia di fotogrammi successivi nella sequenza, il problema è quindi ricavare \mathbf{d} che minimizza il residuo delle differenze quadrate ε in un intorno del punto di interesse:

$$\varepsilon = \iint_W [I(\mathbf{x} - \mathbf{d}) - J(\mathbf{x})]^2 w(\mathbf{x}) d\mathbf{x} \quad (2.3.6)$$

dove W è una finestra quadra centrata in \mathbf{x} , e $w(\mathbf{x})$ che può essere posta a 1, oppure usata per dare maggior importanza ai pixel centrali della finestra rispetto ai periferici (per esempio potrebbe essere una gaussiana). Per semplicità la funzione peso viene posta uguale a 1.

L'inseguimento del punto non è nient'altro che osservare la somiglianza tra l'intorno del punto \mathbf{x} dell'istante corrente con l'istante successivo. Si parla quindi di un inseguimento di caratteristiche, *feature*, ovvero l'insieme di pixel all'interno della finestra W i cui valori

caratterizzano il punto \mathbf{x} . Sempre grazie all'assunzione fatta all'inizio di *frame rate* elevato, è possibile approssimare l'equazione 2.3.5 mediante il suo sviluppo di Taylor del primo ordine:

$$I(\mathbf{x} - \mathbf{d}) \approx I(\mathbf{x}) - \mathbf{g} \cdot \mathbf{d} \quad (2.3.7)$$

dove \mathbf{g} è il vettore delle derivate prime di I rispetto allo spostamento \mathbf{d} .

Il residuo pertanto diventa come:

$$\varepsilon = \int_W [I(\mathbf{x}) - \mathbf{g} \cdot \mathbf{d} - J(\mathbf{x})]^2 w(\mathbf{x}) d\mathbf{x} \quad (2.3.8)$$

Il residuo è quindi una funzione quadratica del vettore \mathbf{d} .

Per minimizzare tale residuo, si pone a zero la derivata rispetto a \mathbf{d} incognito:

$$\varepsilon = \int_W (I(\mathbf{x}) - \mathbf{g} \cdot \mathbf{d} - J(\mathbf{x})) \mathbf{g} w(\mathbf{x}) dA = 0 \quad (2.3.9)$$

Posto $(\mathbf{g} \cdot \mathbf{d}) \mathbf{g} = (\mathbf{g} \mathbf{g}^T) \mathbf{d}$ e assumendo \mathbf{d} costante sulla finestra W si ricava:

$$\left(\int_W \mathbf{g} \mathbf{g}^T w(\mathbf{x}) dA \right) \mathbf{d} = \int_W (I(\mathbf{x}) - J(\mathbf{x})) \mathbf{g} w(\mathbf{x}) dA \quad (2.3.10)$$

che definisce un sistema lineare di due equazioni in due incognite, riscritto in forma matriciale:

$$\mathbf{C} \mathbf{d} = \mathbf{e} \quad (2.3.11)$$

$$\mathbf{C} = \int_W \mathbf{g} \mathbf{g}^T w(\mathbf{x}) dA \quad (2.3.12)$$

$$\mathbf{e} = \int_W (I(\mathbf{x}) - J(\mathbf{x})) \mathbf{g} w(\mathbf{x}) dA \quad (2.3.13)$$

Il sistema lineare può essere risolto in modo affidabile se la matrice \mathbf{C} rimane sopra il livello del rumore dell'immagine e se non soffre del problema del mal condizionamento. Il problema del mal condizionamento si presenta quando piccole perturbazioni negli elementi della matrice quadrata \mathbf{C} provocano grandi variazioni nelle soluzioni \mathbf{d} del sistema.

Il vincolo imposto dal rumore implica che entrambi gli autovalori della matrice \mathbf{C} devono essere grandi, mentre il vincolo imposto dal problema di mal condizionamento richiede che i due autovalori non differiscano di molti ordini di grandezza. Quando il minore dei due autovalori è sufficientemente grande da soddisfare il criterio sul rumore, la matrice \mathbf{C} è solitamente ben condizionata.

Questo è dovuto al fatto che le variazioni d'intensità all'interno della finestra sono legate al massimo valore di intensità consentito; pertanto, l'autovalore maggiore non può essere arbitrariamente grande. La finestra, dunque, non viene scartata se:

$$\min(\lambda_1, \lambda_2) > \lambda \quad (2.3.14)$$

con λ_1 e λ_2 gli autovalori della matrice \mathbf{C} , mentre λ è un valore di soglia fissato, determinato precedentemente osservando una regione ad intensità circa costante dalla quale sono stati ricavati un limite inferiore e superiore per l'autovalore λ .

2.3.2 – CALIBRAZIONE DELLE TELECAMERE

Nei sistemi *motion capture ottici*, è necessaria la *calibrazione* delle telecamere, per ottenere una corretta ricostruzione tridimensionale del movimento. Tale operazione è essenziale per la stima di una serie di parametri del modello ottico della telecamera, la quale permette la correzione delle distorsioni introdotte dalle lenti della telecamera. Per ricostruire dunque la posizione tridimensionale dei marcatori, sono necessarie almeno due telecamere opportunamente sincronizzate, i cui piani focali non siano coincidenti. In altri termini, prima di poter effettuare le acquisizioni è necessario calibrare lo spazio all'interno del quale verrà eseguito il movimento. Dopo aver effettuato questa procedura, il sistema è pronto per acquisire le immagini.

Noti i parametri di calibrazione è possibile:

- eseguire una trasformazione di coordinate dal sistema assoluto al sistema di coordinate della camera localizzato nel suo centro di prospettiva e viceversa;
- eseguire una trasformazione prospettica dallo spazio 3D allo spazio 2D della camera e viceversa;
- eseguire una trasformazione 2D dal sistema di coordinate del piano immagine al sistema di riferimento del sensore e viceversa (Figura 2.4).

I parametri che dovranno essere determinati sono:

- Parametri intrinseci: necessari a specificare le caratteristiche ottiche, geometriche e digitali della telecamera, come sensore, in modo indipendente dalla sua posizione nello spazio.
- Parametri estrinseci: definiscono la posizione del sistema di riferimento della telecamera nello spazio.

Per poter calibrare il sistema, tuttavia, è necessario introdurre un modello matematico dello stesso. Il più utilizzato in letteratura è il modello *pin-hole* (o a *foro stenopeico* o modello della camera oscura).

Tale schema presuppone che tutti i raggi ottici che concorrono alla formazione dell'immagine, ad un certo tempo del loro cammino passano per uno stesso punto: il *punto focale* (*pin-hole*). I raggi, provenienti da un oggetto nello spazio, attraversano il foro e formano sul piano immagine una rappresentazione invertita dell'oggetto stesso.

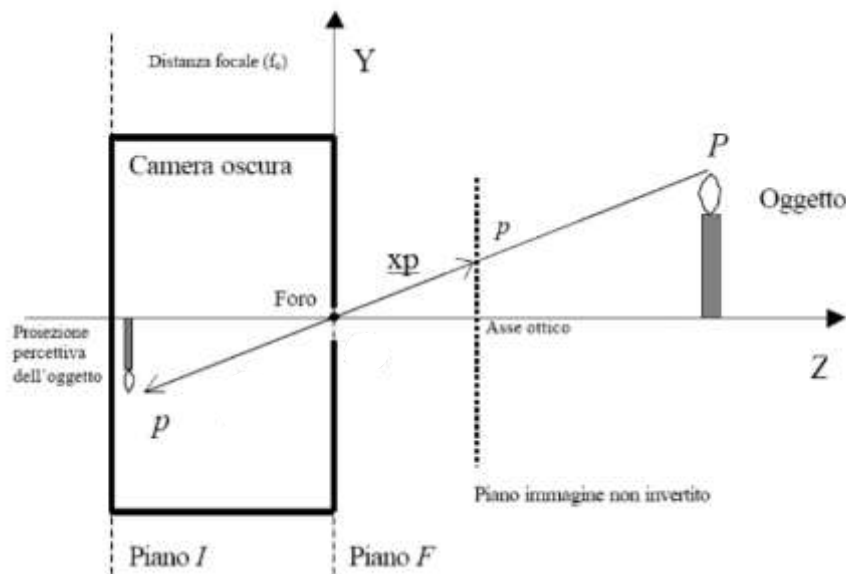


Figura 2.3: modello della camera oscura

Un raggio di luce proveniente da un punto **P** dello spazio attraversa il **foro** e incide sul piano **I**, detto *piano immagine*, in un punto *p*. La distanza tra i piani **F** e **I** è la *distanza focale*. Il punto dove è stato praticato il foro si indica come *centro ottico*, mentre la retta normale al piano focale passante per il foro è detto *asse ottico* (Figura 2.3).

Operando dei prodotti matriciali, si determina il sistema di riferimento di ciascuna telecamera, adottato per riportare le coordinate dei marker secondo il sistema di riferimento assoluto.

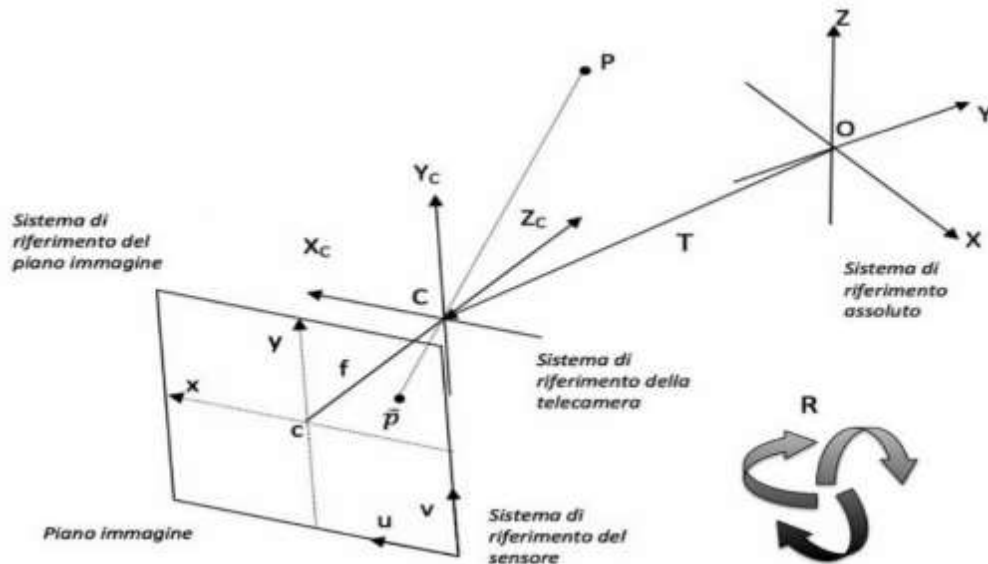


Figura 2.4: Proiezione sul piano immagine di un punto **P** nello spazio tridimensionale [15]

Prendendo in considerazione il sistema di riferimento della telecamera, come da Figura 2.4, definito dalla terna cartesiana (C, X_C, Y_C, Z_C) , con origine nel centro di prospettiva C e con l'asse Z_C coincidente con l'asse ottico (perpendicolare al piano immagine), è possibile ottenere la relazione che intercorre tra le coordinate dell'oggetto P (x_P, y_P) , espresse nel sistema di riferimento del piano immagine, e (X_P, Y_P, Z_P) , espresse invece nel sistema di riferimento della camera mediante la seguente relazione, rispettivamente:

$$\begin{pmatrix} x_P \\ y_P \end{pmatrix} = -f \begin{pmatrix} \frac{1}{Z_P} & 0 \\ 0 & \frac{1}{Z_P} \end{pmatrix} \begin{pmatrix} X_P \\ Y_P \end{pmatrix} \quad (2.3.15)$$

Dove f è la lunghezza focale, ovvero, la distanza del centro di prospettiva della camera dal piano immagine.

La formula (2.3.15) può essere riscritta in coordinate omogenee e diventa:

$$\lambda \begin{pmatrix} x_P \\ y_P \\ 1 \end{pmatrix} = \begin{pmatrix} -f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X_P \\ Y_P \\ Z_P \\ 1 \end{pmatrix} \quad (2.3.16)$$

Il sensore della telecamera ha un proprio sistema di riferimento centrato sull'angolo inferiore destro del piano immagine, nel quale sono espresse le coordinate 2D in uscita dalla camera. Date quindi le dimensioni orizzontale e verticale dei pixel (k_u, k_v) e le coordinate del punto principale $c = (u_0, v_0)$, le coordinate $e = (u, v)$ del punto P nel sistema di riferimento del sensore saranno:

$$\begin{pmatrix} u_P \\ v_P \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{k_u} & 0 & u_0 \\ 0 & \frac{1}{k_v} & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_P \\ y_P \\ 1 \end{pmatrix} \quad (2.3.17)$$

Essendo poi il punto P originariamente espresso rispetto al sistema di riferimento assoluto (O, X, Y, Z) si ha:

$$\begin{pmatrix} X_P \\ Y_P \\ Z_P \\ 1 \end{pmatrix} = \begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X_a \\ Y_a \\ Z_a \\ 1 \end{pmatrix} \quad (2.3.18)$$

Dove il vettore \mathbf{T} e la matrice di rotazione \mathbf{R} esprimono rispettivamente la traslazione e l'orientamento del sistema di riferimento della camera rispetto al sistema di riferimento assoluto.

Il modello pin hole può essere rappresentato da una matrice \mathbf{A} ottenuta come composizione delle tre precedenti trasformazioni:

$$\mathbf{A} = \begin{pmatrix} 1/k_u & 0 & u_0 \\ 0 & 1/k_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix} = \mathbf{HMD} \quad (2.3.19)$$

La limitazione principale di questo approccio è che l'intensità della luce che attraversa il foro e che giunge sugli elementi fotosensibili del piano immagine è troppo piccola per essere rilevata, per cui è necessario allargare il foro (diaframma) e ricorrere ad un'ottica per mettere a fuoco l'immagine.

Riassumendo, i parametri di calibrazione estrinseci stabiliscono una trasformazione di coordinate dal sistema assoluto al sistema di coordinate della telecamera localizzato nel suo centro di prospettiva (matrice \mathbf{D}), mentre i parametri intrinseci stabiliscono una trasformazione prospettica dallo spazio 3D a quello 2D e una trasformazione 2D dal sistema di coordinate del piano dell'immagine al sistema di riferimento del sensore (matrici \mathbf{H} e \mathbf{M}) [20].

2.3.2.1 – *Camera calibration toolbox for Matlab*

Grazie al Toolbox di calibrazione per Matlab di J.Y. Bouguet [21] viene fornito un possibile modo per poter calibrare le telecamere, nel quale metodo è stato assunto come modello di distorsione il modello Brown-Conrady, o a filo a piombo, introdotto nel 1966 da D.C. Brown [22]. Il procedimento di calcolo si articola in diverse fasi, successivamente descritte, e prevede l'uso di una classica scacchiera rigida (Figura 2.5).

Mediante il procedimento, vengono determinati i parametri intrinseci ed estrinseci delle telecamere.

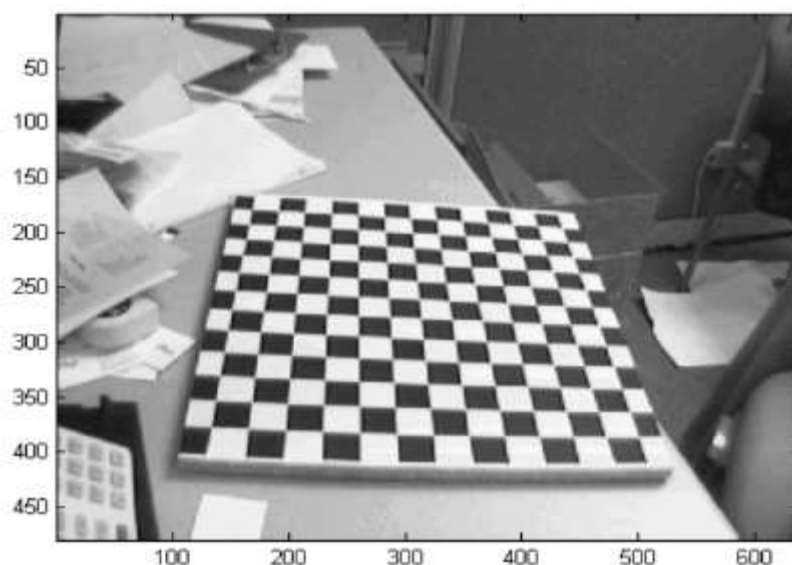


Figura 2.5: scacchiera utilizzata per la calibrazione [23].

La procedura per ricavare i parametri interni (lunghezza focale, punto principale, coefficiente asimmetrico (*skew*), distorsioni) è nota come *calibrazione intrinseca* del sistema, nella quale si procede acquisendo delle immagini della scacchiera (di dimensioni note) ripresa da angolazioni diverse per ogni telecamera. È necessario fornire in input all'algoritmo il numero di tasselli per riga (*wintx*) o per colonna (*winty*) presenti nella scacchiera (per default il valore di *wintx* = *winty* è uguale a 5); vengono quindi digitalizzati i bordi e mappate le distanze fra i vertici di ogni quadrato.

In alcuni casi, l'algoritmo può conteggiare un valore errato di tasselli presenti causando un errore nel processo di calibrazione, soprattutto in presenza di obiettivi caratterizzati da distorsioni molto elevate (Figura 2.6).

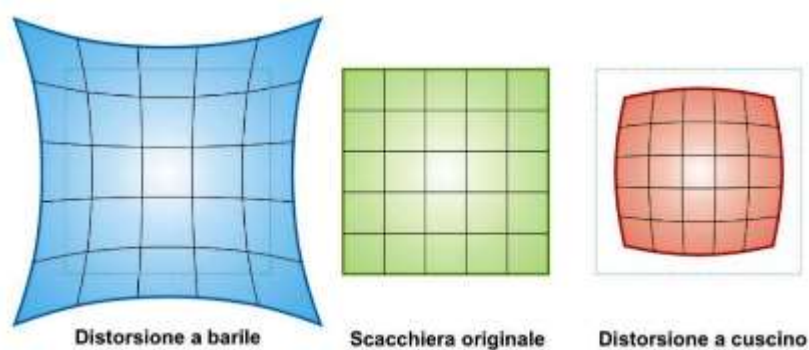


Figura 2.6: tipi di distorsione

Terminata questa fase, si prosegue determinando la *calibrazione estrinseca*, con la quale viene definito un sistema di riferimento globale; i singoli sistemi di riferimento di ciascuna telecamera vengono ricondotti ad un unico sistema di riferimento. Anche per questa calibrazione, oltre ai parametri rilevati precedentemente, è necessario l'uso della scacchiera. Si colloca quindi la scacchiera al centro del volume di acquisizione in modo tale che sia visibile completamente da ogni telecamera e viene richiesto di procedere alla selezione manuale dei quattro vertici della scacchiera.

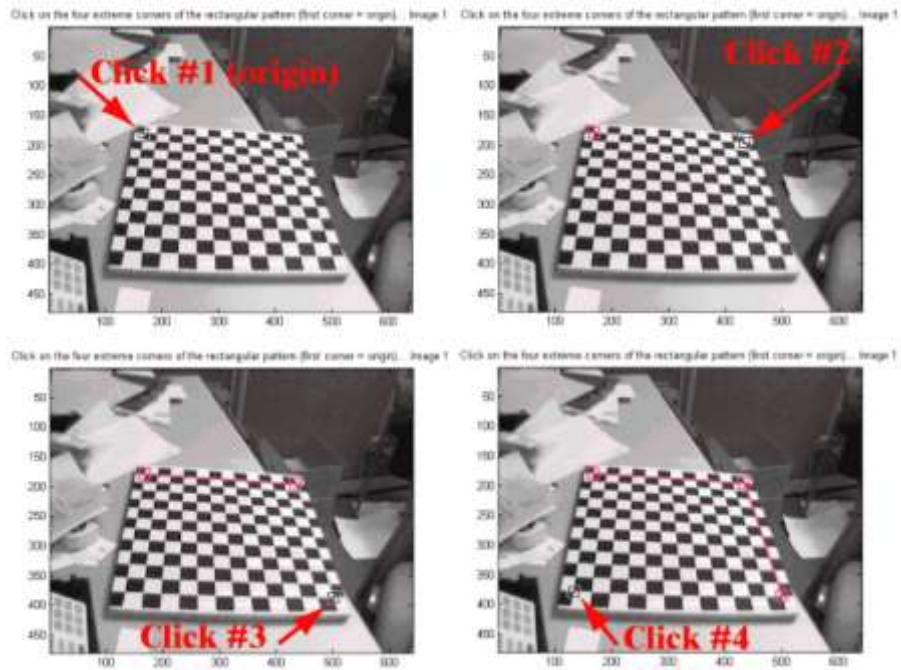


Figura 2.7: esempio di selezione dei vertici [23]

L'ordine di selezione dei quattro vertici non è casuale. Il primo punto cliccato, infatti, viene associato al punto di origine del sistema di riferimento collegato alla singola immagine. Gli altri tre, invece, possono essere selezionati in qualsiasi ordine in quanto il software è in grado di ricostruire la terna di riferimento locale basandosi solo sulla posizione dell'origine e sull'identificazione del piano della scacchiera, definito dai quattro punti totali.

Una volta che la calibrazione estrinseca è stata fatta, le coordinate di un punto nel sistema di riferimento della telecamera possono essere espresse secondo il sistema di riferimento globale.

2.3.3 – TRIANGOLAZIONE

Dopo aver terminato la fase di calibrazione, utilizzando i dati ottenuti si prosegue nella fase successiva: la triangolazione. In questo stadio, il sistema, che deve essere composto da almeno due telecamere, è in grado di combinare le immagini bidimensionali provenienti da ogni telecamera e ricostruire la posizione tridimensionale dei marcatori in ogni istante di tempo. In altri termini per la ricostruzione 3D sono necessari almeno due punti 2D, dove il punto di partenza è rappresentato dalle coordinate delle proiezioni nei sistemi di riferimento del piano immagine (x,y) , mentre il punto di arrivo corrisponde alle coordinate dei marker nello spazio 3D (X,Y,Z) .

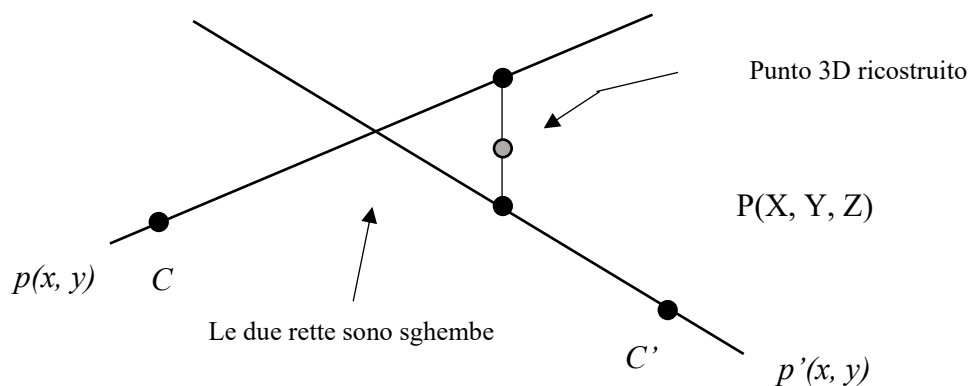


Figura 2.8: ricostruzione 3D tramite triangolazione

Come illustrato in figura 2.8, a causa del rumore le rette che congiungono le coordinate del punto P all'origine del sistema di riferimento delle telecamere, sono sghembe e possono non intersecarsi tra loro; la soluzione alla massima verosimiglianza richiede di minimizzare la somma al quadrato degli errori tra il punto osservato e il punto predetto [15].

2.4 – SISTEMI OTTICI MARKERLESS

La possibilità di ricostruire il movimento umano a partire unicamente da sequenze di immagini è un argomento che, negli ultimi anni, ha riscosso un notevole interesse, soprattutto grazie all'ampio raggio di applicazioni che sarebbe possibile sviluppare con questa tecnica. Le tecniche markerless puntano ad estrarre i parametri relativi al movimento a partire da determinate caratteristiche o silhouettes del soggetto nelle immagini (come, ad esempio, i bordi) [24]. Tuttavia, pur consentendo tempi di preparazione del soggetto ridotti e non risentendo degli artefatti presenti con l'applicazione dei marker, non sono ancora giunti ad un livello di precisione tale da essere comparabile con i sistemi marker-based.

2.4.1 – ACQUISIZIONE E SEGMENTAZIONE DELLE IMMAGINI

Come per l'approccio marker-based, anche in questo tipo di tecnica l'acquisizione dei dati avviene utilizzando delle telecamere opportunamente calibrate e sincronizzate. Questa fase richiede di prestare attenzione alla quantità di informazioni da acquisire in modo tale da limitare le fonti di errore. Al fine di ottenere risultati di qualità, è opportuno che le telecamere siano in numero adeguato al tipo acquisizione che si andrà ad effettuare.

Vari metodi sono stati sviluppati per eseguire l'analisi del movimento con tecnica markerless; uno dei più diffusi è un metodo model-based che sfrutta il matching di un modello del soggetto sui dati sperimentali che assumono la forma tridimensionale di visual hull [25,26,27]. Per visual hull si intende una approssimazione del volume occupato dal soggetto in un dato istante, esso rappresenta una sorta di involucro (letteralmente hull = contenitore) che descrive la superficie corporea del paziente in movimento.

La fase successiva consiste nel ricavare dai risultati ottenuti durante l'acquisizione, le informazioni necessarie per svolgere un'analisi del movimento con lo scopo di indentificare il soggetto d'interesse nella sequenza di immagini; questa fase è chiamata segmentazione. Con il termine segmentazione si indica la sottrazione da ogni canale RGB (figura 2.9) dell'immagine la porzione relativa allo sfondo, in questo modo, al termine della procedura, si ottiene la sola figura del soggetto privata dello sfondo. Successivamente si effettua un'operazione di soglia: se la somma dei valori dei tre canali per un pixel nell'immagine così generata è inferiore a una certa soglia, allora quel pixel viene considerato appartenente allo sfondo.

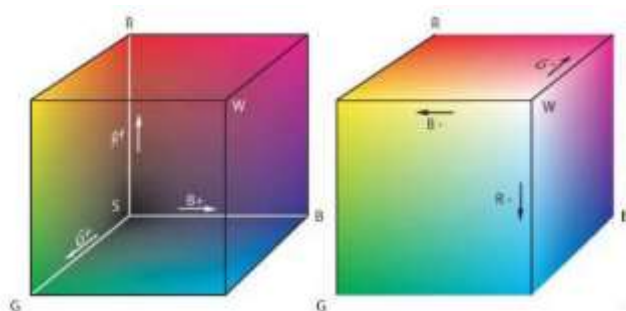


Figura 2.9: rappresentazione dello spazio RGB [20]

Il risultato della segmentazione è una binarizzazione dell'immagine, assegnando ad esempio il valore 0 (nero) a tutti i pixel del background (sfondo) e il valore 1 (bianco) a tutti i pixel del foreground (primo piano). Solo l'informazione relativa alla sagoma del soggetto viene dunque mantenuta. Spesso, al termine della classificazione vengono applicati anche altri algoritmi correttivi che vanno a migliorare il risultato riempiendo eventuali lacune rimaste nella sagoma ed eliminando così pixel spuri erroneamente classificati come foreground.

Le immagini segmentate del soggetto così ottenute vengono chiamate *silhouettes* e sono utili per generare un visual hull (Figura 2.10).



Figura 2.10: Esempio di frame originale dell'acquisizione ed estrazione della silhouette tramite segmentazione [14]

2.4.2 – VISUAL HULL E MODELLO

Note la posizione e l'orientamento delle telecamere, a partire da ogni *silhouette* un generico cono può essere retroproiettato nello spazio; il volume compreso nell'intersezione di tutti i coni costituisce la *visual hull* (Figura 2.11). Per ricostruire la *visual hull* si considerano i voxel dello spazio di lavoro, ovvero elementi di volume che rappresentano un valore di intensità di colore in uno spazio tridimensionale, e si confrontano con le *silhouette* corrispondenti. Nel caso in cui tutte le proiezioni del voxel sul piano delle telecamere appartengono alle silhouette, il voxel fa parte della *visual hull*.

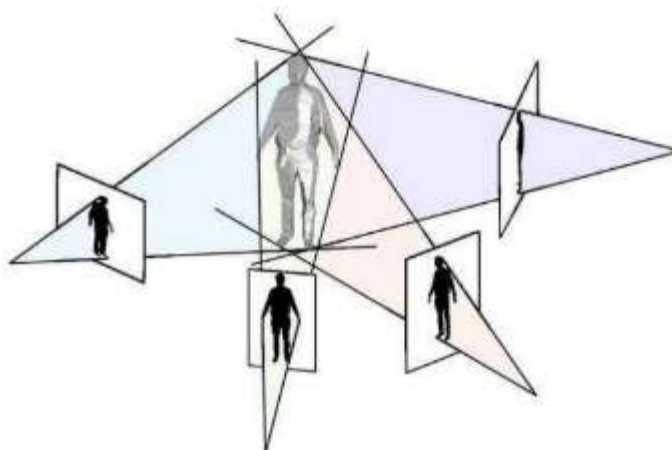


Figura 2.11: Generazione di visualhull a partire da 4 silhouettes [28]

Tuttavia, uno dei principali problemi che affliggono l'approccio mediante *visual hull* è la presenza di volumi fantasma (*phantom volumes*), artefatti che appaiono quando un'area del volume di lavoro è occlusa dalla vista di tutte le telecamere. Per esempio, possono essere registrati quando gli arti del corpo del soggetto occupano una posizione indesiderata in un momento specifico durante l'acquisizione, come in Figura 2.12. È stato provato che un parametro critico nel controllo di questo problema è proprio il numero di telecamere impiegate: un numero maggiore di telecamere permette una ricostruzione più accurata del *visual hull* riducendo la possibile insorgenza di tale artefatto [27].



Figura 2.12: esempio di visualhull affetto da phantom volumes

A questo punto, si rende necessaria la definizione di un modello per il soggetto che guidi l'interpretazione dei dati ottenuti sperimentalmente in modo da poter classificare i segmenti corporei.

Nella realizzazione del modello, l'informazione anatomica è in genere ottenuta tramite una scansione del soggetto mediante un laser scanner: il risultato è un oggetto tridimensionale rappresentante la superficie esterna del soggetto, descritta al calcolatore come mesh poligonale. In genere, una mesh poligonale è espressa da una lista ordinata di punti 3D rappresentanti i vertici dei poligoni e dalla lista di poligoni che indicano come i vertici sono tra loro connessi. Ogni vertice viene univocamente identificato dalla sua posizione nella lista e definito dalle sue tre coordinate rispetto al sistema di riferimento globale, mentre ogni poligono è descritto dagli identificatori dei suoi vertici indicati in senso orario, guardando il poligono dall'esterno.

Le informazioni sulla cinematica possono essere invece rappresentate da un modello articolato del corpo umano. Il modello articolato per l'analisi del movimento di tutto il corpo consiste generalmente in quindici segmenti rigidi, modellanti gli altrettanti segmenti corporei principali (testa, busto, bacino, braccia, avambracci, mani, cosce, gambe, piedi), connessi tra loro tramite articolazioni a sei gradi di libertà (tre traslazioni e tre rotazioni).

Le registrazioni delle telecamere hanno portato alla generazione di *visual hull* consecutive. Tuttavia, questo non è sufficiente per avere un tracciamento affidabile dei movimenti del soggetto, dovuto al fatto che non viene fatta alcuna correlazione tra il tempo e i punti generati nella *visual hull*; inoltre non sappiamo quali punti appartengono a quale arto. Come risultato di queste osservazioni, possiamo dire che la generazione della *visual hull* è utilizzata in questo approccio come una guida di approssimazione che dovrebbe essere migliorata per rispettare una serie di criteri anatomici.

A questo punto del lavoro occorre generare il modello complessivo, specifico per il soggetto in modo da adattarlo ai quindici segmenti rigidi del mesh poligonale specificamente creato sul soggetto, identificando sul mesh i quindici segmenti corporei e le posizioni dei centri articolari. Si riuscirà in questo modo nell'intento di creare un modello specifico per il soggetto e utilizzabile per l'interpretazione dei dati, che contenga al suo interno le informazioni di tipo sia morfologico che cinematico. Si mostra in Figura 2.13 un esempio di mesh triangolare di un soggetto acquisita con un laser scanner e il modello completo della segmentazione in quindici segmenti corporei, specifico per il soggetto dopo l'applicazione dell'algoritmo di generazione automatica del modello [29].

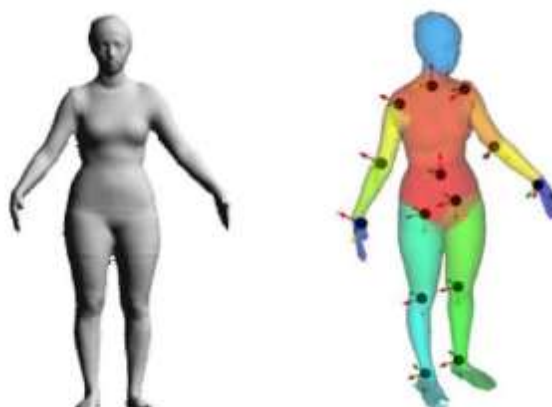


Figura 2.13: rappresentazione della superficie esterna del soggetto con mesh triangolare acquisita con laser scanner (a sinistra) e modello completo del soggetto dopo l'applicazione dell'algoritmo di generazione automatica del modello (a destra) [14]

2.5 – MARKERLESS ED INTELLIGENZA ARTIFICIALE

Negli ultimi anni sono stati sviluppati metodi di intelligenza artificiale che fanno uso di tecniche di deep learning. Grazie a questa tecnica, con una quantità sufficiente di dati, il sistema è in grado di apprendere la rappresentazione corretta e risolvere problemi di apprendimento automatico senza la necessità di un pre-processamento dei dati. Alcuni lavori sviluppati in ambiente subacqueo utilizzano questa tecnologia, come il 3D *silhouette-based* in [30] o la soluzione esposta in [31], ma i risultati sono solo previsioni 2D.

Il fatto di considerare questo tipo di ambiente ricade sulla difficoltà di ottenere un tracker di articolazioni senza marcatori che stima accuratamente le pose naturali nell'ambiente originario.

Nello studio proposto in [31] mediante una videocamera è stato registrato il movimento degli arti inferiori sul piano sagittale. Una rete neurale (deep neural network) è stata “addestrata” utilizzando i dati di diciassette individui, e quindi utilizzata per prevedere le posizioni dei marcatori che approssimano i centri delle articolazioni. Questo tipo di approccio potrebbe facilmente essere modificato consentendone ampie applicazioni in diversi campi, per esempio consentendo l'analisi del movimento senza marcatori da eseguire durante la riabilitazione, l'allenamento o anche la competizione.

Tuttavia, la maggioranza di questi metodi utilizza ancora un approccio marker-based [32, 33, 34].

Capitolo 3

MATERIALI E METODI

Il seguente lavoro, svolto presso il laboratorio di Bioingegneria del Movimento dell'Università di Padova, propone un confronto tra due tecniche di acquisizione per la stima della cinematica articolare del cammino; una tecnica di ricostruzione di coordinate anatomiche marker-based ed una tecnica markerless. Per questo studio sono state analizzate le acquisizioni di camminate di tre soggetti di controllo ed è stato utilizzato un protocollo sperimentale, lo IOR-Gait.

In questo capitolo verranno descritti materiali e metodi utilizzati durante l'intero progetto.

3.1 – SOGGETTI

A questo studio hanno partecipato tre soggetti maschi sani, ai quali è stato chiesto di eseguire una serie di camminate a velocità naturale ciascuna per poter ricavare rispettivamente tre cicli del passo destro e tre cicli del passo sinistro; inoltre è stata eseguita per ciascuno di essi un'acquisizione in posizione statica.

Le caratteristiche dei tre soggetti di controllo presi in considerazione per questo lavoro di tesi vengono riportate nella tabella sottostante.

Caratteristiche	Primo soggetto	Secondo soggetto	Terzo soggetto	MEDIA ± SD
<i>Età [anni]</i>	10	10	10	10 ± 0
<i>Altezza [cm]</i>	117	121	137	125 ± 10,58
<i>Peso [kg]</i>	30	30	30	30 ± 0
<i>Numero di scarpe</i>	30	35	35	33,3 ± 2,89
<i>BMI [kg/m²]</i>	21,92	20,49	15,98	19,46 ± 3,1
<i>Sesso</i>	M	M	M	

Tabella 3.1: dati soggetti

3.2 – STRUMENTAZIONE

Le video acquisizioni sono state realizzate mediante la strumentazione presente nel laboratorio di Bioingegneria del Movimento dell'Università di Padova (Figura 3.2), ossia un sistema stereofotogrammetrico a sei telecamere (60-120 Hz – BTS S.r.l., Padova, Italia), ed infine un set di markers applicati sui soggetti da analizzare. Il sistema è stato appositamente calibrato prima delle acquisizioni come esposto nel paragrafo 2.3.2.



Figura 3.2: Laboratorio di Bioingegneria del movimento presso l'Università di Padova

3.3 – PROCEDURA MARKER-BASED

Per poter ricavare in modo adeguato i dati relativi al tipo di esigenze di studio, è opportuno adottare un protocollo anatomico adatto. Per questo progetto è stato adottato il protocollo *FullBody* IOR-Gait, versione modificata dallo IOR-Gait al quale è stato aggiunto il tronco come in Sawacha et al 2009.

Come descritto precedentemente nel paragrafo 1.4, un protocollo consente di acquisire, elaborare ed analizzare i dati permettendo dunque di andare a valutare i parametri cinematici e dinamici dei vari segmenti corporei. Il protocollo *FullBody*, descritto al paragrafo 1.4.3, prevede l'applicazione di 24 marker direttamente su punti di repere anatomici mentre altri 6 vengono impiegati per la calibrazione.

Il marker-set considerato è:

- Tronco: acromion destro e sinistro (RA e LA), settima vertebra cervicale (C7), quinta vertebra lombare (L5);
- Pelvi: spina iliaca anteriore destra e sinistra (RASIS e LASIS), spina iliaca posteriore destra e sinistra (RPSIS e LPSIS);
- Coscia: gran trocantere destro e sinistro (RGT e LGT), epicondilo laterale destro e sinistro (RLE e LLE), epicondilo mediale destro e sinistro (RME e LME);
- Gamba: tuberosità tibiale destra e sinistra (RTT e LTT), testa della fibula destra e sinistra (RHF e LHF), malleolo laterale destro e sinistro (RME e RMM);
- Piede: calcagno destro e sinistro (RCA e LCA), prima testa metatarsale destra e sinistra (RFM e LFM), seconda testa metatarsale destra e sinistra (RSM e LSM), quinta testa metatarsale (RVM e LVM).

Dopo aver applicato i markers, prima di poter procedere con l'acquisizione vera e propria, è stato necessario operare una calibrazione anatomica dei punti di repere durante una fase di statica (Figura 3.3), ossia mentre il soggetto rimane fermo in piedi per alcuni secondi. Questa procedura di calibrazione permette di ottenere le coordinate dei marker posti sul soggetto nel sistema di riferimento locale associato a ciascun segmento, inoltre consente successivamente di analizzare in dinamica la deviazione degli angoli rispetto alla statica.

3.4 – PROCEDURA MARKERLESS

A partire dalle stesse acquisizioni di cui si è parlato nel paragrafo precedente, si sono acquisite nuove sequenze e si è proceduto con l'elaborazione markerless dei video, puntando ad estrarre i parametri relativi al movimento a partire da determinate caratteristiche morfologiche del soggetto in analisi corrispondenti ai punti in cui si andrebbe a collocare i marker, come mostrato in figura 3.3.

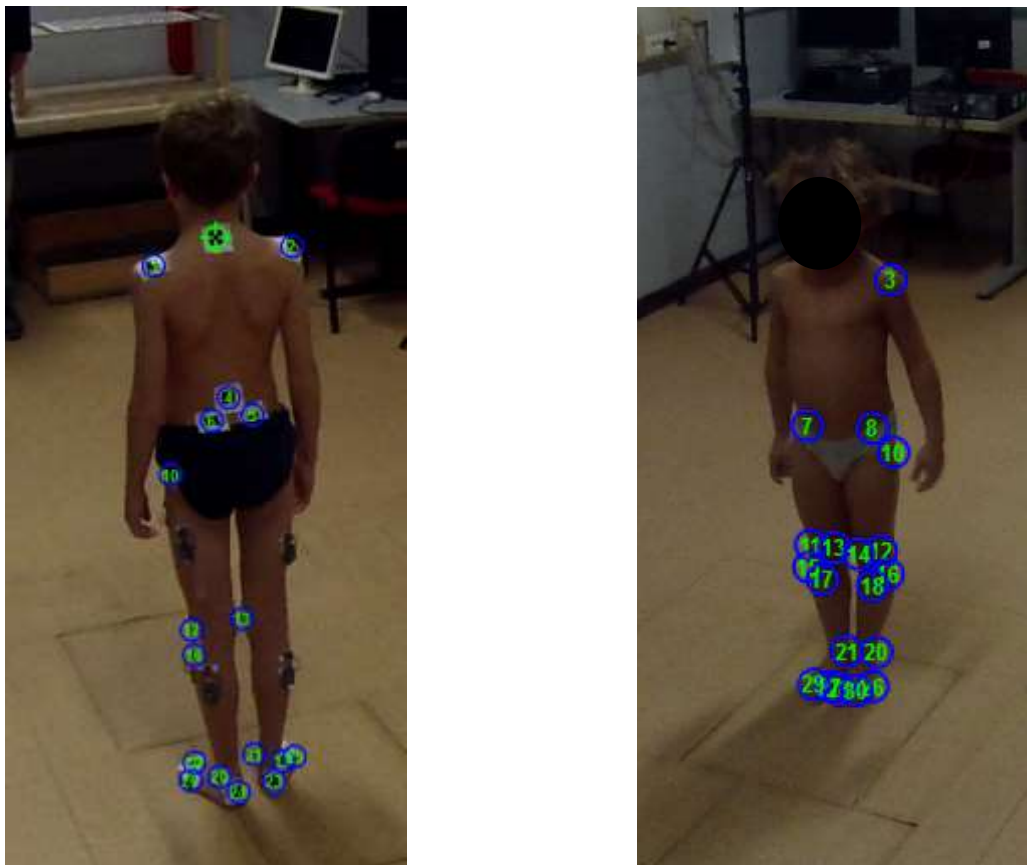


Figura 3.3: a sinistra un esempio di calibrazione statica marker-based; a destra un esempio di calibrazione statica markerless.

3.5 – VIDEO PROCESSING

La ricostruzione delle coordinate 3D dei marker si è ottenuta con il software per il tracking di feature, *TrackOnField* (BBSof S.r.l.), considerando le sequenze dei fotogrammi che rappresentano un ciclo completo del passo. Questo programma è dotato di un algoritmo di tracking automatico, descritto nel paragrafo 2.3.1.1, il quale individua in modo autonomo la posizione dei marker sul soggetto di studio; al fotogramma successivo propone di volta in volta la nuova posizione, ed è compito dell'operatore decidere se accettarla o se correggerla manualmente in caso di errore. Tuttavia, nonostante la funzione svolta, alcune volte è stato necessario disattivare l'algoritmo in quanto, per le ragioni legate alla non visibilità dei marker esposte nel paragrafo 2.3.1, in questi casi venivano rilevate posizioni non coerenti rispetto ai movimenti del corpo.

All'avvio del programma (vedi Figura 3.4) è necessario fornire alcune informazioni e settare dei parametri, specificando inoltre il frame iniziale e finale della sequenza da analizzare. Successivamente, selezionando la finestra che mostra il primo frame del video da analizzare, si specificano le posizioni delle *feature* della sequenza scelta. Ad ogni punto è stato fatto corrispondere un numero identificativo in modo tale da collegare il marker al corrispondente punto descritto dal protocollo IOR-Gait, come mostrato in Figura 3.5.

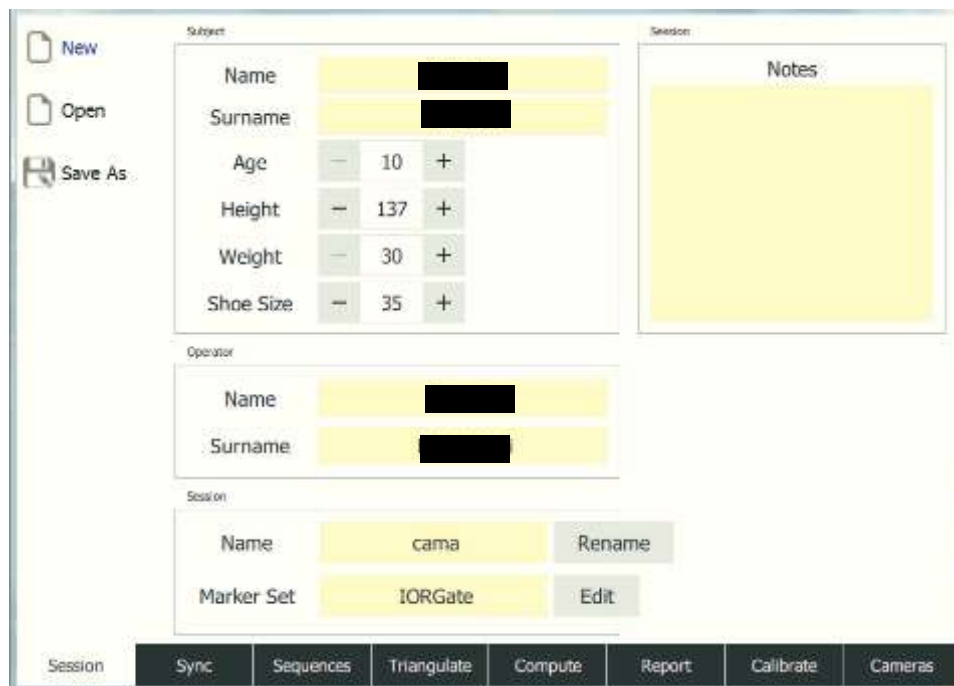


Figura 3.4: esempio di schermata iniziale di *TrackOnField*



Figura 3.5: esempio di feature tracking con *TrackOnField*

Al termine della procedura si ottengono in output delle tabelle, dette *feature table*, contenenti le coordinate 2D digitalizzate dei marker appartenenti al protocollo adottato. Queste feature table sono poi state utilizzate durante la fase di triangolazione, al termine della quale viene prodotto un file .c3d.

Un ulteriore miglioramento dell'analisi e ricostruzione della corretta posizione dei marker, in particolar modo nei frame nei quali non è stato possibile individuarli, è stato reso possibile sfruttando l'interpolazione, consentendo quindi di ricostruire la posizione dei markers mancanti.

Per verificare che la triangolazione e la disposizione dei punti anatomici sia avvenuta correttamente, è possibile visualizzare tali punti, presenti nei file .c3d, tramite il programma Mokka; un software gratuito che consente grazie alla sua interfaccia grafica di visualizzare ogni marker tridimensionalmente da ogni sua angolazione, come presentato in Figura 3.6. E' stato progettato appositamente per applicazioni di analisi del movimento, come *gait analysis*, analisi biomeccanica e analisi sportiva. È un software compatibile con molti tipi di formato presenti in biomeccanica, anche di software proprietari (BTS, Vicon, Motion Analysis, Ascension Technology Corporation etc).

In conclusione, utilizzando i dati contenuti nei file .c3d precedentemente verificati, si procede con il calcolo degli angoli articolari mediante una Toolbox scritta in codice in codice MatLab disponibile presso il BiomovLab: Processing Dynamics V3.0, attraverso la quale è stato possibile ricostruire gli angoli articolari, necessari al confronto di tali tecniche, visibili grazie a dei plot contenenti i grafici utili.

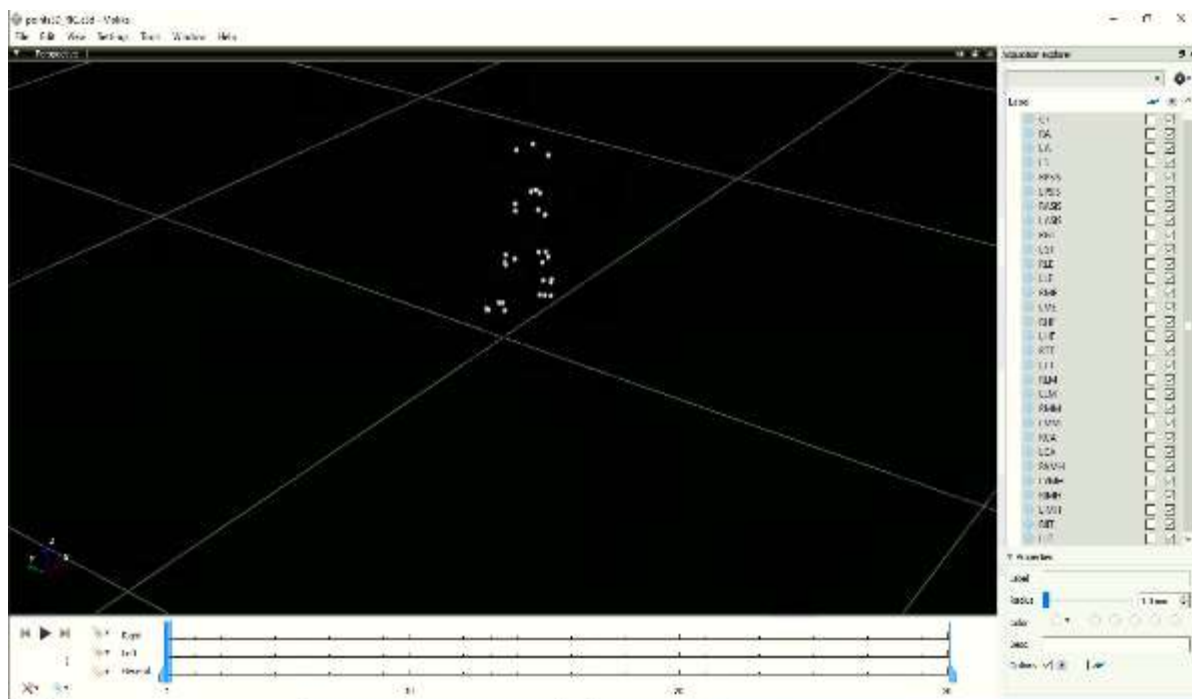


Figura 3.6: esempio di visualizzazione dei markers tramite il software Mokka

3.6 – INDICI DI CONFRONTO

L'obiettivo di questa tesi è quello di effettuare il confronto tra la ricostruzione di coordinate anatomiche ottenuta mediante un protocollo marker-based e una tecnica di analisi markerless. Per ottenere dei risultati che mostrano le differenze tra le due tecniche, sono utilizzati degli indici: *Root Mean Square Error* (RMSE) e *Coefficient of Multiple Correlation* (CMC).

Root Mean Squared Error (RMSE): il valore RMSE (errore quadratico medio) è una misura di errore assoluta in cui le deviazioni vengono elevate al quadrato per evitare che valori positivi e negativi possano annullarsi l'uno con l'altro; indica dunque la discrepanza quadratica media che si verifica sia tra valori dei dati osservati ed i valori dei dati stimati, sia tra due osservazioni fatte con diversi metodi. Con questa misura, inoltre, gli errori di valore maggiore vengono amplificati, caratteristica che può agevolare l'eliminazione dei metodi che presentano gli errori più significativi.

Coefficient of Multiple Correlation (CMC): è una misura numerica del grado di adattamento di un modello di regressione lineare a un insieme di dati. Tecnicamente parlando, è il semplice coefficiente di correlazione per i valori delle variabili dipendenti e valori previsti che si ottengono con la regressione lineare multipla dei minimi quadrati. Questo fornisce una misura statistica del grado di similarità complessiva delle forme d'onda risultanti dai dati delle tecniche utilizzate. Come valore di soglia in questo progetto di tesi è stato assunto 0.7, ovvero vengono considerate molto simili tra loro due curve che superano questo valore.

Capitolo 4

DISCUSSIONE DEI RISULTATI

In questo capitolo saranno mostrati i risultati ottenuti sulle prove eseguite in laboratorio, come illustrato nel capitolo precedente. L'analisi di tali dati sarà condotta per mettere in luce le differenze risultanti tra le due tecniche adottate confrontandole ulteriormente con la stereofotogrammetria, caratterizzata da un data set raccolto da vari progetti eseguiti presso il BiomovLab nel corso degli anni. Verranno quindi riportati i grafici più significativi raffiguranti andamenti di particolare interesse. Il materiale presente in questo capitolo, si riferisce agli angoli articolari dei cicli del passo analizzati da uno stesso operatore, precisamente tre cicli destri e tre cicli sinistri per ciascuno dei soggetti.

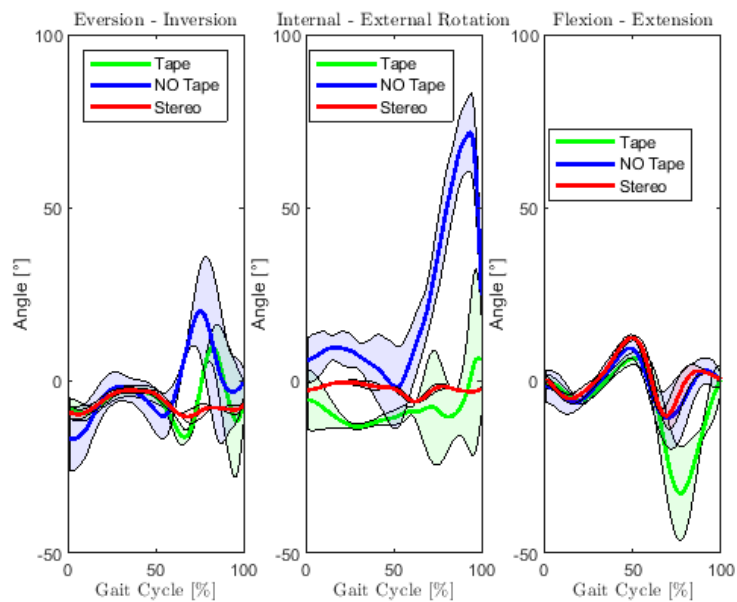


Figura 4.1: LEFT ANKLE – CAVIGLIA SINISTRA. Confronto angoli articolari tra le tecniche utilizzate e la stereofotogrammetria

In **figura 4.1** sono rappresentate la media e la rispettiva deviazione standard delle tre tecniche messe a confronto tra loro, in riferimento ai movimenti della caviglia sinistra. Si nota subito un andamento inizialmente molto simile alla stereofotogrammetria fino a metà movimento, per quanto riguarda il movimento articolare di *eversione-inversione* della caviglia.

Nella fase finale del movimento, ambo le tecniche si discostano dalla fascia rossa di riferimento (stereofotogrammetria). Questo particolare andamento apparentemente errato, in realtà, dopo una successiva analisi nel dettaglio dei singoli trial, si è constatato seguire correttamente il movimento fisiologico del soggetto preso in esame. I trial esaminati, pur riferendosi allo stesso soggetto, possono comunque prevedere un grado di variabilità accettabile intersoggetto. Per quanto riguarda i movimenti di *intra rotazione-extra rotazione*, in questo caso è facile vedere che la tecnica senza i marcatori risulti molto meno correlata alla stereofotogrammetria a differenza della tecnica tape, ma, nonostante il pattern segua l'andamento correttamente, i valori degli angoli in gradi risultano eccessivi. Una considerazione interessante emerge se si osserva la deviazione standard. Nella tecnica no tape risulta essere generalmente minore rispetto alla tecnica che sfrutta i markers, questo è dovuto ad una leggera variabilità del movimento di camminata nella fase di tracking di quest'ultima, come illustrato in **figura 4.2**.

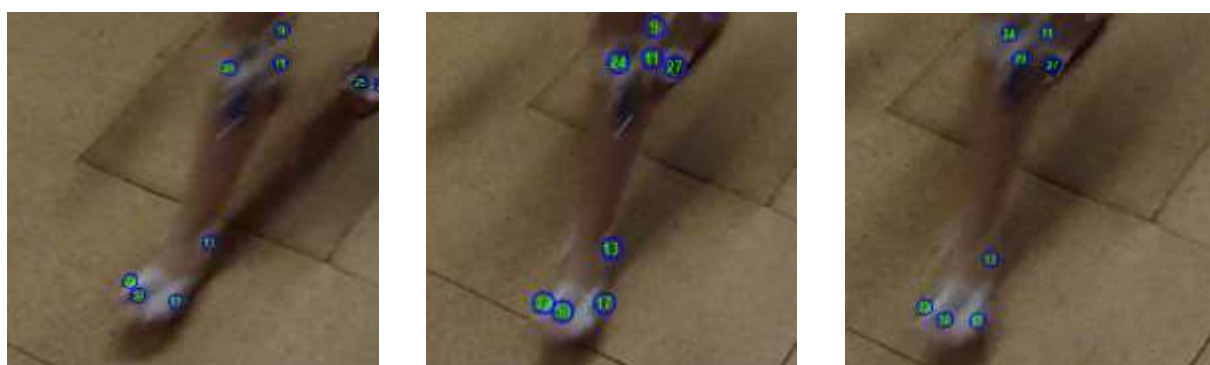


Figura 4.2: movimento di intra-extra rotazione della caviglia sinistra

Nel movimento di *flessione-estensione* (figura 4.1) le tecniche rilevano in modo molto fedele il gesto motorio. Anche qui notiamo uno scostamento da parte della tecnica tape, giustificato dal movimento accentuato del soggetto che eseguiva la camminata, come mostrato nella successiva **figura 4.3**, infatti si nota come il piede risulti più flesso nel frame relativo primo ciclo del passo illustrato a sinistra, rispetto agli altri due cicli.

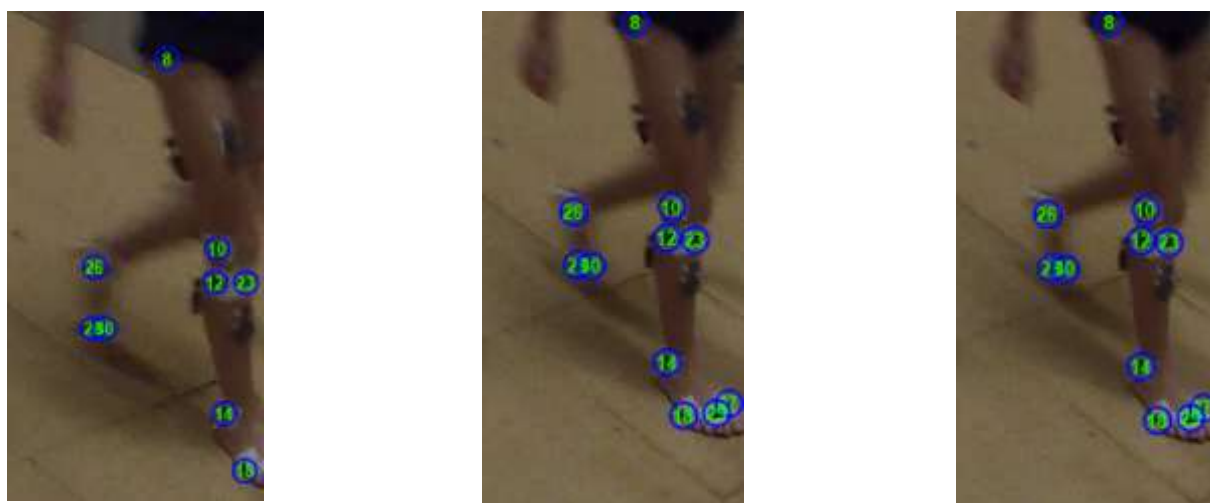


Figura 4.3: movimento di flesso-estensione della caviglia sinistra

Sul plot rappresentato in **figura 4.4** invece troviamo i grafici relativi alla caviglia destra. Nel movimento di *eversione-inversione*, le due tecniche, tape e no tape, sono abbastanza correlate tra loro, rilevando in modo corretto il movimento del cammino. Si può vedere anche qui com'è giusto aspettarsi dal metodo utilizzato, che la tecnica no tape presenta nel tratto finale una maggior variabilità, dovuta giustamente dall'assenza dei markers che fungono da guida, come confermato dai singoli trial visibili in **figura 4.5**, considerazione valida anche per i successivi grafici. In accordo con la letteratura gli angoli che risultano più affidabili se ottenuti con le tecniche video in assenza di tape sono quelli di flesso-estensione, mentre si discostano maggiormente quelli di intra-extra rotazione, seguiti da quelli di inversione-eversione.

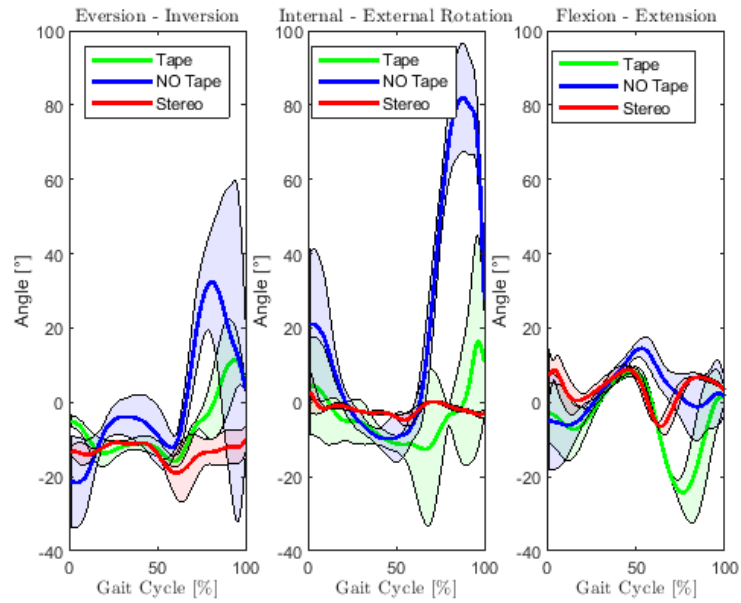


Figura 4.4: RIGHT ANKLE – CAVIGLIA DESTRA. Confronto angoli articolari tra le tecniche utilizzate e la stereofotogrammetria

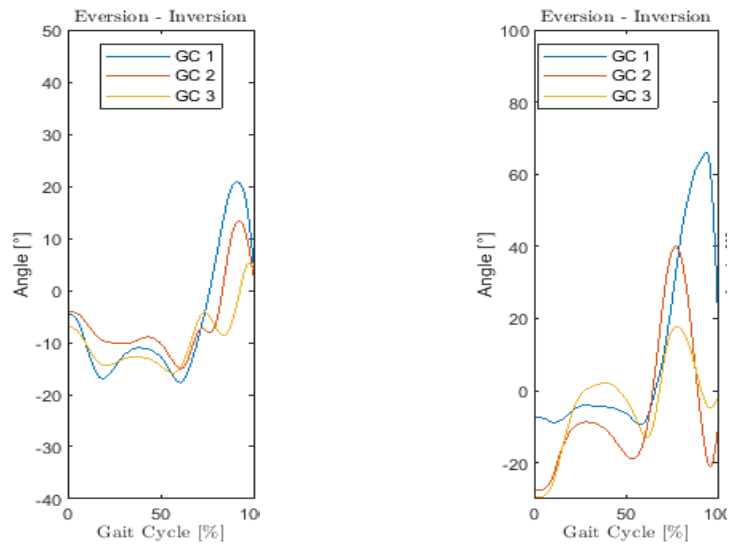


Figura 4.5: *eversione-inversione* TAPE (sinistra) e NOTAPE (destra)

Per quanto riguarda la tecnica no tape nella *intra-extra rotazione*, il pattern di movimento viene rilevato correttamente, nonostante i valori in gradi non risultino fisiologici. Lo stesso discorso fatto precedentemente per la caviglia sinistra lo possiamo estendere anche alla caviglia destra. Ulteriore considerazione da fare, valida anche per la caviglia sinistra, riguarda il tracking della tecnica no tape; la minor variabilità visibile dal grafico *dell'intra-extra rotazione* deriva dall'individuazione dei punti di reperi da parte dell'operatore, il che attribuisce una forte componente soggettiva al metodo, a scapito però della precisione. In **figura 4.6** è mostrato un esempio a riguardo, dove emerge come la tecnica tape possa sembrare a prima vista imprecisa, al contrario, risulta ricalcare bene il movimento. È importante precisare che la qualità dell'immagine spesso non permette un rilevamento semplice dei tape e ancor meno dei punti di reperi in assenza di tape.

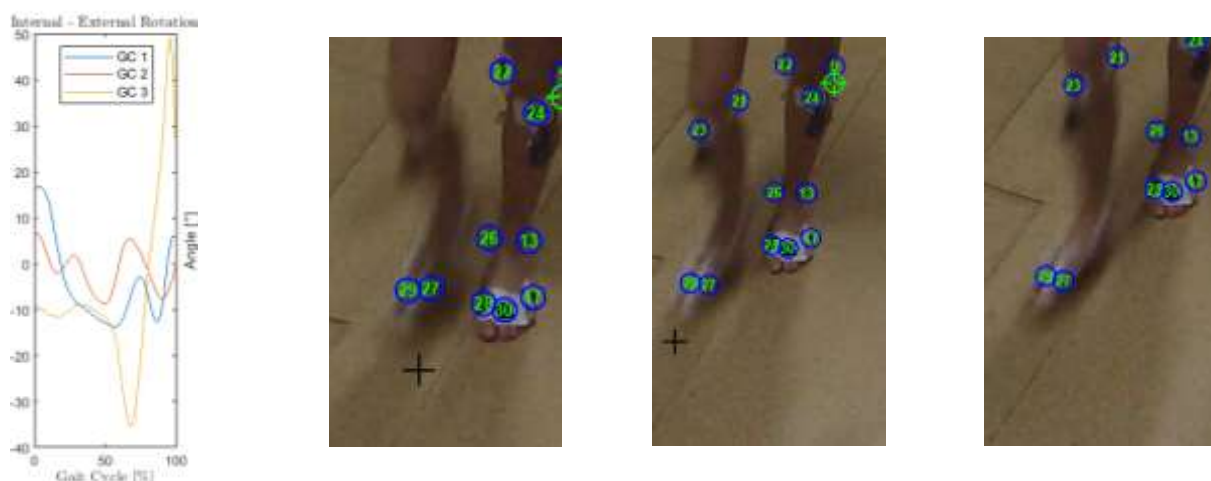


Figura 4.6: a sinistra singoli trials tecnica TAPE, a destra frame corrispondenti alla traiettoria del terzo ciclo del passo che presenta picchi notevoli

In questa sequenza di frame si può vedere come inizialmente il piede venga intra ruotato per poi proseguire sempre più in posizione di extrarotazione. La *flessione-estensione* viene rilevata discretamente da entrambe le tecniche, dove, come detto per la caviglia sinistra, anche nella caviglia destra è presente una maggior precisione nella tecnica tape, confermata da una minor deviazione standard e da un andamento simile alla stereofotogrammetria.

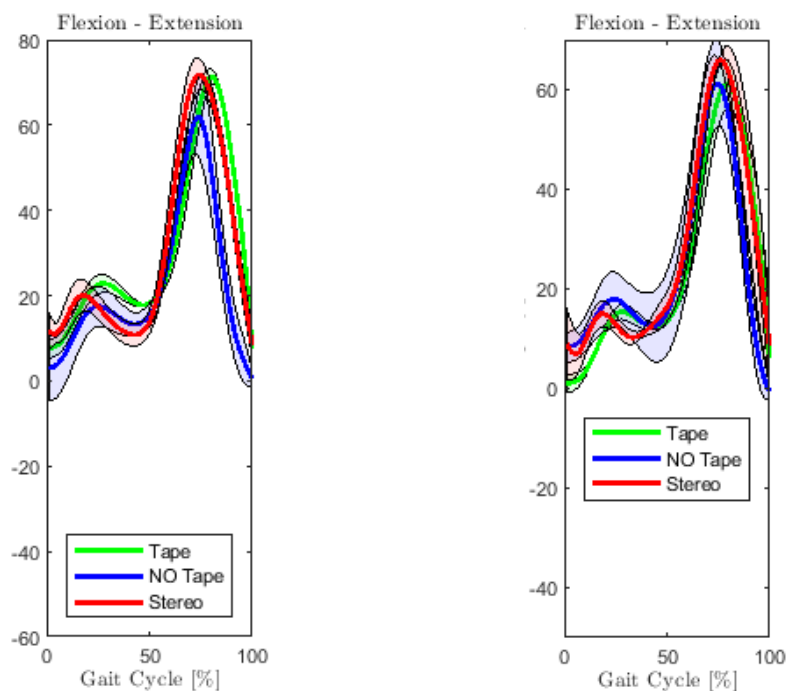


Figura 4.7: RIGHT KNEE / LEFT KNEE – GINOCCHIO DESTRO / GINOCCHIO SINISTRO

Troviamo in **figura 4.7** i movimenti di *flesso-estensione* delle ginocchia, seguiti molto bene dalle tecniche utilizzate per questo progetto. Anche qui si nota una lieve variabilità nella tecnica senza marcatori com'è giusto aspettarsi dal metodo. Per convenzione, si considera solamente il movimento di *flesso-estensione* dato che le tecniche attualmente in uso non sono in grado di rilevare in modo accettabile il movimento sugli altri piani (coronale e trasverso).

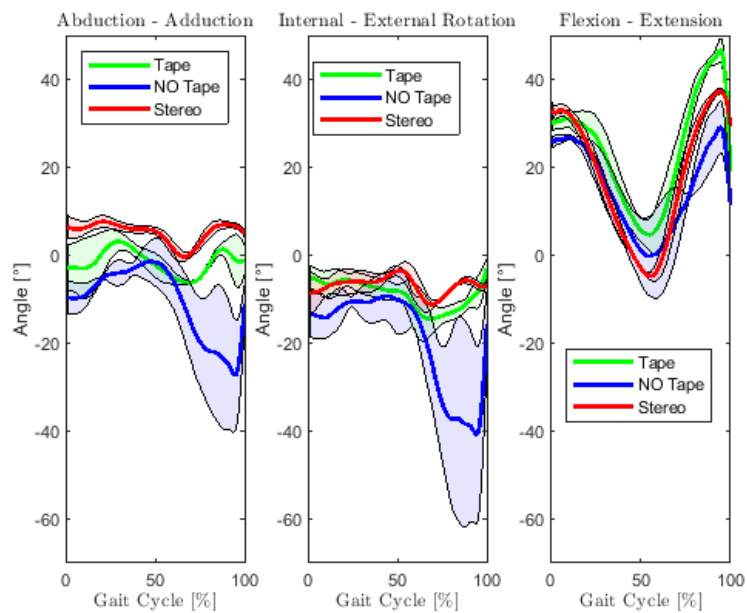


Figura 4.8: LEFT HIP – ANCA SINISTRA. Confronto angoli articolari tra le tecniche utilizzate e la stereofotogrammetria

Nei grafici riportati qui sopra (**figura 4.8**) vengono esaminati i movimenti relativi all'anca sinistra. Salta subito all'occhio un andamento migliore e fedele, più simile alla stereofotogrammetria da parte della tecnica tape, con un picco accentuato da parte della tecnica no tape. A causa del vestiario e di una forte componente di soggettività nella tecnica markerless, risulta difficile stimare la corretta posizione dei punti di repere. È riportato un esempio in **figura 4.9**, dove si evince tale difficoltà.



Figura 4.9: frame relativi all'anca sinistra; visibile difficoltà causata da vestiario

Per quanto riguarda invece la flessione-estensione c'è un buon rilevamento da parte di entrambe le tecniche, ancora una volta però viene confermata la precisione della tecnica tape visibile anche da una minor deviazione standard rispetto alla tecnica concorrente.

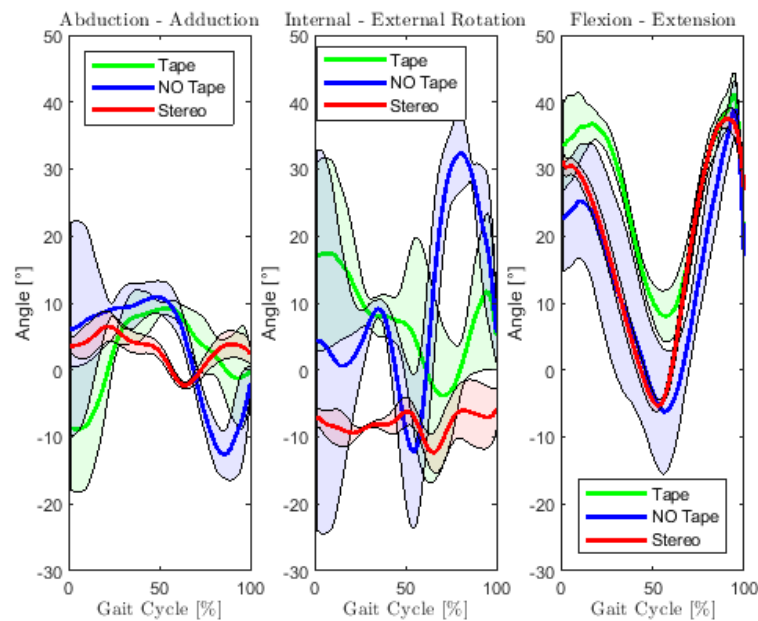


Figura 4.10: : RIGHT HIP – ANCA DESTRA. Confronto angoli articolari tra le tecniche utilizzate e la stereofotogrammetria

In **figura 4.10** si evince subito una fortissima variabilità dei risultati, soprattutto nella tecnica notape per gli angoli di intra-extra rotazione e ab-adduzione. Tale scompenso si può ricondurre alle problematiche riscontrate precedentemente, ovvero assenza dei marker, presenza del vestiario ed anche l'inquadratura delle telecamere, in quanto alcune sequenze scelte vengono prese da angolazioni poco apprezzabili per poter individuare in modo chiaro i punti di repere. In questo caso il soggetto in analisi ha eseguito una camminata a passo abbastanza sostenuto, altro fattore che induce all'errore durante la procedura di tracking. In accordo con la letteratura anche in questo caso gli angoli sul piano sagittale (Flessione-estensione) risultano quelli ricostruiti in modo affidabile con le tecniche video diverse dalla stereofotogrammetria.

Come emerge dalla sequenza di frame in **figura 4.11**, si può vedere una variabilità causata da vestiario e movimento dell'arto superiore che va a nascondere il punto di repere anatomico creando ulteriore confusione nel corretto posizionamento.



Figura 4.11: : frame relativi all'anca destra; visibile difficoltà causata da arto superiore destro

Nel plot sottostante (**figura 4.12**) si vedono gli andamenti delle traiettorie relative al bacino. Si può osservare come la tecnica tape riesca meglio a seguire l'andamento della stereofotogrammetria oltre ad imitare in modo fedele il gesto motorio a differenza del metodo notape. In quest'ultimo, infatti, il rilevamento dei punti anatomici da parte dell'operatore presenta una notevole variabilità come si può osservare dalla deviazione standard dei grafici legati ai movimenti del bacino. Come riportato ad esempio nei frame in **figura 4.13**, è marcata la differenza tra le due tecniche durante il tracking dei punti di repere. Si osserva come il triangolo delineato dai punti (5, 20, 6) si deformi notevolmente in due frame successivi nel metodo markerless mentre nel metodo marker-based risulti stabile. Anche in questo caso la ricostruzione più affidabile è sul piano sagittale.

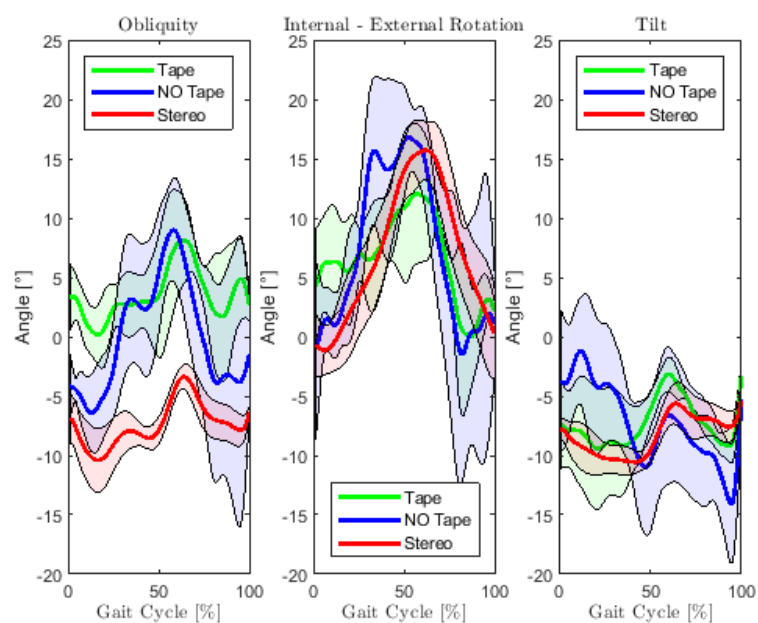


Figura 4.12: : PELVIS – BACINO. Confronto angoli articolari tra le tecniche utilizzate e la stereofotogrammetria

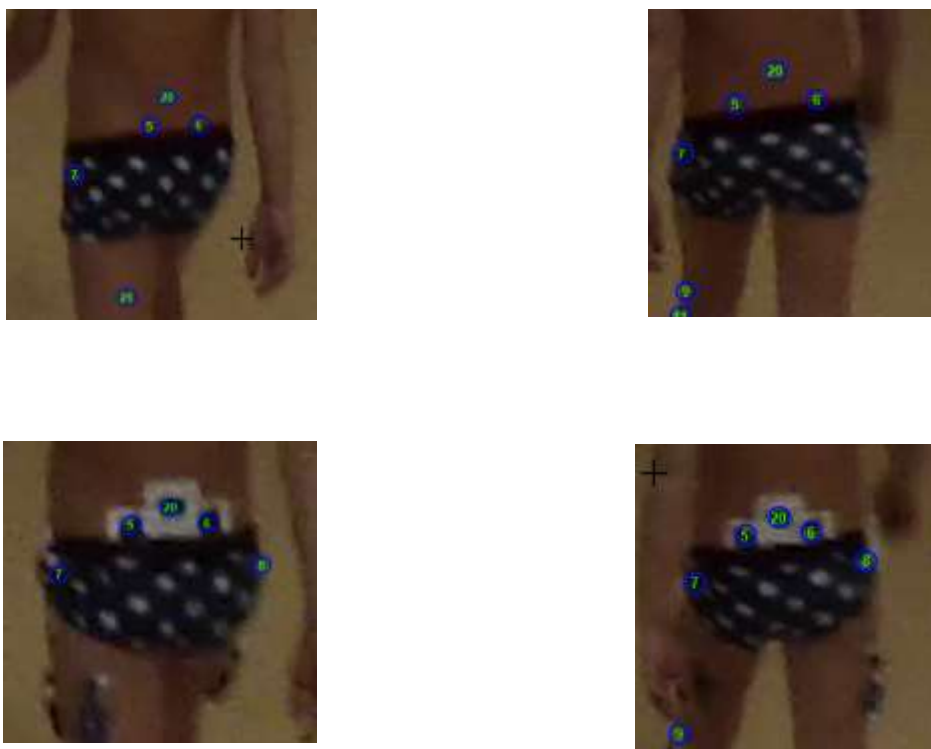


Figura 4.13: : variabilità posizionamento marker tra tecnica NOTAPE sopra e tecnica TAPE sotto

Un aspetto interessante da prendere in esame riguarda la similarità tra le due curve, i cui parametri per la valutazione vengono forniti dal coefficiente di correlazione multipla (CMC) e dall'errore quadratico medio (RMSD). Il primo ci consente di valutare la correlazione tra le curve, mentre con il secondo parametro si va a valutare l'errore tra le traiettorie, calcolandovi massimo, minimo, media e deviazione standard. Questa analisi si può estendere ad ogni traiettoria di ogni soggetto esaminato per questa tesi confrontando ulteriormente le tecniche tra loro, ma diverrebbe un procedimento troppo laborioso. A tal proposito vengono inseriti i risultati più pertinenti.

	LHIP			RHIP			LKNEE	RKNEE
	AA	InEx	FE	AA	InEx	FE	FE	FE
MEAN RMSD (DEG)	12,9551	17,6205	8,4725	8,2587	7,4819	23,5363	17,2012	15,0416
SD RMSD (DEG)	7,2209	13,3335	6,5731	2,6123	3,4841	6,4553	6,1164	15,4182
MIN RMSD (DEG)	1,3211	6,3079	1,1306	3,1015	2,5921	6,6848	5,3099	1,8792
MAX RMSD (DEG)	24,9551	42,4684	21,6105	12,8581	14,1801	30,4236	25,7654	46,3933

Tabella 4.14: : RMSD anca e ginocchio destra e sinistra: sono riportati la media (Mean), deviazione standard (SD), minimo (Min) e massimo (Max). I movimenti presi in considerazione sono: Abduzione- Adduzione (AA), Flesso-Estensione (FE), Intra- Extra Rotazione (InEx).

	RANKLE			LANKLE			PELVIS		
	InEv	InEx	FE	InEv	InEx	FE	OB	InEx	Tilt
MEAN RMSD (DEG)	16,1938	14,6822	11,0311	9,5481	8,1560	14,4583	4,8168	9,7396	4,6263
SD RMSD (DEG)	14,4472	7,7986	4,7484	7,8397	3,2636	11,9658	1,9887	2,7241	2,6757
MIN RMSD (DEG)	2,7798	6,1519	1,2005	0,7905	2,9321	3,9810	0,8382	4,3732	0,0622
MAX RMSD (DEG)	46,4308	34,7659	21,5764	27,4623	13,8229	42,1864	7,5901	14,3807	11,0797

Tabella 4.15: : RMSD caviglia destra e sinistra e bacino: sono riportati la media (Mean), deviazione standard (SD), minimo (Min) e massimo (Max). I movimenti presi in considerazione sono: Inv-Eversione (InEv), Flesso-Estensione (FE), Intra- Extra Rotazione (InEx), Obliquità (OB), Tilt Pelvico (Tilt).

I confronti riportati nelle tabelle 4.14 e 4.15 vengono eseguiti tra le due tecniche utilizzate (TAPE e NOTAPE). I valori inseriti si riferiscono agli andamenti visibili sui grafici precedentemente commentati in quanto presentano comportamenti più interessanti rispetto agli altri grafici degli altri soggetti. Si osservano valori medi dell’RMSD maggiori in corrispondenza delle più accentuate discrepanze nel tracking commentate in precedenza, come ad esempio in figura 4.8 nel movimento di intra-extra rotazione, dimostrando le difficoltà già descritta durante l’analisi.

Per quanto riguarda il CMC invece, vale lo stesso principio adottato per l’indice RMSD, illustrando quindi solamente alcuni dei principali aspetti che presentano un comportamento interessante. Notiamo come alcune traiettorie delle tecniche abbiano un buon grado complessivo di similarità nonostante la diversità del metodo di rilevamento, come si evince in tabella 4.16.

Articolazione	Movimento articolare	CMC
LANKLE	Eversione - Inversione	0.8330
	Intrarotazione - Extrarotazione	0.95563
	Flesso-Estensione	0.7848
RANKLE	Eversione - Inversione	0.76384
	Intrarotazione - Extrarotazione	0.73809
	Flesso-Estensione	---
RKNEE	Flesso-Estensione	0.98074
LKNEE	Flesso-Estensione	0.98137
LHIP	Abduzione - Adduzione	0.97742
	Intrarotazione - Extrarotazione	0.92345
	Flessione - Estensione	0.95318
RHIP	Abduzione - Adduzione	0.97826
	Intrarotazione - Extrarotazione	0.9528
	Flessione - Estensione	0.9483
PELVIS	Obliquità	0.93438
	Intrarotazione - Extrarotazione	0.84899
	Tilt pelvico	0.78927

Tabella 4.16: : valori CMC articolazioni

Si osserva dai valori del coefficiente di correlazione multipla la similarità massima complessiva raggiunta tra le varie articolazioni. La caviglia sinistra presenta valori maggiori a rispetto a quella destra, dove non risulta un valore del CMC sufficiente a stabilirne una buona correlazione (CMC maggiore di 0.7 come descritto nel paragrafo 3.6). Questa differenza la si deve in parte anche al fatto del posizionamento delle telecamere, in quanto non sono situate in corrispondenza dei quattro angoli della stanza, bensì sono leggermente spostate. Osservando invece le ginocchia, si può vedere una notevole correlazione complessiva tra le due tecniche. Si può dire lo stesso di anca destra, anca sinistra e bacino, in quanto si raggiungono valori buoni abbastanza superiori alla soglia; il tilt pelvico invece evidenzia alcune difficoltà nella replicabilità dei risultati come dimostra il valore al di poco sopra la soglia.

Capitolo 5

CONCLUSIONI

Lo scopo di questa tesi è stato quello di confrontare a livello cinematico due diverse tecniche di acquisizione ed analisi del movimento umano; una utilizza appositi marcatori (marker-based) sulla pelle, contraddistinta da una maggiore accuratezza delle misure, ma con una lunga fase di preparazione del soggetto, e una che sfrutta la sola ripresa di immagini da telecamere (detta quindi markerless), risultando dunque più veloce, ma presentando ancora oggi delle limitazioni dal punto di vista della precisione, essendo oltretutto affetta da una maggiore componente soggettiva legata a chi esegue la procedura di tracking. Il metodo sviluppato consiste nel confronto degli angoli articolari del ciclo del passo di tre soggetti, ricavati con le tecniche precedentemente descritte, con gli angoli articolari della stereofotogrammetria. Procedendo in questo modo è stato possibile evidenziare le principali divergenze tra le due metodologie utilizzate, mettendo anche in luce quale delle due risulti più fedele e riproducibile.

Considerando che, essendo esseri umani e che quindi la camminata non avviene mai in modo uguale e preciso ogni volta, ma al massimo può essere simile, possiamo affermare quindi, sulla base di ciò che abbiamo analizzato, che sino ad ora la tecnica marker-based è quella che ha dato risultati migliori. Nonostante alcuni movimenti articolari siano stati fedelmente riprodotti anche dalla tecnica markerless, si nota una precisione minore ed una variabilità maggiore; la tecnica appena citata è ulteriormente penalizzata dalla scarsa qualità di immagine, causata a volte anche da un movimento troppo esasperato da parte del soggetto in esame.

La possibilità di adottare la tecnica di acquisizione del movimento markerless nel campo dell'analisi dell'andatura è molto vantaggiosa se consideriamo che può fornire una riduzione del tempo di preparazione dei soggetti e l'assenza di marcatori che potrebbero modificare la naturalezza del movimento del soggetto.

Tuttavia, se si considera che nella pratica reale, l'analisi del cammino è utilizzata principalmente in ospedali e cliniche ortopediche, il metodo qui sviluppato può essere usato per valutare l'affidabilità e l'interpretabilità della cinematica articolare senza marcatori applicando una precisa determinazione dell'orientamento dei fotogrammi anatomici.

È auspicabile, nel corso degli anni, che si possa arrivare a migliorare la tracciabilità riguardante l'abd – adduzione ed intra – extra rotazione del ginocchio, in modo tale da riprodurre in modo esatto la naturale cinematica articolare di ogni soggetto che si andrà ad analizzare, favorendo quindi ulteriori analisi legate al movimento del ginocchio stesso. Oltre a ciò, è facile pensare ad un netto miglioramento complessivo delle tecniche di tracking presenti al giorno d'oggi, in termini di affidabilità, invasività e precisione, garantendo dunque un minor tempo di preparazione del soggetto con un conseguente minor impiego di materiale applicabile sul paziente ed una maggiore efficacia nell'elaborazione dei dati.

Una problematica che potrebbe essere ulteriormente migliorata è quella legata alla precisione nel reperire i punti anatomici inerenti ad anca e bacino, facendo indossare un vestiario più adatto alle misurazioni, ad esempio evitando indumenti poco aderenti.

BIBLIOGRAFIA

- [1] <https://www.policlinicoitalia.it/laboratorio-analisi-movimento/>
- [2] Università di Cagliari, Dipartimento di Ingegneria Meccanica, Chimica e dei Materiali, Stereofotogrammetria elettronica per l'analisi quantitativa del movimento
- [3] Corso di Bioingegneria del movimento e riabilitazione, docente Ing. Zimi Sawacha – 'Lezione 7 – Protocolli.' (A.A. 2019-2020)
- [4] A. Leardini, Z. Sawacha, G. Paolini, S. Inghosso, R. Nativo, M. G. Benedetti. 'A new anatomically based protocol for gait analysis in children.' 2007, Gait and posture, Vol. 26, p. 560-571.
- [5] <https://www.lascienzainpalestra.it/caviglia-anatomia-e-biomeccanica/>
- [6] <https://www.riabilitazionecampania.it/dislivello-del-bacino/>
- [7] J. Bray, Markerless Based Human Motion Capture: A Survey, Master's Thesis - Department Systems Engineering Brunel University, 2001
- [8] D. Roetenberg, H. Luinge, P. Slycke, Xsens MVN: Full 6DOF Human Motion Tracking Using Miniature Inertial Sensors, Xsens Technologies, 2013
- [9] M. Donati, 3D Reconstruction of The Human Skeleton During Motion, Ph.D. Thesis - University of Bologna, 2006
- [10] R. Parent, Computer Animation: Algorithms and Techniques, Morgan Kaufmann, 2001
- [11] L. Chiari, U. Della Croce, A. Leardini, A. Cappozzo, Human Movement Analysis Using Stereophotogrammetry - Part 2: Instrumental Errors, Gait & Posture, 2004
- [12] U. Della Croce, A. Leardini, L. Chiari, e A. Cappozzo, «Human movement analysis using stereophotogrammetry: Part 4: assessment of anatomical landmark misplacement and its effects on joint kinematics», Gait & Posture, vol 21, n° 2, pagg 226–237, Feb 2005
- [13] A. Leardini, L. Chiari, U. Della Croce, A. Cappozzo, e others, «Human movement analysis using stereophotogrammetry. Part 3. Soft tissue artifact assessment and compensation.», Gait & posture, vol 21, n° 2, pag 212, 2005.

- [14] A. Mantoan, Underwater Gait Analysis: A Markerless Approach, Thesis - University of Padua, 2011
- [15] A. Cappello, A. Cappozzo, P.E. di Prampero, Bioingegneria della Postura e del Movimento, Pàtron Editore, 2003
- [16] B. D. Lucas, T. Kanade , An Iterative Image Registration Technique with an Application to Stereo Vision, 1981
- [17] C. Tomasi, T. Kanade, Detection and Tracking of Point Features. Shape and Motion from Image Streams: a Factorization Method. Part III, Technical Report, 1991
- [18] Luigi Amato, Maurizio Mondelli, (2021), Progettazione e sviluppo di un framework per l'ottimizzazione delle postazioni di lavoro in fabbrica basato sull'analisi posturale
- [19] Valeria Mondini, (2011-2012), Sviluppo e caratterizzazione di algoritmi di segmentazione per l'analisi del cammino in acqua con tecnica markerless
- [20] Filippo Minelle, (2013), Sviluppo di un metodo per confrontare l'analisi del cammino markerless e l'analisi del cammino marker-based in ambiente acquatico
- [21] Jean-Yves Bouguet, Camera Calibration Toolbox for Matlab® (http://www.vision.caltech.edu/bouguetj/calib_doc/), 2010
- [22] D.C. Brown, Close Range Camera Calibration, DBA Systems, 1971
- [23] http://www.vision.caltech.edu/bouguetj/calib_doc/htmls/example.html
- [24] G. Guerra-Filho, «Optical motion capture: theory and implementation», Journal of Theoretical and Applied Informatics, vol 12, n° 2, pag 61–89, 2005.
- [25] Lars Mündermann, Stefano Corazza and Thomas P Andriacchi (2006), The evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical applications, Journal of NeuroEngineering and Rehabilitation 2006, 3:6 doi:10.1186/1743-0003-3-6.
- [26] Stefano Corazza, Emiliano Gambaretto, Lars Mündermann, and Thomas P. Andriacchi (2010), Automatic Generation of a Subject-Specific Model for Accurate Markerless Motion capture and biomechanical application.
- [27] S. Corazza, L. Mündermann, A. M. Chaudhari, T. Demattio, C. Cobelli, T. P. Andriacchi (2006), A Markerless Motion Capture System to Study Musculoskeletal Biomechanics: Visual

Hull and Simulated Annealing Approach, *Annals of Biomedical Engineering*, Vol. 34, No. 6, June 2006 pp. 1019–1029.

[28] Benjamin Petit, Jean-Denis Lesage, Clement Menier, Jeremie Allard, Jean-Sebastien Franco, Bruno Raffin, Edmond Boyer, Francois Faure, Multicamera Real-Time 3D Modeling for Telepresence and Remote Collaboration.

[29] S. Corazza, E. Gambaretto, L. Mundermann, e T. P. Andriacchi, «Automatic Generation of a Subject-Specific Model for Accurate Markerless Motion Capture and Biomechanical Applications», *IEEE Transactions on Biomedical Engineering*, vol 57, n° 4, pagg 806 –812, Apr 2010.

[30] E. Ceseracciu, Z. Sawacha, S. Fantozzi, M. Cortesi, G. Gatta, S. Corazza, and C. Cobelli, “Markerless analysis of front crawl swimming,” *Journal of biomechanics*, vol. 44, no. 12, pp. 2236–2242, 2011.

[31] N. J. Cronin, T. Rantalainen, J. P. Ahtiainen, E. Hynynen, and B. Waller, “Markerless 2d kinematic analysis of underwater running: A deep learning approach,” *Journal of biomechanics*, vol. 87, pp. 75–82, 2019.

[32] S. Ceccon, E. Ceseracciu, Z. Sawacha, G. Gatta, M. Cortesi, C. Cobelli, and S. Fantozzi, “Motion analysis of front crawl swimming applying cast technique by means of automatic tracking,” *Journal of Sports Sciences*, vol. 31, no. 3, pp. 276–287, 2013.

[33] F. A. Magalhaes, Z. Sawacha, R. Di Michele, M. Cortesi, G. Gatta, and S. Fantozzi, “Effectiveness of an automatic tracking software in underwater motion analysis,” *Journal of sports science & medicine*, vol. 12, no. 4, p. 660, 2013.

[34] D. Volpe, D. Pavan, M. Morris, A. Guiotto, R. Ianssek, S. Fortuna, G. Frazzitta, and Z. Sawacha, “Underwater gait analysis in parkinson’s disease,” *Gait & posture*, vol. 52, pp. 87–94, 2017.

[35] Capozzo A., Catani F., Della Croce U., Leardini A. Position and orientation in space of bones during movement: anatomical frame definition and determination. *Clinical Biomechanics* 1995;10(4): 171-8