

POLITECNICO DI TORINO

Laurea Magistrale in Ingegneria del Cinema e dei
Mezzi di Comunicazione



Tesi Laurea Magistrale

Sistemi di motion capture a confronto: benefici e implicazioni nella virtual production

Supervisor

Prof. Andrea Bottino
Ing. Roberto Iacoviello

Candidato

Alessandro Chiarle

2024

Indice

1	Introduzione	1
2	Stato dell'arte	4
2.1	Virtual Production	4
2.2	Motion Capture	4
2.3	Motion capture e produzione cinematografica	6
2.4	Catena di produzione e pipeline di motion capture	6
2.5	Mocap usata per forme non umanoidi	10
2.6	Tipologie di tracking:	10
2.7	Range, latenza, jitter, accuracy, frequency, dof (position, rotation)	18
2.8	Studio della letteratura scientifica	19
3	Metodologia	25
4	Strumentazione Hardware	27
4.1	Sistema Htc	27
4.2	Sistema Xsens	31
4.3	Sistema Rokoko	36
4.4	Sistema Move One	39
5	Strumentazione Software	40
5.1	Unreal Engine 5.3 e 5.4	40
5.1.1	LiveLink	41
5.1.2	Metahuman	43
5.1.3	Quixel Bridge	46
5.1.4	OpenXR	46
5.1.5	Xsens	47
5.1.6	Rokoko	47
5.1.7	Json Utilities	47
5.2	Rokoko Studio	48
5.3	Xsens MVN 2024	50
5.4	Move One	52
5.5	SteamVR	53
6	Procedura	55
6.1	Descrizione setup fisico	55
6.2	Descrizione setup software	58
6.3	Fase di cattura e porting	63
6.3.1	Porting Rokoko	63
6.3.2	Porting Xsens	65

6.3.3	Porting Move One	66
6.3.4	Possibilità di correzioni	67
6.3.5	Conversione movimento Htc su controller IK	68
6.4	Acquisizione dati	69
7	Analisi dati	76
8	Conclusioni	92
8.1	Sviluppi futuri	95
Appendice		97
	Tabelle deviazione standard con tuta Xsens	97
	Tabelle deviazione standard con tuta SmartSuitPro 2 Rokoko	98
Bibliografia		105

1 Introduzione

La motion capture, o mocap, è una tecnologia rivoluzionaria che consente di registrare e digitalizzare i movimenti umani per creare animazioni realistiche. Essa ha trovato una enorme applicazione sia nella virtual production che nello sviluppo di videogiochi, permettendone una rapida evoluzione e lasciando una creatività sempre maggiore a registi e animatori, i quali possono concentrarsi sullo stile da dare al movimento più che sul movimento in sé stesso, migliorando così l'immersività e la qualità delle produzioni audiovisive.

La motion capture si sta specializzando su due fronti in particolare: la cattura del movimento umano e la virtual production. La mocap legata al movimento umano resta di sicuro uno degli argomenti principali, e si sta ampliando negli ultimi anni anche alla cattura del movimento animale, anche se questo ambito resta molto più complesso da realizzare per ovvie ragioni. Possiamo considerare per esempio il gioco Stray, dove il giocatore si identifica in un gatto virtuale che deve compiere varie missioni. I suoi movimenti virtuali sono stati inizialmente catturati da video riprese dal vero usando un sistema di motion capture ottico per velocizzare il processo e poter ottenere dei risultati di partenza delle animazioni da gatto molto più realistiche rispetto a quelle che fossero state create solamente al computer.

In campo cinematografico, oltre alla cattura del movimento umano, la motion capture ha trovato un sempre maggiore sviluppo nel tracciamento delle camere direttamente sul set. Si sono aperte innumerevoli possibilità riguardo scelte stilistiche e registiche su possibili tagli e inquadrature nel mondo virtuale potendo utilizzare camere reali trasposte e ricreate nel set virtuale. In questo modo il regista e gli operatori possono utilizzare vere camere sul set e vedere il risultato in real time rispetto all'ambiente virtuale. Questo ha portato a nuove possibilità prima inimmaginabili, poiché, effettuando una modellazione più ampia della scena, questa tecnica rende possibile effettuare uno scouting direttamente all'interno del mondo virtuale, tramite, per esempio, dei visori, e scegliere l'inquadratura sul momento in base alla regia, andando oltre il limite tecnico di dover stabilire per forza a priori il taglio dell'inquadratura.

Dal punto di vista storico bisogna ricordare che gli inizi della motion capture risalgono agli anni '60, con Lee Harrison III, il quale creò una prima rudimentale tuta dotata di potenziometri che permettevano di animare figure 3D in tempo reale su schermo. Questo sistema prese il nome di ANIMAC.

Nel 1984 viene fondata l'azienda Vicon, viene così accelerata la ricerca nel settore della motion capture ottica in diversi centri universitari (Glasgow e Giappone). Alla fine degli anni 80, grazie a una crescita e maggiore maturità dell'azienda, insieme a un notevole progresso tecnologico dell'epoca, sono stati in grado di portare un sistema di motion capture ottico per il film Titanic (1995) e l'anno successivo a effettuare la motion capture per Ghost di Michael Jackson. Questo ha permesso all'azienda di affermarsi nel settore e stabilire il nuovo standard, spingendo avanti il progresso.

Negli anni '90, la motion capture ha visto un'ulteriore evoluzione con l'introduzione di nuove tecnologie e metodologie. L'industria dell'intrattenimento ha iniziato a sperimentare maggiormente con la motion capture per creare effetti visivi più realistici e coinvolgenti. L'uso della motion capture non si limitava più solo ai movimenti del corpo, ma iniziava a includere anche le espressioni facciali, migliorando significativamente la qualità delle performance digitali.

Una pietra miliare agli inizi degli anni 2000 in cui viene usata la motion capture è sicuramente la trilogia del Signore degli Anelli (2001-2002-2003). In questo caso è stata applicata in due modi diversi: sia per ricreare il movimento di personaggi fantastici, così da sfruttare un movimento di partenza molto più preciso, sia per avere una preview di possibili inquadrature per la virtual production, per esempio riuscendo a capire l'angolazione e il taglio da dare alla telecamera durante le riprese per rendere al meglio le scene. In questo periodo si inizia a parlare anche di "performance capture", poiché oltre al movimento del corpo se ne riesce a catturare anche l'espressione facciale, portando il risultato virtuale ad una performance attoriale completa.

Negli ultimi anni la motion capture si è sviluppata per migliorare e aprire nuove possibilità nella virtual production e nel mondo videoludico. Per esempio, nella serie "The Mandalorian", sono stati sfruttati enormi led wall insieme al motore grafico Unreal Engine e a sistemi di camera tracking così da creare ambienti

virtuali immersivi in tempo reale. In questo modo è stato possibile avere un render di alta qualità già sul set, permettendo agli attori di vedere immediatamente il risultato virtuale di ogni loro movimento cambiato nella realtà. Tramite il camera tracking è stato possibile, inoltre, sfruttare il motore grafico per simulare e adattare il mondo virtuale alla prospettiva della camera, potendo effettivamente vedere lo sfondo direttamente dal vero (sul led wall). Questa innovazione ha permesso inoltre di aumentare l'efficienza di produzione, poiché ha introdotto la possibilità di cambiare al momento elementi virtuali senza aspettare i tempi di render, e di avere un'illuminazione sull'attore reale migliore, poiché i led con lo sfondo effettivo dell'ambiente virtuale permettono un'illuminazione realistica sugli attori e sugli oggetti di scena.

La motion capture costituisce insomma un mezzo tecnologico sempre più innovativo per l'industria cinematografica e televisiva. Permette di catturare e utilizzare movimenti via via più complessi, sia in caso di registrazione che in caso di utilizzo in tempo reale, portando a modalità di elaborazione del movimento sempre più efficienti e a buon mercato, consentendo anche a una piccola azienda o alla singola persona di crearsi il proprio sistema di motion capture.

Il focus della motion capture risiede proprio nell'ottimizzazione e integrazione di questi sistemi al fine di massimizzare i risultati da un lato e selezionare modalità proporzionate agli scopi produttivi dall'altro tenendo sempre presenti i relativi costi.

L'obiettivo di questa tesi è proprio quello di effettuare una comparazione dei diversi sistemi di motion capture (MoveOne, Xsens e Rokoko) messi a disposizione dalla Rai, in modo da poter indicare alla produzione il sistema migliore in base alle esigenze del set in questione. La valutazione finale, cioè, non stabilisce una rigida classifica in cui un sistema prevale su un altro, ma un'indicazione relativa, diversa caso per caso, in base alle esigenze della produzione, alle specifiche tecniche hardware e software, alle attrezzature necessarie, ai tempi di realizzazione, ai risultati qualitativi che si vogliono ottenere e ai budget disponibili per sostenerne i costi.

2 Stato dell'arte

2.1 Virtual Production

La virtual production è un insieme di tecniche che vengono utilizzate per creare contenuti multimediali, film e serie tv. Questa racchiude l'utilizzo di strumenti come motori di rendering in real time, realtà virtuale, realtà aumentata, motion capture. Tutti questi strumenti hanno rivoluzionato il workflow, permettendo di visualizzare e sperimentare gli effetti visivi e le inquadrature già durante la fase di ripresa e non solamente avere un feedback nella parte di post-produzione. Un classico esempio sono i led wall utilizzati per "the Mandalorian", che hanno permesso di caricare l'ambiente circostante al personaggio sul set e di poter recitare avendo i riflessi e gli sfondi dell'ambiente circostanti corretti, non dovendoli poi applicare in post.

2.2 Motion Capture

La motion capture è una tecnica che permette la cattura e registrazione dei movimenti di persone o oggetti tramite l'utilizzo di sensori e telecamere. Questi movimenti vengono poi utilizzati per animare modelli digitali in modo più coerente e meno meccanico, andando a ridurre di molto le tempistiche di animazione rispetto a un metodo di animazione tradizionale.

I sistemi disponibili possono variare in termini di tecnologia e applicazione. Alcuni utilizzano telecamere multiple per rilevare marker riflettenti o a LED posizionati sui soggetti, mentre altri impiegano sensori inerziali (IMU) o sensori magnetici montati sul corpo per misurare accelerazioni e rotazioni. Negli ultimi anni si stanno sviluppando notevolmente sistemi basati sull'intelligenza artificiale, mediante il solo utilizzo di camere. Ogni tipo di sistema presenta vantaggi specifici, come l'alta precisione, la portabilità e la facilità di configurazione, rendendoli adatti a diversi contesti di utilizzo, dalla produzione cinematografica e dei videogiochi all'ambito medico e sportivo.

La motion capture può essere di diversi tipi, principalmente è full body [figura 1], dove viene registrato il movimento completo di tutto il corpo del soggetto e viene trasferito completamente al modello digitale. Questa acquisizione comprende

anche le due sottocategorie della face capture e del finger tracking. Nel dettaglio la face capture si concentra sul viso e sulle espressioni facciali. È possibile effettuare anche una cattura separata dalla body capture e poi aggiungerla in un secondo momento. Si può notare, per esempio, un'evoluzione interessante della face capture presente in Avatar 1 rispetto a quella di Avatar 2. Gli attori nel primo film avevano una camera puntata sul viso con una ventina di punti di controllo, mentre in Avatar 2 sono passati a 2 camere, permettendo di avere oltre 50 punti di controllo nel sequel. Questi punti servono a tracciare il movimento della muscolatura facciale permettendo l'animazione del viso dell'avatar tramite il movimento dei muscoli sottopelle, e non occupandosi di animare direttamente la parte superficiale.



Figura 1 Esempio motion capture full body

Spesso si possono anche animare personaggi con proporzioni diverse dall'attore reale, ed è reso possibile grazie all'adattamento delle proporzioni alle varie esigenze.

Nella full body capture è spesso presente il finger tracking. Nello specifico è la parte della cattura del movimento delle dita tramite principalmente l'utilizzo di guanti con sensori di diverso tipo per permettere un tracciamento molto preciso e permettere quindi una maggiore fluidità nella successiva pulizia e animazione.

Tra i film più famosi che hanno fatto ricorso a questa tecnica possiamo ricordare Avatar, Il Signore degli Anelli (Gollum per citarne uno), film della Marvel, Lo Hobbit (Smaug), Star Wars, Pirati dei Caraibi (Davy Jones).

2.3 Motion capture e produzione cinematografica

La tecnologia della motion capture si è diffusa sempre di più nella produzione cinematografica grazie alla grande riduzione e semplificazione del flusso di lavoro per gli animatori. Questi ultimi possono partire da una base di movimenti estremamente già dettagliata e precisa, ottenuta attraverso la registrazione dei movimenti reali degli attori invece di doverla creare da zero. Questa tecnica permette una migliore esperienza visiva per il pubblico. Il tempo risparmiato dalla creazione dei movimenti da zero permette agli animatori di dedicare più tempo e risorse ai dettagli più fini e alle sfumature espressive dei movimenti.

2.4 Catena di produzione e pipeline di motion capture

Nel mondo della produzione cinematografica e dell'animazione, pitchvis, previs e postvis rappresentano tre passaggi essenziali del processo tecnico e creativo, consentendo ai creatori di contenuti di visualizzare, organizzare e ottimizzare le loro idee prima di procedere alla produzione effettiva.

Tecniche e fasi:

- **Pitchvis:** Breve trailer che mostra su quale base vanno stabiliti budget e necessità. Deve creare aspettativa e invogliare studi e produttori a investire nel progetto.
- **Previs:** Nella fase di riproduzione serve a esplorare e visualizzare tutte le possibilità di come potrebbe risultare il prodotto finale. Consiste nel realizzare artwork, storyboard, modelli 3D che verranno poi usati in fase di postvis.
- **Techvis:** A partire dalla previs, il reparto tecnico progetta movimenti di camera, dei personaggi, in modo che si muovano nel mondo reale

esattamente come si muoveranno in quello virtuale. Questo passaggio, insieme alla previs hanno avuto un ruolo fondamentale, per esempio, in Avatar 1 e 2, permettendo al regista di poter girare nello studio con una sorta di camera virtuale che faceva da ponte con l'ambiente generato in CG, permettendogli di scegliere in modo più naturale e intuitivo le inquadrature, le ottiche e i movimenti di camera.

Questa si può definire una sorta di modalità in AR (realtà aumentata), dove la fotocamera deve essere accuratamente tracciata e ricreata nel mondo virtuale.

Gli strumenti usati per tracciare il movimento della camera e degli altri elementi della scena sono marcatori ottici, encoder, IMU (inertial motion tracking), computer vision o una combinazione di queste tecniche. Il render in real time viene compositato con il plate girato in live action per creare un feedback con il sistema Simulcam. In aggiunta, perché la sincronizzazione sia efficace, possono essere aggiunti sensori, LIDAR e altri elementi fisici tracciati.

- **Virtual production:** Permette al regista, tramite l'utilizzo di motori grafici come Unreal Engine e Unity, di visualizzare in tempo reale l'ambiente virtuale e le possibili interazioni con esso da parte di personaggi sul quale è stato effettuato il tracking tramite motion capture. Allo stesso modo questa tecnica può essere applicata per le inquadrature, potendo effettuare il tracking di camere reali e vedere in tempo reale le inquadrature nell'ambiente virtuale. Questo materiale verrà poi usato come footage di reference per i tagli di montaggio e per la produzione dei vfx. Un esempio recente è Inside out 2, dove la Pixar ha ricreato stanze intere anche se non vengono inquadrare completamente, così da permettere una creatività maggiore per quanto riguarda la scelta delle inquadrature.
- **Postvis:** è la fase in cui il girato live – action viene compositato con il footage prodotto in fase di previs. È un utile strumento per ridurre i costi di produzione, perché permette di capire se una scena funzioni o meno e dunque modificarla a costi minimi.

Andando nel dettaglio della motion capture, si può notare che indipendentemente dalla tipologia di mocap venga utilizzata, si possono individuare degli step comuni utili per utilizzare al meglio il sistema.

Motion Capture Pipeline

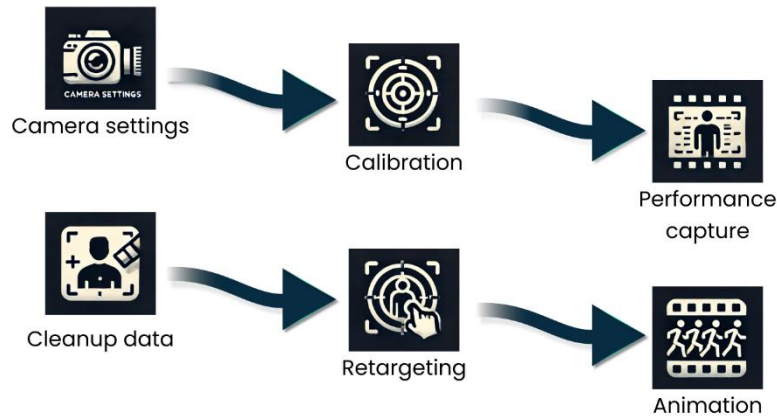


Figura 2 Schema relativo alle diverse fasi della pipeline di motion capture

- **Camera setting:**

Si controllano tutte le camere in uso nell'ambiente di cattura con l'ausilio di alcuni marker posti su corpi rigidi. Si definisce l'orientamento delle camere nello spazio relativamente alle posizioni dei marker, e da questo momento è possibile triangolare la posizione dei marker che vengono visti da almeno due strumenti di ripresa.

- **Calibration:**

Viene definita la posizione iniziale del soggetto nello spazio. Si invita il performer a stare in piedi in una posizione standard (solitamente A pose o T pose) mentre il software calcola le relazioni tra i marker posti sull'attore e i riferimenti del sistema. Il procedimento solitamente è rapido (ordine dei secondi) e come best practice è meglio ripeterla più volte durante la sessione di cattura, così da avere tutti i marker e sensori funzionanti e allineati. Per quanto riguarda i sistemi inerziali la calibrazione del soggetto è fondamentale poiché è in quel momento

che viene stabilito il centro del sistema di riferimento, dal quale verranno calcolati gli spostamenti rispetto a quella posizione iniziale.

- **Performance capture:**

Fase di cattura effettiva dei movimenti del performer tramite programmi dedicati (ad esempio Unreal Engine, software proprietari dei vari sistemi).

- **Cleanup Data:**

Fase di post processing, che comprende l'analisi dei dati acquisiti nella performance capture e pulizia dei dati numerici. In questa fase solitamente si risolvono i problemi di drifting, specialmente degli arti inferiori rispetto al terreno. Questo problema viene riscontrato solitamente con sistemi di cattura non ottici, poiché meno precisi.

- **Retargeting:**

È la fase in cui ci si assicura che il modello digitale di destinazione finale dei movimenti registrati sia in grado di prendersi in carico tutte le animazioni registrate con il sistema di motion capture. Lo scheletro del modello digitale viene associato allo scheletro creato nel software dedicato alla fase di performance capture, facendo molta attenzione alla nomenclatura delle ossa e alla gerarchia (ogni sistema ha una sua gerarchia e spesso differiscono tra diversi sistemi). Solitamente sono presenti dei plug-in che permettono di effettuare il retargeting in modo automatico.

La situazione resta diversa nel caso si voglia effettuare il retarget su uno scheletro custom (magari non di forma umanoide), caso nel quale si rende necessario creare la gerarchia in modo da associare le ossa e i movimenti nel modo desiderato.

- **Animation:**

Layer di pulizia dei dati sul modello finale, aggiunta e correzione di animazioni. È possibile quindi correggere i movimenti senza andare a distruggere tutto il movimento di partenza, in modo additivo (correggere l'impugnatura di una spada, la direzione della testa ecc..).

Questa pipeline può essere applicata allo stesso modo anche alla face capture. La differenza principale è sul retargeting, poiché esso avviene nello scheletro relativo al viso e ai suoi controlli per la deformazione.

2.5 Mocap usata per forme non umanoidi

La motion capture è stata sviluppata non solo per migliorare e semplificare l'acquisizione del movimento umano, ma si è estesa anche alla cattura di movimenti non umanoidi, come ad esempio quelli degli animali, reali o inventati. Questo ha portato a nuove sfide tecniche, poiché ogni tipologia di animale ha una struttura anatomica differente, bisogna sviluppare ogni volta un rig ad hoc in modo da poterlo utilizzare nel migliore dei modi. Lo stesso procedimento può essere adottato anche per la realizzazione dei movimenti di animali non reali. Si può considerare come esempio Avatar 2. Per creare le cavalcature (le varie tipologie di pesci) dei Na'vi, in questo caso, hanno utilizzato due persone per fare l'interpretazione del movimento dell'animale: uno per la testa e uno per la coda. In mezzo a queste due figure era presente l'attore principale che interpretava il Na'vi che cavalcava la creatura. Oltre a questo esempio nel cinema si possono citare numerosi altri esempi, come il Balrog, il troll, Gollum nel Signore degli Anelli, i clicker in The Last of Us (serie tv).

2.6 Tipologie di tracking:

Nella motion capture sono state sviluppate diverse tecnologie di tracking per sopperire alle più diverse esigenze. Si passa dalla tecnologia meccanica, considerata quella più invasiva e limitante nei movimenti, all'AI, sviluppata negli ultimi anni, che rende minima l'invasività del sistema. Di seguito saranno spiegate le diverse tipologie:

- Meccanici

I tracker meccanici sono dispositivi che usano strutture meccaniche collegate tramite giunti forniti di sensori in grado di calcolare la posizione e l'orientamento del punto finale della catena detto end effector. Sono stati i primi tipi di tracker

sviluppati. I vantaggi dell'utilizzo di questa tecnologia risiedono sicuramente in un'accuratezza elevata e costante per tutta l'area di lavoro e a bassa latenza; inoltre, sono dei sensori privi di deriva e immuni da interferenze ambientali (ottiche, magnetiche, sonore). Il problema maggiore riguarda l'essere meccanici, e quindi hanno bisogno di essere ancorati da qualche parte per scaricare la forza di reazione. Un altro problema è l'area attiva di lavoro, la quale è legata all'estensione della struttura meccanica. Queste strutture sono spesso particolarmente pesanti o ingombranti e dunque non ergonomiche.

- Magnetici

Consistono in una serie di sensori (receiver), tipicamente tra la dozzina e la ventina, che posti sull'attore misurano la relazione spaziale rispetto a un campo magnetico prodotto da una sorgente fissa (emitter).

Il campo magnetico è creato tramite una spira alla quale è applicato un campo elettrico. Il receiver misura il campo magnetico, e il valore di questa misura dipende dalla sua posizione e orientamento rispetto alla sorgente.

Per avere tutti i dati necessari a determinare la posizione, si devono generare in sequenza (per non creare interferenze) tre campi magnetici differenti orientati verso gli assi x,y,z.

Il receiver ha tre bobine metalliche ortogonali ognuna delle quali restituisce come misura del campo magnetico un voltaggio in funzione di posizione e orientamento. Al termine del ciclo, al ricevitore sono stati calcolati 9 voltaggi diversi: 9 equazioni in 3 variabili di posizione e 3 di orientamento. Il problema principale è dato dalle interferenze magnetiche che si generano nell'ambiente. I trasmettitori AC generano correnti di Foucault nei metalli presenti nell'ambiente che producono a loro volta dei piccoli campi magnetici che generano interferenze. Il vantaggio più importante è che possono lavorare in real-time calcolando la componente di traslazione e rotazione dei sensori relativamente allo spazio. I sensori sono tipicamente a 6 DOF (degrees of freedom) e si possono ottenere dei risultati molto precisi se si accostano ai sistemi ottici che richiedono un numero più ampio di trackers.

- Ottici

Il sistema di motion capture ottico utilizza una serie di telecamere che si interfacciano con i marker posizionati sulla tuta in prossimità dei giunti. Questa soluzione è molto nota per l'elevata accuratezza e capacità di catturare movimenti rapidi e complessi. È utilizzata per qualsiasi campo, partendo dal campo cinematografico e arrivando al campo biomedico.

Il tracking può essere di due tipologie diverse: "dall'esterno verso l'interno" e "dall'interno verso l'esterno".

I tracker utilizzati nella prima tipologia di questo sistema dipendono dalla risoluzione della camera e possono essere:

- **Marker passivi:** sono composti da materiale altamente retroriflettente di forma sferica, che una volta illuminati, creano degli spot luminosi nell'immagine ripresa dalle telecamere. La posizione del marker è ottenuta tramite triangolazione.
- **Marker attivi:** in questo caso i marker sono propriamente dei LED ad infrarossi che inviano dei segnali luminosi alle camere. È presente un sistema di illuminazione in contemporanea e serve controllare l'ampiezza e la frequenza del segnale per distinguere i diversi marker.

Per ottenere informazioni sull'orientamento di un oggetto si devono utilizzare costellazioni rigide di 3 o più marker.

La seconda categoria "inside looking out" funziona grazie a sensori posizionati sugli oggetti. Questi sensori sono RGB o di profondità, e spesso per esempio sono utilizzati nei sistemi di tracciamento per la realtà virtuale, permettendo una ricostruzione di una porzione dell'ambiente 3D partendo da immagini, analizzando il movimento dei punti 3D e ricostruendo di conseguenza i 6 DOF del dispositivo. Questo sistema è utilizzato spesso negli HMD, ed è semplificato da camere RGBD.

Il problema principale legato a questo sistema è legato all'ambiente di cattura, che deve essere un ambiente controllato, con illuminazione e materiale dei

pavimenti possibilmente non riflettenti e una configurazione precisa di telecamere per la cattura. Un ulteriore fattore negativo è il costo di tutto il sistema, molto elevato sia per quanto riguarda il lato hardware che software.

Sistema Optitrack¹:

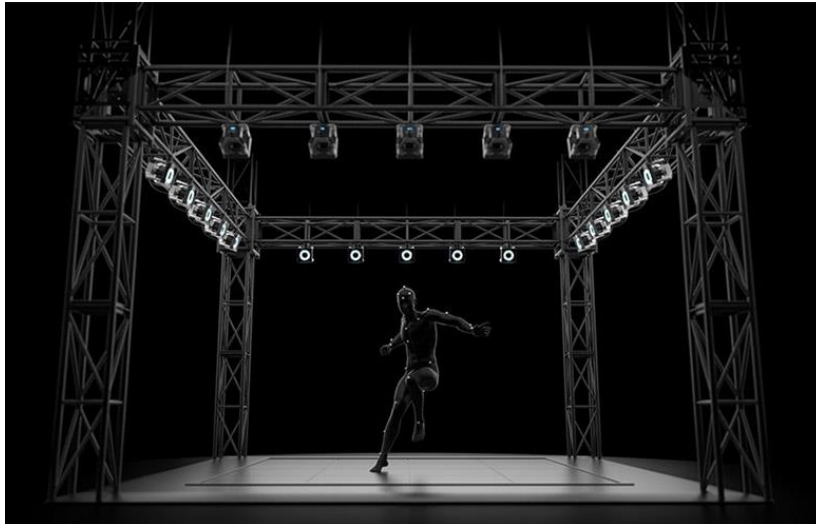


Figura 3 Sistema Optitrack

Accuratezza: +/- 0.2mm

Latenza: <10 ms

Accuratezza rotazione: +/- 0.1 gradi

Con l'ultima cam sviluppata da loro Prime^x sono arrivati a una risoluzione 4K con 12 MP e un frame rate di 300 fps a risoluzione massima, potendolo aumentare anche a 1000 fps con una risoluzione 2048 x 896.²

Il range di questa camera passa da 36 a 91 metri.

Il sistema Optitrack supporta una doppia tecnologia di marker, sia passivi che attivi. I marker attivi di questo sistema vengono utilizzati per via della bassa latenza, utile nella realtà virtuale e mixed reality.

¹ Dati presi dal sito ufficiale: <https://optitrack.com>

² <https://optitrack.com/support/hardware/primex-120.html>

Questo sistema, per esempio, è stato utilizzato sul set di Star Wars: The Mandalorian season 1, effettuando il tracking delle camere principali durante il film. (ILM + Epic Games, creazione del virtual set).

Sistema Vicon³:

Questo sistema si basa sull'uso di marcatori passivi, riflettono la luce emessa da telecamere ad alta velocità, consentendo al sistema di calcolare la posizione tridimensionale di ciascun punto.

Una caratteristica principale di questo sistema è la sua grande precisione, grazie a una frequenza di campionamento di oltre 100 fps, in base al tipo di camera utilizzata. Si passa da sistemi di 2.2MP con frame rate molto elevati (330 fps) all'ultimo loro sistema rilasciato, Valkyrie⁴, che abbassa di poco il frame rate, passando a 150 fps, ma aumentando notevolmente la risoluzione del sensore, portandola a 26MP. Le prestazioni di cattura possono aumentare considerevolmente se si effettua un subsampling, arrivando anche a 2000 fps.

È stato usato per numerose produzioni, per esempio su produzioni Marvel. Un'altra collaborazione è stata fatta direttamente con Epic Games.

Motion capture acquatica⁵

Le camere utilizzate per la motion capture sono di tipo ottico e si basano sulla tecnologia a infrarossi, solo che questi non si propagano in modo ottimale sott'acqua. Per risolvere questo problema il team R&D della produzione ha sviluppato delle camere che sfruttano i raggi ultravioletti. Questo ha permesso di risolvere quel problema ma a causa della superficie riflettente dell'acqua, hanno dovuto creare un nuovo stratagemma per permettere una corretta cattura dei movimenti in acqua. Hanno riempito la superficie di piccole sfere galleggianti semitrasparenti. Questo ha permesso la corretta cattura in acqua poiché non vi

³ <https://www.vicon.com/>

⁴ https://info.vicon.com/hubfs/hardware/valkrie-spec-sheet.pdf?_ga=2.89082843.1944137124.1720428877-323246896.1720428877

⁵ Da Slimdogs: <https://youtu.be/fkiVOqfDQe4?si=Kgq7i58m1s61Rqdk>

era più la superficie riflettente, ma quella tecnologia non funzionava a dovere quando era fuori dall'acqua.

Per risolvere questo ulteriore problema, è stata utilizzata la tecnologia ottica classica a infrarossi per la parte in superficie. Per tutte le azioni che prevedevano di tuffarsi, nuotare, emergersi e immergersi in acqua, erano presenti entrambe le tecnologie sul set della piscina, con un volume formato dalla cattura a infrarossi per la parte superiore e un volume formato dalla cattura a ultravioletti per l'immersione. I dati di acquisizione di entrambe venivano inviati a un computer che aveva il compito di fondere le due tecnologie e permettere lo switch tra superficie e immersione in tempo reale.

- Inerziali

Sono sistemi microelettromeccanici MEMS (accelerometri, giroscopi e magnetometri) e arrivano a 9 DOF: 3 gradi di libertà per la posizione, 3 gradi di libertà per la rotazione e 3 gradi di libertà per il campo magnetico. Questi sensori calcolano velocità angolari (attraverso tre giroscopi ortogonali) per integrazione temporale, ottenendo tre angoli di orientamento, e accelerazioni spaziali (tre accelerometri solidali con i giroscopi). Conoscendo l'orientamento dell'oggetto (dai giroscopi) e sottraendo la forza di gravità, si ottengono le accelerazioni lungo gli assi e, per doppia integrazione, la posizione dell'oggetto. I tracker sono spesso integrati in dispositivi indossabili che possono essere facilmente applicati a diverse parti del corpo. I punti di forza di questi sensori sono: l'assenza di interferenze da segnali esterni, la vestibilità, portabilità e facilità di utilizzo, poiché non richiedono un ambiente specifico o una configurazione complessa.

Il problema è che le misure sono ottenute per integrazione o doppia integrazione; quindi, il rumore sulle misure di velocità angolare e accelerazione non si elimina, ma si somma. Il problema è ancora più grave nel caso di doppia integrazione. Inoltre, l'errore cresce geometricamente nel tempo (la misura deriva) e il drift è più elevato per la posizione, a causa della doppia integrazione e del drift introdotto nel calcolo dell'orientamento. La deriva può raggiungere i 40mm in 2 secondi. Una possibile soluzione consiste nell'utilizzare altri tipi di tracker per resettare periodicamente l'output dei tracker inerziali. Soffrono spesso di problemi relativi la precisione, che può diminuire nel tempo a causa di problemi

di drifting. Per sopperire a questo problema occorrono numerose calibrazioni per mantenere una buona accuratezza.

Un sistema che si basa su questa tecnologia e ne fa il suo punto di forza è il sistema di Movella, Xsens. Questo sistema in particolare verrà trattato in seguito nei prossimi capitoli, ma si può già accennare una grande precisione e libertà di movimento, priva di interferenze di alcun genere. Un ulteriore vantaggio di questo sistema risiede nei buffer integrati nei sensori, che permettono la registrazione anche con una breve perdita temporanea del segnale.

- Markerless Ai

Negli ultimi anni si stanno diffondendo e crescendo in modo esponenziale sistemi di motion capture basati sull'intelligenza artificiale. Il vantaggio di questi sistemi è relativo alla possibilità di cattura markerless, quindi permettendo un movimento più naturale e libero. Possono essere utilizzati ovunque e non richiedono per ora un hardware troppo costoso. Tutto questo processo è delocalizzato su cloud, quindi va a favore per poter effettuare la cattura su dispositivi poco costosi, ma d'altro canto si è costretti ad affidarsi alla risoluzione in remoto sui server proprietari del programma. Un problema relativo a questo sistema deriva dal fatto della qualità delle camere e degli algoritmi usati. Un secondo problema è relativo alla cattura video, dipende dal frame rate delle camere usate, spesso non molto elevata, avendo difficoltà su movimenti rapidi e complessi. Sono soggette a occlusioni.

Move.ai:

Sistema che utilizza videocamere standard o iPhone montati su treppiedi. Si possono usare più iPhone (da 2 fino a 6). Il funzionamento si può suddividere in due parti: Acquisizione ed elaborazione dati. Nella prima parte il software tramite le camere registra video per ogni inquadratura contemporaneamente, sfruttando anche le informazioni relative al Lidar (nel caso di iPhone). Nella seconda parte, i file vengono inviati sui server proprietari in cloud, dove vengono applicati algoritmi di AI, Computer vision, Biomeccanica e fisica. Una volta analizzati i dati vengono poi trasformati in file compatibili per poterli successivamente leggere ed elaborare (Fbx,Bvh,Usd ecc). La stessa azienda ha pubblicato da poco

un'applicazione, sempre per iPhone chiamata MoveOne, il cui principio di fondo resta lo stesso, ma è necessario utilizzare solamente un iPhone ed effettuare l'upload e analisi di un solo video. Questo ovviamente estrarrà dei risultati di livello inferiore rispetto alla versione con più iPhone ma resta comunque un risultato molto buono in una prima analisi.

Wonder Dynamics:

Sistema differente rispetto al precedente, permette il semplice upload di un solo video da parte dell'utente, senza il multicamera, e in remoto tramite algoritmi viene effettuata una scansione del video con rilevamento automatico degli attori e dei loro movimenti. Un ulteriore differenza rispetto al software precedente è relativa al video: con Wonder Dynamics non è necessario che il punto camera sia fisso; quindi, è possibile effettuare carrellate e avere una sostituzione del soggetto anche con una camera in movimento. Questo software permette anche una sostituzione diretta sul video finale dell'attore con un personaggio realizzato in CG, andando a ricreare anche l'intorno dello sfondo del personaggio stimando il contenuto dalle parti precedenti e successive del video. L'analisi viene effettuata frame by frame per garantire una cattura accurata dell'azione. Successivamente una volta che si sceglie il personaggio, l'AI mappa i movimenti rilevati sul personaggio in cg e cerca di riflettere in modo accurato la performance reale. Infine, gli utenti possono esportare gli elementi analizzati e elaborati dal sistema in formato fbx, insieme anche a clean plates e maschere alpha.⁶

⁶ <https://wonderdynamics.com/>

2.7 Range, latenza, jitter, accuracy, frequency, dof (position, rotation)

Per poter valutare i diversi sistemi di motion capture, sono state ideate delle metriche in modo da poter evidenziare eventuali problemi e caratteristiche del sistema. Le metriche più importanti sono: accuracy, jitter, drift e latenza.

ACCURACY

L'accuracy è la differenza tra la posizione attuale dell'oggetto e quella misurata dal sensore. È diversa per posizione/orientamento e varia con la distanza dal centro del sistema di riferimento. La distanza alla quale l'accuratezza è accettabile definisce l'area di lavoro del tracker.

JITTER

Il jitter consiste nella valutazione del valore misurato da parte di un sensore quando l'oggetto sul quale è stato effettuato il tracking (e di conseguenza lo stesso tracker) è stazionario. I dati ricavati dal sensore non coincidono con la realtà effettiva ma si discostano leggermente per valori di posizione e orientamento, facendo risultare la posizione virtuale dell'oggetto mossa rispetto alla realtà. Più questo valore è elevato, peggio è il tracciamento.

DRIFT

Questa metrica valuta il discostamento totale del sistema dall'origine rispetto all'inizio della misurazione. I sensori di tracking, soprattutto quelli non ottici, soffrono spesso di drifting, avendo un aumento dell'errore fisso introdotto nella misura dal sensore in funzione del tempo. Durante sessioni di motion capture con sistemi che soffrono di questa deriva, è utile effettuare frequenti calibrazioni, per poter riportare sempre al minimo queste derive.

LATENCY

La latenza riguarda il ritardo temporale di risposta del sensore nel misurare una variazione di posizione o orientamento dell'oggetto. Questo può provocare errori di sincronismo nel caso vengano effettuate sessioni di tracking in real time, dove si vuole avere un riscontro in tempo reale il più preciso possibile.

La latenza del sistema non dipende solo da quella dei sensori, può essere dovuta a diversi fattori:

- Latenza del sistema di tracking;
- Latenza di pipeline di rendering;
- Latenza del display;
- Latenza del sistema di sincronizzazione;

La latenza sotto i 100 ms non è percepibile direttamente (ma una scena in movimento sembra instabile).

2.8 Studio della letteratura scientifica

Per capire l'accuratezza dei diversi sistemi di mocap sono stati utilizzati diversi approcci. La maggior parte degli studi sono stati condotti su software esterni dedicati all'analisi biomeccanica dei movimenti, soprattutto per poter valutare l'utilizzo di questi sistemi in ambiente biomedico e riabilitativo [1]. Molti studi si sono concentrati sullo studio dell'accuratezza di tecnologie a relativo basso costo, come Oculus Quest 2, SteamVr, Xsens, e sistemi di cattura più innovativi, come il LiDAR [2], la cattura volumetrica [3] e tramite l'utilizzo di camere RGB insieme ad algoritmi [4].

Il sistema di tracker di SteamVR (Htc Vive tracker 3.0) è stato confrontato con il sistema Vicon, considerato come gold standard tra i sistemi di mocap [5]. I due sistemi hanno due frequenze differenti di cattura: 200Hz Vicon, 90Hz Htc Vive. Il sistema Vicon è un sistema basato su tracker ottici. In questo caso è stato effettuato un esperimento per capire in un tracciamento dinamico quanta differenza ci fosse tra i due sistemi. L'esperimento si è svolto in un ambiente controllato, e consisteva nell'analisi del moto della parte inferiore del corpo di un atleta durante una pedalata su una cyclette a diverse velocità concordate in anticipo (80giri/minuto, 120giri/minuto, 160giri/minuto). Nello studio sono stati posti 3 tracker Htc Vive fissati all'anca, al ginocchio e alla caviglia, oltre ai tracker del sistema Vicon. Dopo aver fatto le misurazioni, hanno calcolato un coefficiente di Pearson $r < 0,5$ per tutte le condizioni, e una differenza in termini di millimetri

accettabile tra i due sistemi. Hanno notato infine che la differenza sostanziale tra i due sistemi va a notarsi quando vengono raggiunte velocità di movimento più elevate, con il sistema Htc che tende a non essere così accurato (questo dovuto anche alla frequenza di cattura, massima a 90Hz).

L'accuratezza e la precisione del sistema Htc Vive sono state studiate anche tramite il confronto con un braccio robotico [6]. In questo esperimento sono state posizionate due base stations di prima generazione per Htc Vive e un tracker Htc Vive (non specificato di quale generazione) alla fine di un braccio robotico. Successivamente è stato introdotto un secondo tracker a terra così da rendere più precisa la misurazione. Potendo replicare il movimento con il braccio robotico, hanno calcolato e confrontato gli errori quadratici medi tra i valori dei due sistemi (tracker e braccio robotico). Infine, hanno ottenuto una notevole accuratezza dei risultati, con valori di deviazione standard nell'ordine del millimetro e inferiore al grado per gli angoli. Inoltre, con l'aggiunta del tracker a terra, è stato notato un incremento di accuratezza per quanto riguarda il tracciamento sull'asse Y (altezza nel sistema di riferimento locale).

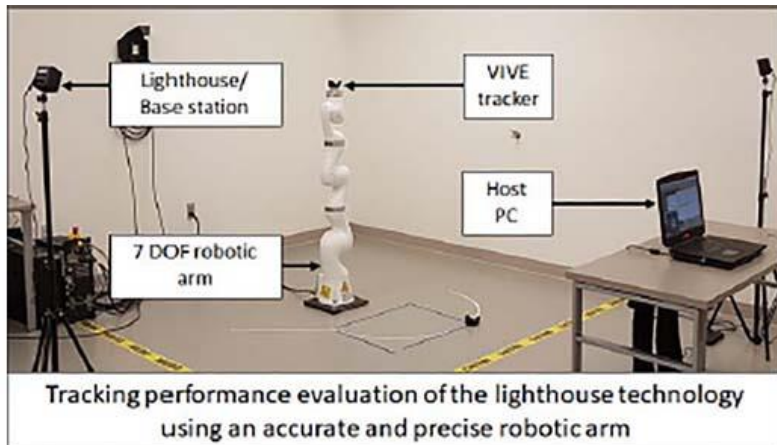


Figura 4 Setup esperimento descritto nel paper

In un altro studio [7] viene confrontato il sistema Htc Vive Tracker 2.0 con il sistema Oculus Quest 2, sempre per valutare l'accuratezza e il jitter dei diversi sistemi. In questo caso è stato effettuato un esperimento in un ambiente controllato, ponendo dei marker a terra a distanze fisse tra di loro (coprendo un'area di 5x5 metri). Hanno posizionato 4 base station SteamVR a disposizione

quadrata attorno all'area, due tracker Htc Vive a due angoli di un carrellino e hanno messo in mezzo ai due tracker l'HMD Quest 2 su una testa di un manichino. Per effettuare la cattura del movimento è stato utilizzato SteamVR desktop per entrambi i sistemi. Il carrellino veniva mosso ogni 20 secondi e posizionato su ogni marker, così da poter calcolare l'RMS del jitter relativo a ogni posizione. Hanno trovato un'accuratezza generale di 0,06mm per l'Oculus Quest 2 e di 0,18mm per Htc Vive. Infine, non è stato effettuato uno studio per quanto riguarda la latenza del sistema e l'inclusione dei controller per il Quest 2.

Passando a un altro studio [8], è stato confrontato anche il sistema Xsens con il sistema Vicon per l'analisi della corsa, concentrandosi sugli arti inferiori. In questo caso, l'atleta ha indossato entrambi i sistemi di motion capture e ha corso su un tapis roulant a tre velocità differenti: 10 km/h, 12 km/h, 14km/h. I dati sono stati acquisiti a 120Hz per entrambi i sistemi e sono stati elaborati con diversi programmi, sia di cinematica diretta (Plug-in Gait) e inversa (OpenSim). Hanno calcolato i coefficienti di correlazione di Pearson e le differenze quadratiche medie RMSD, ottenendo ottimi risultati per quanto riguarda la correlazione $\rho > 0,96$ per gli angoli di articolazione sul piano sagittale, mentre i piani frontali e trasversali mostrano differenze maggiori. Questo li ha portati a consigliare l'utilizzo del sistema Xsens nel caso di analisi sul piano sagittale, mentre viene sconsigliato per il piano frontale e trasversale. In generale Xsens viene considerato una valida alternativa per registrare il movimento in un ambiente naturale.

Il sistema di motion capture Xsens MTw Awinda [9] è stato sviluppato con dei tracker wireless a batteria, così da rispondere a possibili esigenze sul campo dove risultava difficile applicare una tuta con i relativi cavi. Questi sensori sono soggetti a interferenze date da materiali ferromagnetici (possono essere anche una scrivania, una sedia, una macchina o qualsiasi dispositivo elettronico portatile). Viene introdotta inoltre da parte di Xsens anche una nuova pipeline brevettata di elaborazione di un segnale. È stato introdotto anche un protocollo per evitare la perdita dei dati in caso di perdita di segnale. I dati sono salvati su un buffer nel singolo tracker e mandati non appena viene ristabilita la connessione con l'host. Il protocollo proprietario fornisce una sincronizzazione temporale accurata fino a 20 tracker mantenendo i 10 microsecondi.

Frequenza:

Tabella 1 Rapporto frequenza con numero di sensori

TABLE I
MAXIMUM OUTPUT FRAME RATE VS. MAXIMUM NUMBER OF MTW
WIRELESSLY CONNECTED

Number of MTw	Maximum output frame rate f_R (Hz)
1-5	120 Hz
6-9	100 Hz
10	80 Hz
11-20	60 Hz

Le frequenze variano in base alla quantità di tracker utilizzati.

Latenza:

- 1 MTw = 95 ms circa
- 20 MTw = 19 ms circa

Ritrasmissione pacchetto: Il buffer di ogni unità può contenere fino a 1000 pacchetti dati (10 secondi di dati mancanti con tracker che operano a 100Hz).

Overflow del buffer: Tramite il protocollo si previene tramite l'incremento degli intervalli temporali.

Sincronizzazione temporale tra tracker: La trasmissione in broadcast da parte del master contiene un timestamp che viene confrontato con ogni orologio interno dei singoli tracker, permettendo un'accuratezza entro i 5 microsecondi.

Sono stati effettuati due test: uno corto (ordine del minuto, in ambiente controllato e privo di interferenze) e uno lungo (ordine della decina di minuti, in un parcheggio, pieno di macchine e quindi con molte possibili interferenze).

Nel secondo test hanno rilevato dei valori di accuratezza per rollio e beccheggio molto buoni rispetto al valore di accuratezza standard ($0,51^\circ$ e $0,59^\circ$ rispetto a $0,75^\circ$). L'angolo di imbardata invece è risultato leggermente superiore al valore standard ($1,65^\circ$ rispetto a $1,55^\circ$). È interessante la leggerezza e il possibile utilizzo in ambiente reale. Tramite questo sistema è stato possibile monitorare l'intero corpo di un atleta durante una maratona, senza mai perdere il segnale.

Una volta applicati tutti i tracker sulla persona, hanno seguito quest'ultima con una bicicletta, un laptop e l'host così da ricevere i dati in tempo reale per tutta la gara.

Per confrontare i diversi sistemi di motion capture è stato creato da parte di un team di ricercatori un framework nel quale si possono ricreare gli errori dei tracker direttamente in un ambiente virtuale [10].

Hanno implementato VRIK dal pacchetto FinalIK su Unity perché veloce e adatto a eseguire simulazioni multiple in successione. Il kit è stato successivamente personalizzato per essere adattato alle esigenze del caso studio. Hanno sviluppato anche un algoritmo di analisi e calcolo di errori posizionali, rotazionali e di angoli anatomici. Come caso studio hanno preso come modello diverse animazioni di ballo da Mixamo, e, una volta importate nell'ambiente virtuale, hanno posizionato i tracker virtuali che simulano Htc Vive sul manichino nelle posizioni desiderate.

Riguardo lo studio del sistema di Xsens, è stato rilasciato da poco sul sito ufficiale un tutorial per l'integrazione del sistema di tracker Htc Vive all'interno del software proprietario di Movella. Una volta configurato, è possibile avere un tracciamento più preciso di tutto il sistema, grazie al motore MVN fusion che combina entrambi i sistemi di tracking per migliorare la posizione globale. Questo sistema può essere esteso anche a più sistemi nella stessa cattura, migliorando l'allineamento tra diversi attori. Questo tracciamento è permesso nel single level, nel multi level e nel no level.

Queste tre tipologie si riferiscono alla possibilità di tracciamento dell'altezza e posizione: nel single level si viene vincolati parzialmente a terra e non viene effettuata la registrazione dell'altezza, il multi level risulta ideale se si vuole effettuare un tracciamento multi piano, come per esempio la salita o la discesa di un piano di scale o la semplice azione di sedersi su una sedia. Infine, il no level permette di non registrare l'interazione col pavimento e il cambiamento della posizione nello spazio, esprimendo tutto il movimento della persona riferendosi all'osso pelvico invece che alla radice, in modo da potersi concentrare sull'analisi degli angoli del corpo umano [11]. Questa soluzione è spesso utile in campo

biomedico.⁷ Sulla documentazione ufficiale⁸ viene mostrato di quanto viene migliorata la posizione globale con i due sistemi integrati: due attori si avvicinano e battono il cinque e grazie al position aiding (integrazione di entrambi i sistemi Xsens e Htc), non è presente compenetrazione nelle mani, mantenendo la posizione globale corretta.

⁷ Riferimento sito: https://base.movella.com/s/article/Scenarios-in-MVN?language=en_US

⁸ https://base.movella.com/s/article/HTC-Vive-Setup-Guide?language=en_US

3 Metodologia

Lo studio propone un approccio il più oggettivo possibile per effettuare un confronto tra diversi sistemi di motion capture con il sistema considerato ground truth Htc, poiché è stata dimostrata un'elevata precisione di tale sistema confrontandolo con sistemi ottici gold standard nel mondo cinematografico.⁹ Il confronto avverrà registrando lo stesso movimento con diversi sistemi contemporaneamente, così da poter avere un riferimento oggettivo della cattura.

In particolare, verranno utilizzati i sistemi:

- Htc Vive Pro, con Hmd, controller e 4 tracker 3.0
- MoveOne (sistema di motion capture basato su tecnologia di Intelligenza Artificiale)
- Xsens Mvn Awinda Starter
- Rokoko Smartsuit Pro 2

Nel prossimo capitolo verranno spiegati nel dettaglio i vari sistemi utilizzati.

Per effettuare un confronto approfondito sono stati scelti dei movimenti a diverse velocità, così da poter osservare e verificare il comportamento di ogni singolo sistema nelle differenti situazioni, per poterli successivamente analizzare seguendo diverse metriche, come l'accuratezza del movimento, la deviazione standard e la portabilità.

Sono stati scelti 3 movimenti: la classica T-pose, quindi con velocità di movimento trascurabile, in questo modo è possibile avere una stima dell'errore di ogni sistema applicato direttamente su una persona, tenendo conto di un range dato da micromovimenti intrinseci della persona. È stato scelto poi il movimento di una camminata a bassa velocità su un percorso stabilito e un salto raggruppando le ginocchia al petto, con relativa velocità maggiore.

Questi movimenti sono stati registrati con 3 sessioni diverse, in modo da poter effettuare tra una registrazione e l'altra la calibrazione, così da avere risultati il più precisi possibili.

⁹ [<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10490571/>]

Movimenti:

- T-pose
- Camminata su un percorso tracciato
- Salto gambe al petto e mani in basso

Per effettuare la cattura sono intervenute due persone: il performer e l'addetto a far partire le registrazioni. È stato possibile applicare questo metodo poiché in post si possono modificare le registrazioni, tagliandole tutte al momento del ciak.

Una possibile analisi, oltre a quelle di base, riguarda l'errore relativo a differenti ossa. Ci possono essere due vie per analizzarlo:

- Analisi delle posizioni più statiche (T-pose) per tutta la durata dell'animazione, così da vedere di quanto si scosta (analisi magari relativa solo ai piedi, ossa che possono stare quasi del tutto ferme).
- Analisi relativa a diverse metriche come il coefficiente di Pearson e il calcolo dell'RMSD (Root Mean Square Deviation).

Per impostare il caso studio e realizzare i confronti, verranno seguiti vari passaggi chiave. Inizialmente, è necessario stabilire i sistemi di acquisizione appropriati per la raccolta dei dati. Successivamente, si procede con il setup della postazione, assicurandosi che tutte le attrezzature e i software necessari siano configurati correttamente. Una volta pronta la postazione, si procede con la registrazione dei dati. Dopo aver completato la registrazione, è fondamentale verificare che questa sia andata a buon fine. Una volta confermata la qualità della registrazione, i dati vengono esportati su Unreal Engine. Successivamente, si effettua il retarget dei dati per adattarli ai modelli o animazioni previsti. Il passaggio successivo consiste nell'utilizzare un Blueprint realizzato appositamente per questo studio chiamato Animation Baker per salvare i dati incapsulati nel modo corretto in un file *Json*. Una volta completate queste operazioni, i dati vengono immessi su Excel per una fase di elaborazione dettagliata analizzando l'indice di correlazione di Pearson [12], l'RMSD, la deviazione standard e verranno creati i relativi grafici per interpretare e visualizzare i risultati ottenuti.

4 Strumentazione Hardware

Di seguito verranno descritti tutti gli apparati hardware e programmi software utilizzati nel caso studio.

4.1 Sistema Htc



Figura 5 Htc Vive Pro 2

Il sistema è il Vive 2 Pro, sviluppato da Valve, ed è composto da un head mounted display con due controller e due base stations per il sistema di tracking. Inoltre, per avere più punti di controllo sono stati utilizzati 4 tracker htc 3.0.

È stato utilizzato il sistema Vive 2 Pro completo poiché senza di esso non si sarebbero potuti collegare i tracker 3.0 ad Unreal, non permettendo quindi la cattura del movimento.

Questo sistema si basa sulla tecnologia lighthouse 2.0 che permette un tracciamento a bassa latenza, di soli 22ms [13] e elevata accuratezza [14], [15].

Base Station 2.0



Figura 6 Base Station 2.0

Emettono scansioni laser e infrarossi a intervalli regolari. Queste scansioni servono per sincronizzare i sensori presenti sui tracker, controller e visore. In questo modo ogni dispositivo sincronizza il proprio orologio interno con quello della base station, in modo da avere un timestamp comune. Gli impulsi a infrarossi sono emessi con scansioni orizzontali e verticali attraverso pattern noti, in modo che, una volta rilevati dai sensori messi sui tracker, essi possano rilevare il momento esatto di questo incontro.

Tramite i dati temporali ottenuti dall'incrocio tra scansioni laser e infrarossi, il sistema può calcolare la posizione e l'orientamento dei dispositivi tracciati nello spazio tridimensionale, tramite triangolazione, dove la distanza e l'angolo di ogni impulso laser rilevato vengono utilizzati per determinare la posizione esatta del dispositivo.

La sincronizzazione di diverse base station avviene tramite infrarossi, non tramite cavo fisico, in modo che possa funzionare in modo indipendente e senza ingombri dovuti a cavi.

Le base station operano a una frequenza di 90 Hz.

Best practices:

Per massimizzare il risultato di tracciamento bisogna posizionare le base station ad un'altezza superiore a 2 metri, con un angolo tra i 25° e 35°. Dopodiché conviene posizionare il performer / oggetto sul quale effettuare il tracking a una distanza minima di 0,5 metri dalle base station e assicurarsi che l'area di cattura non sia estremamente luminosa, altrimenti l'eccessiva luminosità potrebbe interferire negativamente con il tracciamento.

L'area di copertura con 4 base station è sull'ordine dei 10 metri quadri.

L'area minima consigliata per avere la massima resa di tracking è di 2 metri x 1,5 metri nel caso di due sole base station.

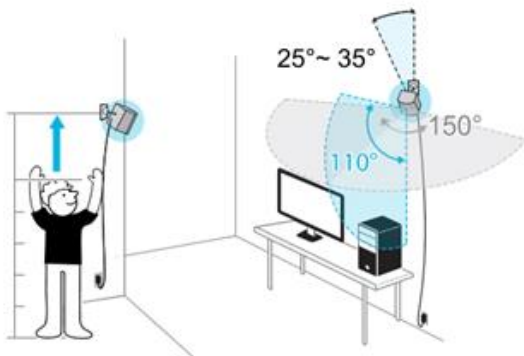
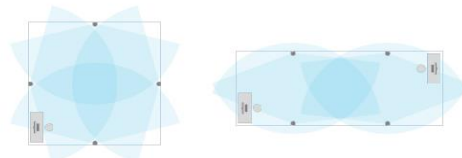


Figura 7 Consigli sugli angoli per le Base Station 2.0

Single user setup



Multi-user setup

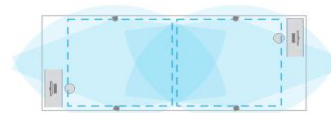


Figura 8 Consigli sulla disposizione per le Base Station 2.0

Htc Vive Tracker 3.0



Figura 9 Htc Vive Tracker 3.0

Sono dispositivi compatti, leggeri, con un peso di soli 75gr possono essere facilmente usati per tracciare oggetti o parti del corpo. Il tracker ha una serie di sensori a infrarossi che lavorano insieme alle base station per effettuare la triangolazione della posizione e orientamento. Questi sensori sono disposti in modo ottimale per massimizzare la qualità del tracciamento. Essi comunicano con il sistema principale tramite connessioni wireless a bassa latenza. Hanno una dimensione contenuta di 70,9 x 79 x 44,1 mm e un FOV di 240°.

4.2 Sistema Xsens

Il primo sistema di cattura studiato è stato l'Xsens Awinda Starter, il sistema entry level (a livello professionale) della Movella, azienda specializzata in sistemi di motion capture basati su tecnologia IMU.



Figura 10 Tuta Xsens Awinda Starter

Il sistema di cattura è racchiuso in una valigia rigida e non è propriamente una tuta, ma un insieme di sensori dotati di un velcro resistente, una serie di straps di diverse grandezze sui quali è possibile attaccare le singole unità IMU e 3 stazioni di ricarica da 6 slot per tutti i sensori. Una stazione di queste funziona da host e si collega tramite cavo proprietario al computer, così da poter permettere la cattura dei movimenti dei sensori. Il singolo host può sincronizzare e registrare i movimenti fino a 20 sensori differenti (una tuta completa più 3 accessori). Questo permette di avere una grande vestibilità, potendosi adattare maggiormente a corporature differenti. Il contro principale resta il fatto di dover attivare singolarmente tutti i singoli sensori e doverli posizionare nelle posizioni corrette, per avere il massimo della prestazione. Per semplificare questa procedura, è possibile inserire manualmente tutti i sensori sulle loro stazioni di carica, in modo che si accendono in automatico tutti assieme.

Sono dotati anche di uno standby automatico dopo un breve periodo di inattività, per risparmiare batteria, e si riaccendono in automatico col più piccolo

movimento. Se i sensori restano in standby per 12 ore, si spengono automaticamente per risparmiare batteria.

Questo problema si aggiunge alla necessità di ogni sensore di avere una propria batteria, al posto di averne una centralizzata.

Di seguito i dati tecnici relativi al sistema utilizzato:

Peso singolo tracker: 16 gr;

Dimensione: 47mm x 30mm x 13mm;

Durata batteria: 6h di utilizzo;

Frequenza:

Tabella 2 Frame Rate acquisizione del sistema Xsens

TABLE I
MAXIMUM OUTPUT FRAME RATE VS. MAXIMUM NUMBER OF MTW
WIRELESSLY CONNECTED

Number of MTw	Maximum output frame rate f_R (Hz)
1-5	120 Hz
6-9	100 Hz
10	80 Hz
11-20	60 Hz

Le frequenze variano in base alla quantità di tracker utilizzati.

Latenza:

- 1 MTw = 9,5 ms circa
- 20 MTw = 19 ms circa

Ritrasmissione pacchetto: Il buffer di ogni unità può contenere fino a 1000 pacchetti dati (10 secondi di dati mancanti con tracker che operano a 100Hz).

Overflow del buffer: Tramite il protocollo si previene tramite l'incremento degli intervalli temporali.

Sincronizzazione temporale tra tracker: La trasmissione in broadcast da parte del master contiene un timestamp che viene confrontato con ogni orologio interno dei singoli tracker, permettendo un'accuratezza entro i 5 microsecondi.

All'interno di ogni unità sono presenti un giroscopio, un accelerometro, un magnetometro, un barometro e un termometro. Di seguito lo schema riportato dal paper ufficiale.

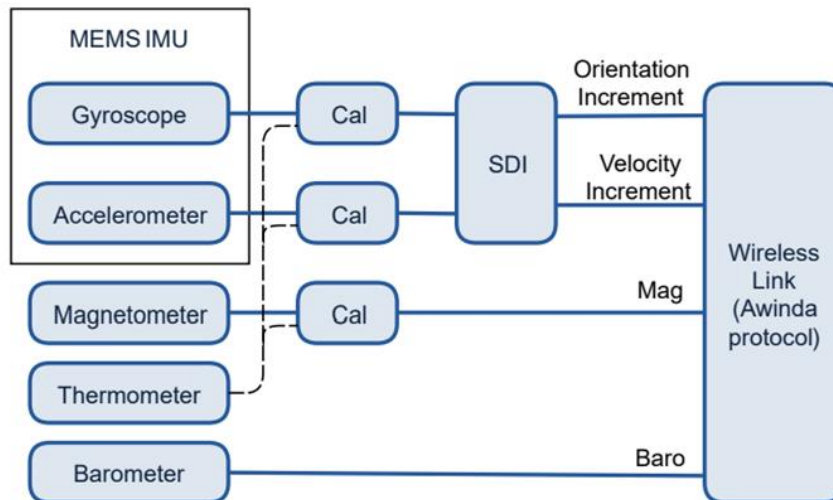


Figura 11 Schema architettura singolo sensore Xsens

Il sistema è sincronizzato dalla stazione di carica master, che permette di avere una precisione nell'ordine dei 10 microsecondi.

La stazione di carica host ha dimensioni: 148mm x 104mm x 31,9mm, e include un'antenna per migliorare la ricezione del segnale delle unità, arrivando a 50 metri in linea d'aria (senza ostacoli).



Figura 12 Host che permette il collegamento di Xsens e computer



Figura 13 Basi di ricarica sensori Xsens

Pipeline di cattura del segnale:

Cattura dati

- **Giroscopio:** sensore inerziale utile per registrare le velocità angolari e per dare una stima dell'orientamento nel tempo. Questo sensore può accumulare un errore nel tempo portando a episodi di drifting.
- **Accelerometro:** sensore inerziale utilizzato per misurare l'accelerazione lineare. Quando il sensore non è in movimento, il valore dell'accelerazione equivale alla accelerazione di gravità (9,8 m/s²).
- **Magnetometro:** Serve a misurare i campi magnetici vicini, usato spesso come riferimento per la direzione.
- **Termometro:** Sensore utilizzato per aiutare a compensare le dipendenze dalla temperatura degli altri sensori durante il rilevamento.
- **Barometro:** Viene utilizzato come sensore di supporto per avere informazioni relative all'altezza.

Integrazione Strap – Down (SDI)

I sensori internamente registrano a una frequenza molto elevata, nell'ordine dei 1000 Hz, e vengono filtrati con un filtro passa basso avente larghezza di banda di 184 Hz. Questo permette di avere comunque dati di qualità elevata. Questo permette all'utente di avere anche il controllo sulla frequenza di cattura, mantenendo però una qualità elevata anche se con una risoluzione temporale ridotta.

Protocollo di comunicazione wireless Awinda

Questo protocollo è stato progettato e brevettato da Xsens per ottimizzare lo scambio di dati e la comunicazione con i sensori.

I sensori sono collegati tutti alla stessa rete wireless creata dalla stazione host e inviano i loro dati con una certa frequenza. Ogni unità ha all'interno anche un piccolo data buffer, per la registrazione locale dei pacchetti. Questo protocollo permette di rilevare e gestire una breve perdita dei dati, potendo inviarli

nuovamente aumentando le finestre di tempo per la ritrasmissione mantenendo un'elevata precisione. Il buffer può contenere 1000 pacchetti di dati (10 secondi di dati mancanti con il sistema operante a 100 Hz). In caso di overflow, il protocollo può combinare i singoli pacchetti insieme in modo da non scartare completamente i dati più vecchi ma avendo solamente una risoluzione di misura ridotta. Inoltre, ogni tracker ha un orologio interno, e l'host tramite questo protocollo è in grado di effettuare una sincronizzazione nell'ordine della precisione di 5 microsecondi.

Filtro Kalman per l'orientamento.

È un filtro riadattato specificatamente per Xsens, in grado di elaborare i dati provenienti da ogni sensore in modo da ricreare una stima di orientamento e posizione ottimale.

Data output

Vengono elaborati e forniti al sistema i dati calibrati relativi a posizione e orientamento del relativo sensore nell'ambiente 3D, in modo da poter stimare la posa corretta nell'ambiente virtuale.

4.3 Sistema Rokoko

La Smartsuit pro 2, sviluppata da Rokoko è una tuta per la motion capture basata su 19 sensori IMU.



Figura 14 Rokoko Smartsuit Pro 2

Ogni sensore include:

1 giroscopio: misura la velocità angolare, permettendo di rilevare la rotazione e l'orientamento del sensore;

1 accelerometro: misura l'accelerazione lineare nelle 3 dimensioni;

1 magnetometro: misura l'orientamento rispetto al campo magnetico terrestre, cercando di correggere le deviazioni dovute a interferenze per i giroscopi.

Il sistema utilizzato è anche provvisto di guanti, permettendo così anche il tracciamento delle dita. I guanti hanno un sensore per ogni dito, un sensore per il polso e un sensore nell'hub.



Figura 15 Rokoko SmartGloves

Ogni sensore è collegato a un hub posto nella parte posteriore della tuta, il quale ha il compito di raccogliere i dati e di inviarli al pc dove è presente il software proprietario dedicato all'acquisizione della motion capture. La connessione con il computer può avvenire in due modi: wireless o via cavo, in base alle proprie esigenze.

La tuta è alimentata tramite powerbank generica, ed è consigliata una che eroghi 5V/3A da una singola porta Usb-A al fine di avere una prestazione ottimale (non oltrepassare quella potenza altrimenti potrebbe danneggiarsi l'hardware della tuta).

Un aspetto molto importante dal punto di vista hardware relativo al setup della Smartsuit pro 2 riguarda il Wifi.

È altamente consigliabile usare un router dedicato a 5 GHz, con solamente le connessioni della tuta e del computer usato per la cattura. Inoltre, per avere la maggiore banda possibile e migliore stabilità, risulta una migliore scelta tenerlo scollegato da internet.

Riguardo la configurazione iniziale del wifi, non è molto intuitiva e non ha funzionato subito il collegamento tra la tuta e il computer. Nonostante fossero stati effettuati correttamente tutti i passaggi della configurazione, la Smartsuit pro 2 continuava a non essere riconosciuta dal computer. Questo problema nasce

dalla porta utilizzata dal wifi, che deve essere la numero 36, altrimenti non viene riconosciuto dalla tuta. Dopo aver provveduto a impostare la porta corretta, la tuta è stata riconosciuta senza alcun problema.

Inoltre, sempre per avere una connessione e certezza di funzionamento, ho notato che, se la batteria viene collegata alla tuta dopo aver collegato anche i guanti, il pairing avviene nel modo corretto.

Il setup della tuta resta molto veloce da fare, dell'ordine dei 5 minuti in caso di aggiornamenti. È anche presente la possibilità di fare tracking in real time e sempre tramite il software proprietario è possibile fare elaborazioni semplici di clean up.

Possibilità di effettuare tracking facciale tramite app proprietaria e il tracking delle dita tramite appositi guanti. Un altro punto di forza è dato dalla possibilità di tracciare fino a 5 persone contemporaneamente.

4.4 Sistema Move One

Sistema di motion capture basato su intelligenza artificiale. Fa parte dell'azienda Move.ai, e consiste in un'applicazione realizzata ad hoc per iPhone (compatibile con iOS 16.4 o superiori).

Il funzionamento è molto semplice, consiste in una registrazione del movimento tramite la camera RGB, senza l'integrazione del lidar, la quale, caricata sui server online dell'azienda e tramite algoritmi e AI addestrate proprietarie, viene analizzata. A differenza delle tute, il movimento non viene creato in real time. Una volta effettuata la registrazione, essa viene caricata sui server proprietari dell'azienda dove avviene l'elaborazione del video, la risoluzione e ricostruzione del movimento. Viene creato in automatico un manichino (proprietario, con una propria struttura scheletrica), e viene reso disponibile al download solo in un secondo momento dopo aver concluso l'elaborazione, convertendo il movimento nei formati .fbx o .usdc e associando il movimento al manichino. Vengono date due ulteriori scelte per poter scaricare anche il video originale della performance.

L'hardware fisico utilizzato dal sistema è semplicemente un iPhone fissato su un cavalletto, così da essere più stabile. Inoltre, la ripresa deve essere fissa, così da poter massimizzare il risultato dell'elaborazione.

5 Strumentazione Software

5.1 Unreal Engine 5.3 e 5.4

Il motore grafico utilizzato è Unreal Engine sviluppato da Epic Games, nelle versioni 5.3 e 5.4. Il motivo di questa scelta deriva da un'esigenza pratica di utilizzo. Ogni sistema di motion capture ha un software proprietario per l'acquisizione del movimento. Ogni azienda ha sviluppato quindi un rig proprietario per poter rendere al meglio la performance di cattura. Questo porta però a una discrepanza se si mettono a confronto sistemi diversi, poiché risulta difficile comparare sistemi diversi aventi rig con strutture anche solo leggermente diverse e con nomenclatura diversa. Per ovviare a questo problema è stato scelto di appoggiarsi a Unreal Engine, poiché per via dei suoi recenti sviluppi molto importanti nel campo della virtual production, rende possibile effettuare un retargeting dell'animazione, passando dallo scheletro proprietario a uno scheletro comune. In questo caso è stata usata la tecnologia sviluppata sempre da Epic Games chiamata MetaHumans che consente la creazione di personaggi umani ultrarealistici, con possibilità di avere rig avanzati sia per il controllo del corpo che del viso e delle dita. Epic Games ha investito molto su questa tecnologia e le aziende di mocap si sono mosse di conseguenza. Ogni azienda di motion capture presa in considerazione ha sviluppato degli appositi plug in che permettono una connessione tra il software proprietario e Unreal, semplificando il passaggio dell'animazione tra un software e l'altro. Hanno reso anche disponibili dei tutorial e dei file per spiegare e permettere il retarget dallo scheletro proprietario al MetaHuman in modo semi-automatico, così da rendere più user friendly e veloce questo procedimento. La versione 5.4 di Unreal Engine ha portato ulteriori aggiornamenti alla parte di animazione, rigging e virtual production. Partendo dalla parte di animazione e rigging, da questa versione è possibile effettuare dei rigging ad hoc in modo più semplice e intuitivo in modo modulare su qualsiasi tipo di personaggio. Inoltre, è possibile effettuare l'operazione di retargeting automatico senza doversi appoggiare a retargeter esterni. Questo sistema resta molto buono ma non ancora affidabile al 100%, poiché nel caso ci sia qualche particolarità nello scheletro di partenza, è possibile

che il retarget non funzioni a dovere, andando a rompere l'animazione e il rig di arrivo.

Un altro aspetto molto importante di Unreal Engine è relativo alla possibilità di utilizzare plugin per migliorare l'esperienza con i vari sistemi di motion capture, permettendo la connessione in tempo reale e il retargeting del movimento da un software all'altro. In particolare, questo aspetto è stato utilizzato per poter importare i movimenti di Xsens all'interno di Unreal. Gli altri software invece non hanno permesso uno streaming in real time su Unreal, poiché quella funzionalità richiedeva abbonamenti aggiuntivi di cui non si poteva disporre in quel momento. Questi software hanno comunque sempre la possibilità di, una volta registrata l'animazione, poterla esportare in un file .fbx, in modo da poterla importare in un qualsiasi software di 3D.

Di seguito verranno trattati tutti i plugin utilizzati.

5.1.1 LiveLink

Live Link permette di effettuare lo streaming di animazioni da fonti esterne (nel nostro caso di motion capture) all'interno di Unreal. Questo resta possibile poiché è possibile lo sviluppo a terze parti. Lo streaming è possibile farlo in live, quindi questo rende possibile una live preview dei movimenti, dando inoltre una maggiore possibilità di controllo e correzione in tempo reale.

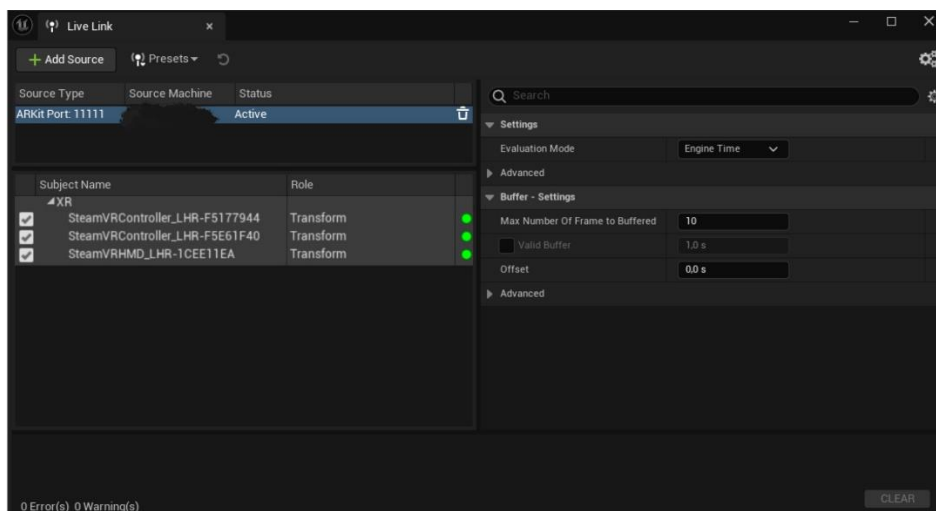


Figura 16 Live Link plugin

L'interfaccia è semplice e funzionale. È composta da 3 finestre: la prima finestra è quella delle fonti esterne, dove è possibile selezionare quelle che si vogliono utilizzare tra quelle disponibili. La finestra sottostante visualizza tutti i dati e configurazioni relativi a quel determinato sistema che sono disponibili. In particolare, nel nostro caso si può vedere il sistema Htc collegato con tutti i sensori spuntati. Per semplificare la selezione dei sensori relativi ad ogni sistema che si vogliono utilizzare, hanno diviso la finestra in 4 colonne: la prima, quella più importante, contiene il tick di spunta, per deciderne l'utilizzo o meno. La seconda e la terza sono relative al nome del tracker / sistema e al ruolo disponibile, mentre la quarta colonna è relativa a una spia colorata in base alla qualità della connessione, per vedere in modo rapido se sono presenti eventuali anomalie o sensori in standby. Per poter visualizzare le risorse relative a una determinata tecnologia (Rokoko, Xsens, Htc), è necessario attivare i plugin proprietari sviluppati dalle relative aziende. Questo permette un corretto dialogo tra i software esterni e Unreal.

L'utilizzo in contemporanea di fonti diverse permette quindi di poter combinare il controllo tra le fonti: per esempio è possibile prendere la motion capture full body su una tuta Rokoko e poter applicare al viso dello stesso personaggio la motion capture facciale da un software differente, per esempio da Live Link Face, app proprietaria di Epic che permette una ricostruzione precisa delle espressioni facciali, basandosi sulle API di Apple relative alla Face recognition, tutto in real time.

È presente anche il plugin LiveLinkXR, il cui principio alla base resta identico al LiveLink normale, ma permette un migliore riconoscimento dei sistemi di realtà virtuale. In particolare, è stato usato per riconoscere il sistema Htc, permettendo di riconoscere l'hmd, i controller e tutti i sensori disponibili per il tracking.

5.1.2 Metahuman

Tecnologia sviluppata da Epic, permette a chiunque abbia un account Epic di creare un avatar rappresentante una persona in modo semplice ed efficace. Per permettere questo hanno creato una piattaforma online chiamata MetaHuman creator, dalla quale, dopo aver effettuato l'accesso, è possibile modificare numerosissimi parametri dell'avatar, sia del viso che del corpo.

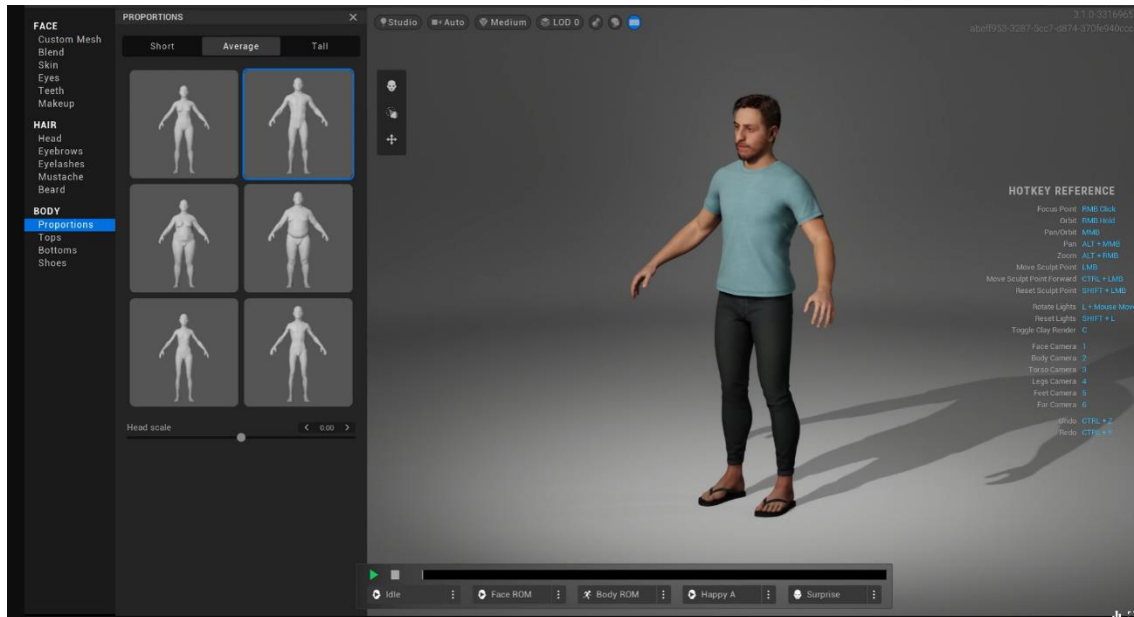


Figura 17 Schermata editor Metahuman

Una volta creato il proprio avatar, per importarlo su Unreal basterà fare due passaggi: attivare il plugin MetaHuman così da poter usare tutte le funzioni relative ad esso e poterlo utilizzare all'interno di Unreal, e aprire Quixel Bridge, portale di dialogo tra Unreal e librerie proprietarie di Epic. Nella finestra di dialogo, dopo aver effettuato l'accesso con lo stesso account Epic, comparirà una tab relativa ai MetaHuman creati e disponibili al download. È inoltre possibile scegliere la risoluzione di download tra bassa, media, alta e massima.

Il plugin di MetaHuman permette inoltre di ricreare il proprio avatar cercando di copiare il viso reale della persona tramite i seguenti passaggi:

1. Effettuare due video con LiveLinkFace app. Nel primo video bisogna prima guardare frontalmente la camera, poi effettuare rotazioni leggere del capo verso destra e verso sinistra seguendo con lo sguardo il movimento, e infine bisogna fare un sorriso molto enfatizzato. Nel secondo video si

dovrà realizzare la performance voluta da inserire come animazione nel Metahuman.

2. Importazione del video su Unreal
3. Nella cartella metahuman aggiungere una capture source tramite il metahuman animator
4. Andare nel tool capture manager e selezionare la source creata poco prima
5. Click destro nella cartella metahuman, crea metahuman ID.

Da qui si apre una schermata dalla quale bisogna scegliere 3 frame dal video importato poco prima: il primo per la vista frontale, e i due successivi per le viste laterali. Dopodiché si seleziona il corpo desiderato (per scegliere a grandi linee il genere e l'altezza).

Successivamente si sceglie la posa con il sorriso (chiamata teeth pose).

Si va avanti con la procedura guidata fino all'ultimo passaggio: Prepare to animate. Questo passaggio è quello più oneroso dal punto di vista computazionale e ha come requisito minimo 64gb di RAM, altrimenti non può portare a buon fine la creazione del personaggio.

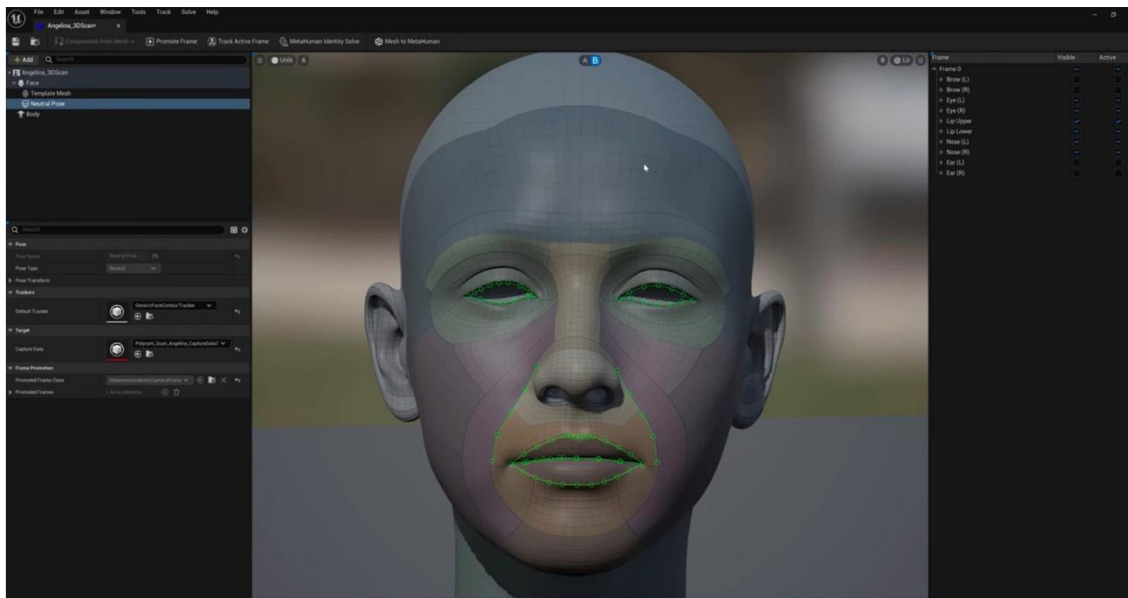


Figura 18 Schermata processo mesh to Metahuman

Dopodichè, una volta creato o selezionato il proprio avatar, si può creare l'animazione registrata con Live Link Face app: si seleziona il Metahuman ID, si seleziona Head Movement Mode, control rig e si esegue il process dell'animazione.

Infine, si procede con l'export dell'animazione processata ed è possibile aggiungerla nel Metahuman desiderato.

Spesso si può verificare un problema legato alla testa, che nella viewport risulta staccata dal corpo (il Metahuman è diviso tra corpo e testa). Per risolvere questo problema, bisogna spuntare delle opzioni sul body: details animation tab keyframe animation mode e spuntare il tick su disable post process blueprint. Bisogna fare la stessa cosa per la testa. Dopodiché, occorre andare nell'anim graph della testa e aggiungere una blend mask sul movimento, così da poter recuperare il movimento della testa dato dal corpo e non dalla cattura facciale. Questo porta a tenere la testa attaccata al corpo, mantenendo l'animazione facciale catturata con l'app e l'animazione di tutto il corpo e del collo data dalla mocap del corpo.

5.1.3 Quixel Bridge

Portale che permette l'utilizzo di numerosi asset grafici di altissima qualità tutti foto-scansionati, come per esempio mesh o materiali di qualsiasi tipo. È molto usato a livello di creazione di ambienti, poiché ricco di scansioni fotogrammetriche relative a rocce, alberi e molta natura, tutto di altissima qualità, permettendo anche di scegliere una qualità della texture inferiore alla risoluzione massima in caso ci siano vincoli progettuali.

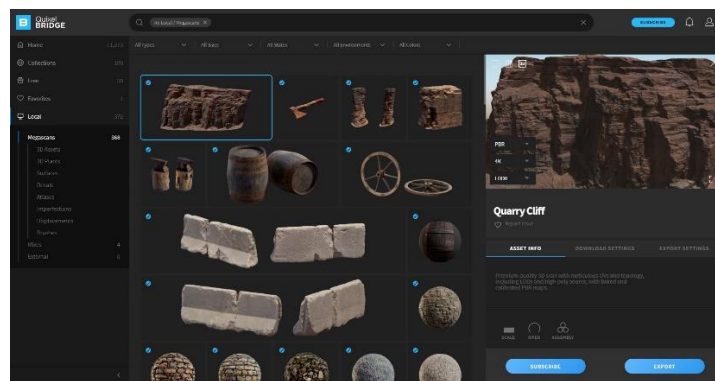


Figura 19 Quixel Bridge, schermata principale

5.1.4 OpenXR

Questo plugin si occupa di supportare diverse tipologie di head mounted display e controller, andando a creare un collegamento ogni qualvolta si colleghi una di queste periferiche. Si occupa di creare una compatibilità tra diverse tipologie di dispositivi, andando ad abbinare i controlli del dispositivo esterno con quelli che più sono simili nella sua libreria. Questo plugin ha sostituito quello proprietario di Steam per l'integrazione dell'headset all'interno di Unreal.

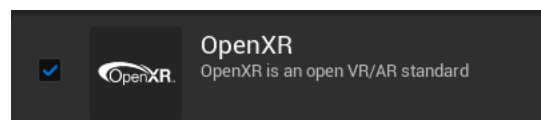


Figura 20 Plugin OpenXR

5.1.5 Xsens

Permette di effettuare il live link direttamente dal programma proprietario di Movella, potendo retargettare il movimento direttamente su un metahuman (o sull' avatar desiderato) in tempo reale. Per fare questo bisogna impostare la stessa porta di host sia sul programma di Xsens che su Unreal. Dopodiché è possibile visualizzare il movimento e registrarlo anche direttamente su Unreal.

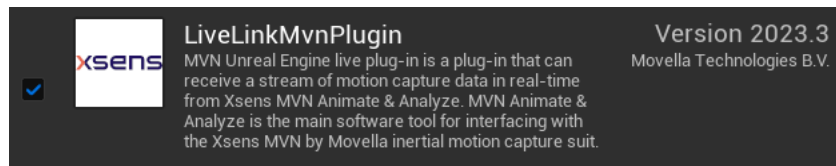


Figura 21 Plugin Xsens

5.1.6 Rokoko

Permette di effettuare il live streaming dal programma proprietario Rokoko Studio direttamente su Unreal in real time, ma bisogna disporre di un abbonamento a pagamento per sbloccare questa funzione.

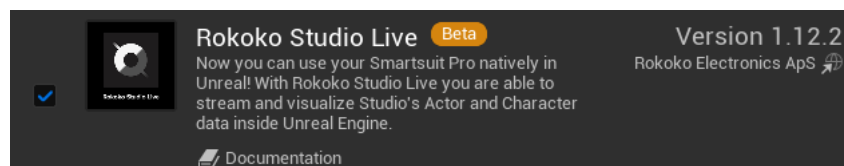


Figura 22 Plugin Rokoko Studio

5.1.7 Json Utilities

Attivando questo plugin, viene importata una libreria relativa alle operazioni con i file json, alla quale è possibile accedere attraverso i nodi in un blueprint. Questa libreria si è rivelata fondamentale per poter salvare e organizzare le informazioni nel modo corretto riguardo lo studio effettuato.



Figura 23 Plugin Json Utilities

5.2 Rokoko Studio

Software proprietario di Rokoko per la motion capture. Permette grazie a un'interfaccia grafica semplice e intuitiva di gestire la cattura o streaming in real time di una performance. Una volta effettuato il login, si può creare subito una nuova scena e un nuovo personaggio che avrà le misure del performer.

La schermata è molto semplice e pulita, con 4 aree principali: a sinistra è presente il finder, l'area dove sono presenti eventuali clip, attori e hardware collegato. Al centro è presente una grande viewport, così da poter visualizzare al meglio i movimenti del performer. A destra è presente un'altra colonna con possibili opzioni che possono essere spuntate. Queste permettono un primo controllo sulla motion capture, in modo da poter semplificare al meglio il lavoro, riducendo i tempi di elaborazione. Per esempio, è possibile cambiare in real time il tracking dell'altitudine (in caso di salti o salite e discese da scale), oppure anche attivare l'opzione "treadmill", cioè, effettuare la registrazione del personaggio bloccata all'origine del mondo virtuale anche se nella realtà il performer si sposta dal luogo di partenza.



Figura 24 Rokoko Studio schermata principale

Tuttavia, le opzioni riguardanti la limitazione del movimento conviene attivarla solamente dopo la registrazione, poiché non verrebbero salvati i relativi dati di

movimento, mentre se si registra il più possibile, successivamente è sempre possibile riprocessare i dati andando ad attivare opzioni limitanti il movimento.

Insieme al pannello dei dettagli, sono presenti altri due pannelli, relativi al livestreaming e all'export. Per quanto riguarda l'ultimo pannello, è possibile esportare l'animazione in due formati: fbx e bvh. Nella versione a pagamento è possibile scegliere anche numerosi parametri per poter esportare nel migliore dei modi l'animazione, mentre nella versione gratuita è possibile l'esportazione solo tramite fbx con un preset standard.

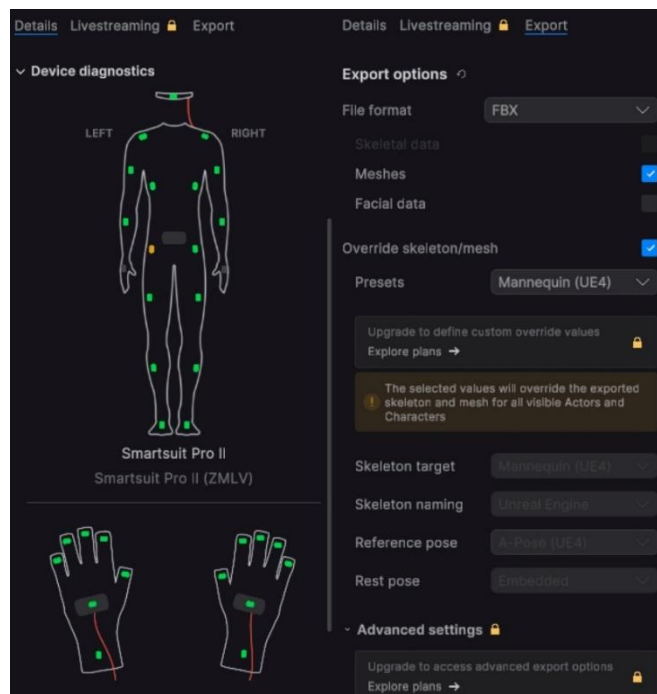


Figura 25 Rokoko Studio, dettagli e Export

Subito sotto a questo pannello si può vedere uno schema molto minimale ma ottimale che permette di visualizzare in tempo reale, e anche in un secondo momento, lo stato di ogni sensore cambiando la luce (da verde a rossa), permettendo così di individuare eventuali interferenze o malfunzionamenti.

Infine, la porzione inferiore dello schermo è occupata dalla timeline. In particolare, si possono notare 5 layer: un layer per la cattura del movimento del corpo, due sottolayer relativi alla stima del punto di contatto del piede sinistro e destro con il

pavimento. Gli altri due layer relativi alla cattura dei movimenti da parte dei guanti della tuta.

Successivamente può essere collegata la tuta tramite cavo usb, in modo da poter fare eventuali aggiornamenti. Fatto questo, se tutto funziona a dovere, è possibile collegarla tramite Wifi.

Questo programma supporta la cattura anche di più attori contemporaneamente, permettendo di effettuare scene più complesse migliorando quindi la produttività e la qualità delle animazioni.

5.3 Xsens MVN 2024

Software proprietario della Movella, permette di registrare e fare livestreaming di motion capture su diversi software, in base alla licenza in possesso. La schermata principale è pulita e funzionale, con un design leggermente datato, e si divide in 3 sotto-finestre, con al di sopra la barra degli strumenti, essenziale per poter usare il software.

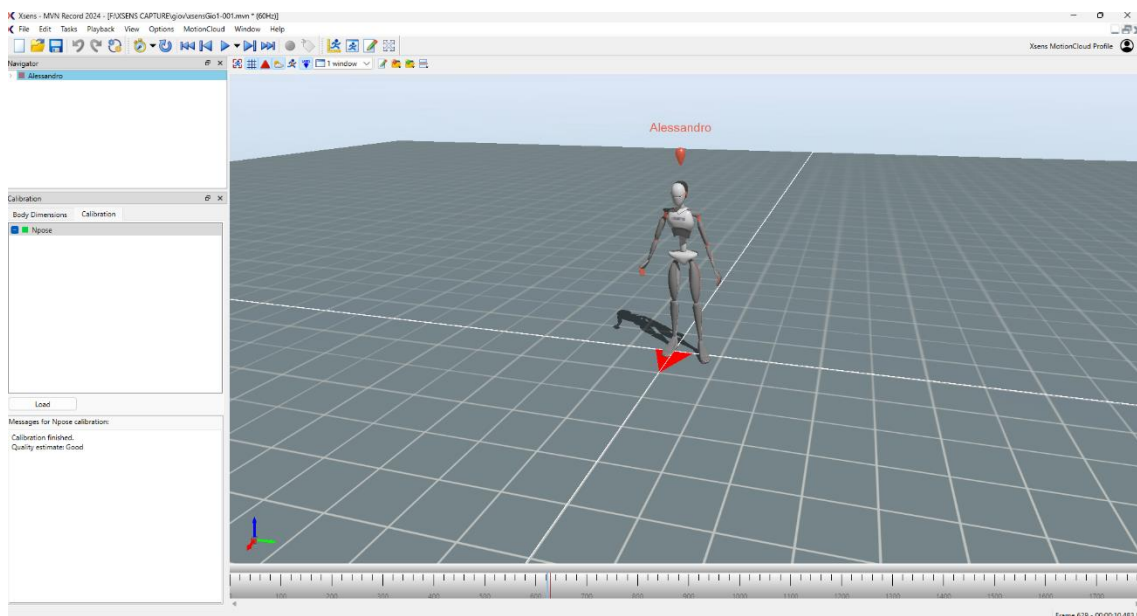


Figura 26 Xsens schermata principale

Al di sotto di essa sulla sinistra sono presenti due finestre una sotto l'altra: la prima rappresenta i sistemi in uso, fornendo tutti i dati specifici di ogni sistema, mentre la seconda rappresenta la finestra attraverso la quale è possibile

effettuare la calibrazione della tuta. Sulla destra è presente la viewport attraverso la quale è possibile effettivamente vedere l'azione del performer attraverso la ricostruzione su un manichino rispetto anche al "centro del mondo".

Per quanto riguarda la configurazione del manichino sono necessarie solo alcune misure del performer, come altezza, lunghezza del piede e del palmo della mano.

La calibrazione è molto semplice da effettuare e può essere fatta anche da una singola persona in modo autonomo.

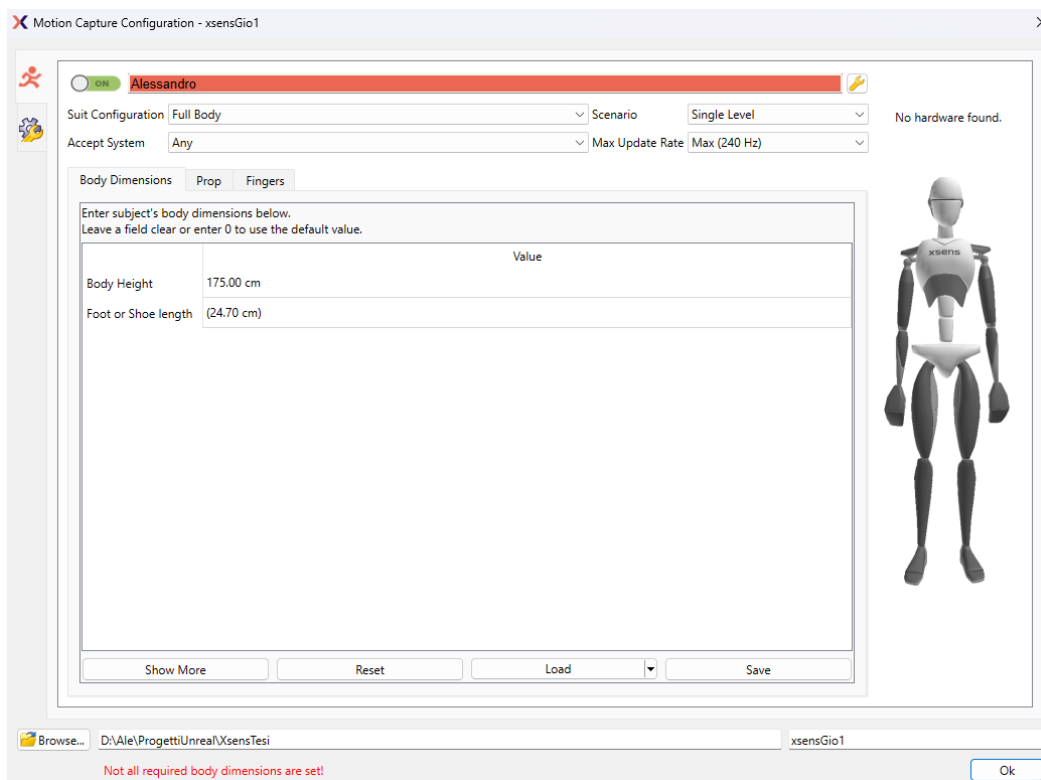


Figura 27 Xsens schermata configurazione

È possibile utilizzare anche i guanti MANUS gloves, sempre della stessa azienda ma venduti separatamente, in modo da poter effettuare contemporaneamente sia la body capture che la finger capture. In questo caso bisogna specificare nella configurazione del manichino se sono presenti dei guanti per la mocap, e nel caso selezionarli dalla apposita tendina.

Con la licenza utilizzata si è potuto registrare il movimento in un primo momento e effettuare il livestreaming su Unreal in un secondo momento.



Figura 28 Guanti Quantum Manus

5.4 Move One

Il software è una semplice applicazione per iPhone (non ancora disponibile su Android), la quale registra tramite camera rgb (e tramite Lidar se il telefono ne è provvisto) il video del performer. La durata del video è preimpostata su dei valori fissi di 10s, 20s, 30s o 60 secondi (se si ha un abbonamento).

Il sistema, essendo online, è a crediti: l'utente può avere 30 crediti per 30 secondi di prova, dopodiché deve acquistare un abbonamento se vuole effettuare altri take di motion capture. I crediti vengono scalati solamente quando vengono elaborati dal sistema.

Ha una semplice interfaccia: una volta aperta l'applicazione, si passa in automatico alla fotocamera, con un tasto in basso per far partire la registrazione e di fianco un tasto per passare alla galleria dei video già registrati. In alto a destra è presente il tasto relativo al profilo dell'utente con giusto poche opzioni essenziali, come il login, la quantità di crediti disponibili e la possibilità di acquistarne altri.

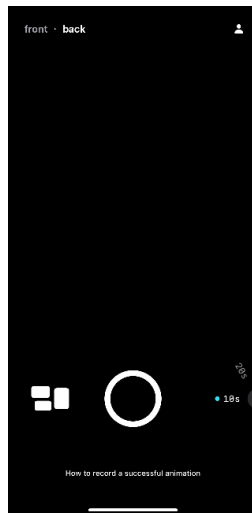


Figura 29 Move One schermata principale

Dopo aver inviato la registrazione della performance al sistema online bisogna aspettare l'elaborazione, e una volta terminata, si riceve una notifica con un aggiornamento del video nella galleria, al quale è stato associato un manichino di MoveOne con il movimento sincronizzato al video di partenza. L'applicazione permette di effettuare il download dell'animazione creata tramite il formato fbx, e permette inoltre di poter scaricare il video di partenza sul quale hanno girato gli algoritmi proprietari di AI per ricreare il movimento sul manichino.

5.5 SteamVR

Software essenziale per poter interfacciare il sistema utilizzato come riferimento di tracking Htc Vive Pro 2 e permettere quindi il dialogo con Unreal Engine. Il programma è presente nella libreria di Steam, ed è gratuito per il pubblico. Una volta aperto, si apre in automatico il mirroring con il visore collegato e una piccola finestra recante il nome della versione del programma con anche dei suggerimenti subito al di sopra.

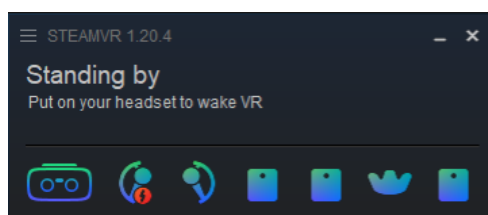


Figura 30 Schermata principale SteamVR

Nonostante la piccolezza della schermata, aprendo il menù è possibile accedere a tutte le impostazioni relative al visore. Le più voci più importanti prese in considerazione e utili per interfacciare correttamente il sistema sono: configurazione della stanza e dispositivi.

La prima opzione serve per poter effettuare la calibrazione della stanza configurando il centro del mondo virtuale. Per svolgere questa operazione il software da due possibilità: fare la configurazione dell'intera stanza, o farla in modo autonomo. La seconda opzione è quella scelta per poter operare senza doversi spostare dal computer.

La seconda voce molto importante riguarda la gestione dei dispositivi: una volta accesi i device fondamentali, cioè le base stations, il casco e i controller, è possibile accendere anche i tracker aggiuntivi. Dopodiché questi vengono rilevati subito dal sistema, ed è possibile visualizzarli in un'apposita schermata, la quale permette di scegliere per ogni tracker un ruolo da un elenco (telecamera o qualsiasi parte del corpo). Una volta stabiliti i ruoli, vestendo l'hmd è possibile vedere la rappresentazione virtuale di tutti i tracker già operativi e funzionanti nell'ambiente virtuale.

Queste operazioni permettono di visualizzare tutti i tracker nel modo corretto una volta aperto il progetto e il Live Link su Unreal Engine.

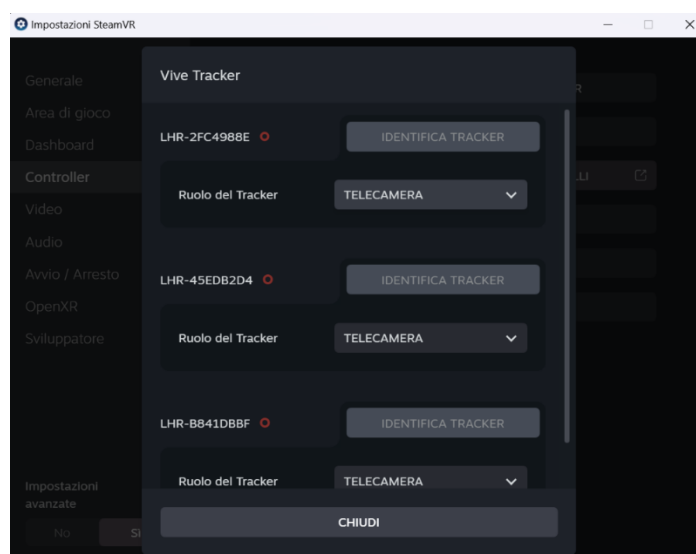


Figura 31 SteamVR schermata di gestione tracker

6 Procedura

6.1 Descrizione setup fisico

Il setup è formato da tre cavalletti, due base station, un iPhone, una sedia, un HTC Vive Pro 2, due controller HTC, quattro tracker HTC 3.0, una Smartsuit Pro 2, un Xsens Awinda Starter, due computer (Z8 e Asus Zephyrus G16) e due prolunghe per collegare le base station.



Figura 32 Setup della stanza

Di seguito una piantina rappresentate il floor plan dell'area utilizzata per effettuare i test:

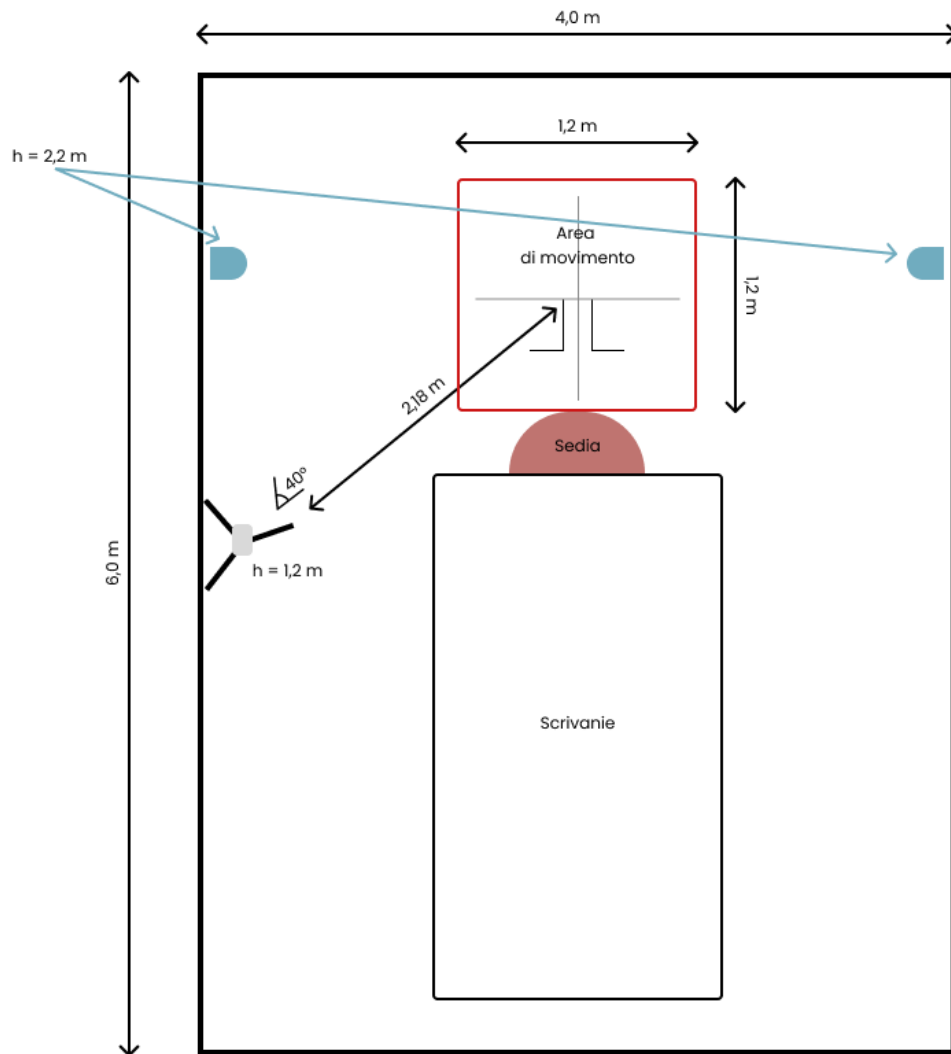


Figura 33 Piantina setup cattura

L'area di calpestio è stata delimitata come un quadrato di 1,2 metri di lato con un nastro nero, e al suo centro sono state tracciate delle linee guida centrali per i piedi, in modo che potesse essere facilitato il posizionamento della persona. Per la parte superiore del busto, è stata usata una sedia con lo schienale alto, poiché conforme con una possibile posizione rilassata di partenza. L'altezza per l'appoggio delle mani è di 110 centimetri. La larghezza standard del supporto è

di 30 centimetri. La distanza dello schienale, quindi dell'area di appoggio per le mani, è di 60 centimetri dal centro dell'area delimitata dal nastro [figura 27].

Le base station sono state montate su due cavalletti identici seguendo le best practices a un'altezza di 2,2 metri, con un orientamento di circa 40 gradi verso l'area di movimento. Il terzo cavalletto, dotato di porta telefono, è stato disposto nel modo migliore possibile considerando gli spazi, non in modo proprio frontale ma con un angolo di circa 45 gradi a 2,18 metri di distanza rispetto al centro della stanza. Questo ha permesso di poter avere una migliore ripresa della persona intera per la maggior parte del tempo. L'iPhone è stato utilizzato in modo verticale, così da poter sfruttare al massimo il formato 9:16 per riprendere a figura intera.

Per i sistemi di motion capture, sono state necessarie due sessioni di registrazione, così da poter avere il migliore risultato possibile con entrambe le tecnologie. Inizialmente ho indossato la Smartsuit Pro 2 della Rokoko, dopodiché ho indossato il sistema Htc Vive Pro 2, infine mi sono messo in posizione per le registrazioni.

Allo stesso modo per effettuare le registrazioni del sistema Xsens, ho prima posizionato tutti i sensori della tuta della Movella, poi ho indossato tutti i sensori del sistema Htc e ho ripetuto il tutto.

Il sistema Htc, nonostante l'headset avesse un cavo collegato costantemente al computer fisso Z8, visto il range limitato dei movimenti scelti per il confronto, non ha creato alcun tipo di problema.

Dal punto di vista pratico, la tuta di Xsens è risultata un poco più scomoda rispetto alla Smartsuit Pro 2: il dover ricordare le posizioni dove sistemare ogni tracker per ogni parte del corpo risulta molto tediosa e lunga rispetto al dover indossare semplicemente una tuta. Al contrario però una volta indossata, risulta molto più leggera e permette un movimento più ampio e più "libero" di quello effettuabile con la tuta intera (Rokoko). Indubbiamente il sistema migliore dal punto di vista del comfort resta MoveOne, che, consistendo in una semplice ripresa, non richiede di indossare alcun tipo di tracker.

6.2 Descrizione setup software

Dal punto di vista software il setup è stato un poco più complesso. Per prima cosa è stato creato in anticipo un progetto vuoto su Unreal Engine, versione 5.3. In seguito, sono stati attivati tutti i plugin descritti in precedenza, poi è stato importato il third person setup, in modo da avere la cartella con i preset di manichini attraverso cui fare i retargeting in un secondo momento. È stato importato un Metahuman, così da avere lo scheletro comune per il confronto eseguito in seguito. In particolare, ho prima creato il Metauman utilizzando una ripresa di me stesso, (ho seguito la procedura di creazione del Metahuman descritta in precedenza) e poi l'ho importato nel progetto, posizionandolo al “centro del mondo” nella viewport, così da avere un primo riferimento per distanze e dimensioni.

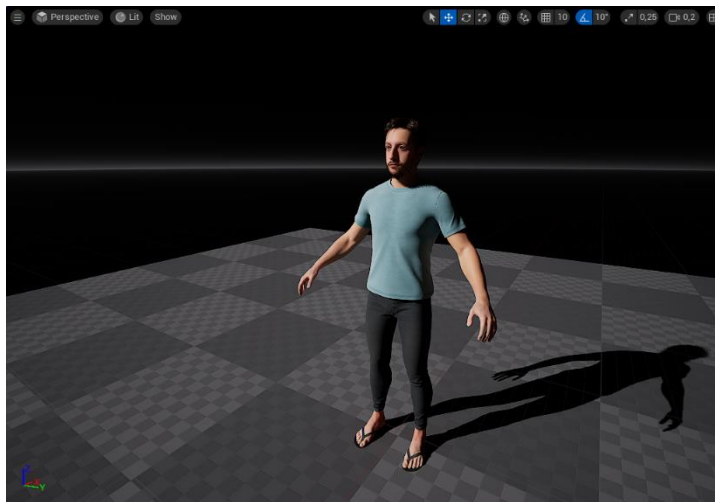


Figura 34 Metahuman importato nell'ambiente

Ho aggiunto poi 7 cubi rappresentanti testa, mani, gomiti e piedi, in modo da poter simulare visivamente i tracker del sistema di riferimento Htc. A ogni cubo, nella tab dei dettagli, è stato possibile aggiungere il component Live Link Controller, attraverso il quale ho potuto collegare la mesh al movimento effettivo nello spazio del tracker andando a cambiare nel menù a tendina in “Subject Representation” il tracker assegnato.

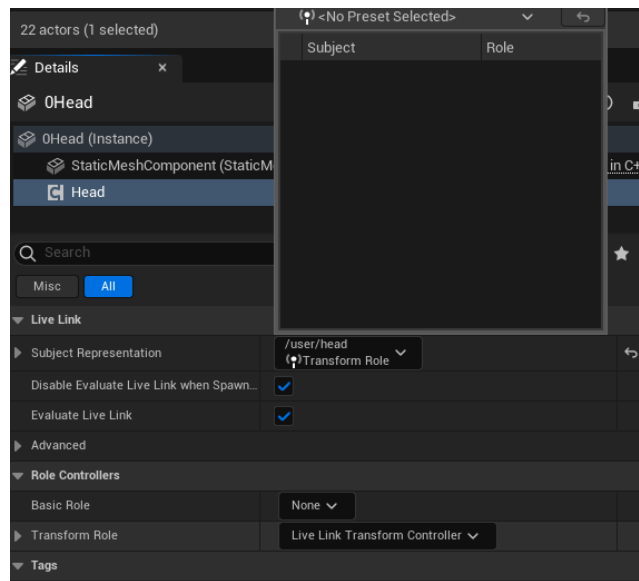


Figura 35 LiveLink component per Htc

È stata adottata questa soluzione per la registrazione del sistema poiché non è ancora presente un sistema proprietario di Htc in grado di integrare i sensori Htc in un manichino completo. Tutti i 7 cubi sono stati imparentati come figli di un piano Global di una determinata superficie, così da poter effettuare in seguito traslazioni globali senza andare a perdere o rovinare il movimento registrato in partenza.

Prima di avviare la registrazione, ho dovuto inizialmente attivare SteamVR, così da poter riconoscere il visore Vive Pro 2, ho effettuato la calibrazione della stanza e successivamente ho acceso tutti i tracker. Una volta verificato il corretto funzionamento di tutti i dispositivi, l'engine grafico Unreal Engine è stato eseguito tramite linea di comando con la seguente stringa aggiunta in coda: "-xrtrackingonly", poiché è risultato l'unico modo funzionante per poter far leggere correttamente il visore e tutti i sensori dal progetto.

Dopodiché, con il progetto aperto, è stato possibile aprire la finestra di Live Link, scegliere il sistema (Htc) e verificare che fossero funzionanti e disponibili tutti i tracker a disposizione. Infine, dopo aver impostato ogni tracker al relativo cubo, il tracciamento è iniziato in real time, potendo quindi vedere la disposizione effettiva dei tracker nello spazio.

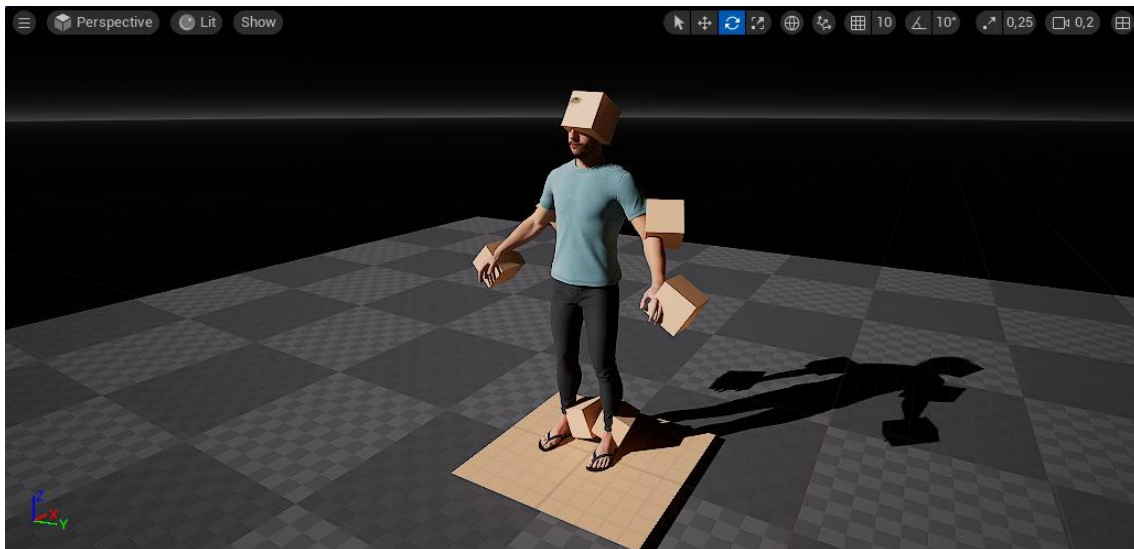


Figura 36 Viewport con sistema Htc

Per registrare il tracking del sistema Htc ho creato una cartella dentro la quale ho aggiunto 6 level sequence nominate "Htc_nomeTuta" in modo da poter registrare e passare da un livello all'altro con ordine, e poter individuare subito il movimento da registrare. La lunghezza di ogni level sequence è stata impostata in modo approssimativo sui 600 frame, così da essere sicuri di registrare il movimento completo. In un secondo momento è stato possibile editare la lunghezza per renderle conformi con tutte le altre animazioni registrate. Per comodità e compattezza i movimenti sono stati tradotti nel modo seguente: T-pose con 1, la camminata in cerchio con 2, il salto con 3.

Il setup con la Smartsuit Pro 2 è stato molto semplice, una volta collegata la tuta, ho creato in modo molto semplice una nuova scena, ho attivato il manichino e ho effettuato la calibrazione, che consiste nel seguire un countdown restando in n-pose per qualche secondo. Come best practice consigliata dalla casa madre, dopo aver effettuato la calibrazione, ho attivato il tasto boost di registrazione della tuta, forzandola a registrare il movimento al frame rate più alto disponibile (200 fps).



Figura 37 Subito prima del processo di calibrazione

Per il setup con la tuta Xsens, una volta accesi e posizionati tutti i tracker correttamente, è possibile aprire il programma sul desktop MVN Studio ed effettuare la calibrazione. Il software è realizzato molto bene e permette di eseguire la calibrazione anche in autonomia, senza dover essere in più persone. Essa consiste nel mettersi in n-pose per qualche decina di secondi, poi effettuare una breve camminata in modo da verificare che tutti i sensori siano attivi e funzionanti, infine tornare in n-pose nel punto di partenza per un'altra decina di secondi. Questa calibrazione determina anche il "centro del mondo". Per rendere più comoda e accessibile la calibrazione è stata introdotta anche una guida sonora, così da poter seguire tutti i passaggi senza dover guardare lo schermo direttamente. Questo ha reso possibile effettuarla sul luogo dei test al posto di doverla fare di fronte al computer (oltretutto con molto meno spazio a disposizione).

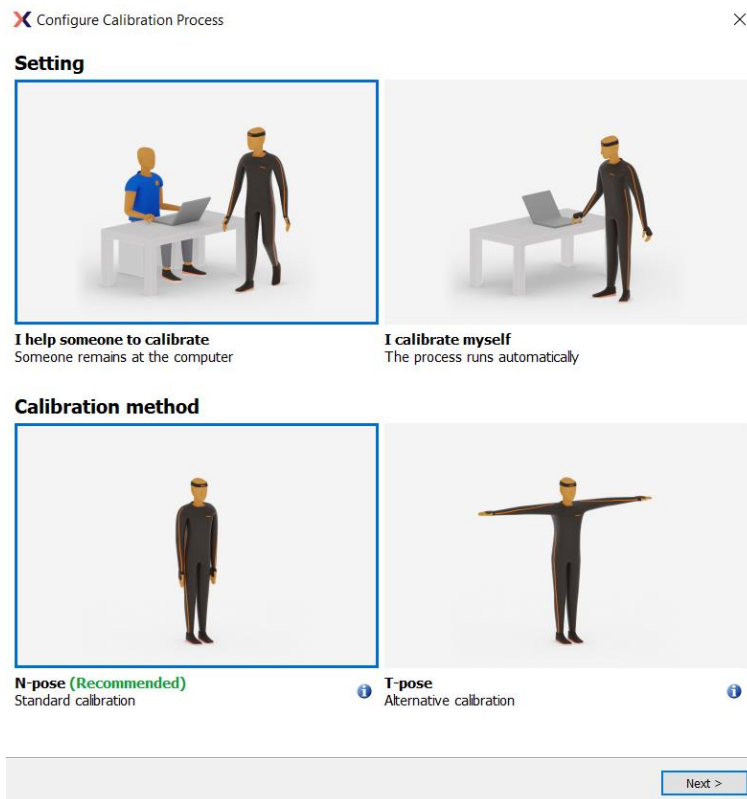


Figura 38 Xsens schermata per la scelta della tipologia di calibrazione

Il setup relativo a MoveOne è rimasto quello più semplice: essendo una semplice applicazione per iPhone, è bastato aprirla e verificare che l'inquadratura fosse corretta. In questo caso sono stato aiutato da una collega per verificare che tutto fosse posizionato nel modo corretto, provando tutti i movimenti in anticipo. In generale ho notato un grosso problema nell'applicazione: non è possibile, una volta aperta la galleria interna, riorganizzare o eliminare le riprese (nel caso il take fosse risultato da scartare). Per quanto riguarda la calibrazione di MoveOne, l'azienda afferma che non sia necessaria (infatti non è presente una vera e propria calibrazione) ma all'inizio della ripresa è consigliato posizionarsi in A - pose così da facilitare il successivo lavoro da parte degli algoritmi di intelligenza artificiale per la ricostruzione dello scheletro. Un'attenzione utile per effettuare una migliore registrazione con MoveOne è relativa alle scarpe: gli algoritmi di intelligenza artificiale tendono a dare un risultato più approssimativo sulla posizione se sono utilizzate scarpe chiare con soles chiare su pavimento chiaro e riflettente. Risulta un migliore risultato se presente un contrasto elevato tra pavimento e scarpa.

6.3 Fase di cattura e porting

Una volta impostati tutti i dispositivi nel modo corretto, sono stato aiutato da un collega e una collega per avviare tutte le registrazioni. In rapida sequenza è stata avviata inizialmente la tuta (Smartsuit Pro 2 o Xsens) poi il record su Unreal del sistema Htc, e per ultimo il sistema MoveOne, poiché essendo basato su tempi standard e crediti utilizzati, non ci si poteva dilungare troppo né commettere errori. Per poter sincronizzare le registrazioni, ho effettuato all'inizio di ogni ripresa un movimento simile a un un ciak, battendo le mani sopra la testa, in modo che risultasse il più chiaro possibile e facilmente individuabile.

Per ogni sessione di cattura le tempistiche sono state di circa 10 secondi a movimento.

Un aspetto negativo relativo a MoveOne è dato dal fatto che non sia stata sviluppata la possibilità di effettuare una registrazione in autonomia. Non sono presenti né una guida sonora né un timer, e questo non permette di far partire la registrazione da soli, obbligando la persona ad essere aiutata.

Le registrazioni sono avvenute in due giorni differenti: il primo giorno è stata effettuata la registrazione con i sistemi Htc, Rokoko e MoveOne. Il secondo giorno è stata utilizzata la tuta di Xsens, il sistema Htc e MoveOne.

6.3.1 Porting Rokoko

Una volta effettuata la registrazione, ho eseguito un leggero data cleaning grazie ai tool forniti all'interno dal software Rokoko studio: ho potuto aggiustare i punti di contatto dei piedi con il pavimento, ho evitato di spuntare l'opzione elevation tracking in tutte le registrazioni, così che il manichino rimanesse dopo il movimento sempre vincolato al piano 0.

Successivamente ho effettuato l'export nel formato fbx, attivando l'override Skeleton/mesh in modo da poter selezionare lo scheletro compatibile per il retargeting su Unreal.



Figura 39 Rokoko Studio esportazione animazione

Su Unreal ho importato gli fbx con i movimenti corretti, tolto il tick dall'import della mesh, e impostato lo skeleton per l'importazione SK_Mannequin_Skeleton (preso dal third person pack importato nella prima fase di creazione del progetto). In seguito, ho preso il retargeter presente nella cartella character (sempre importata precedentemente nel third person pack) e ho aperto il file "Rtg__UE4Manny_UE5Manny".

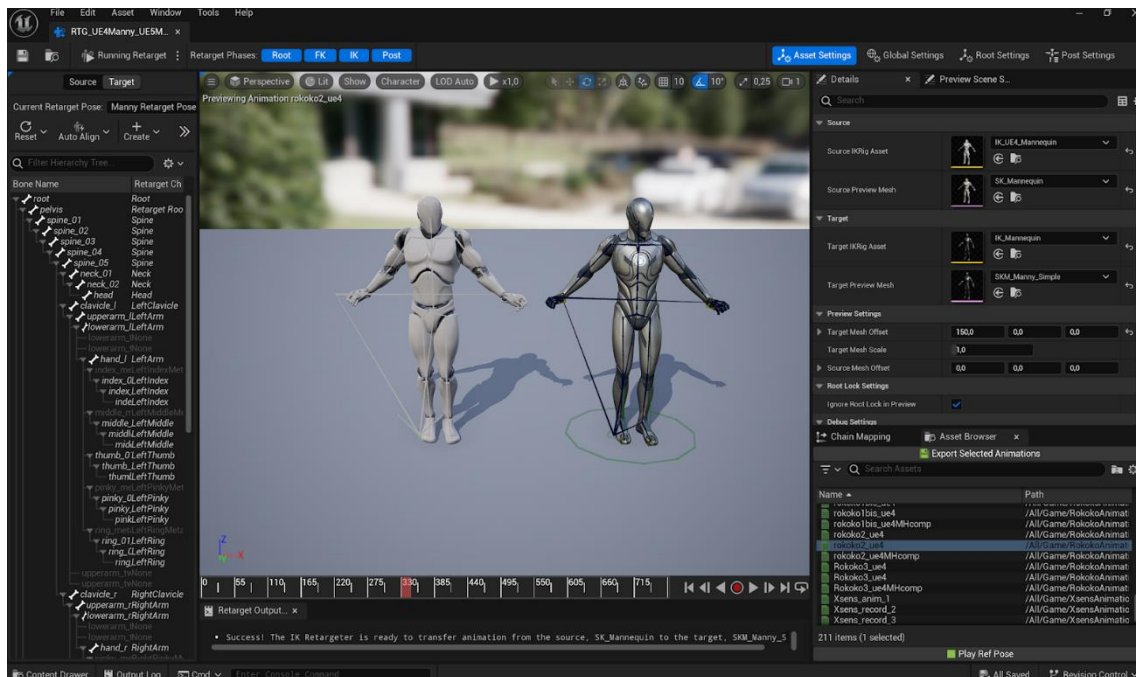
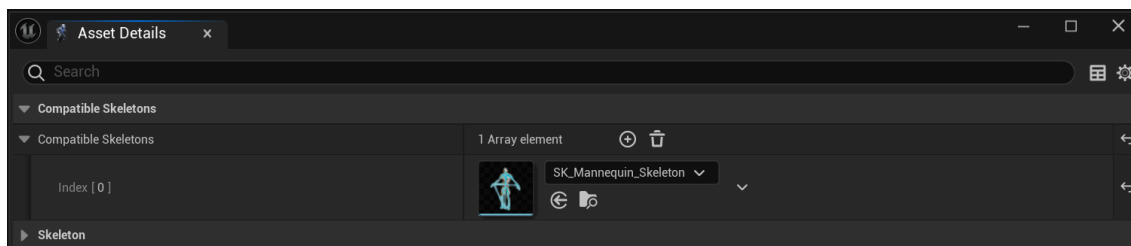


Figura 40 Retargeting animazione su scheletro Metahuman

Nella finestra in basso a destra ho selezionato le animazioni sulle quali effettuare il retarget e ho esportato le animazioni nella cartella RokokoAnimations. Tuttavia, persiste ancora una questione: il software non è in grado di determinare se le nuove animazioni esportate siano effettivamente compatibili con il sistema Metahuman. Per risolvere questo problema, bisogna aprire lo scheletro utilizzato dal Metahuman, windows, asset details e aggiungere lo scheletro SK_Mannequin_Skeleton.



Con questa operazione l'avatar può leggere l'animazione esportata con il retargeter nel modo corretto. A questo punto, grazie all'editor interno di Unreal, è stato possibile editare le animazioni nel modo corretto, andando a tagliare la parte precedente al ciak, così da poter sincronizzare l'animazione con le altre effettuate dagli altri sistemi.

6.3.2 Porting Xsens

Per effettuare il porting dell'animazione da Xsens su Unreal, il procedimento è stato leggermente differente dal precedente. Il software di per sé, senza alcuna licenza, non permette lo streaming in real time su altre piattaforme. Per risolvere questo problema, ho dovuto esportare tutto il progetto di Unreal, caricarlo su un pc avente la licenza per lo streaming di Xsens e aprirlo nuovamente insieme al software della Movella, attivando inoltre il plugin su Unreal. Prima di far dialogare il live streaming di Xsens con il Metahuman su Unreal, ho dovuto preparare il Metahuman.

Inizialmente ho dovuto trovare lo scheletro base sul quale si appoggia il blueprint del Metahuman nel content browser, una volta trovato, ho aperto il menù contestuale con il tasto destro, e sotto la voce di Xsens ho creato un asset di remap nella stessa cartella. Una volta aperto il nuovo file ho selezionato la name convention (Unreal/Metahuman) e ho selezionato una T-pose dall'elenco di animazioni disponibili (le quali sono state importate in precedenza con il third

person pack). Se non fosse disponibile nessuna T-pose, occorrerebbe farsene una ad hoc per il modello che si vuole animare. Dopodiché bisogna andare sulla voce LiveLinkMVNRemap e selezionare Metahuman base skeleton.

Una volta effettuate queste operazioni, si crea un animation blueprint con lo skeleton di partenza del Metahuman e si inserisce nell'anim graph il nodo Live Link Pose, selezionando come input il sistema Xsens rilevato nel Live Link. Una volta compilato il blueprint e salvato, è possibile vedere lo streaming da parte del software di Xsens tramite Live Link su Unreal in tempo reale, selezionando la stessa porta tra i due programmi, potendo effettivamente vedere l'animazione su entrambi i software aperti in simultanea. In questo modo ho registrato l'animazione direttamente su Unreal e in un secondo momento ho effettuato il taglio sul ciak allo stesso modo della tuta Smartsuit Pro 2.

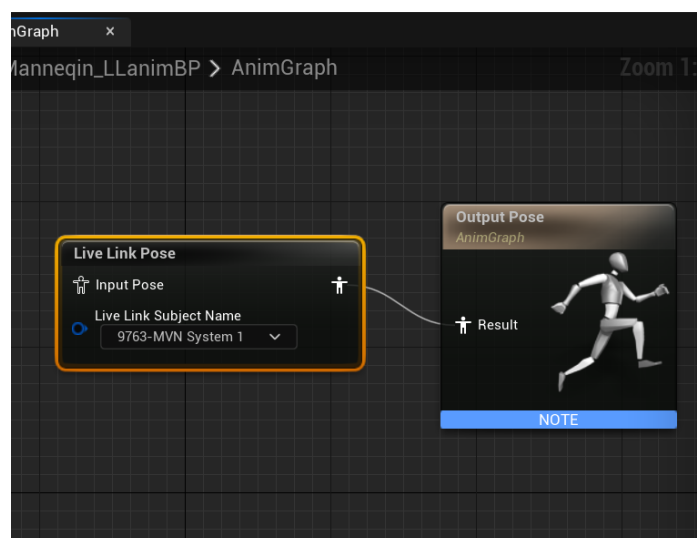


Figura 41 LiveLink Pose di Xsens per effettuare lo stream in tempo reale su Metahuman

6.3.3 Porting Move One

Una volta scaricato l'fbx dai server di MoveOne, è necessario innanzitutto creare una cartella nella Content Folder, nominarla MoveOne e inserire (tramite esplora risorse esterno, non drag & drop nell'engine) le risorse date direttamente dall'azienda. Creare all'interno una sottocartella chiamata animation nella quale inserire i file fbx con le animazioni per il confronto. Nella finestra di dialogo per l'importazione, spuntare la casella per importare le animazioni.

Dopodiché bisogna trovare il file IKMetahuman (su Unreal evidenziato giallo poiché è un rig), aprire il menù contestuale con il tasto destro, asset action e selezionare “edit selection in property matrix”. Una volta aperta la schermata, bisogna selezionare nella preview mesh lo scheletro del Metahuman che si vuole utilizzare, inserirlo e salvare.

Infine, bisogna aprire il file blu “MoveOne_Metahuman” (ik retargeter), selezionare gli scheletri corretti in source e target, selezionare l’animazione che si vuole convertire, effettuare il play, verificare che sia tutto corretto, e poi cliccare su Export Selected Animation, selezionare la cartella di salvataggio e procedere con l’esportazione.

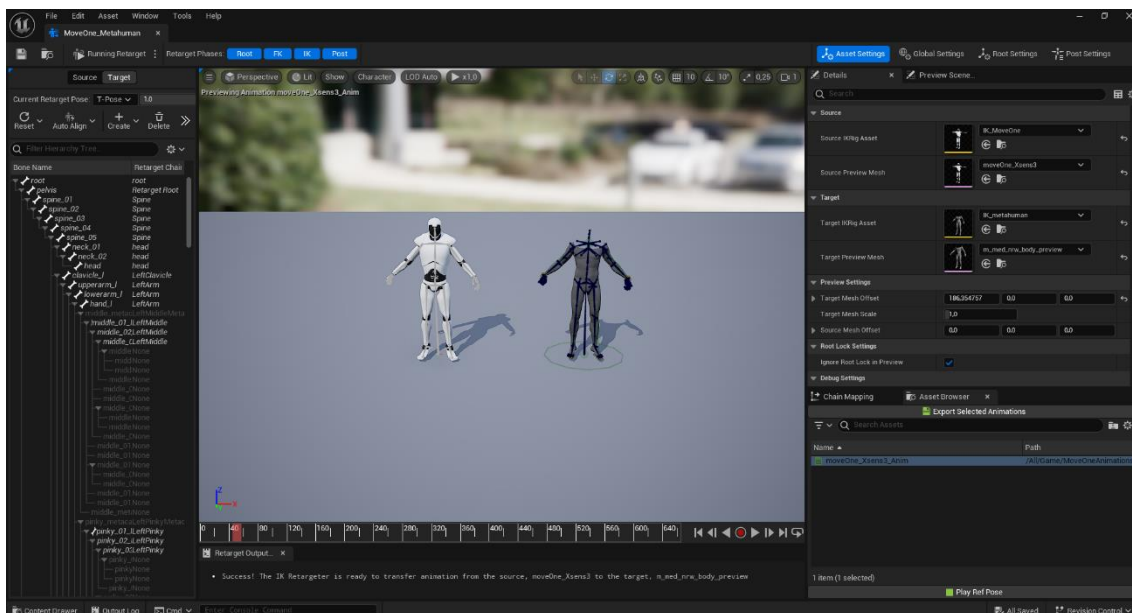


Figura 42 MoveOne processo di retarget

6.3.4 Possibilità di correzioni

Una volta effettuati tutti i porting delle animazioni, è possibile tramite un workflow molto leggero e non invasivo, effettuare delle correzioni sull’animazione, in modo da correggere o modificare la posizione di un singolo osso, errori di camminata o di qualsivoglia tipo.

Inizialmente si crea un level sequence, nel quale si inserisce il Metahuman con l’animazione da modificare. Si seleziona il [+] nel body, bake to control rig. Si

cancella l'animazione di partenza e si aggiunge un additive layer sull'animazione creata tramite il bake. In questo modo si possono effettuare correzioni senza perdere l'animazione di partenza, ma aggiungendole "sopra", permettendo di avere un blending di diverso tipo tra un'animazione e l'altra. Questo processo è stato utilizzato non per correggere un qualche movimento specifico ma solamente per spostare l'offset totale dell'animazione, in modo che fosse più centrale e quindi più sovrapponibile possibile con le altre.

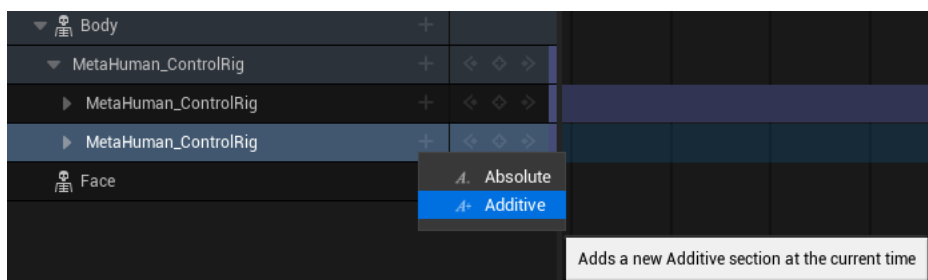


Figura 43 Modifiche additive nel Level Sequence

6.3.5 Conversione movimento Htc su controller IK

Abbiamo successivamente anche trasferito i movimenti dei cubi rappresentanti il sistema Htc su un IK rig di un manichino, per poter effettivamente vedere come si poteva riuscire a ricreare il movimento dell'intera persona solamente con 7 punti di controllo. Per ottenere questo risultato abbiamo esportato i movimenti dei cubi da Unreal su Blender; lì abbiamo prima scaricato un manichino da Mixamo; quindi, abbiamo creato un rig IK automatico con l'add-on di Mixamo su Blender. Dopodiché è stato effettuato il primo retarget, sovrapponendo i cubi di partenza con i controller IK del manichino, andando ad aggiustare tutti gli errori derivanti dall'utilizzo di sistemi di coordinate differenti. Una volta effettuate le correzioni, il manichino si muoveva a dovere, ma, senza l'animazione dell'osso radice, non poteva muoversi propriamente con tutto il corpo nello spazio. Per sopperire a questo problema, non avendo un tracker adibito alla cattura del movimento del bacino, abbiamo optato per trasferire il movimento sull'asse X,Y della posizione della testa, potendo quindi sfruttare la posizione, ricreando un movimento completo.

Abbiamo esportato il manichino con l'animazione in formato “.fbx” su Unreal. Per far avvenire l'importazione su Unreal nel modo corretto, è stato importato in un primo momento solamente il manichino senza l'animazione, e solo successivamente le diverse animazioni, così da poter associare il manichino proprietario all'animazione. Infine, abbiamo sfruttato l'aggiornamento della versione 5.4 di Unreal, la quale, data un'animazione su un certo manichino, è in grado di effettuare un retarget su un'altra tipologia di scheletro. Questa funzione può essere sfruttata al meglio con scheletri umanoidi. In questo modo l'animazione è stata trasferita sullo scheletro del Metahuman dove sono stati effettuati tutti i paragoni con i diversi sistemi di motion capture.

Le animazioni risultanti da questo processo sono risultate non delle migliori rispetto agli altri sistemi di motion capture, con evidenti errori riguardo la risoluzione del movimento delle ginocchia nel salto e nella posizione del bacino (non essendoci un vero e proprio tracker per quella porzione di scheletro).

Considerando però una possibile ottimizzazione del procedimento e di pulizia dell'animazione, è possibile lo stesso utilizzare questo sistema per effettuare semplici animazioni.

6.4 Acquisizione dati

In questo capitolo, analizzeremo dettagliatamente la procedura di acquisizione e analisi dei dati. Inizieremo con una spiegazione della struttura di acquisizione dati, descrivendo i metodi utilizzati. Successivamente, illustreremo il Blueprint level per HTC, evidenziando i passaggi necessari per configurare e utilizzare correttamente l'attrezzatura HTC. Poi, passeremo all'AnimationBaker Blueprint per poter effettuare leggere nel modo corretto le animazioni registrate con le tute, sottolineando le differenze rispetto all'HTC, come le specifiche modalità di acquisizione e gestione dei dati. Proseguiremo con la descrizione del metodo usato per la generazione del file json e il salvataggio delle posizioni su di esso, fornendo un esempio per chiarire il formato e il contenuto dei dati salvati. Infine, discuteremo la modularità e la scalabilità del sistema, evidenziando come la

struttura possa essere facilmente adattata e ampliata per includere ulteriori dispositivi o per gestire diverse esigenze di acquisizione dati.

Una volta allineati tutti gli scheletri rappresentanti i diversi sistemi di motion capture con il centro del mondo, in modo da far coincidere il più possibile la posizione di ogni singolo scheletro (in base alle registrazioni effettuate, senza alterare le posizioni delle singole ossa), ho potuto procedere con la parte di cattura dei dati.

L'obiettivo da raggiungere è stato quello di salvare, per ogni frame, le informazioni relative a location e rotation di determinate ossa dello scheletro del Metahuman. Per poter descrivere nel dettaglio le modalità di acquisizione delle posizioni, è necessario prima spiegare la struttura su come verranno salvati i dati.

In una cartella sono stati creati 3 file Blueprint di tipo Structure. Ho scelto di utilizzare questo tipo di file poiché risultano molto utili per raggruppare diverse variabili sotto un'unica entità. In questo modo ho potuto creare un'organizzazione a 3 livelli di struttura: il livello superiore è stato chiamato AnimationData, al suo interno è presente un unico array chiamato Frame Struct Data, al cui interno sono presenti due variabili: una variabile Int e un array Struct chiamato Bone Data, al cui interno sono presenti una variabile String contenente il nome del singolo osso, un vettore Location contenente i valori di posizione per ogni asse rispetto al centro del mondo, e un vettore Rotation contenente le rotazioni locali per ogni asse, sempre relativo al singolo osso. Di seguito uno schema per poter comprendere meglio.

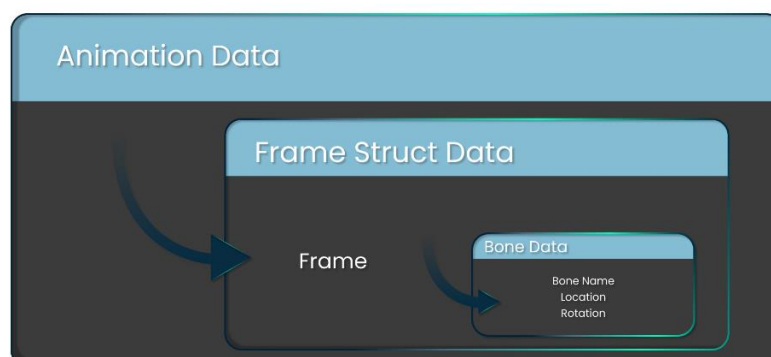


Figura 44 Schema della struttura di acquisizione dati

Un secondo passaggio necessario in preparazione dell'acquisizione dati vera e propria è stato dato dalla creazione di 4 Animation Blueprint. Questa tipologia di file permette di leggere e mandare in play su uno scheletro una determinata animazione compatibile con quello scheletro.

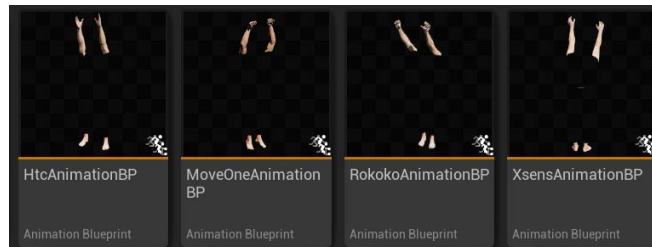


Figura 45 Animation Blueprint

Essa è risultata fondamentale, poiché tramite questa tipologia di blueprint è stato possibile eseguire l'animazione all'interno dell'AnimationBaker Blueprint, e ottenere un'organizzazione più semplice e lineare.

La registrazione dei dati avviene direttamente su Unreal Engine 5, passando dalla game mode. Questo passaggio insieme al comando tramite linea di comando "t.MaxFPS 30" permette di uniformare tutte le animazioni a 30 fps.

Una volta creata la struttura di acquisizione dati, per la registrazione dei dati ho utilizzato un actor blueprint personalizzato nominato AnimationBaker.

Questo blueprint consente di registrare le seguenti informazioni dettagliate:

1. **Frame di riferimento:** La posizione temporale all'interno dell'animazione.
2. **Nome dell'osso corrispondente:** Identificazione specifica di ciascun osso all'interno dello scheletro.
3. **Vettore di posizione (Location):** Coordinate X, Y, Z della posizione dell'osso rispetto all'origine del mondo.
4. **Vettore di rotazione (Rotator):** Angoli di rotazione interni dell'osso espressi in termini di pitch, roll e yaw.

Il blueprint permette di effettuare sia una cattura completa delle ossa dello scheletro, sia la creazione di un array personalizzato dove si possono scegliere i nomi degli ossa di interesse, evitando così la necessità di una cattura totale. Inoltre, il blueprint offre la possibilità di selezionare il numero di frame

dell'animazione da salvare, rendendo il processo di registrazione flessibile e adattabile alle diverse esigenze sperimentali.

I dati raccolti vengono immediatamente salvati in un file json, strutturato in modo corretto per garantire l'integrità e la facilità di accesso alle informazioni. Questa configurazione consente una raccolta dati efficiente e accurata, facilitando l'analisi successiva delle animazioni scheletriche. Nel dettaglio lo schema è principalmente formato da due funzioni e l'event graph.

L'*event graph* è la sezione principale del blueprint, nella quale è possibile una programmazione grafica tramite nodi.

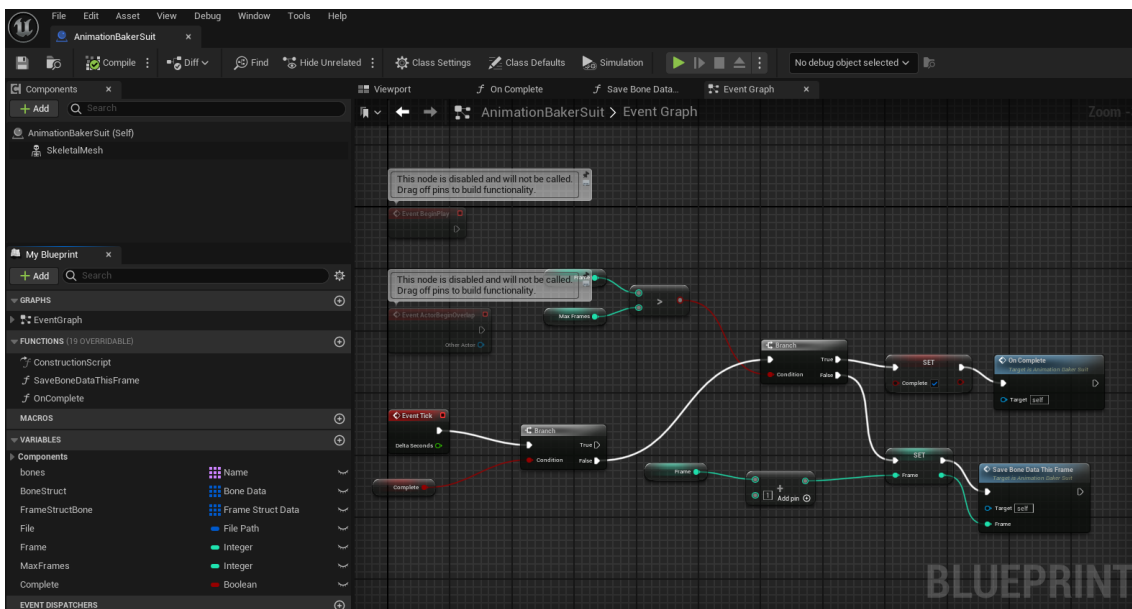


Figura 46 Blueprint event graph

Una volta eseguito il blueprint, per ogni tick (corrispondente a ogni frame) che cambia nella game mode, viene attivata la funzione SaveBoneDataThisFrame, la quale riceve in ingresso il frame corrente. Per controllare che la funzione venga eseguita per tutti i frame necessari, è stato creato un piccolo contatore con relativo: “if contatore > Max Frames”, se il risultato è negativo, allora procede al frame successivo e continua il salvataggio, mentre se il risultato è positivo, quindi è stato raggiunto l’obiettivo di frames da registrare, nel nostro caso 300, lo switch passa alla funzione OnComplete, la quale va a completare la struttura dati creata in precedenza convertendola e salvandola in un file json.

La prima funzione creata si chiama **SaveBoneDataThisFrame**:

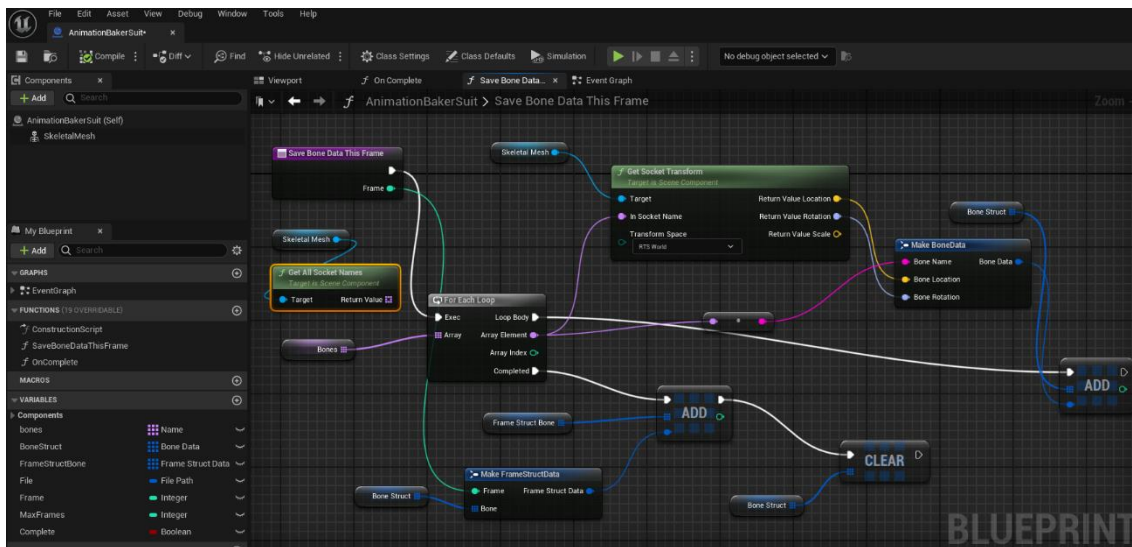


Figura 47 Blueprint funzione *SaveBoneDataThisFrame*

Si può dividere la schermata dei nodi in due parti, la parte superiore e la parte inferiore. Nella parte superiore è presente il primo nodo di input, con un'unica variabile frame che viene passata dall'esterno. Inizia subito un loop che passa tutti gli elementi di un determinato array. È possibile scegliere se analizzare tutte le ossa della skeletal mesh attraverso get All Socket names, oppure scegliere di analizzare solo determinate ossa aggiunte manualmente nell'array Bones.

Il nodo For Each Loop ha 4 uscite. L'uscita Array element è quella più importante che funziona da filtro per il nodo successivo Get Socket Transform. Questo nodo ha come input la skeletal mesh (sulla quale viene effettuata l'analisi) e il Socket name (nome dell'osso da analizzare). In output restituisce 3 vettori, uno di posizione, uno di rotazione e uno di scala. Nel nostro caso interessano solamente i valori di posizione (vengono salvati comunque anche i dati di rotazione per completezza del file). Da qui si passa al nodo MakeBoneData, nodo dove si va a popolare la prima struttura Bone Data per il singolo osso, salvando il nome e i due vettori descritti in precedenza. Dopodiché si procede ad aggiungere nella struttura superiore Frame Struct Data il frame di riferimento (quello ricevuto in ingresso nella funzione) insieme alla struttura Bone Data, in modo da creare un array. Per poter salvare correttamente tutti i valori è necessario aggiungere alla

fine una funzione che elimini i valori appena scritti sulla struttura Bone Data, così da poter salvare nuovi dati relativi a un nuovo osso.

Funzione **OnComplete**:

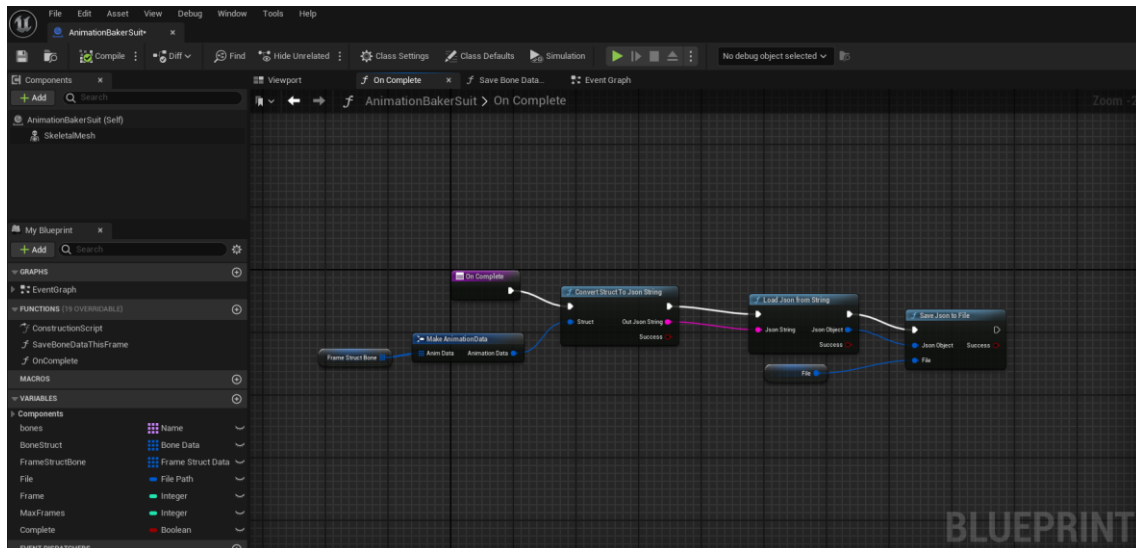


Figura 48 Blueprint funzione OnComplete

Questa funzione permette di finalizzare la scrittura della struttura e salvarla su un file json. Per questo motivo sfrutta il plugin json Blueprint Utilities, permettendo di usare questi nodi altrimenti non presenti. Viene incapsulata la struttura Frame Bone Struct in Animation Data. Viene richiamato quindi un nodo per tradurre inizialmente tutta la struttura in una json String, così da poter usare subito dopo un nodo in grado di convertire questa stringa json in un oggetto json. Questo passaggio è necessario poiché il nodo Save json to File ha due input: un json object e un file. Il primo è dato dalla conversione del nodo precedente, mentre il secondo è il percorso sul computer dove si intende salvare il file.

Blueprint level per Htc:

Per salvare i dati relativi alle posizioni dei cubi approssimanti il sistema Htc si è dovuto intervenire direttamente nel Blueprint del livello di gioco. Questo ha permesso di poter importare le singole shapes approssimanti determinate ossa dello scheletro direttamente dentro il blueprint. Il funzionamento dell'acquisizione differisce per due aspetti rispetto al classico Animation Baker: il motivo principale

risiede nel fatto della differenza del sistema in utilizzo. In questo caso sono dei movimenti di cubi di riferimento registrati in un level sequence. Questo ha portato a utilizzare il nodo Play Current Level Sequence, così da poter avviare all'inizio dell'evento blueprint una determinata sequenza (contenente le animazioni registrate in precedenza con il sistema Htc). In secondo luogo, è presente anche una differenza secondaria nella funzione equivalente Save Bone Data, qui rinominata Save Shape Data. Non essendoci un vero e proprio scheletro e non dovendo effettuare l'analisi di tutte le ossa ma solo di quelle principali, ho deciso di aggiungere manualmente tutte le shapes all'interno del blueprint e di crearne un array. Da questo punto in poi il blueprint funziona allo stesso modo dell'animation baker spiegato nel paragrafo precedente.

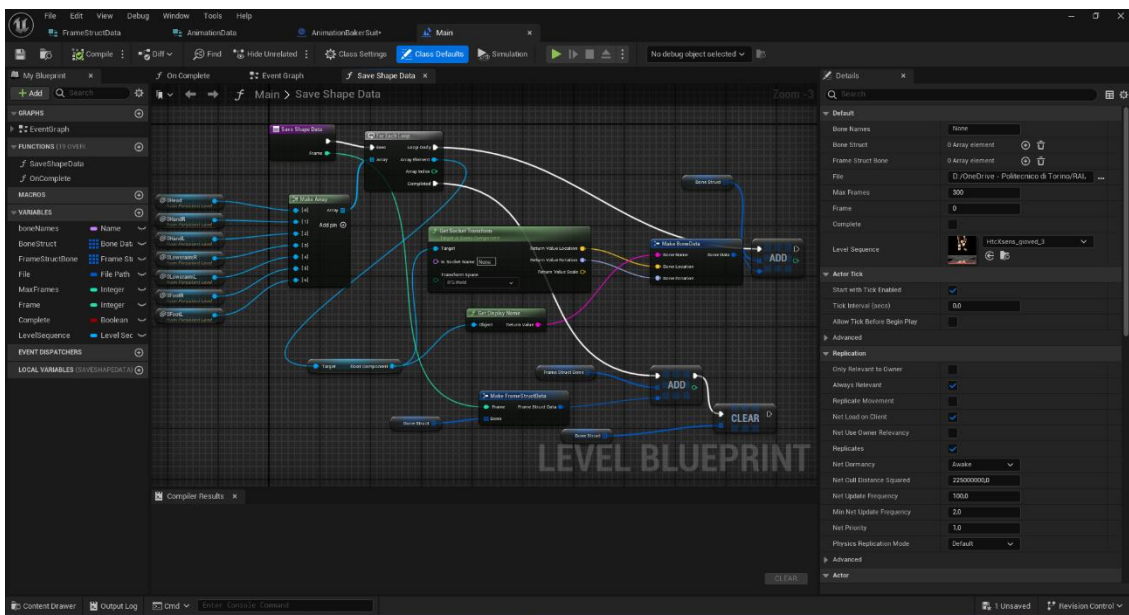


Figura 49 Blueprint level per acquisizione Htc

In generale il punto di forza di questo sistema di paragone consiste nel poter intercambiare molto facilmente animazioni e scheletri differenti. Questo permette di poter sfruttare le stesse strutture di acquisizione con sistemi di motion capture differenti, potendo registrare e salvare i dati nel modo corretto. È possibile, inoltre, scalare e modulare il sistema, aggiungendo o diminuendo il numero di AnimationBakerBP. Il salvataggio su file json permette inoltre di avere un file rapido da leggere e facilmente analizzabile con strumenti esterni. Nel capitolo seguente verranno illustrati e analizzati i file json.

7 Analisi dati

Analisi dell'errore

In questo capitolo si presenta l'analisi dei dati rilevati dalle misurazioni effettuate sul campo.

Ho suddiviso i risultati in 3 parti, una parte per ogni tipologia di test effettuato: la prima parte si concentrerà su una misura statica, come la T-pose e due dinamiche come camminata e salto.



Figura 50 Schema delle ossa prese in considerazione

Per quanto riguarda la prima parte relativa alla posizione della Tpose mi sono concentrato sull'analisi dell'errore relativo di ogni singolo osso dei diversi sistemi, così da poter avere un'idea della qualità del tracking di ogni sistema.

Il campione di frame preso in analisi è di 90 frames (corrispondenti a 3 secondi) per entrambe le registrazioni effettuate (Xsens e Rokoko). Inizia al frame 90 e finisce al frame 180 per la comparazione dei sistemi con Xsens, mentre inizia dal frame 110 e termina al frame 200 per quanto riguarda la registrazione con

Rokoko. Questo range di frame corrisponde al momento in cui la posizione dell'attore è stata mantenuta statica in T-pose in entrambi i casi.

L'analisi del movimento effettuata, nonostante sia presente un errore intrinseco di movimento per ogni sistema dovuto al fattore umano (basti pensare che nonostante una persona voglia stare immobile, in qualsiasi caso effettua dei micromovimenti per rimanere in equilibrio) permette comunque di effettuare un confronto. I sistemi sono stati posizionati in modo da registrare il movimento in contemporanea, permettendo quindi di registrare a loro volta i micromovimenti della persona.

Come da aspettative, i grafici relativi al sistema Htc sono approssimabili a una linea retta, notando un effettivo migliore tracciamento, considerato punto di riferimento per questi paragoni essendo un sistema di tracking ottico. In particolare viene preso ad esempio il tracciamento dell'osso relativo alla testa.

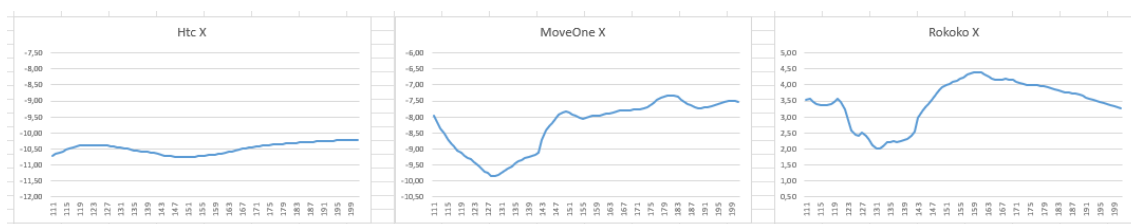


Grafico 1 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse X dell'osso della testa

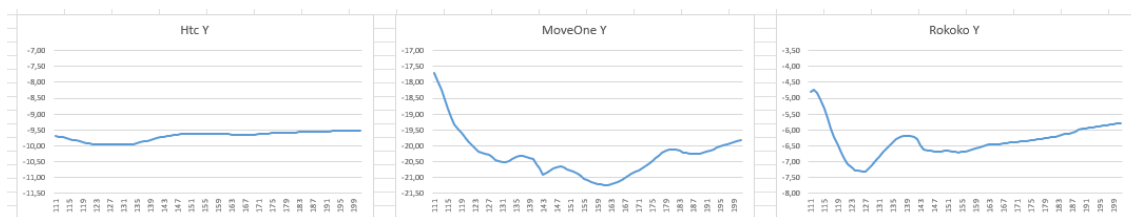


Grafico 2 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse Y dell'osso della testa

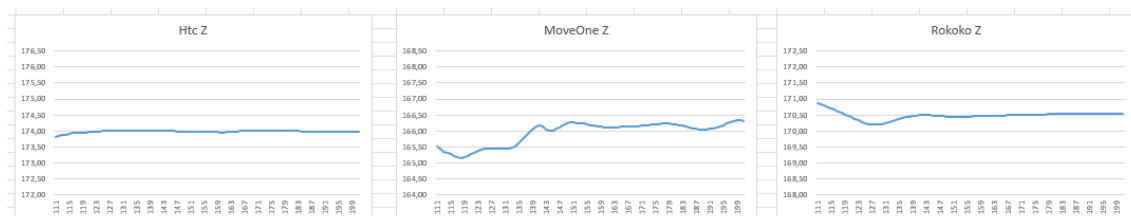


Grafico 3 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse Z dell'osso della testa

I grafici che andiamo a prendere in considerazione hanno sull'asse orizzontale la successione dei frame nei quali è stata catturata l'animazione sincronizzata per tutti i sistemi, e sull'asse verticale l'effettivo valore registrato, espresso in centimetri. Ogni grafico rappresenta un singolo asse di movimento relativo ad ogni sistema di tracciamento.

Si può notare dai grafici successivi (Grafico 5, 6) che è presente però un abbassamento relativo della qualità del sistema di riferimento Htc e di conseguenza un rialzo del rumore relativo nella parte bassa del tracking, con i piedi. In particolare questo aumento dell'errore è avvenuto durante la comparazione con la tuta Rokoko del piede destro.

Questo rumore per il sistema Htc è causato principalmente dal materiale del pavimento dove sono stati effettuati i test, il quale non era ottimale e leggermente riflettente, ragione per cui ha portato a lievi interferenze in fase di triangolazione del movimento. Tale fenomeno però non è stato riscontrato in modo costante: per esempio nella stessa misurazione ma con il piede sinistro non è stata rilevata nessuna interferenza di rilievo, rendendo così molto stabile il tracciamento.

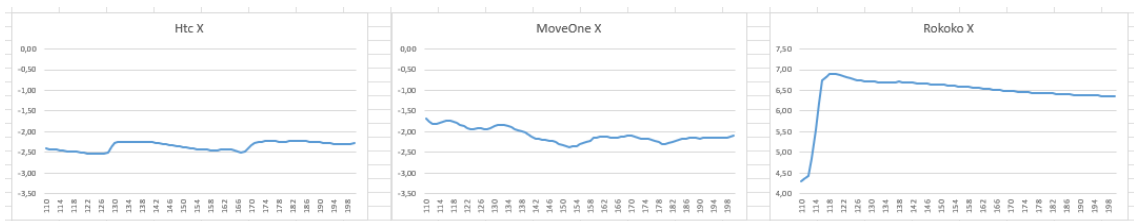


Grafico 4 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse X dell'osso del piede destro

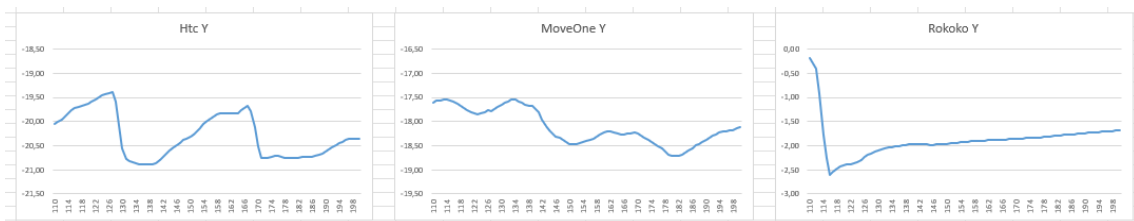


Grafico 5 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse Y dell'osso del piede destro

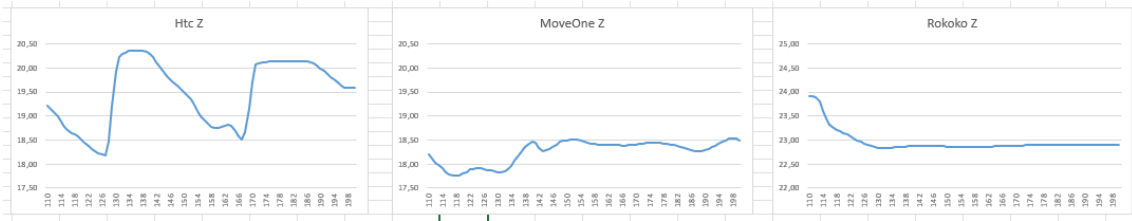


Grafico 6 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse Z dell'osso del piede destro

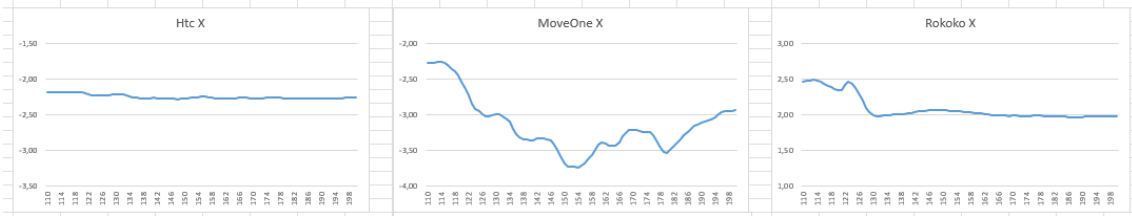


Grafico 7 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse X dell'osso del piede sinistro

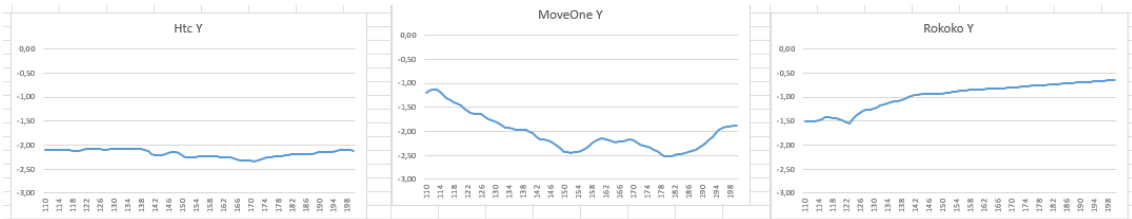


Grafico 8 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse Y dell'osso del piede sinistro

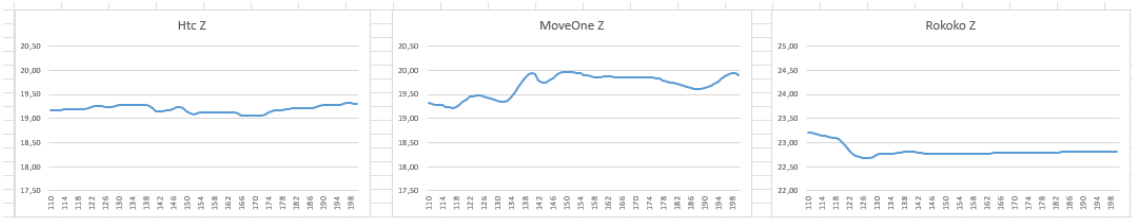


Grafico 9 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse Z dell'osso del piede sinistro

Prendendo in analisi la mano destra come esempio per gli arti superiori, si può notare una riduzione notevole dell'errore per quanto riguarda il sistema Htc, riuscendo a mantenere un tracciamento molto lineare e costante rispetto ai corrispettivi MoveOne e Rokoko.

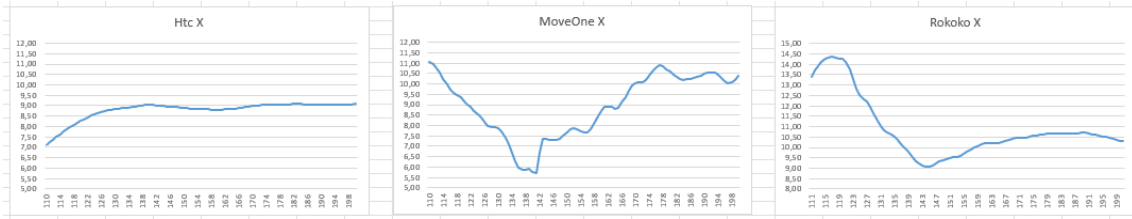


Grafico 10 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse X dell'osso della mano destra

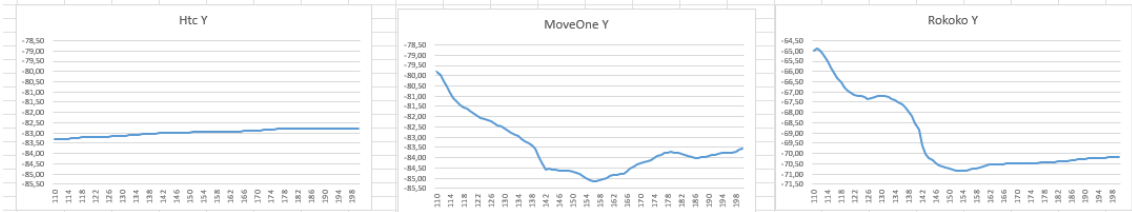


Grafico 11 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse Y dell'osso della mano destra

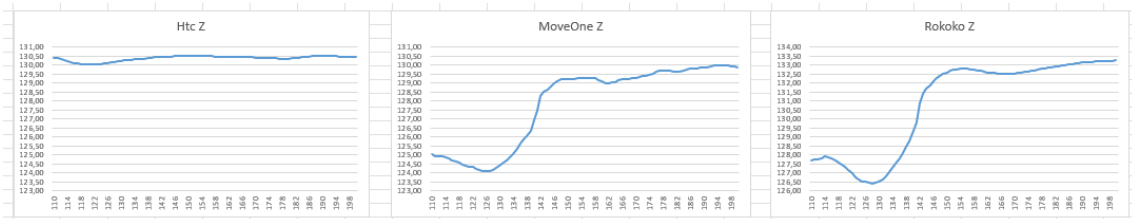


Grafico 12 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse Z dell'osso della mano destra

Resta molto interessante osservare il paragone anche con la tuta di Xsens, che riesce a raggiungere risultati molto buoni, riducendo al minimo il rumore, arrivando a una curva molto somigliante con Htc. Nei grafici si può notare il comportamento nel caso del piede sinistro e della mano destra:

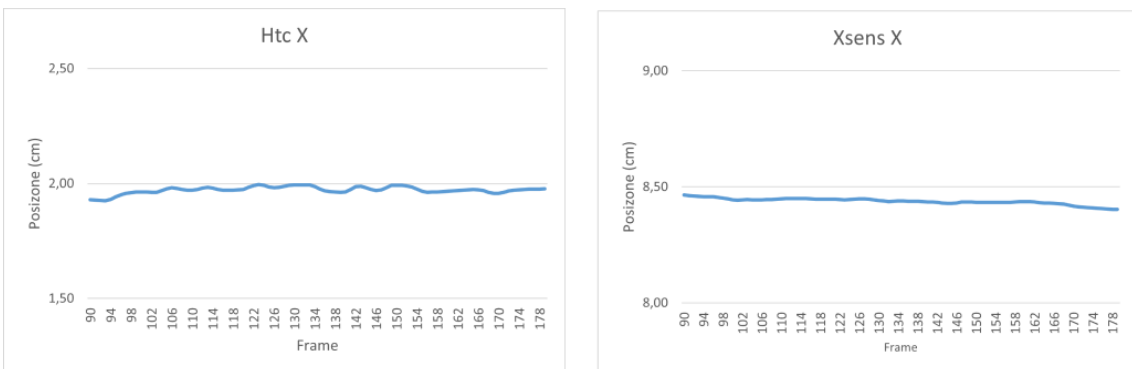


Grafico 13 Comparazione sistemi Htc e Xsens relativo all'asse X dell'osso del piede sinistro

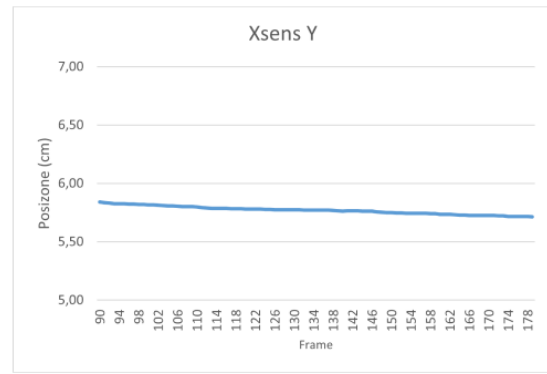
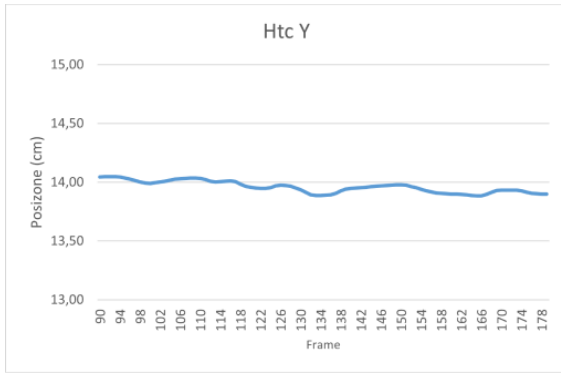


Grafico 14 Comparazione sistemi Htc e Xsens relativo all'asse Y dell'osso del piede sinistro

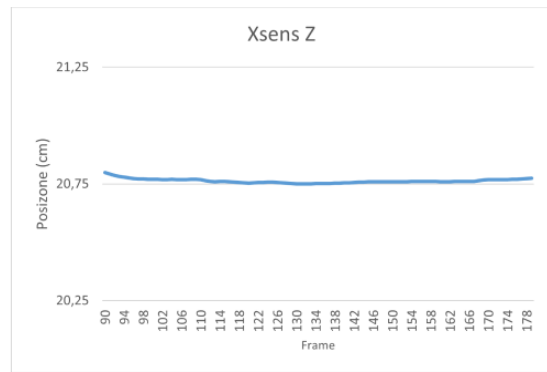
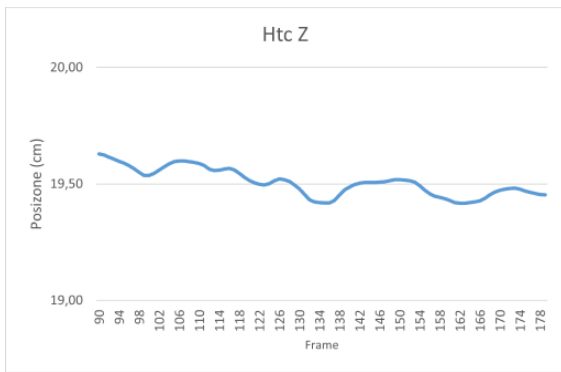


Grafico 15 Comparazione sistemi Htc e Xsens relativo all'asse Z dell'osso del piede sinistro

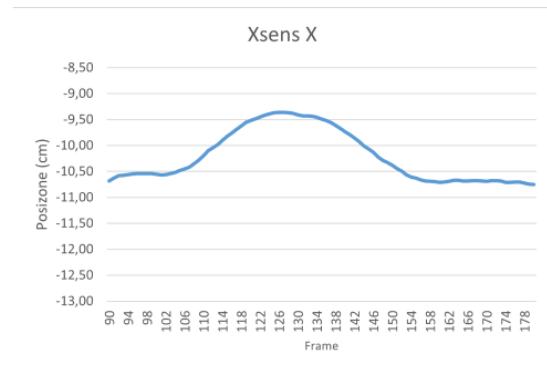
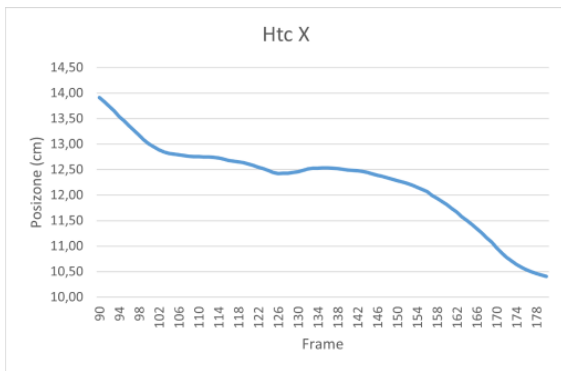


Grafico 16 Comparazione sistemi Htc e Xsens relativo all'asse X dell'osso della mano destra

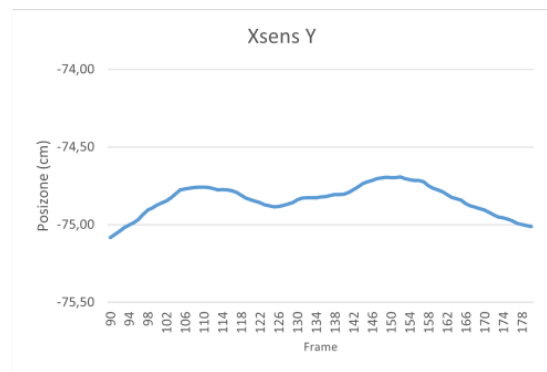
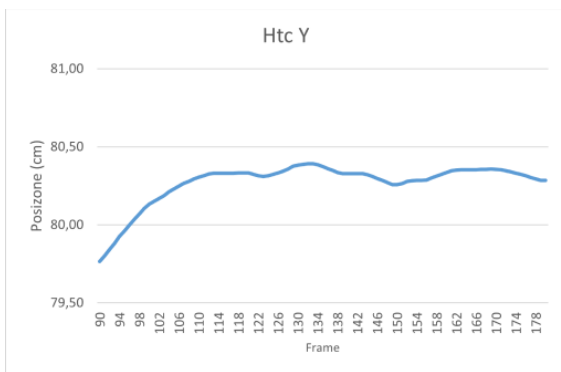


Grafico 17 Comparazione sistemi Htc e Xsens relativo all'asse Y dell'osso della mano destra

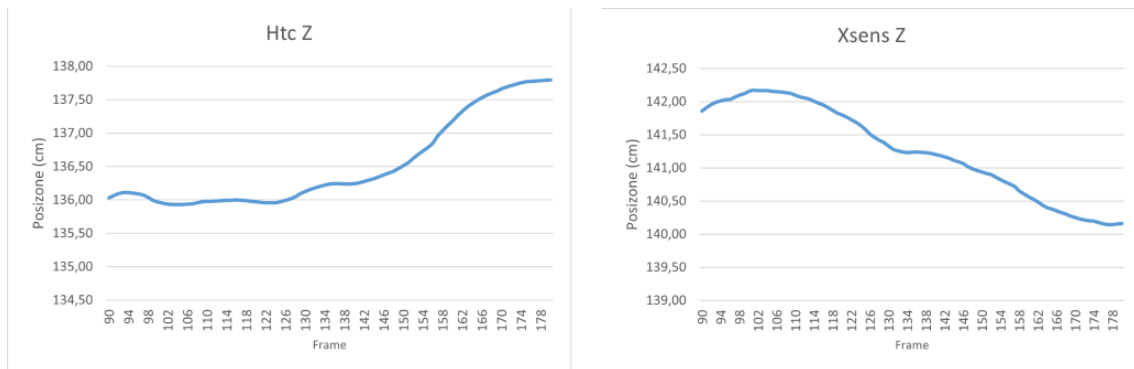


Grafico 18 Comparazione sistemi Htc e Xsens relativo all'asse Z dell'osso della mano destra

Per semplicità e per non essere ridondanti sono riportati solo alcuni assi a scopo di esempio. Tutti gli altri assi sono stati comunque analizzati e confrontati, ma avendo risultati simili a questi abbiamo ritenuto opportuno per una spiegazione più fluida non inserirli.

Passando all'analisi del tracciamento del sistema MoveOne, si può notare un maggiore rumore generale in tutti i tracciamenti effettuati. Da notare una fluttuazione rilevante di tutti gli assi durante la rivelazione assieme alla tuta Xsens, (andando a prendere per esempio il grafico sempre della mano destra):

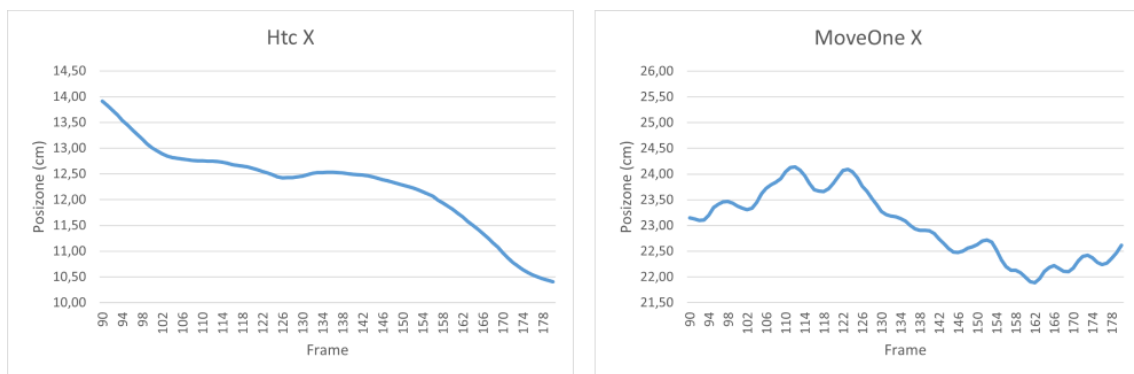


Grafico 19 Comparazione sistemi Htc e MoveOne relativo all'asse X dell'osso della mano destra sessione 1

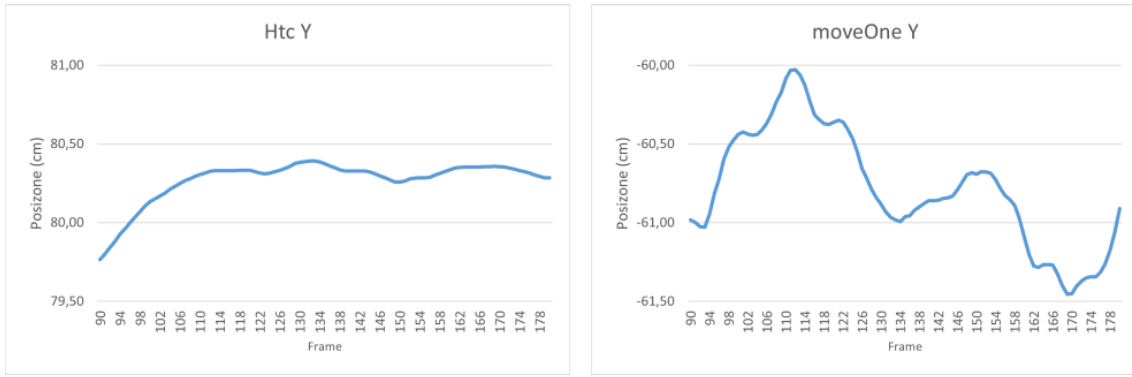


Grafico 20 Comparazione sistemi Htc e MoveOne relativo all'asse Y dell'osso della mano destra sessione 1

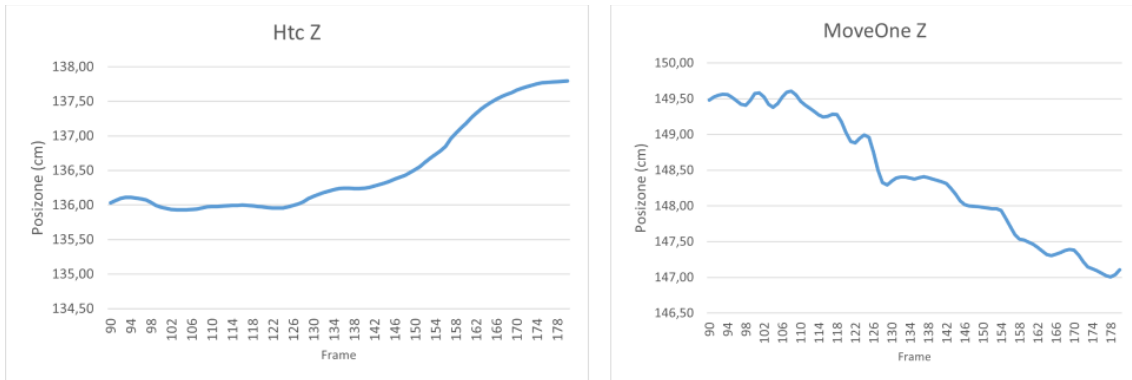


Grafico 21 Comparazione sistemi Htc e MoveOne relativo all'asse Z dell'osso della mano destra sessione 1

Questo risultato di MoveOne è confermato anche durante la sessione di registrazione con la tuta Rokoko come si può vedere dai grafici (Grafico 22, 23, 24), ottenendo un rumore molto evidente su diversi assi, i grafici riportati sono sempre relativi alla mano destra:

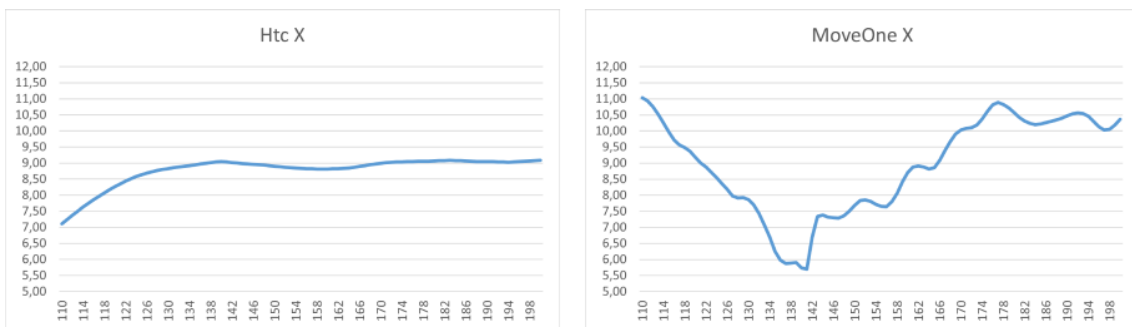


Grafico 22 Comparazione sistemi Htc e MoveOne relativo all'asse X dell'osso della mano destra sessione 2

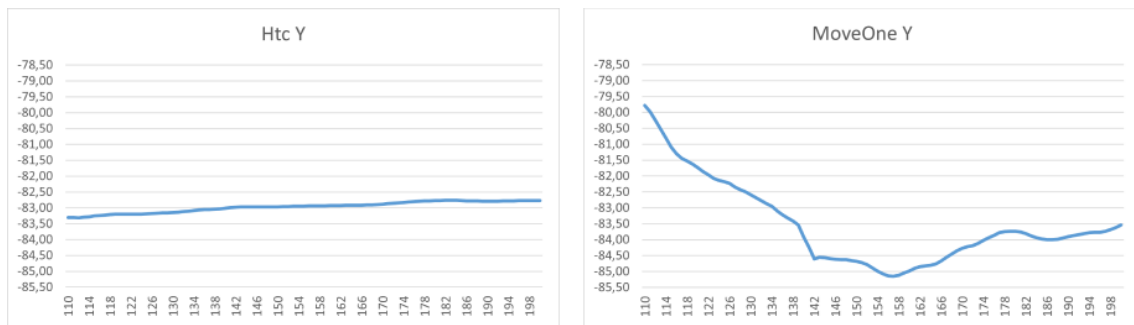


Grafico 23 Comparazione sistemi Htc e MoveOne relativo all'asse Y dell'osso della mano destra sessione 2

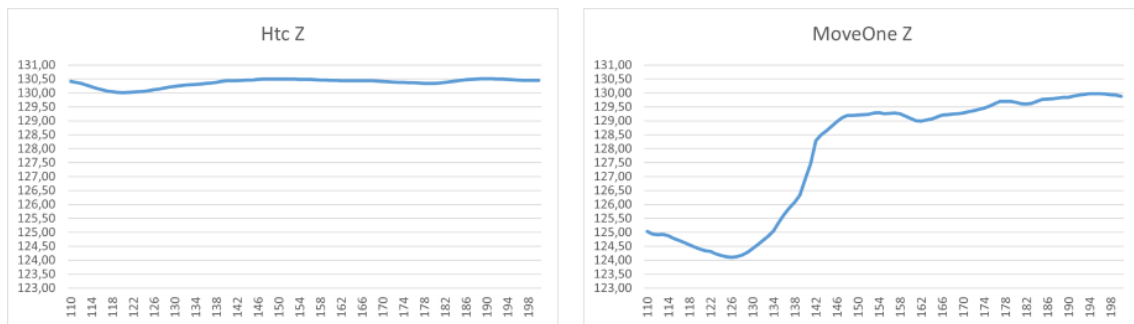


Grafico 24 Comparazione sistemi Htc e MoveOne relativo all'asse Z dell'osso della mano destra sessione 2

Concludendo l'analisi dell'errore stimato per tutti i sistemi e per le diverse ossa prese in considerazione, si può affermare una quantità di rumore inferiore delle tute Xsens e Rokoko rispetto al sistema Htc nella parte degli arti inferiori. Questo è dovuto principalmente alla tipologia del sistema: basandosi su sensori IMU, le tute restano immuni a interferenze dovute a condizioni di luce o di riflessione dei materiali, cosa che invece si è verificata con il sistema Htc. In tutti gli altri casi la soglia di errore del sistema risulta minima e molto attenuata rispetto agli altri sistemi.

Un'altra dinamica che è stata tenuta in considerazione riguarda l'indice di correlazione di Pearson (Equazione 1), che riguarda una possibile relazione lineare tra due insiemi di dati, espressa con un indice che varia tra 1 e -1. Se il risultato si avvicina a 1 allora è presente una forte correlazione lineare tra i due insiemi di dati; se si avvicina a -1, è presente anche in questo caso una forte correlazione lineare ma stavolta negativa. Se il risultato si avvicina allo 0 allora la correlazione lineare è di minore rilievo. Se coincide con 0 allora i due insiemi di dati sono completamente indipendenti. È formato dal rapporto tra la Covarianza

tra i due insiemi di dati e la moltiplicazione delle varianze dei singoli insiemi. Il coefficiente è calcolato sull'insieme delle ossa prese in analisi. Nel nostro caso abbiamo usato questo coefficiente per paragonare i diversi sistemi con Htc, risultato oggettivamente più preciso come spiegato in precedenza (Paragrafo 4.1).

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \cdot \sigma_Y}$$

Equazione 1 Indice di correlazione di Pearson

Inoltre, nelle tabelle seguenti relative all'indice di correlazione di Pearson, è stato messo a confronto il sistema Htc con la ricostruzione del suo movimento sul manichino tramite l'IK rig, ottenendo risultati di Pearson conformi e simili con gli altri sistemi dal punto di vista numerico ma non ottenendo una qualità adeguata dal punto di vista grafico, non riuscendo a ricostruire a dovere tutta la cinematica delle ossa, rendendo l'animazione molto imprecisa e necessaria di ulteriori correzioni.

In generale è stato riscontrato un coeff di Pearson molto buono per tutti i sistemi sull'asse Z, con valori intorno a 0,98 e 0,99. La situazione cambia notevolmente sugli altri assi. È stato riscontrato un valore negativo per quanto riguarda l'asse Y nel confronto con il sistema MoveOne e la tuta Xsens, mentre per il confronto dei sistemi Htc con MoveOne e Rokoko, non ha presentato valori negativi. Da segnalare solo un valore molto basso di 0,17 per quanto riguarda il coefficiente relativo all'asse delle X di Rokoko.

È stato anche effettuato un confronto del coefficiente di Pearson di tutte le ossa considerando il modulo dei vettori senza suddividerli nelle varie componenti, costituito dalla combinazione delle tre componenti X,Y,Z. Nelle tabelle è stato nominato **Pearson tot** e risulta molto elevato in tutte le misurazioni e in tutti i sistemi, partendo da un valore di 0,96 e arrivando a un valore di 0,99.

Tabella 3 Correlazione di Pearson dei sistemi MoveOne e Xsens rispetto a Htc (T-pose)

	MoveOne	Xsens	Htc_IK
Pearson X	0,906882075	0,716798907	0,892110101
Pearson Y	-0,386462142	-0,385148234	-0,396434362
Pearson Z	0,995055006	0,993149348	0,992733132
Pearson tot	0,996473798	0,993267077	

Tabella 4 Correlazione di Pearson dei sistemi MoveOne e Rokoko rispetto a Htc (T-pose)

	MoveOne	Rokoko	IK_Htc
Pearson X	0,881868133	0,17199934	0,917044298
Pearson Y	0,891207869	0,508485674	0,935001532
Pearson Z	0,987236905	0,990509072	0,991071942
Pearson tot	0,97544227	0,969382545	

Un altro valore preso in considerazione è il Root Mean Square Deviation (RMSD), indice usato per paragonare i valori effettivi misurati rispetto al sistema di riferimento Htc. In tutti i casi si è ottenuto un ottimo valore, suggerendo un buon livello di accuratezza riguardo il tracciamento del movimento da parte di tutti i sistemi, non sorpassando mai il valore 0,21.

Tabella 5 Valori RMSD di MoveOne e Xsens rispetto a Htc (T-pose)

T-pose RMSD

<i>MoveOne</i>	0,13
<i>Xsens</i>	0,16

Tabella 6 Valori RMSD di MoveOne e Rokoko rispetto a Htc (T-pose)

T-pose RMSD

<i>MoveOne</i>	0,21
<i>Rokoko</i>	0,21

Un'ulteriore analisi è stata compiuta per capire la dispersione dei valori misurati rispetto alla media: abbiamo calcolato le deviazioni standard. I risultati ottenuti per quanto riguarda questi primi dati relativi alle T-pose sono molto buoni,

mantenendo una deviazione standard per ogni asse molto bassa. Ad esempio, possiamo notare nell'analisi dell'osso della mano destra una deviazione mai superiore a 0,86 centimetri, raggiungendo un valore molto basso di 0,09 centimetri con la tuta Xsens sull'asse Y (Grafico 25). Inoltre, andando a considerare la sessione di tracking con la tuta Rokoko, il risultato di MoveOne risulta ben peggiore rispetto al confronto con Xsens, aumentando fino a 2,23 centimetri la deviazione standard sull'asse Z.

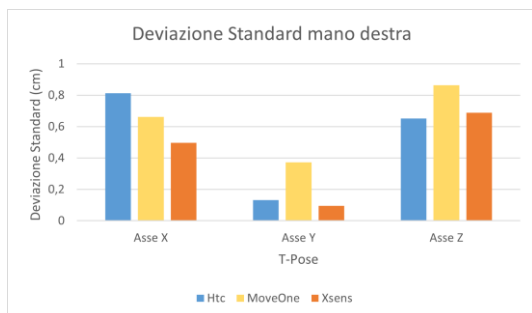


Grafico 25 Deviazione standard mano destra sessione 1 (T-pose)

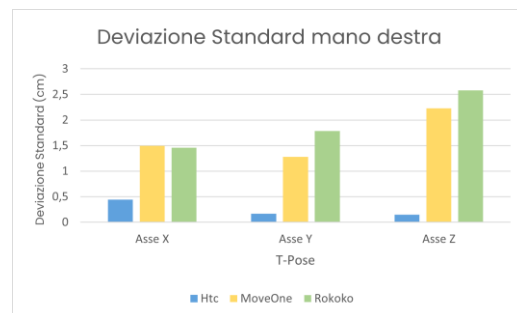


Grafico 26 Deviazione standard mano destra sessione 2 (T-pose)

Questo risultato risulta conforme con gli altri dati presi in analisi, andando a evidenziare il problema di tracciamento dell'altezza sull'asse Z da parte del sistema basato su intelligenza artificiale MoveOne.



Figura 51 Illustrazione assi cartesiani

Si può notare anche un problema di tracciamento che si può presentare ogni tanto sulle tute, come per esempio in questo caso nella tuta Rokoko, facendo salire notevolmente la deviazione standard relativa all'asse Z, arrivando fino a 2,58 centimetri. Di seguito vengono lasciati ulteriori dati di deviazione standard relativi ai diversi sistemi. In generale si può osservare una buona qualità con deviazioni minime in particolare sull'asse Z, andando a superare (Grafico 27) anche il sistema Htc sull'asse Z, assumendo un valore di 0,018 centimetri contro gli 0,245 centimetri di Htc. Per quanto riguarda gli altri assi si mantengono risultati

per la maggior parte conformi alle aspettative, con una deviazione standard spesso nel range del centimetro. In appendice si possono trovare tutti i dati relativi alla deviazione standard per le diverse ossa e tute.

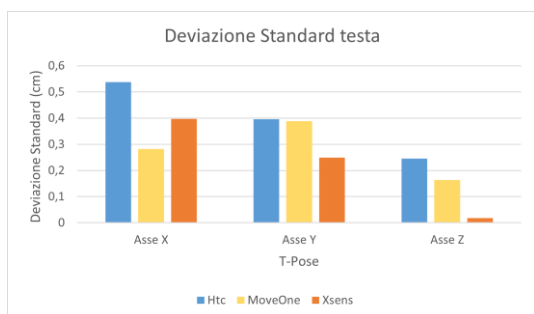


Grafico 27 Deviazione standard relativa alla testa nella sessione 1 (T-pose)

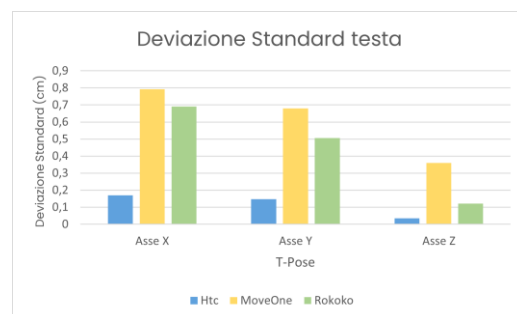


Grafico 28 Deviazione standard relativa alla testa nella sessione 2 (T-pose)

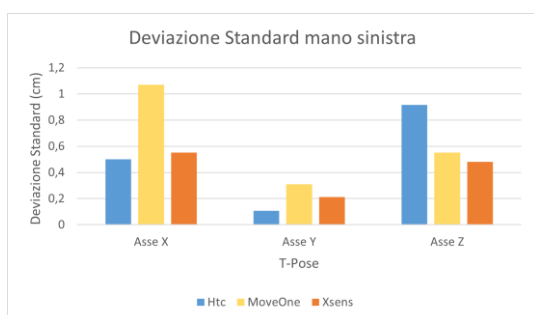


Grafico 29 Deviazione standard relativa alla mano sinistra nella sessione 1 (T-pose)

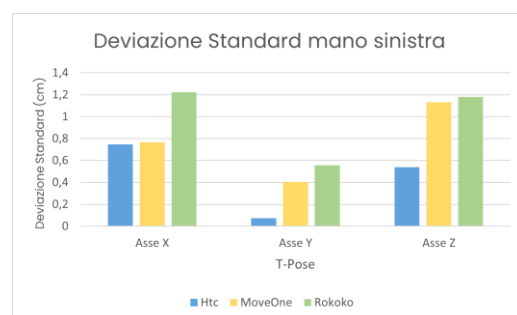


Grafico 30 Deviazione standard relativa alla mano sinistra nella sessione 2 (T-pose)

Procedendo con l'analisi relativa alla sessione effettuata per la camminata, emergono dati buoni per il Root Mean Square Deviation, con valori di confronto con il sistema Htc di soli 0,25 centimetri e 0,36 centimetri tra Htc e MoveOne. D'altro canto, risulta un valore simile ma leggermente più elevato per le due tute, con valori di 0,35 centimetri per il confronto con la Smartsuit Pro 2 di Rokoko e di 0,53 centimetri nel confronto con la tuta Xsens. Questi valori si possono notare molto simili negli ultimi test effettuati sul movimento del salto, con un incremento del valore di RMSD da parte della tuta Rokoko e una diminuzione di valore da parte della tuta Xsens, arrivando a un valore di 0,33 centimetri. Di seguito le tabelle riassuntive dei valori RMSD trovati:

Tabella 7 Valori RMSD di MoveOne e Xsens rispetto a Htc (Walk)

	RMSD
MoveOne	0,36
Xsens	0,53

Tabella 8 Valori RMSD di MoveOne e Rokoko rispetto a Htc (Walk)

	Walk RMSD
MoveOne	0,25
Rokoko	0,35

Tabella 9 Valori RMSD di MoveOne e Xsens rispetto a Htc (Jump)

	Jump RMSD
MoveOne	0,36
Xsens	0,53

Tabella 10 Valori RMSD di MoveOne e Rokoko rispetto a Htc (Jump)

	Jump RMSD
MoveOne	0,25
Rokoko	0,35

Dal punto di vista del coefficiente di Pearson, i risultati si sono rivelati ottimi nella camminata per MoveOne, ottenendo punteggi superiori rispetto a Rokoko. Possiamo notare che la Smartsuit Pro 2 ha un ottimo indice di correlazione per l'altezza, ma in questo test ha avuto un risultato molto basso sull'asse x con 0,17, indicando il paragone con l'asse X del sistema Htc quasi come del tutto scorrelato.

Tabella 11 Correlazione di Pearson dei sistemi MoveOne e Rokoko rispetto a Htc (Walk)

	MoveOne	Rokoko	IK_Htc
Pearson X	0,881868133	0,17199934	0,917044298
Pearson Y	0,891207869	0,508485674	0,935001532
Pearson Z	0,987236905	0,990509072	0,991071942
Pearson tot	0,97544227	0,969382545	

Questo valore risulta coerente alle aspettative, poiché in questo test specifico è stato riscontrato un problema di interferenza magnetica sui sensori degli arti inferiori, che è andato a influire sul tracciamento della prova specifica di camminata. È stato possibile individuare facilmente il problema tramite software, poiché avendo una legenda e una mappa di tutti i sensori, questi si sono illuminati di arancione, rilevando l'anomalia. Questo può essere dovuto a diversi fattori: sia perché erano presenti molti materiali metallici vicino all'area di test, sia perché sono stati indossati più dispositivi contemporaneamente. Probabilmente i sensori inferiori della tuta sono entrati in conflitto con i sensori dell'Htc, andando a rovinare il tracciamento dei dati.

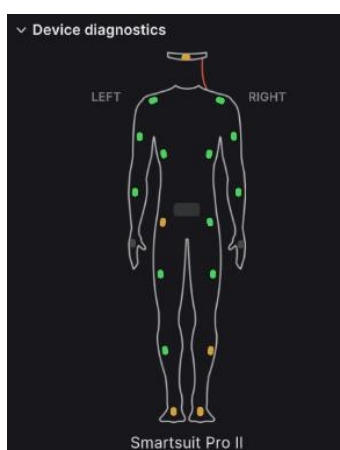


Figura 52 Diagnostica sensori SmartsuitPro 2 Rokoko

Andando ad analizzare il confronto del movimento catturato con la tuta Xsens, possiamo notare valori superiori per l'asse X e Y della tuta rispetto al sistema

MoveOne, ottenendo valori di 0,54 e di 0,41. Vediamo inoltre una forte diminuzione della correlazione per il sistema basato su intelligenza artificiale, arrivando a valori molto bassi come 0,33 riguardo l'asse delle ascisse e 0,28 per l'asse delle ordinate.

Tabella 12 Correlazione di Pearson dei sistemi MoveOne e Xsens rispetto a Htc (Walk)

	MoveOne	Xsens	IK_Htc
Pearson X	0,334492	0,538525	0,537129
Pearson Y	0,280108	0,414815	0,448294
Pearson Z	0,975401	0,97537	0,980328
Pearson tot	0,938203	0,92737	

Nell'ultimo test effettuato relativo al salto, possiamo notare dei valori del coefficiente di Pearson in generale più elevati rispetto ai test precedenti, ad eccezione di alcuni singoli assi. Per esempio, nell'asse Y del confronto con la Smartsuit Pro 2 è presente un valore negativo, andando a significare che il valore della Y tende lievemente a comportarsi in modo opposto nel paragone tra i due sistemi con Htc. Da notare inoltre un valore estremamente basso relativo l'asse Y sia del sistema MoveOne che Xsens, con valori di 0,08 e 0,05.

Tabella 13 Correlazione di Pearson dei sistemi MoveOne e Rokoko rispetto a Htc (Jump)

	MoveOne	Rokoko	IK_Htc
Pearson X	0,6120138	0,60466	0,585497
Pearson Y	-0,2257729	-0,23956	-0,20351
Pearson Z	0,9389838	0,947352	0,933744
Pearson tot	0,8992855	0,916155	

Tabella 14 Correlazione di Pearson dei sistemi MoveOne e Xsens rispetto a Htc (Jump)

	MoveOne	Xsens	IK_Htc
Pearson X	0,923768	0,847009	0,876974
Pearson Y	0,079811	0,050504	0,08345
Pearson Z	0,978655	0,961107	0,979617
Pearson tot	0,970998	0,946968	

8 Conclusioni

I risultati ottenuti nel corso di questa analisi permettono di analizzare e avere una panoramica più chiara dei diversi sistemi di motion capture che possono essere utilizzati in ambito di produzione Rai, permettendo di scegliere in modo coerente in base alle diverse esigenze.

Dai dati osservati si può notare in generale un comportamento sorprendentemente competitivo per quanto riguarda il sistema di motion capture MoveOne rispetto alle tute: portando a volte i risultati sono stati superiori. Un problema riscontrato rispetto a questo sistema è relativo al rumore del mantenimento della posizione poiché questo risulta essere percepibile e fastidioso alla vista, obbligando a una lavorazione in post-produzione di pulizia allo scopo di attenuazione del disturbo. Questo problema però probabilmente potrà essere risolto direttamente dall'azienda andando a migliorare l'algoritmo di valutazione e risoluzione del movimento di motion capture. Il limite principale di questa tecnologia resta comunque l'inaffidabilità relativa all'operabilità e la poca solidità del tracciamento in più sessioni di cattura: si è vincolati infatti all'utilizzo dell'applicazione su cellulare iPhone, a dei tempi prestabiliti, e a dover effettuare l'upload della cattura sui server proprietari, non potendo effettuare l'elaborazione del video in locale. Inoltre, nell'applicazione non è stato sviluppato un sistema avanzato di gestione dei video nella libreria; quindi, per il possibile utilizzo in lato produzione andrebbe a risultare poco intuitivo e poco efficiente, non permettendo né una rinomina dei file, né una loro suddivisione in cartelle una volta registrati. Per questi motivi l'utilizzo di MoveOne per la produzione potrebbe concentrarsi sulla cattura di movimenti umani brevi (con una durata massima di 60 secondi) e

semplici, così da sfruttare al massimo il vantaggio della portabilità e della non invasività del sistema pur a discapito di una minore precisione e a una breve pulizia in post-produzione.

La tuta Xsens dai dati catturati risulta essere la soluzione più vicina al sistema ottico di riferimento Htc, superandolo anche dal punto di vista di portabilità e operabilità.

Mentre il sistema Htc deve essere allestito in un ambiente chiuso, adottando tutti gli accorgimenti del caso descritti in precedenza, le tute Xsens e Smartsuit Pro 2 possono essere utilizzate anche in esterna, andando così ad ampliare le possibilità di cattura di movimenti direttamente sul set di produzione. Vi sono comunque notevoli differenze tra le due tute: il costo di accesso è molto differente.

Il sistema di Movella Xsens Awinda Starter: ha un costo iniziale per l'acquisto della tuta di 4230 euro¹⁰ al quale bisogna aggiungerci il costo annuale di 2990 euro del software di cattura e gestione della tuta Mvn Animate Plus.¹¹

Volendo si può salire anche di sistema di motion capture sempre della stessa azienda, aumentando notevolmente il costo ma anche la qualità. Per la produzione questa tuta permette di avere una grossa affidabilità e solidità: le batterie dei sensori durano a lungo, permettono lunghe sessioni di registrazione e di live streaming sui principali motori grafici come Unreal e Unity, permettendo quindi di velocizzare di molto i tempi di cattura, di setting e di concentrarsi più sull'obiettivo finale della produzione e non sul funzionamento del mezzo di cattura. Risulta un sistema ideale nelle produzioni cinematografiche e televisive più complesse, avendo sia un ampio range di configurazioni hardware e software che un'affidabilità riconosciuta a livello industriale. Lo svantaggio sorge in produzioni brevi e piccole: a meno che non sia richiesta una precisione molto elevata, può risultare inefficiente l'utilizzo di Xsens, poiché i tempi di setup iniziale potrebbero portare a ritardi e potrebbe risultare un costo eccessivo. Adottare un

¹⁰ Aggiornato a luglio 2024 [<https://shop.movella.com/product-lines/motion-capture/products/xsens-mvn-awinda-starter>]

¹¹ Aggiornato a luglio 2024 [<https://shop.movella.com/product-lines/motion-capture/products/mvn-animate-plus/mvn-am-plus>]

sistema troppo professionale in una piccola realtà potrebbe insomma non essere la scelta migliore e si potrebbe risolvere con sistemi di motion capture meno costosi e con tempistiche di setup più rapide.

La Smartsuit Pro 2 della Rokoko, adatto per una possibile produzione con budget più limitato. Il costo di questo sistema di motion capture, infatti, resta molto più contenuto rispetto la tuta Xsens. Si parte con un costo di soli 1995 euro.¹²

Anche qui, per poter sfruttare al meglio tutte le potenzialità e poter effettuare il live streaming su software esterni come Unreal e Unity, è richiesto un abbonamento annuale dal costo di 600 euro per l'abbonamento Pro. Il vantaggio del sistema della Rokoko è la sua modularità. E'è possibile infatti acquistare anche solamente i guanti con un costo di soli 1495 euro. È presente, inoltre, un accessorio chiamato Coil Pro dal costo di 1995 euro, con il quale l'azienda promette di diminuire drasticamente il drift e di aumentare così la precisione massima. Nel caso si voglia effettuare l'acquisto completo di tutti i sistemi Rokoko(escluso l'abbonamento software), è presente anche un bundle al prezzo di 5645 euro.¹³

Questa tuta può avere un vasto campo di utilizzo e risulta essere un ottimo compromesso tra vestibilità, praticità, portabilità e utilizzo. Risulta ideale per le catture in esterno e in situazioni dove è necessario spostarsi frequentemente. Inoltre, la facilità d'uso può essere un buon motivo di scelta: dopo aver effettuato la prima configurazione e aver dedicato un router per una connessione sicura e stabile al massimo della banda tra tuta e computer, questa tuta risulta essere quasi "plug and play", poiché richiede solamente una batteria per alimentare tutti i sensori e si connette automaticamente al router dedicato. In questo modo vi è un drastico abbassamento dei tempi di setup. La vestibilità risulta buona permettendo un ottimo range di movimenti. Come abbiamo visto dai risultati anche questo sistema di motion capture permette di catturare il movimento con un margine di errore molto basso e tante possibilità di configurazione, permettendo di adattarsi a diversi tipi di produzione cinematografica, dalla più

¹² Aggiornato a luglio 2024 [https://eu.store.rokoko.com/products/smartsuit-pro-ii?_gl=1*1dh32g6*_gcl_au*MjA0OTMyNDE2My4xNzlwMjk5NTQw*_ga*MTA2Njc3ODcwOC4xNzlwMjk5NTQw*_ga_2TG4CJCP6E*MTcyMDI5OTUzOS4xLjAuMTcyMDI5OTUzOS42MC4wLjA]

¹³ Aggiornato a luglio 2024 [<https://eu.store.rokoko.com/pages/full-performance-capture-coil-pro>]

limitata alla più complessa. L'unica differenza notevole rilevata tra le due tute riguarda le interferenze magnetiche: ogni sensore Xsens è dotato interamente di 3 magnetometri, uno per ogni asse, che permettono di effettuare correzioni in tempo reale sugli altri sensori. La tuta della Rokoko, invece, non si affida a dei componenti hardware per la correzione magnetica, ma a degli algoritmi che vanno a correggere e stimare l'errore dal punto di vista software, ottenendo quindi risultati inferiori rispetto alla Movella.

8.1 Sviluppi futuri

Questa tesi mi ha permesso di esplorare a fondo la motion capture, permettendomi di approfondirne i diversi sistemi.

Potrebbe essere interessante proseguire il presente lavoro con la Rai attraverso lo sviluppo di un software in grado di gestire in contemporanea i live streaming di diversi flussi in ingresso di motion capture, con la possibilità di visualizzare in real time la qualità della cattura, fornendo così dei dati statistici come quelli già usati in questa tesi per il confronto tra i diversi sistemi (come il coefficiente di Pearson e l'RMSD).

Inoltre bisogna ricordare che è stato da poco annunciato¹⁴ un nuovo software dall'azienda Move.ai chiamato MoveLive, in grado di effettuare la motion capture in tempo reale senza dover effettuare l'elaborazione in post.¹⁵

Un ulteriore sviluppo chiave potrebbe essere trovare un metodo per ottenere la massima precisione nella ripetitività dei movimenti, consentendo l'esecuzione di test sequenziali senza la necessità di sovrapporre i sensori di ogni sistema. L'impiego di un manichino robotico potrebbe aiutare il confronto tra le tute, consentendo di intervenire direttamente sul movimento.

Un'altra possibilità interessante è costituita dall'integrazione del sistema del controllo di qualità in live della cattura con il sistema di volumetric capture. Questo inizialmente richiederebbe lo sviluppo di un apposito software per il riconoscimento e la ricostruzione della volumetric capture, la relativa conversione

¹⁴ Giugno 2024

¹⁵ Dal sito ufficiale [<https://www.move.ai/real-time>]

e il retargeting su un effettivo scheletro. In questo modo potrebbe essere possibile integrare anche questa tecnologia.

Il progredire della motion capture sta permettendo un enorme progresso nel settore in termini di efficienza, precisione, portabilità e tempistiche. Questo consente ai registi e agli animatori di concentrarsi su un'animazione sempre più ad alto livello, potendosi occupare maggiormente della rifinitura del movimento più che del movimento in sé. Inoltre, l'utilizzo dell'intelligenza artificiale per effettuare la motion capture sta aumentando in modo esponenziale, permettendo una performance sempre più accurata e annullando l'ingombro del sistema tradizionale. Il futuro della motion capture appare molto promettente, offrendo sempre nuove opportunità creative e tecniche che continueranno a rivoluzionare il settore della virtual production.

Appendice

Tabelle deviazione standard con tuta Xsens

Tabella 15 Confronto deviazione standard osso della testa sessione 1

head		Htc		MoveOne		Xsens
	σ_x	0,537488244		σ_x	0,282163	σ_x 0,397654
	σ_y	0,395662544		σ_y	0,389033	σ_y 0,24945
	σ_z	0,244993896		σ_z	0,163609	σ_z 0,018253

Tabella 16 Confronto deviazione standard osso della mano destra sessione 1

hand_r		Htc		MoveOne		Xsens
	σ_x	0,812922522		σ_x	0,661243	σ_x 0,496168
	σ_y	0,131732644		σ_y	0,372218	σ_y 0,094712
	σ_z	0,652026496		σ_z	0,863093	σ_z 0,687727

Tabella 17 Confronto deviazione standard osso della mano sinistra sessione 1

hand_l		Htc		MoveOne		Xsens
	σ_x	0,500888732		σ_x	1,07087	σ_x 0,552688
	σ_y	0,10765591		σ_y	0,309258	σ_y 0,211927
	σ_z	0,916214761		σ_z	0,552075	σ_z 0,4815

Tabella 18 Confronto deviazione standard osso del gomito destro sessione 1

lowerarm_r		Htc		MoveOne		Xsens
	σ_x	0,276626414		σ_x	0,340809	σ_x 0,420197
	σ_y	0,137956985		σ_y	0,273696	σ_y 0,114023
	σ_z	0,456161736		σ_z	0,510026	σ_z 0,429758

Tabella 19 Confronto deviazione standard osso del gomito sinistro sessione 1

lowerarm_l		Htc		MoveOne		Xsens
	σ_x	0,397172236		σ_x	0,573116	σ_x 0,387169
	σ_y	1,097303905		σ_y	0,28745	σ_y 0,133843
	σ_z	0,561367155		σ_z	0,248572	σ_z 0,287842

Tabella 20 Confronto deviazione standard osso del piede destro sessione 1

foot_r		Htc		MoveOne		Xsens
	σ_x	0,028137712		σ_x	0,104605	σ_x 0,02928
	σ_y	0,111000584		σ_y	0,138922	σ_y 0,018757
	σ_z	0,161247788		σ_z	0,177967	σ_z 0,015112

Tabella 21 Confronto deviazione standard osso del piede sinistro sessione 1

foot_l		Htc		MoveOne		Xsens
	σ_x	0,015150999		σ_x	0,169507	σ_x 0,013881
	σ_y	0,049415266		σ_y	0,275789	σ_y 0,033792
	σ_z	0,058348613		σ_z	0,147735	σ_z 0,009023

Tabelle deviazione standard con tuta SmartSuitPro 2 Rokoko

Tabella 22 Confronto deviazione standard osso della testa sessione 2

head		Htc		MoveOne			Rokoko
	σ_x	0,1683		σ_x	0,791963		σ_x 0,691082
	σ_y	0,14622		σ_y	0,679287		σ_y 0,505946
	σ_z	0,03474		σ_z	0,359981		σ_z 0,121111

Tabella 23 Confronto deviazione standard osso della mano destra sessione 2

hand_r		Htc		MoveOne			Rokoko
	σ_x	0,43687		σ_x	1,492737		σ_x 1,457064
	σ_y	0,16672		σ_y	1,281316		σ_y 1,786304
	σ_z	0,14243		σ_z	2,225784		σ_z 2,576689

Tabella 24 Confronto deviazione standard osso della mano sinistra sessione 2

hand_l		Htc		MoveOne			Rokoko
	σ_x	0,74216		σ_x	0,766045		σ_x 1,222655
	σ_y	0,07465		σ_y	0,403761		σ_y 0,555281
	σ_z	0,53688		σ_z	1,131059		σ_z 1,181275

Tabella 25 Confronto deviazione standard osso del gomito destro sessione 2

lowerarm_r		Htc		MoveOne			Rokoko
	σ_x	0,22798		σ_x	0,957915		σ_x 1,112397
	σ_y	0,20589		σ_y	0,952708		σ_y 1,140709
	σ_z	0,07012		σ_z	0,673046		σ_z 1,081645

Tabella 26 Confronto deviazione standard osso del gomito sinistro sessione 2

lowerarm_l		Htc		MoveOne			Rokoko
	σ_x	0,47429		σ_x	1,193271		σ_x 1,264406
	σ_y	0,11551		σ_y	0,54216		σ_y 0,526468
	σ_z	0,22654		σ_z	0,480934		σ_z 0,791435

Tabella 27 Confronto deviazione standard osso del piede destro sessione 2

foot_r		Htc		MoveOne			Rokoko
	σ_x	0,10366		σ_x	0,173522		σ_x 0,466833
	σ_y	0,46124		σ_y	0,361039		σ_y 0,375026
	σ_z	0,68092		σ_z	0,242707		σ_z 0,234217

Tabella 28 Confronto deviazione standard osso del piede sinistro sessione 2

foot_l		Htc		MoveOne			Rokoko
	σ_x	0,02976		σ_x	0,382833		σ_x 0,165604
	σ_y	0,07787		σ_y	0,377776		σ_y 0,27765
	σ_z	0,07219		σ_z	0,229406		σ_z 0,119529

Indice delle figure

Figura 1 Esempio motion capture full body	5
Figura 2 Schema relativo alle diverse fasi della pipeline di motion capture.....	8
Figura 3 Sistema Optitrack.....	13
Figura 4 Setup esperimento descritto nel paper.....	20
Figura 5 Htc Vive Pro 2	27
Figura 6 Base Station 2.0.....	28
Figura 7 Consigli sugli angoli per le Base Station 2.0	29
Figura 8 Consigli sulla disposizione per le Base Station 2.0	29
Figura 9 Htc Vive Tracker 3.0.....	30
Figura 10 Tuta Xsens Awinda Starter.....	31
Figura 11 Schema architettura singolo sensore Xsens	33
Figura 12 Host che permette il collegamento di Xsens e computer.....	33
Figura 13 Basi di ricarica sensori Xsens	33
Figura 14 Rokoko Smartsuit Pro 2	36
Figura 15 Rokoko SmartGloves	37
Figura 16 Live Link plugin.....	41
Figura 17 Schermata editor Metahuman	43
Figura 18 Schermata processo mesh to Metahuman.....	44
Figura 19 Quixel Bridge, schermata principale.....	46
Figura 20 Plugin OpenXR	46
Figura 21 Plugin Xsens	47
Figura 22 Plugin Rokoko Studio	47
Figura 23 Plugin Json Utilities	47
Figura 24 Rokoko Studio schermata principale.....	48
Figura 25 Rokoko Studio, dettagli e Export	49
Figura 26 Xsens schermata principale	50
Figura 27 Xsens schermata configurazione	51
Figura 28 Guanti Quantum Manus	52
Figura 29 Move One schermata principale.....	53
Figura 30 Schermata principale SteamVR	53
Figura 31 SteamVR schermata di gestione tracker	54
Figura 32 Setup della stanza.....	55

Figura 33 Piantina setup cattura.....	56
Figura 34 Metahuman importato nell'ambiente	58
Figura 35 LiveLink component per Htc.....	59
Figura 36 Viewport con sistema Htc.....	60
Figura 37 Subito prima del processo di calibrazione	61
Figura 38 Xsens schermata per la scelta della tipologia di calibrazione.....	62
Figura 39 Rokoko Studio esportazione animazione	64
Figura 40 Retargeting animazione su scheletro Metahuman	64
Figura 41 LiveLink Pose di Xsens per effettuare lo stream in tempo reale su Metahuman	66
Figura 42 MoveOne processo di retarget	67
Figura 43 Modifiche additive nel Level Sequence	68
Figura 44 Schema della struttura di acquisizione dati	70
Figura 45 Animation Blueprint	71
Figura 46 Blueprint event graph	72
Figura 47 Blueprint funzione SaveBoneDataThisFrame	73
Figura 48 Blueprint funzione OnComplete	74
Figura 49 Blueprint level per acquisizione Htc.....	75
Figura 50 Schema delle ossa prese in considerazione	76
Figura 51 Illustrazione assi cartesiani	87
Figura 52 Diagnostica sensori SmartsuitPro 2 Rokoko	90

Indice delle tabelle

Tabella 1 Rapporto frequenza con numero di sensori.....	22
Tabella 2 Frame Rate acquisizione del sistema Xsens	32
Tabella 3 Correlazione di Pearson dei sistemi MoveOne e Xsens rispetto a Htc (T-pose).....	86
Tabella 4 Correlazione di Pearson dei sistemi MoveOne e Rokoko rispetto a Htc (T-pose).....	86
Tabella 5 Valori RMSD di MoveOne e Xsens rispetto a Htc (T-pose)	86
Tabella 6 Valori RMSD di MoveOne e Rokoko rispetto a Htc (T-pose)	86
Tabella 7 Valori RMSD di MoveOne e Xsens rispetto a Htc (Walk)	89
Tabella 8 Valori RMSD di MoveOne e Rokoko rispetto a Htc (Walk)	89
Tabella 9 Valori RMSD di MoveOne e Xsens rispetto a Htc (Jump).....	89
Tabella 10 Valori RMSD di MoveOne e Rokoko rispetto a Htc (Jump)	89
Tabella 11 Correlazione di Pearson dei sistemi MoveOne e Rokoko rispetto a Htc (Walk).....	90
Tabella 12 Correlazione di Pearson dei sistemi MoveOne e Xsens rispetto a Htc (Walk).....	91
Tabella 13 Correlazione di Pearson dei sistemi MoveOne e Rokoko rispetto a Htc (Jump)	91
Tabella 14 Correlazione di Pearson dei sistemi MoveOne e Xsens rispetto a Htc (Jump)	92
Tabella 15 Confronto deviazione standard osso della testa sessione 1	97
Tabella 16 Confronto deviazione standard osso della mano destra sessione 1	97
Tabella 17 Confronto deviazione standard osso della mano sinistra sessione 1	97
Tabella 18 Confronto deviazione standard osso del gomito destro sessione 1	97
Tabella 19 Confronto deviazione standard osso del gomito sinistro sessione 1	97
Tabella 20 Confronto deviazione standard osso del piede destro sessione 1 ..	97
Tabella 21 Confronto deviazione standard osso del piede sinistro sessione 1.	97
Tabella 22 Confronto deviazione standard osso della testa sessione 2	98
Tabella 23 Confronto deviazione standard osso della mano destra sessione 2	98

Tabella 24 Confronto deviazione standard osso della mano sinistra sessione 2	98
Tabella 25 Confronto deviazione standard osso del gomito destro sessione 2	98
Tabella 26 Confronto deviazione standard osso del gomito sinistro sessione 2	98
Tabella 27 Confronto deviazione standard osso del piede destro sessione 2 ..	98
Tabella 28 Confronto deviazione standard osso del piede sinistro sessione 2.	98

Indice dei grafici

Grafico 1 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse X dell'osso della testa	77
Grafico 2 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse Y dell'osso della testa	77
Grafico 3 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse Z dell'osso della testa	77
Grafico 4 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse X dell'osso del piede destro	78
Grafico 5 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse Y dell'osso del piede destro	78
Grafico 6 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse Z dell'osso del piede destro	79
Grafico 7 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse X dell'osso del piede sinistro	79
Grafico 8 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse Y dell'osso del piede sinistro	79
Grafico 9 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse Z dell'osso del piede sinistro	79
Grafico 10 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse X dell'osso della mano destra	80
Grafico 11 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse Y dell'osso della mano destra	80
Grafico 12 Comparazione sistemi Htc, MoveOne e Rokoko relativo all'asse Z dell'osso della mano destra	80
Grafico 13 Comparazione sistemi Htc e Xsens relativo all'asse X dell'osso del piede sinistro	80
Grafico 14 Comparazione sistemi Htc e Xsens relativo all'asse Y dell'osso del piede sinistro	81
Grafico 15 Comparazione sistemi Htc e Xsens relativo all'asse Z dell'osso del piede sinistro	81
Grafico 16 Comparazione sistemi Htc e Xsens relativo all'asse X dell'osso della mano destra	81

Grafico 17 Comparazione sistemi Htc e Xsens relativo all'asse Y dell'osso della mano destra	81
Grafico 18 Comparazione sistemi Htc e Xsens relativo all'asse Z dell'osso della mano destra	82
Grafico 19 Comparazione sistemi Htc e MoveOne relativo all'asse X dell'osso della mano destra sessione 1	82
Grafico 20 Comparazione sistemi Htc e MoveOne relativo all'asse Y dell'osso della mano destra sessione 1	83
Grafico 21 Comparazione sistemi Htc e MoveOne relativo all'asse Z dell'osso della mano destra sessione 1	83
Grafico 22 Comparazione sistemi Htc e MoveOne relativo all'asse X dell'osso della mano destra sessione 2	83
Grafico 23 Comparazione sistemi Htc e MoveOne relativo all'asse Y dell'osso della mano destra sessione 2	84
Grafico 24 Comparazione sistemi Htc e MoveOne relativo all'asse Z dell'osso della mano destra sessione 2	84
Grafico 25 Deviazione standard mano destra sessione 1 (T-pose)	87
Grafico 26 Deviazione standard mano destra sessione 2 (T-pose)	87
Grafico 27 Deviazione standard relativa alla testa nella sessione 1 (T-pose) ..	88
Grafico 28 Deviazione standard relativa alla testa nella sessione 2 (T-pose) ..	88
Grafico 29 Deviazione standard relativa alla mano sinistra nella sessione 1 (T-pose)	88
Grafico 30 Deviazione standard relativa alla mano sinistra nella sessione 2 (T-pose)	88

Bibliografia

- [1] M. C. Wibowo, S. Nugroho e A. Wibowo, «The Use of Motion Capture Technology in 3D Animation,» *International Journal of Computing and Digital System*, 2024.
- [2] J. Li, J. Zhang, Z. Wang, S. Shen, C. Wen, Y. Ma, L. Xu, J. Yu e C. Wang, «LiDARCap: Long-range Markerless 3D Human Motion Capture with LiDAR Point Clouds,» *IEEE*, 2022.
- [3] Y. Zhu, C. Detig, S. Kane e G. Lourie, «Kinematic Motion Analysis with Volumetric Motion Capture,» *IEEE*, 2022.
- [4] T. S. K. U. L. O. A. K. Y. I. S. F. S. Y. Nobuyasu Nakano, «Evaluation of 3D Markerless Motion Capture Accuracy Using OpenPose With Multiple Video Cameras,» *Front. Sports Act. Living*, vol. 2, 2020.
- [5] S. P. D. B. A. S. K. W. Sebastian Merker, “Measurement Accuracy of the HTC VIVE Tracker 3.0 Compared to Vicon System for Generating Valid Positional Feedback in Virtual Reality,” *Sensor (Basel)*, 2023.
- [6] A. K. L. F. C. S. Soumitra P. Sitole, «Application and Evaluation of Lighthouse Technology for Precision Motion Capture,» *IEEE Sensor Journal*, vol. 20, 2020.
- [7] J. G. C. H. A. K. Valentin Holzwarth, «Comparing the Accuracy and Precision of SteamVR Tracking 2.0 and Oculus Quest 2 in a Room Scale Setup,» 2021.
- [8] M. G. G. B. E. M. J. R. B.-J. F. V. B. P. H. V. Frank J. Wouda, «On the Validity of Different Motion Capture Technologies for the Analysis of Running,» *IEEE*, 2018.

- [9] M. S. N. R. a. G. B. Monique Paulich, *Xsens MTw Awinda: Miniature Wireless Inertial - Magnetic Motion Tracker for Highly Accurate 3D Kinematic Application*, Xsens Technologies.
- [10] S. K. a. P. L. Florian Herrmann, *MoCaCo: A Simulator Framework for Motion Capture Comparison*, Science and Technology Publications.
- [11] M. G. a. G. B. Martin Schepers, *Xsens MVN: Consistent Tracking of Human Motion Using Inertial Sensing*, Xsens Technologies.
- [12] S. Ciklacandir, S. Ozkan e Y. Isler, «A Comparison of the Performances of Video-Based and IMU Sensor-Based Motion Capture Systems on Joint Angles,» *IEEE*, 2022.
- [13] D. C. Niehorster, «The Accuracy and Precision of Position and Orientation Tracking in the HTC Vive Virtual Reality System for Scientific Research,» *Iperception*, 2017.
- [14] M. Borges, A. Symington, B. Coltin, T. Smith e R. Ventura, «HTC Vive: Analysis and Accuracy Improvement,» *IEEE/IROS*, 2018.
- [15] M. B. P. E. P. C. R. F. a. J. S. T. Susanne M. van der Veen, «Agreement Analysis between Vive and Vicon Systems to Monitor Lumbar Postural Changes,» *Sensor (Basel)*, 2019.