



**Politecnico
di Torino**

Politecnico di Torino

Corso di Laurea Magistrale in

**INGEGNERIA DEL CINEMA E DEI MEZZI DI
COMUNICAZIONE**

a.a. 2024/2025

Sessione di Laurea Ottobre 2025

Tesi di Laurea Magistrale

**Progetto ItARTS (Italy for Arts:
Resource, Time and Space)**

Progettazione e Implementazione di un Sistema VR immersivo e
adattivo per l'Opera Musicale Generativa SHARED

Relatori:

Tatiana Mazali
Andrea Bottino

Candidato:

Alice Cavasino

Indice

PARTE I	1
1 INTRODUZIONE E CONTESTO DEL PROGETTO	1
1.1 Progetto ItARTS: Contesto e finalità	1
1.2 Finalità del progetto	2
1.3 Architettura del progetto e contributo personale	3
1.4 Obiettivi della tesi	5
2 SHARED VR.....	6
2.1 Concept e obiettivi artistici	6
2.2 Esperienza dell'utente	7
2.3 Struttura narrativa	8
3 SHARED VR: CRITICA ESTETICA.....	10
3.1 Teorie dei nuovi media e cultural heritage	10
3.2 Dall'evoluzione dei musei virtuali alle installazioni immersive VR per l'arte	13
3.3 Multiutenza, gesto e generatività	18
4 EMBODIMENT E PRESENZA IN VR	22
4.1 Fondamenti fenomenologici e cognitivi dell'embodiment	22
4.2 Corpo, ambiente e interfacce digitali	24
4.3 Presenza e senso di "essere nel luogo"	25
PARTE 2	27
5 ARCHITETTURA DEL SISTEMA SHARED VR.....	28
5.1 Requisiti del sistema e scelte tecnologiche	28
5.2 Architettura software e hardware	29
6 IMPLEMENTAZIONE HAND TRACKING	32
6.1 Librerie	32
6.2 Sperimentazione	34
6.3 Sviluppo prototipo	41
7 IMPLEMENTAZIONE NETWORKING.....	45
7.1 Netcode for Game Object (NGO)	46
7.2 Test preliminari	49
8 TESTING E VALIDAZIONE	55
8.1 Metodologia e strumenti di valutazione	55
8.2 Analisi dei risultati	57
8.3 Discussione e osservazioni qualitative	61
9 CONCLUSIONI E FUTURE WORKS	63
9.1 Valutazione critica e contributo alla ricerca artistica	63
9.2 Prospettive e sviluppi futuri	64
9.3 Considerazioni finali	64
RINGRAZIAMENTI	65
ELENCO FIGURE	67
BIBLIOGRAFIA	69

PARTE I

1 Introduzione e contesto del progetto

1.1 Progetto ItARTS: Contesto e finalità

Il progetto *ItARTS – Italy for Arts: Resources, Time and Space*, descritto nella scheda progetto del MUR [1] e promosso da una rete di istituzioni dell'AFAM¹ insieme a partner universitari e centri di ricerca, nasce con l'obiettivo di creare un ambiente di sperimentazione transdisciplinare che metta in dialogo arti, scienze e tecnologie. Consiste nel sostenere la contemporaneità della creazione artistica attraverso l'integrazione di un approccio **STEAM**² e la valorizzazione del patrimonio culturale.

Uno degli aspetti principali del progetto consiste nella creazione di un *Living Learning Lab* a L'Aquila, concepito come spazio aperto per l'innovazione e la creatività. Questo ambiente ha consentito lo sviluppo di tre progetti artistici multidisciplinari presentati nell'ambito dell'*ItARTS Festival*, una vetrina internazionale delle migliori pratiche artistiche e musicali italiane, che viaggerà tra istituzioni europee ed extraeuropee prima di concludersi nel capoluogo abruzzese, designato Capitale Italiana della Cultura per il 2026. L'iniziativa vede la partecipazione di cinque istituzioni partner: il Conservatorio "A. Casella" dell'Aquila in veste di leader, il Saint Louis College of Music di Roma, l'Accademia di Belle Arti dell'Aquila, l'Università dell'Aquila e il Politecnico di Torino. La collaborazione tra le diverse istituzioni mira a costruire un legame sinergico tra linguaggi e metodologie eterogenee, con l'intento di sviluppare nuove forme di espressione artistica multimediale che integrino arti visive, tecnologie interattive e immersive, oltre a competenze di tipo imprenditoriale e comunicativo.

L'approccio innovativo e multidisciplinare di *ItARTS* si concretizza attraverso diversi obiettivi: da un lato creare un contesto in cui i giovani artisti possano acquisire competenze tecniche e creative, ma anche capacità imprenditoriali utili per affrontare le sfide del mercato culturale attuale; dall'altro stimolare la produzione di opere multimediali capaci di affrontare temi sociali

¹ L' AFAM (Alta Formazione Artistica, Musicale e coreutica) è una locuzione che indica l'istruzione superiore artistica appartenente al sistema universitario italiano, disciplinato dalla Legge 21 dicembre 1999, n. 508 che lo ha equiparato al sistema universitario. È sottoposto al Ministero dell'Università e della Ricerca e include istituti statali e autorizzati, che rilasciano titoli accademici legalmente riconosciuti, articolati secondo il modello 3+2 (triennio + biennio specialistico) con riconoscimento di crediti (CFA = CFU) analoghi a quelli universitari. [2]

² L'approccio STEAM nasce come estensione del tradizionale STEM (Science, Technology, Engineering, Mathematics), integrando le arti per promuovere un apprendimento interdisciplinare basato sulla creatività, sul pensiero critico e innovazione. Il termine viene teorizzato per la prima volta da Georgette Yakman (2008) [3] che ne definisce un modello educativo fondato sull'integrazione tra competenze scientifico-tecnologiche e capacità creative, volto a promuovere un apprendimento interdisciplinare e basato sull'esperienza. Nel saggio "STEM + Art = STEAM" (2013) [4], John Maeda ne evidenzia la rilevanza strategica per l'innovazione e lo sviluppo economico, sottolineando come l'arte unita alle discipline scientifiche stimoli pensiero critico e creatività, competenze centrali per affrontare le sfide della contemporaneità.

e questioni globali, offrendo così una prospettiva artistica che si confronta con la realtà e ne interpreta i cambiamenti. Un ulteriore elemento chiave è la volontà di favorire la circolazione internazionale delle opere, incrementandone la visibilità e rafforzando il ruolo dell'Italia nei circuiti della cultura contemporanea.

L'iniziativa non si limita alla produzione artistica, ma punta anche a coinvolgere attivamente il pubblico, trasformandolo da semplice spettatore a parte integrante del processo creativo. A questo scopo, il progetto prevede attività di formazione e momenti partecipativi per aumentare la consapevolezza e l'interesse verso la cultura e la creatività. Tale struttura si traduce in una proposta fortemente multidisciplinare, che comprende ambiti diversi come design, arti applicate, tecnologie digitali, teatro, danza, cinema e musica. L'eterogeneità delle istituzioni coinvolte permette di mettere in relazione linguaggi e metodologie differenti, creando un ambiente fertile per l'innovazione e la ricerca artistica.

In questa tesi si analizzerà in particolare il laboratorio *SHARED* nato all'interno del contesto ItARTS. In particolare, l'elaborato si concentrerà sulla sua declinazione in realtà virtuale, al cui sviluppo ho avuto la possibilità di prendere parte in prima persona.

1.2 Finalità del progetto

Il laboratorio *SHARED* ideato da Michelangelo Lupone, compositore e ricercatore presso il Centro Ricerche Musicali³ (CRM) di Roma, indaga il rapporto tra suono, gesto e spazio, con l'intento di superare la fruizione passiva dell'opera coinvolgendo lo spettatore come parte attiva del processo creativo. Se nell'arte convenzionale lo spettatore ricopre un ruolo prevalentemente passivo, in *SHARED* l'utente viene invece coinvolto come co-autore dell'esperienza: le sue azioni non solo generano effetti immediati nell'ambiente sonoro e visivo, ma si accumulano e si stratificano nel tempo, contribuendo a un processo narrativo collettivo e dinamico.

L'opera si basa su tre principi cardine: partecipazione condivisa, memoria collettiva e adattività.

- La **partecipazione condivisa** implica che ogni gesto del visitatore diventi parte integrante della composizione artistica, eliminando la separazione tra chi crea e chi fruisce.
- La **memoria** collettiva si concretizza nell'archiviazione e nella sedimentazione delle interazioni: ogni azione lascia una traccia che continua a influenzare l'evoluzione dell'opera, creando una sorta di "palinsesto vivo" in costante trasformazione.

³ Il Centro Ricerche Musicali (CRM) è un'associazione senza scopo di lucro fondata a Roma nel 1988 da Laura Bianchini e Michelangelo Lupone, dedicata alla ricerca musicale nei suoi aspetti artistici e scientifici. Riconosciuto nel 1990 dal Ministero dell'Università e della Ricerca come centro di ricerca musicale, sviluppa tecnologie innovative per la sintesi del suono, la spazializzazione e l'interazione, ed è promotore del festival internazionale ArteScienza. [5]

- L'**adattività**, infine, è la capacità del sistema di reagire in tempo reale ai comportamenti degli utenti e di rimodellare continuamente il paesaggio sonoro e visivo, generando esperienze sempre nuove e irripetibili.

La struttura di *SHARED* si declina in tre forme complementari, ciascuna delle quali interpreta questi principi in modo specifico:

- **Performance dal vivo:** musicisti e performer interagiscono con oggetti plastici modulari sensibili al suono, trasformando lo spazio scenico in un ambiente compositivo in continua evoluzione.
- **Installazione interattiva:** il pubblico può entrare direttamente nello spazio dell'opera e manipolarne gli elementi attraverso gesti e movimenti, senza mediazioni strumentali.
- **Versione in realtà virtuale (VR):** oggetto di questa tesi, che trasferisce i principi fondanti dell'opera in un ambiente digitale immersivo e multiutente, permettendo di sperimentare la stessa logica generativa in un contesto tecnologico rinnovato.

L'introduzione della realtà virtuale non snatura il progetto originale, bensì ne amplifica le possibilità espressive. L'esperienza VR, infatti, consente a più utenti di condividere lo stesso spazio virtuale e di interagire tra loro e con l'opera in tempo reale. L'impiego del **riconoscimento gestuale** consente di eliminare i controller, favorendo un'interazione diretta e naturale che intensifica il senso di **embodiment**, ovvero la percezione del proprio corpo come parte integrante dell'ambiente digitale. Questo aspetto risulta fondamentale per la coerenza con la poetica di Lupone, che ha sempre posto al centro la fisicità del gesto e la sua valenza espressiva.

Un ulteriore vantaggio della declinazione VR è la sua **portabilità**: l'opera può essere presentata in contesti museali, festival e installazioni temporanee senza i vincoli logistici e materiali, raggiungendo così un pubblico più ampio e diversificato. Al contempo, la versione virtuale preserva la **memoria collettiva** delle interazioni, traducendola in un archivio digitale che continua a evolversi con il contributo dei partecipanti.

La finalità ultima di **SHARED VR** è quindi duplice: da un lato **preservare e potenziare** i valori artistici e concettuali dell'opera originaria, incentrati su memoria e adattività, e dall'altro **sperimentare nuove modalità di fruizione** che possano aprire scenari futuri per la conservazione e la diffusione del patrimonio artistico attraverso media immersivi. In questo senso, la versione VR non rappresenta una semplice trasposizione tecnica, ma un vero e proprio sviluppo evolutivo del progetto, capace di proiettare SHARED verso nuove forme di partecipazione e di esperienza collettiva.

1.3 Architettura del progetto e contributo personale

L'architettura originaria del laboratorio SHARED nasce come sistema fisico modulare e adattivo, capace di integrare musica e arti visive in un'unica esperienza interattiva.

La configurazione fisica dell'opera prevede moduli plastici triangolari collegati fra loro, pensati per deformarsi e rispondere ai movimenti del pubblico. Ogni modulo integra sensori di movimento e di contatto, attuatori interni e diffusori sonori, che consentono di generare risposte audio-visive in tempo reale. L'interazione con le forme avviene attraverso movimenti del corpo e gesti delle mani ispirati alla gestualità della direzione orchestrale: questa scelta progettuale riflette la volontà di utilizzare un linguaggio corporeo immediatamente riconoscibile e al tempo stesso ricco di sfumature espressive.

La declinazione in realtà virtuale, oggetto di questa tesi, parte da questa struttura concettuale per trasporla in un ambiente digitale immersivo e multiutente. L'obiettivo non è riprodurre fedelmente la fisicità dei moduli, ma preservarne la logica di interazione e di risposta adattiva, sfruttando le potenzialità offerte dalla VR.

La progettazione della versione VR ha comportato una serie di scelte tecnologiche mirate. L'intero sviluppo è stato realizzato con il motore **Unity**, scelto per la sua flessibilità e per la compatibilità con i dispositivi standalone come *Oculus Quest 3*. Per la gestione dell'interazione si è adottato l'utilizzo di *XR Interaction Toolkit* in combinazione con *XR Hands*, due librerie che consentono di implementare il tracciamento delle mani e di tradurre i gesti in azioni interattive, eliminando la necessità di controller fisici. La parte di rete implementata con *Netcode for GameObjects*, soluzione nativa di *Unity* che semplifica la sincronizzazione in tempo reale tra i client e garantisce una migliore scalabilità per applicazioni multiutente.

Per la componente sonora si è mantenuto il legame con la poetica originaria di Lupone attraverso l'integrazione con **Max/MSP**, utilizzato per la generazione audio dinamica e spazializzata. La comunicazione tra Unity e Max/MSP avviene tramite protocollo *OSC (Open Sound Control)*, che permette lo scambio di dati in tempo reale e consente all'ambiente virtuale di reagire immediatamente alle azioni degli utenti. Infine, l'utilizzo del *Universal Render Pipeline (URP)* ha garantito un rendering ottimizzato e prestazioni elevate sui visori standalone, mantenendo una resa visiva coerente con le esigenze artistiche del progetto.

Il mio contributo si è concentrato proprio sullo sviluppo tecnico della versione VR, infatti ho realizzato il progetto Unity, integrato le tecnologie necessarie e definito i flussi di comunicazione tra i diversi moduli del sistema. L'attività ha richiesto non solo competenze di programmazione e di gestione della rete multiutente, ma anche un costante confronto con la poetica di SHARED, per garantire che ogni scelta tecnica fosse coerente con i principi di partecipazione, memoria e adattività che definiscono l'opera.

L'analisi tecnica delle scelte implementative e delle soluzioni adottate sarà approfondita nella seconda parte di questa tesi, dedicata alla progettazione e allo sviluppo del prototipo.

1.4 Obiettivi della tesi

Questo elaborato si propone di documentare e analizzare approfonditamente lo sviluppo di *SHARED VR*, mettendo in relazione dimensioni tecniche e concettuali del progetto. L'obiettivo principale è mostrare come sia stato possibile **trasporre in ambiente virtuale** i principi fondanti dell'opera preservandone la coerenza artistica e, al tempo stesso, valorizzare le opportunità offerte dallo strumento della realtà virtuale.

Il lavoro si articola in due parti principali:

- la prima introduce il **contesto teorico e artistico**, con particolare attenzione al rapporto tra arte immersiva e *cultural heritage*, i concetti di *embodiment* e presenza e anche le tendenze della New Media Art;
- la seconda descrive la **progettazione e lo sviluppo del prototipo VR, a partire** dall'architettura del sistema, all'implementazione del *hand-tracking* e del *networking multiutente*. Inoltre, raccoglie i **risultati del testing e della validazione**, offrendo una riflessione critica sull'esperienza e analizzando possibili evoluzioni del progetto.

Oltre alla descrizione tecnica, l'elaborato si propone di riflettere criticamente sul ruolo della realtà virtuale come mezzo per la creazione artistica e per la condivisione del patrimonio culturale.

In questo modo la tesi non si limita solo a raccontare il processo di realizzazione dal punto di vista tecnico, ma contribuisce al dibattito contemporaneo sulle potenzialità dei media immersivi nella New Media Art, offrendo una chiave di lettura differente sulla funzione della realtà virtuale nel processo artistico e sulle sue potenzialità per arricchire l'esperienza culturale della collettività.

2 SHARED VR

Questo capitolo presenta **SHARED VR**, la declinazione in realtà virtuale del progetto SHARED, con l'obiettivo è fornire al lettore una comprensione chiara dell'esperienza. I dettagli tecnici (architettura software/hardware, implementazione del tracciamento mani e del networking) saranno trattati nella seconda parte di questa tesi.

All'interno di SHARED possiamo distinguere il **Regno Fisico**, che comprende la *Concert Performance* e l'*Installazione Partecipativa* con gesti acustici e interfacce tangibili, e il **Regno Virtuale**, che trasferisce e amplifica queste dinamiche all'interno di un ambiente VR simulato. [6]

Nel concerto, un ensemble composto da violino, violoncello e tuba, esegue una partitura aperta in cui i gesti strumentali costituiscono l'input primario. Una patch Max/MSP cattura parametri timbrici e li traduce in controlli di sintesi, spazializzazione ambisonica e deformazioni meccatroniche dei pannelli, mentre i medesimi flussi alimentano visual generativi proiettati nello spazio scenico.

Nell'installazione partecipativa, invece, sono i visitatori a intervenire: un sistema di hand-tracking a infrarossi riconosce gesti statici dati da delle pose della mano (pinch, palmo su/giù e pugno) e gesti tracciati (cerchi, curve, angoli, segni di infinito). I primi sono interpretati come eventi discreti per la selezione o l'attivazione di texture sonore, i secondi come traiettorie continue che modulano spazializzazione e timbro. Una patch Max/MSP ricalibrata calcola velocità e curvature dei movimenti per controllare sintesi e spazializzazione, mentre un modulo multi-agente fonde i contributi simultanei e registra le interazioni, costruendo una memoria persistente. Anche qui, i pannelli flessibili rispondono deformandosi, in coerenza con il paesaggio sonoro e visivo generato.

L'etica interattiva di SHARED si estende in un ambiente VR multiutente, preservando e amplificando le potenzialità co-creative. Il regno virtuale rispecchia gli elementi essenziali, come vocabolario gestuale, motore audio adattivo, feedback multisensoriale, introducendo al contempo affinamenti specifici della VR.

2.1 Concept e obiettivi artistici

SHARED si pone in antitesi al paradigma mimetico della logica del *digital twin*, adottando un approccio di *extensive XR* che non riproduce semplicemente il reale ma lo reinterpreta, offrendo nuovi spazi di agency e co-creazione. [6] L'obiettivo non è riprodurre fedelmente i moduli dell'installazione fisica, ma utilizzare la realtà virtuale come **medium autonomo**, capace di generare nuove condizioni di esperienza.

SHARED VR si propone dunque come un'installazione artistica procedurale multiutente, che privilegia la costruzione di un'esperienza **simbolica e relazionale**, in cui il senso nasce dalle azioni dei partecipanti e dalle loro interazioni reciproche.

Il concept pone al centro l'idea di un oracolo, un'entità reattiva che memorizza, trasforma e restituisce i gesti del pubblico sotto forma di luce e suono. Questo dispositivo diventa il fulcro della composizione audiovisiva, un organismo vivo che cresce, pulsa e si dissolve in base al contributo degli utenti.

Il concept si fonda principalmente su tre obiettivi artistici:

- **Esplorare la VR come linguaggio artistico:** SHARED VR si costruisce come ambiente simbolico, sospeso e astratto, che sfrutta le caratteristiche proprie del medium immersivo, come immersione, astrazione e manipolazione sensoriale, per generare forme espressive inedite.
- **Rendere l'opera accessibile e diffusa:** la versione VR è stata pensata per essere presentata in musei, festival e installazioni temporanee, rendendo il progetto più portatile e itinerante rispetto alla complessità logistica dei moduli fisici.
- **Sperimentare modelli di co-creazione:** SHARED VR è stato sviluppato attraverso un processo di co-design agile tra artisti, ingegneri e ricercatori, in cui il valore estetico emerge dall'interazione collettiva e non da una partitura predefinita. Questa impostazione riflette l'approccio dell'*artistic research*, che pone la pratica artistica come forma di produzione di conoscenza. [7][8]

In questa prospettiva, SHARED VR si configura come un dispositivo artistico-processuale, un'opera generativa che integra ricerca estetica, innovazione tecnologica e riflessione critica sui media immersivi. Non un adattamento digitale, ma un terreno di sperimentazione in cui corpo, suono e ambiente virtuale ridefiniscono i confini dell'esperienza condivisa.

2.2 Esperienza dell'utente

L'esperienza di SHARED VR è pensata per essere accessibile e intuitiva, in modo che chiunque possa entrarvi senza particolari competenze tecniche o familiarità con i dispositivi immersivi.

La dimensione multiutente costituisce un elemento essenziale, l'esperienza è progettata per tre partecipanti che coabitano lo stesso spazio virtuale. L'utente non è chiamato a interpretare un avatar né a muoversi in scenari complessi: la sua presenza è suggerita in maniera essenziale e simbolica attraverso le mani luminose ed effetti visivi che ne suggeriscono l'esistenza. Ogni utente inoltre è rappresentato da un colore specifico che aiuta nella distinzione tra i partecipanti, associando ogni gesto a una specifica identità visiva e sonora. Questa scelta allontana l'esperienza dal linguaggio del videogioco e la orienta verso una dimensione rituale, in cui ogni partecipante percepisce la presenza altrui senza che questa diventi ingombrante. La co-presenza diventa così una condizione di dialogo silenzioso: non ci sono parole né obiettivi da raggiungere, ma solo un ambiente sensibile da costruire insieme.

Lo spazio virtuale appare come un ambiente atemporale e privo di gravità, dove non esistono confini architettonici o riferimenti reali. Al centro, si colloca l'oracolo, una struttura non solida ma emergente, dalla forma di cilindro a base triangolare le cui facce non sono esplicite, bensì

disegnate da particelle luminose in costante movimento, che ne suggeriscono i volumi, la pulsazione e l'evoluzione. Non si tratta di un semplice elemento scenico, ma di un'entità sensibile che accoglie i contributi dei partecipanti, li memorizza, li trasforma e li restituisce come segni sonori e visivi che arricchiscono l'ambiente condiviso.

L'utente percepisce l'oracolo come un organismo instabile, vivo e responsivo, che si espande, si contrae e si trasforma in base ai contributi gestuali; infatti, ogni suo gesto non si esaurisce nell'istante in cui viene compiuto ma viene assorbito e stratificato.

L'obiettivo artistico del progetto è dunque quello di trasformare l'atto corporeo del gesto in un linguaggio espressivo universale, restituendo al corpo la centralità che spesso le interfacce tecnologiche tendono a ridurre. L'eliminazione dei controller a favore del tracciamento delle mani permette un'interazione più naturale, immediata e intuitiva, in cui le mani diventano il tramite diretto della creatività, con cui l'utente disegna scie luminose, modella il suono, lascia tracce che sopravvivono oltre il momento presente.

Un altro aspetto fondamentale è che ogni esperienza è unica, perché ciò che accade dipende esclusivamente dalle azioni del gruppo che vi prende parte. In questo senso SHARED VR è un'opera generativa e adattiva, che si riscrive continuamente e che non è mai identica a sé stessa.

2.3 Struttura narrativa

L'esperienza di SHARED VR si articola come un percorso drammaturgico che accompagna i partecipanti lungo un arco emozionale ben definito. L'ingresso nello spazio avviene in un vuoto assoluto, un ambiente astratto e privo di riferimenti, che genera da subito un senso di sospensione e di meraviglia. Al centro di questo spazio sorge l'oracolo, la cui apparizione determina l'avvio dell'esperienza e concentra l'attenzione dei partecipanti, guidati da una voce che li introduce con delicatezza all'ascolto e al gesto: in questa fase il tempo sembra fermarsi e tutto è in attesa di essere generato. Gradualmente gli utenti scoprono che le mani costituiscono gli strumenti con cui accedere all'opera. Attorno a sé, ciascun utente, trova degli elementi sonori che può afferrare e attivare: una volta preso un elemento, viene riprodotto il suono associato e i movimenti della mano che lo trattiene ne modulano i parametri.

Attraverso un processo di apprendimento implicito, fatto di osservazione, tentativi ed errori non punitivi, ciascun partecipante impara così a riconoscere le relazioni tra gesto e suono, sviluppando un linguaggio personale e spontaneo di interazione.

Quando l'utente rilascia l'elemento aprendo la mano, dal suo palmo si genera una scia luminosa che viene assorbita dall'oracolo. All'interno di esso si genera una particella di luce del colore corrispondente all'utente, simbolo del suo contributo alla memoria collettiva. I contributi di ciascuno si sommano e vanno a comporre un paesaggio audiovisivo condiviso che cresce e si trasforma nel tempo.

Quando i contributi dei partecipanti superano una certa soglia, l'oracolo si espande, inglobando simbolicamente i tre partecipanti. Lo spazio si satura di luce e suono, le traiettorie si moltiplicano e l'esperienza culmina in un momento corale di co-creazione: questo segna la conclusione del

rito collettivo. Rimane negli utenti una sensazione di distacco malinconico, ma anche la consapevolezza di aver partecipato a un processo di co-creazione che ha lasciato traccia, una memoria viva destinata a sopravvivere oltre la durata dell'esperienza.

Questa struttura narrativa, che procede dalla sospensione iniziale alla dissoluzione finale passando per momenti di scoperta, composizione e libertà creativa, incarna in modo diretto i principi fondativi di SHARED: la **partecipazione condivisa**, che rende ogni utente co-autore dell'opera; la **memoria collettiva**, che conserva e trasforma le tracce dei gesti; e l'**adattività**, che permette all'ambiente di mutare costantemente in base alle azioni. Al tempo stesso, la scelta di un'estetica del vuoto e di una rappresentazione minimale della presenza umana orienta l'esperienza verso una dimensione simbolica e rituale, distante dalle logiche mimetiche della realtà virtuale tradizionale.

Proprio questo equilibrio tra immediatezza del gesto, visibilità della mediazione tecnologica e costruzione di un'opera collettiva apre la strada a una riflessione critica più ampia. SHARED VR non è solo un ambiente interattivo, ma un dispositivo estetico che mette in discussione il ruolo dello spettatore, le modalità di rappresentazione nella VR e i confini stessi dell'opera d'arte generativa.

Nel capitolo successivo è esposta una critica estetica del progetto, che lo colloca all'interno del panorama delle tecnologie immersive e della new media art, analizzando come memoria, gesto e partecipazione diventino i pilastri di una nuova modalità di fruizione e co-creazione.

3 SHARED VR: critica estetica

Nel corso dell'ultimo decennio le tecnologie immersive hanno assunto un ruolo sempre più centrale nel modo di concepire e vivere il patrimonio culturale e le esperienze artistiche contemporanee. Secondo Bekele et al. (2018) [10], l'integrazione di VR, AR e MR rappresenta oggi una strategia efficace per facilitare l'accesso al *cultural heritage* e coinvolgere attivamente i visitatori, non solo offrendo copie virtuali del patrimonio, ma creando esperienze in grado di generare coinvolgimento emotivo e cognitivo.

Carrozzino e Bergamasco (2010) [11] definiscono la realtà virtuale come una tecnologia complessa capace di creare ambienti digitali interattivi che comunicano attraverso feedback sensoriali, superando i tradizionali codici linguistici e rendendo l'esperienza accessibile anche a utenti non specializzati. Tuttavia, gli autori sottolineano come molte istituzioni culturali, in particolare i musei, non sfruttino pienamente il potenziale della VR, limitandosi spesso a impiegarla come strumento di supporto multimodale piuttosto che come linguaggio artistico in grado di ridefinire l'esperienza stessa del patrimonio.

Riprendendo queste considerazioni, Bekele individua nell'immersione e nell'interazione i due pilastri di un'esperienza VR efficace: la possibilità per l'utente di sentirsi parte integrante di uno spazio virtuale e di modificarlo attraverso il proprio gesto. Questi concetti trovano una diretta corrispondenza nei tre principi fondanti di SHARED-ItARTS già presentati in precedenza: partecipazione, memoria collettiva e adattività. Il progetto SHARED si colloca, infatti, in una prospettiva che va oltre la replica del reale, proponendo un luogo virtuale condiviso in cui la memoria sonora e visiva si costruisce e si trasforma attraverso l'interazione di più utenti.

Per comprendere come tali principi si inseriscano nel panorama più ampio delle applicazioni immersive nel *cultural heritage*, questo capitolo analizza sia i concetti su cui si basano le teorie dei nuovi media, sia una serie di casi artistici significativi che hanno esplorato il ruolo del corpo e l'interazione nello spazio virtuale. L'ultima parte del capitolo approfondirà i temi della multiutenza, dell'interazione gestuale e della natura generativa delle opere, che rappresentano il nucleo concettuale e tecnico del prototipo sviluppato per SHARED-ItARTS.

3.1 Teorie dei nuovi media e cultural heritage

Nel contesto del *cultural heritage*, in cui le tecnologie immersive vengono sempre più impiegate per trasformare la fruizione del patrimonio culturale in esperienze partecipative, SHARED VR propone un'interpretazione immersiva del patrimonio sonoro e performativo contemporaneo, in cui la memoria collettiva si costruisce nel tempo attraverso la partecipazione attiva del pubblico.

Per comprendere pienamente le logiche culturali e tecnologiche che guidano questo approccio, è utile richiamare alcune teorie fondamentali dei nuovi media, in particolare il concetto di rimediazione (Bolter e Grusin) [12] e i principi di modularità e variabilità (Manovich) [13], che

forniscono un quadro teorico per analizzare il rapporto tra media digitali, patrimonio culturale e interazione.

Il concetto di rimediazione, introdotto da Bolter e Grusin in *Remediation: Understanding New Media*, descrive come un nuovo medium rappresenta, incorpora e rielabora uno o più media precedenti. *SHARED VR*, infatti, rimedia la performance sonora installativa originaria, trasponendola in un linguaggio immersivo e interattivo. Alla base di questo processo vi è la “doppia logica” che struttura i nuovi media: la logica dell’**immediatezza**, che tende a nascondere la mediazione per offrire un’esperienza il più trasparente possibile, e la logica dell’**ipermediazione**, che al contrario espone la mediazione, moltiplicando i livelli di rappresentazione e richiamando l’attenzione sulla presenza stessa del medium.

Bolter e Grusin descrivono la realtà virtuale come un medium il cui scopo è scomparire, introducendo il concetto dell’**immediatezza trasparente** in cui l’utente riceve un punto di vista in prima persona, che rimane sempre il centro dell’ambiente, e i movimenti della testa sono tracciati in tempo reale dal sistema per restituire una visuale dinamica e coerente. Questo tipo di esperienza favorisce un coinvolgimento immersivo, in cui il soggetto percepisce di trovarsi all’interno dell’ambiente virtuale stesso.

In *SHARED VR* ritroviamo proprio questi principi: l’ambiente virtuale in cui si ritrova l’utente risponde ai suoi movimenti e ai suoi gesti. Il poter interagire con il proprio corpo, senza l’uso di controller, potenzia la **logica di immediatezza** descritta da Bolter e Grusin: l’interfaccia sembra scomparire e l’esperienza appare naturale e trasparente, come se non vi fosse mediazione. Tuttavia, questa trasparenza è continuamente messa in discussione dalla presenza visibile dell’oracolo, dagli effetti audio-visivi generati in tempo reale e dalla co-presenza degli altri utenti. Questi elementi rendono esplicita la mediazione tecnologica in atto, attivando la logica dell’**ipermediazione**, che ricorda costantemente al pubblico che ciò che vive è una costruzione mediale, multilivello e condivisa.

La memoria condivisa costruita all’interno di *SHARED VR* non è solo un archivio digitale, ma un ambiente esperienziale in cui il gesto del singolo influisce sulla percezione collettiva. La logica immersiva della VR viene così combinata con una struttura relazionale che riflette i cambiamenti nella definizione del sé e della partecipazione nei media digitali. Bolter e Grusin parlano infatti di un **sé immerso**, legato alla soggettiva visiva e al coinvolgimento corporeo, e di un **sé connesso**, che si costruisce attraverso la rete delle relazioni digitali. In *SHARED*, queste due dimensioni convivono: ogni utente agisce all’interno di un ambiente in prima persona, ma allo stesso tempo co-costruisce l’esperienza con gli altri partecipanti, contribuendo a formare una memoria collettiva in costante evoluzione.

Per comprendere più a fondo le dinamiche operative di SHARED VR, è utile approfondire i principi dei nuovi media esposti da Lev Manovich, in particolare quelli di modularità e variabilità, che permettono di analizzare la struttura flessibile e generativa del sistema.

In *The Language of New Media* (2001), Lev Manovich propone cinque principi fondamentali dei nuovi media. Tra questi, la modularità descrive una logica compositiva in cui i contenuti siano costituiti da unità indipendenti e riutilizzabili, che possono essere combinate tra loro in forme sempre diverse. Ogni modulo mantiene la propria identità ma si adatta al contesto in cui viene inserito. In SHARED, questa modularità è evidente nella struttura del sistema sonoro e visivo, che è composta da frammenti discreti che possono essere attivati separatamente o in combinazione, producendo configurazioni sempre nuove. Questa struttura modulare rende l'opera flessibile, scalabile e adattabile, inoltre permette al pubblico di interagire senza vincoli rigidi di percorso o sequenza, contribuendo alla costruzione dell'esperienza in modo attivo.

Il principio della variabilità si lega strettamente a quello della modularità e si riferisce alla capacità dei media digitali di generare molteplici versioni di uno stesso contenuto in base agli input ricevuti. SHARED incarna perfettamente questo principio in quanto l'interazione di ogni utente, e ancor più la co-interazione di più utenti simultanei, genera una diversa configurazione dell'opera, una variazione sonora e visiva non predeterminata. La memoria sonora condivisa non è mai uguale, ma si evolve, si modifica, si stratifica, rendendo ogni esperienza unica, irripetibile e in continua trasformazione.

Inoltre, SHARED adotta un modello di interattività aperta, in cui gli utenti determinano liberamente l'ordine e le modalità di accesso ai contenuti, senza dover seguire un percorso prestabilito. Questo tipo di interazione, secondo Manovich, è una sottocategoria del principio di variabilità e rappresenta una delle forme più avanzate di partecipazione nei nuovi media. L'utente, quindi, non solo attiva contenuti, ma trasforma l'opera stessa, intervenendo in un processo generativo che si sviluppa in tempo reale e in relazione agli altri partecipanti. Il risultato è una forma d'arte interattiva e adattiva, in cui l'opera è sempre aperta e mai conclusa, ma si trasforma continuamente attraverso l'esperienza collettiva.

Le dinamiche discusse si inseriscono in una più ampia tendenza artistica contemporanea, in cui le tecnologie immersive sono sempre più impiegate per esplorare nuovi modelli di partecipazione, corporeità e relazione, come si vedrà nei casi studio trattati nel prossimo paragrafo.

3.2 Dall'evoluzione dei musei virtuali alle installazioni immersive VR per l'arte

L'idea di museo virtuale nasce già negli anni '90, grazie alle prime tecnologie di modellazione 3D e rendering in tempo reale, quando diversi musei e centri di ricerca iniziarono a sperimentare forme di digitalizzazione del patrimonio culturale. Le prime esperienze si concentravano sulla ricostruzione fedele di ambienti, opere e reperti, con l'obiettivo primario di garantire la conservazione e l'accesso remoto. Addison (2001) [14] definisce il *virtual heritage* come la combinazione delle tecnologie digitali per documentare, analizzare e diffondere la cultura.

Le prime realizzazioni di musei virtuali si concretizzavano in modelli 3D statici, esplorabili attraverso interfacce rudimentali, dai CD-ROM interattivi basati su QuickTime VR ai primi siti web in VRML⁴. Progetti come il Virtual Museum of Canada o le ricostruzioni archeologiche sviluppate nell'ambito delle conferenze VAST (*Virtual Reality, Archaeology and Cultural Heritage*) testimoniavano la volontà di trasporre fedelmente lo spazio fisico nello spazio digitale.

Tuttavia, in questo contesto l'utente, pur potendo "navigare" nello spazio, assumeva un ruolo sostanzialmente passivo, privo di un reale potere di modifica o influenza sull'ambiente. Le priorità di questa fase erano la precisione geometrica, la fedeltà cromatica e la completezza della ricostruzione. Aspetti fondamentali per la conservazione, ma che lasciavano poco spazio all'interpretazione personale o alla partecipazione attiva.

Con il consolidarsi delle tecnologie interattive e dei dispositivi immersivi, l'attenzione si sposta progressivamente dalla replica del reale alla costruzione di esperienze partecipative, aprendo la strada alle installazioni immersive in VR.

La VR si dimostra particolarmente efficace nel trasmettere emozioni e creare quella che viene definita empatia tra l'utente e il contenuto. L'immersione sensoriale, resa possibile dalla completa esclusione del mondo fisico, favorisce una connessione diretta e multisensoriale con l'opera.

Secondo Uyan Dur (2021) [15] la VR, rispetto a tutti gli altri medium, possiede un potenziale maggiore nell'ampliare la percezione, l'analisi e l'interpretazione da parte degli artisti. Precursori della VR come Brenda Laurel (1995) [16], sottolineano il potenziale della VR come mezzo artistico, anticipando come la tecnologia si sarebbe evoluta in base alle necessità degli artisti:

"Gli artisti, che hanno sempre dovuto riflettere sull'interazione tra le risposte intellettuali e fisiche alle loro opere, possono svolgere un ruolo più centrale nello sviluppo della VR rispetto ai tecnologi, i quali potrebbero accontentarsi di considerare i computer come un mezzo che si rivolge esclusivamente all'intelletto umano disincarnato."

⁴ VRML (Virtual Reality Modelling Language) è un formato di file nato nel 1994 per rappresentare grafica vettoriale 3D interattivi. VRML sfruttava un semplice file di testo per specificare le caratteristiche del poligono desiderato, inoltre era possibile aggiungere agli elementi grafici degli URL in modo da permettere l'apertura di una pagina web o un nuovo file VRML da Internet attraverso un web browser. Animazioni, suoni, illuminazione ed altri aspetti del mondo virtuale potevano interagire con l'utente o attivati da eventi esterni come i timer. I browser odierni non supportano nativamente VRML e per fruirne i contenuti era necessario anche all'epoca fare ricorso a particolari plug-in. Si può considerare a tutti gli effetti un formato obsoleto, non più sviluppato o utilizzato.

L'analisi di Marshall McLuhan in *Understanding New Media: The Extension of Man* (1964) [17], aiuta a collocare la VR in una prospettiva storica, evidenziando come ogni nuovo medium sia legato alle logiche dei media precedenti. La VR, infatti, eredita le logiche dei media precedenti quali cinema e videogiochi, ma grazie alla sperimentazione artistica si arriva oltre il semplice riuso di forme preesistenti.

Il termine "*virtual reality*", viene coniato nel 1987 da Jaron Lanier, figura pionieristica della realtà virtuale. In *Dawn of the New Everything* (2017) [18], Lanier propone 52 diverse definizioni di VR, tra cui:

«una nuova forma d'arte che deve sfuggire alle grinfie del gaming, del cinema, del software tradizionale, delle strutture di potere della New Economy e, forse, anche delle idee dei suoi pionieri» (p. 237)

Secondo Lanier, definire la VR come *nuova forma d'arte* significa considerarla come un campo che va oltre la tecnologia: gli artisti giocano un ruolo fondamentale nell'esplorazione, nello sviluppo e nell'evoluzione delle sue potenzialità.

Lo sviluppo delle tecnologie e delle Creative VR Apps ha portato a un aumento della crescente diffusione della VR tra artisti e designer negli ultimi anni.

Tra le esperienze VR più significative del panorama artistico, vi è *Chalkroom* (2017) [19], installazione artistica in realtà virtuale interattiva co-creata da Laurie Anderson e Hsin-Chien Huang, che rappresenta un caso emblematico di come la scrittura possa diventare elemento spaziale e narrativo in un ambiente immersivo. Vincitore del premio Best VR Experience alla 74ª Mostra del Cinema di Venezia, il progetto nasce con l'obiettivo di scardinare le convenzioni normalmente associate al genere, cercando di sfruttare il medium della realtà virtuale secondo modalità inedite. L'opera consiste in un'enorme struttura, fatta di scritte, parole e disegni, che l'utente può esplorare volando.

Il concetto alla base può essere interpretato attraverso i tre aspetti analizzati da Uyan Dur (2021): tipografia, calligrafia e lettering. Questi sono i concetti chiave che permettono di analizzare come influenzare le emozioni del lettore/utente e modellarne l'attenzione e la percezione.

La tipografia, secondo Ambrose e Harris (2015) [20], è il mezzo attraverso cui un'idea scritta assume una forma artistica. Dunque, la **tipografia** viene intesa come un sistema ripetibile che può combinarsi in infinite modalità e si riferisce alla disposizione delle lettere in una composizione. La **calligrafia** invece è un'arte basata su regole e principi, il cui scopo è dare un'espressione estetica ed elegante alla scrittura. Infine, il **lettering** si basa sulla creazione di una composizione originale e unica.

In *Chalkroom*, una volta immerso nella VR, l'utente si muove verso il centro della *Chalkroom*, percorrendo un corridoio buio, accompagnato dalla narrazione audio di Laurie Anderson. Durante la navigazione, l'utente entra a far parte della narrazione vivendo diverse forme di

interazione in ciascuna stanza e percepisce l'effetto visivo di testi e disegni come un insieme, leggendo in modo casuale i frammenti che attirano la sua attenzione. In questo senso, *Chalkroom* offre possibilità di esplorazione illimitate e consente a ogni visitatore di vivere un'esperienza unica. Come ha dichiarato Laurie Anderson [21], *Chalkroom* «è una biblioteca di storie, e nessuno riuscirà a trovarle tutte».

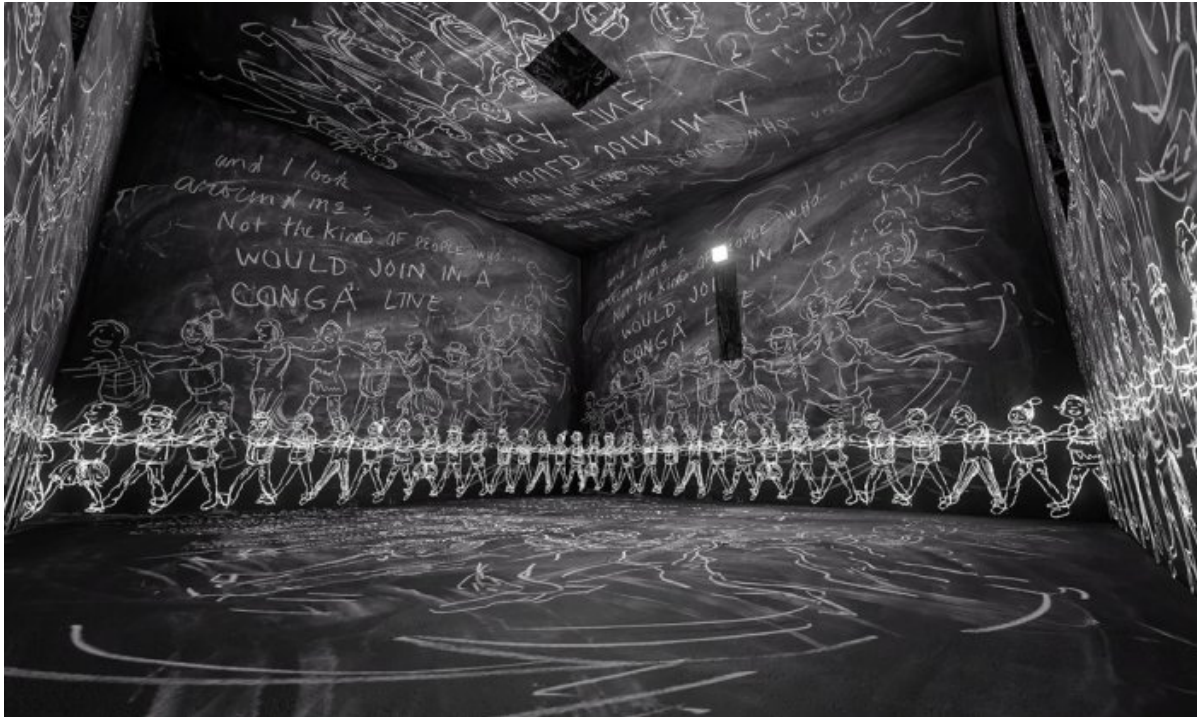


Figura 3.2.1: *Chalkroom* di Laurie Anderson and Hsin-Chien Huang, The Dance Room, 2017 [F3.2.1]

La comprensione libera e intuitiva della scrittura ricorda la natura anomala della tipografia sperimentale. L'irregolarità dei testi, i segni del cancellino di legno, le lettere fluttuanti, le superfici polverose e sporche creano un'atmosfera ben diversa dalle esperienze VR convenzionali.

Chalkroom riesce a combinare con successo il contesto concettuale e intellettuale con le possibilità offerte dalla VR. È un'esperienza che amplia la percezione nei contesti di corpo, essere, mente, linguaggio e narrazione. L'opera non solo dimostra le possibilità uniche che l'ambiente VR può offrire per la narrazione attraverso la tipografia sperimentale, ma evidenzia anche i diversi usi della VR in contrasto con le pratiche convenzionali.

Questa chiave di lettura può essere estesa anche ad altre opere VR. Se in *Chalkroom* il vocabolario è costituito da segni testuali, in *SHARED VR* invece è costituito da suoni forma, unità di contenuto sonoro che si modificano in base alle azioni dell'utente, creando quella che si potrebbe definire una tipografia sonora.

Allo stesso modo, in *We Live in an Ocean of Air* di Marshmallow Laser Feast (2018) [22] i "segni" sono i dati biometrici (battito cardiaco, respiro, movimento) che tracciano una calligrafia

corporea nello spazio virtuale, mentre nell'opera *teamBorderless* (2018) [23] forme, luci e pattern digitali costituiscono un lettering ambientale che viene completamente riscritto in base alla presenza e ai movimenti del pubblico.

Pur adottando linguaggi sensoriali diversi, è possibile affermare che le installazioni immersive VR condividano dunque una logica modulare e combinatoria, in linea con i concetti di modularità e variabilità teorizzati da Manovich e con l'idea di opera generativa che si trasforma attraverso l'interazione.

L'approccio generativo e partecipativo di SHARED VR trova affinità con opere immersive che condividono la logica di un ambiente che si trasforma in base all'azione del pubblico.

Il concept alla base di *We Live in an Ocean of Air* di Marshmallow Laser Feast (2018), è quello di far vivere ai visitatori l'esperienza della fusione totale con uno scenario di natura digitale grazie all'impiego della VR in combinazione a sensori biometrici. Oltre il visore, gli utenti indossano dei sensori sui polsi e uno zaino wireless per la raccolta, trasmissione e rappresentazione di dati biometrici quali frequenza cardiaca, respirazione e rilevamento dei movimenti: tali dati vengono elaborati per influenzare dinamicamente l'ambiente circostante. Come in SHARED, l'esperienza è concepita per essere vissuta collettivamente, ogni utente contribuisce attivamente all'opera. Tuttavia, una differenza sostanziale è che in SHARED l'unico strumento utilizzato è il visore senza l'utilizzo di ulteriori sensori per l'hand tracking. L'uso di un unico strumento riduce la percezione di mediazione tecnologica, aumentando il senso di immersione e di naturalezza dell'interazione.



Figura 3.2.2: *We Live in an Ocean of Air* di Marshmallow Laser Feast [F3.2.2]

TeamLab Borderless (2018) è un museo immersivo situato a Tokyo concepito come uno spazio senza mappa, in cui il visitatore è invitato ad esplorare liberamente senza seguire un percorso prestabilito. Gli ambienti sono generati in tempo reale per garantire un'esperienza unica e irripetibile. Il pubblico è parte attiva dell'esperienza in quanto l'opera si trasforma in base alla presenza e i movimenti degli utenti. Il concetto chiave è la co-costruzione dell'esperienza, aspetto che viene riproposto anche in SHARED VR dove, però, la dimensione collettiva non si sviluppa in un grande spazio fisico condiviso, ma in uno spazio virtuale multiutente, dove le azioni di ciascun partecipante incidono sia sull'ambiente visivo sia sulla memoria condivisa.



Figura 3.2.3: *TeamLab Borderless*, Tokyo [F3.2.3]

In conclusione, SHARED VR si inserisce in un contesto di opere immersive che superano la rappresentazione statica, adottando sistemi modulari e reattivi alle azioni dell'utente. Ciò che caratterizza SHARED VR è l'uso del suono come elemento generativo primario, che si combina con l'interazione gestuale per costruire una memoria condivisa.

3.3 Multiutenza, gesto e generatività

La co-presenza multiutente, l'uso del gesto come interfaccia naturale e la generatività che permette all'esperienza di trasformarsi in base alle azioni dell'utente, sono i tre aspetti principali di SHARED VR.

Le tecnologie per spazi condivisi mirano a creare ambienti elettronici distribuiti, in cui i partecipanti possano sfruttare proprietà spaziali come il contenimento e il movimento per gestire la propria comunicazione. Benford et al. (2001) [24] raggruppa in cinque categorie gli approcci spaziali al CSCW (*Computer Supported Cooperative Work*), tra questi ci sono gli ambienti virtuali collaborativi.

Gli Ambienti Virtuali Collaborativi (CVE) si basano su sistemi di realtà virtuale in rete per supportare il lavoro di gruppo (Benford et al., 1994) [25] e consistono in mondi virtuali condivisi, ovvero spazi generati dal computer i cui occupanti sono rappresentati graficamente tramite avatar. Lo spazio condiviso consiste in un sistema di coordinate accessibile a tutti, che permette di comprendere posizioni e orientamenti degli oggetti: ogni utente può controllare il proprio punto di vista e può interagire con gli altri e con lo spazio virtuale.

I CVE cercano anche di offrire un contesto persistente e integrato per la cooperazione, unendo nello stesso spazio sia i partecipanti sia le informazioni. Un presupposto comune è che i partecipanti "lascino il mondo fisico" per entrare nel mondo virtuale.

Per classificare le tecnologie di spazi condivisi, Benford introduce tre concetti necessari sia per esplorare i compromessi di design coinvolti nell'applicazione di queste tecnologie a supporto di diverse attività cooperative, sia a identificare vuoti nella tecnologia dove nuovi approcci potrebbero essere sviluppati:

- il **trasporto** consiste nella differenza tra i concetti in locale e in remoto. Il locale si caratterizza per la permanenza dei partecipanti nel proprio spazio originario, che diviene il punto di riferimento dell'interazione. Mentre il remoto si definisce come la condizione in cui i partecipanti e gli oggetti abbandonano il proprio spazio originario per accedere a uno spazio altro. Il concetto di trasporto è dunque simile a quello di realtà virtuale ma a differenza dell'immersione, il trasporto include anche la possibilità di introdurre partecipanti e oggetti remoti nell'ambiente locale.
- L'**artificialità** è la misura in cui uno spazio è sintetico o si basa sul mondo fisico. Descrive il continuum tra la sintesi totale di un ambiente, indipendente da ogni realtà esterna, e la trasmissione mediata elettronicamente di un luogo fisico.
- La **spazialità** riguarda il livello di supporto per le proprietà spaziali fisiche fondamentali come contenimento, topologia, distanza, orientamento e movimento. Questo aspetto si caratterizza per le nozioni di **luogo**, un contesto contenitivo per i partecipanti, e di **spazio**,

un contesto che fornisce un quadro spaziale di riferimento coerente, navigabile e condiviso.

Di conseguenza, SHARED VR si configura come un ambiente collaborativo che riprende i concetti presentati da Benford. In termini di trasporto, il presupposto dell'opera è quella di far immergere l'utente in uno spazio virtuale condiviso: la multiutenza e la co-presenza mantengono un legame, seppur mediato, con la dimensione fisica esterna. Per quanto riguarda l'artificialità, lo spazio virtuale di SHARED è totalmente sintetico e segue logiche interattive proprie, senza alcun riferimento diretto a luoghi reali. Infine, sul piano della spazialità, il senso di spazio è quello predominante: l'utente è sempre a conoscenza della posizione degli altri partecipanti e ha riferimenti visivi e sonori per orientarsi.

Un tratto distintivo di SHARED VR è la sua concezione di spazio come ambiente fluido e senza confini definitivi, in cui le soglie tra reale e virtuale, tra individuo e collettività, tra esperienza e memoria risultano costantemente ridefinite dall'azione del pubblico. Negli spazi virtuali, infatti, i confini non devono essere considerati come barriere, ma come soglie dinamiche capaci di modulare il grado di consapevolezza reciproca e di interazione tra domini fisici e sintetici (Benford et al., 2001). Questa visione si collega alla definizione di realtà mista proposta da Milgram e Kishino (1994) [26], intesa come *"fusione di mondi reali e virtuali"* tale che *"oggetti del mondo reale e oggetti del mondo virtuale sono presentati insieme all'interno di un unico display"*.

SHARED VR radicalizza questo approccio, proponendo un'esperienza in cui i confini stessi tendono a dissolversi, dando vita a un ecosistema partecipativo e generativo che non distingue più nettamente la presenza corporea dalla dimensione digitale.

All'interno di questo ambiente, l'utente è chiamato di fatto a interagire usando il proprio corpo: il gesto è usato come interfaccia naturale. Le *Natural User Interfaces* (NUI) si fondano proprio sulla possibilità di interagire con ambienti digitali attraverso movimenti intuitivi e spontanei, riducendo il carico cognitivo e favorendo l'immedesimazione. Lo studio di Manghisi et al. (2018) [27] ha dimostrato come l'impiego delle interfacce gestuali possa incrementare il coinvolgimento degli utenti. Lo studio è stato condotto nel contesto della *cultural heritage*, dove sono state proposte due interfacce per la navigazione virtuale dei siti archeologici della Murgia in Puglia. Per l'interfaccia gestuale è stato definito un vocabolario basato su tre fasi: la raccolta dei compiti rilevanti, l'elicitazione dei gesti spontanei da parte dei partecipanti e la selezione dei gesti più intuitivi. I test condotti hanno evidenziato che gli utenti non solo trovavano l'interazione più naturale rispetto al mouse, ma percepivano anche una maggiore immersione, al punto da sottostimare il tempo realmente trascorso nell'esperienza.

Dunque, le interfacce basate sui gesti migliorano la qualità della fruizione, rendendola più fluida e immersiva. SHARED VR si appropria di questo principio e lo rielabora in chiave artistica: il corpo non è solo mezzo di navigazione, ma diventa strumento espressivo e generativo; infatti, grazie

all'utilizzo dell'hand tracking l'utente interagisce con le proprie mani. Nella prima fase di sviluppo del progetto sono state individuate tre pose di base della mano, successivamente combinate con movimenti nello spazio. Tuttavia, come nello studio di Manghisi, anche in SHARED VR il vocabolario gestuale è stato progressivamente rivisto attraverso cicli di test con gli utenti, che hanno portato a una semplificazione delle azioni per favorire immediatezza e naturalezza.

Il risultato è che ogni gesto non si limita a costituire un comando, ma diventa un **atto generativo**, capace di trasformare lo spazio sonoro e visivo condiviso. Ed è proprio questa dimensione generativa, aperta a molteplici possibilità d'uso e di significato, che trova corrispondenza nelle riflessioni sulla **multistabilità tecnologica** elaborate da De Boer (2021) [28].

De Boer ragiona sulla multistabilità ovvero su come specifiche tecnologie possano essere utilizzate per una vasta gamma di scopi (il "multi"), ma non per tutti (la "stabilità").

Il concetto, già analizzato da Rosenberger (2020a) [29], descrive la possibilità che una tecnologia *"possa sempre essere impiegata per molteplici scopi, possa sempre inserirsi in molteplici contesti, possa sempre avere significati diversi per persone diverse, possa sempre evolvere in modi differenti all'interno di culture differenti [...] Allo stesso tempo [...] qualsiasi tecnologia è sempre limitata riguardo a ciò che può significare e a come può essere utilizzata"* (Rosenberger 2020a, p. 86).

Da questa definizione si deduce che una tecnologia non possa mai essere ridotta a un unico uso particolare: le sue applicazioni sono relative a specifici contesti socio-materiali e il significato che le persone le attribuiscono varia in base alle situazioni e ai valori coinvolti.

Kyle Whyte (2015) [30] distingue due modi diversi in cui la multistabilità viene messa in evidenza:

- La **multistabilità immaginativa** che riguarda l'indagine fenomenologica condotta in prima persona, in cui si formula un'ipotesi empiricamente verificabile su quante configurazioni stabili di uno stesso oggetto possano essere percepite e/o incorporate dal soggetto (ibid., p. 69).
- La **multistabilità pratica** va invece intesa come modo di indagare per mettere in luce come particolari intrecci uomo-tecnologia plasmino determinate forme di significato e di utilizzo all'interno di contesti sociali e culturali.

In entrambi i casi, la multistabilità rivela la capacità degli oggetti tecnologici di assumere ruoli diversi a seconda delle situazioni e delle comunità che li adottano. In questo senso, viene introdotto il concetto di "pivot", ovvero *"ciò che permette alla variazione di avere senso come variazione"* (ibid., p. 75), quindi gli elementi che rendono intelligibili le differenze di significato e di uso.

L'analisi di De Boer si intreccia con la teoria delle affordance di Gibson (1979) [31], intesa come possibilità d'azione offerta dall'ambiente in relazione alle abilità corporee dell'utente. Ogni

tecnologia, come ogni ambiente, presenta un “*paesaggio ricco di affordance*” (Rietveld & Kiverstein, 2014) [32], ma l’insieme di possibilità effettivamente percepite varia a seconda dei contesti comportamentali (Heft, 2001) [33] e delle pratiche sociali. In questo senso, le tecnologie non sono mai neutre né univoche: incorporano *affordance* che orientano l’azione in modi specifici, talvolta invitando o imponendo determinati comportamenti (Tollon, 2020) [34].

Questo quadro concettuale chiarisce come la generatività di SHARED VR sia espressione della multistabilità delle sue *affordance*: ogni gesto compiuto nel sistema non si esaurisce in un’azione determinata, ma apre traiettorie multiple che modificano in tempo reale la memoria sonora e visiva condivisa.

La generatività di questo progetto, quindi, si fonda sulla possibilità che le stesse tecnologie vengano continuamente ri-significate dall’interazione del pubblico. Come sottolinea la teoria della mediazione tecnologica (Verbeek, 2005) [35], le nuove tecnologie non solo modificano l’esperienza degli utenti, ma costituiscono nuove relazioni tra soggetto e oggetto, ridefinendo entrambi. Dunque, l’opera non si limita a ospitare interazioni, ma crea un ambiente in cui i partecipanti e la tecnologia si co-constituiscono reciprocamente, generando un processo estetico aperto, plurale e mai identico a sé stesso.

In questa prospettiva, le *affordance* di SHARED VR non sono percepite in astratto, ma attraverso corpi situati che vedono, sentono e agiscono nello spazio virtuale. È nel rapporto tra struttura tecnologica e capacità percettivo-motorie dei partecipanti che il gesto diventa atto generativo e che la co-presenza acquista senso. Il capitolo successivo approfondisce proprio questa dimensione, discutendo le basi fenomenologiche dell’*embodiment* e della presenza in VR.

4 Embodiment e Presenza in VR

Un'esperienza in VR non si limita a un rapporto cognitivo un ambiente virtuale, ma si fonda sulla capacità del corpo di sentirsi parte integrante di uno spazio che non è fisicamente tangibile. I concetti **embodiment** e **presenza** risultano essere fondamentali per comprendere questa dinamica.

Il concetto di **embodiment** si riferisce all'idea che la percezione e la cognizione siano indissolubilmente radicate nel corpo e nelle sue azioni: dunque, indica il fatto che i feedback sensoriali relativi al corpo virtuale dell'utente hanno un effetto sul sistema cognitivo dell'individuo. L'utente non conosce il mondo dall'esterno, ma attraverso l'esperienza incarnata di esso.

Per **presenza**, invece, si intende il senso soggettivo di "essere nel luogo", ovvero percepire l'ambiente virtuale come se fosse reale. Tale condizione richiede che l'utente sia essere predisposto ad accettare di sospendere l'incredulità, accogliendo la sensazione di trovarsi in un posto diverso da quello reale. In *At the heart of it all: The concept of presence* (1997) [36], Lombard e Ditton analizzano il concetto di presenza, sottolineando come il coinvolgimento dell'utente dipenda dalla capacità della tecnologia di farsi trasparente, permettendo all'individuo di dimenticare il medium e di vivere il contenuto come reale.

L'illusione di non mediazione sarà più completa se il medium è percettivamente e psicologicamente immersivo, in questo senso la VR è uno strumento potente in quanto l'utente non è solo spettatore, ma partecipe.

In questo capitolo si approfondiranno i fondamenti fenomenologici e cognitivi dell'embodiment, il rapporto tra corpo, ambiente e interfacce digitali e, infine, il senso di presenza e co-presenza in VR. L'obiettivo è mostrare come questi quadri teorici, applicati al caso di SHARED VR, permettano di comprendere più a fondo le dinamiche percettive e relazionali che caratterizzano l'opera.

4.1 Fondamenti fenomenologici e cognitivi dell'embodiment

Alla base di ogni esperienza percettiva si trova il corpo, inteso come condizione primaria di accesso al mondo. In *Phenomenology of Perception* (1945) [37], Merleau-Ponty descrive il **corpo proprio** come la matrice originaria della percezione: non vediamo, ascoltiamo o tocchiamo il mondo da una posizione astratta e disincarnata, ma attraverso la trama sensoriale e motoria che ci lega ad esso. Il corpo non è un oggetto tra gli altri, ma un medium che abita lo spazio e lo rende significativo. Attraverso il gesto, la postura e il movimento, l'individuo stabilisce un legame preriflessivo con la realtà che lo circonda, un legame che precede ogni elaborazione concettuale.

Da questo punto di vista, la percezione non è un processo di rappresentazione interna, ma un atto incarnato che si realizza nel contatto diretto con il mondo. Il corpo, come sottolinea Merleau-Ponty, non è un accessorio della coscienza, ma la sua condizione di possibilità. È dunque grazie al corpo che lo spazio si organizza in termini di prossimità e distanza, di possibilità d'azione e di resistenza, costituendo un orizzonte dinamico in cui il soggetto si muove. Come afferma il filosofo, «il corpo è il veicolo dell'essere-nel-mondo, e avere un corpo significa, per un essere vivente,

essere intrecciato in un ambiente determinato, identificarsi con certi progetti ed essere continuamente impegnato in essi». L'esperienza percettiva diventa sempre esperienza situata, legata a un corpo che non osserva dall'esterno ma partecipa dall'interno, radicando la coscienza nella dimensione concreta della vita vissuta: il corpo si rivela insieme come soggetto e come oggetto, come ciò che sente e ciò che è sentito, mostrando la sua natura ambigua e costitutiva dell'esperienza percettiva.

Queste intuizioni fenomenologiche hanno trovato nuova vitalità nella ricerca contemporanea, in particolare nelle scienze cognitive. In *The Embodied Mind* (1991) [38], viene proposta la teoria dell' **embodied cognition** che supera il paradigma computazionale classico: la mente non è un processore che manipola simboli indipendentemente dal contesto, ma un sistema dinamico che emerge dall'interazione tra organismo e ambiente. La cognizione è quindi **enattiva**: conoscere significa agire, esplorare, manipolare. In questa prospettiva, la percezione non è una fotografia passiva della realtà, ma un'azione che costruisce il mondo stesso dell'esperienza. Mente, corpo e ambiente costituiscono così un sistema interdipendente e inseparabile.

Questa visione si contrappone alla tradizionale separazione tra interno ed esterno, sottolineando come la coscienza sia sempre **situata**: l'individuo non elabora il mondo a partire da rappresentazioni astratte, ma attraverso la continua modulazione dei suoi schemi corporei. I movimenti, i gesti, le posture diventano strumenti cognitivi tanto quanto le operazioni logiche o linguistiche. La conoscenza stessa è radicata nella capacità del corpo di orientarsi, di anticipare, di reagire: la mente, in altre parole, è nel corpo e con il corpo.

Con l'avvento delle tecnologie digitali e delle esperienze immersive, questa prospettiva incarnata si è arricchita di nuove implicazioni. Mark Hansen, in *Bodies in Code: Interfaces with Digital Media* (2006) [39], sottolinea che la centralità del corpo non viene meno nell'era tecnologica, ma anzi si rafforza. Egli parla di *body-in-code* per indicare come la corporeità non venga sostituita dai media digitali, bensì riconfigurata attraverso di essi. Secondo Hansen, «*all reality is mixed reality*»: la tecnologia, dunque, non cancella l'esperienza incarnata, ma ne ridefinisce i confini e le modalità. La realtà virtuale, in questo senso, non rappresenta una negazione del corpo, bensì un laboratorio che rende visibile la sua funzione mediatrice: anche in un ambiente interamente simulato, infatti, l'utente continua a percepire, muoversi e orientarsi attraverso schemi corporei radicati nella sua esperienza sensoriale.

La realtà virtuale mostra, dunque, come il corpo sia la **vera interfaccia** dove ogni gesto diventa un atto di comunicazione con l'ambiente digitale. Anche quando il corpo fisico viene ridotto a un avatar schematizzato o a semplici tracciati luminosi, l'esperienza rimane incarnata: i feedback visivi, sonori e tattili sono recepiti e rielaborati attraverso lo schema corporeo dell'utente. Ciò produce un'**illusione di continuità** che rafforza il senso di presenza. In altre parole, il corpo non si dissolve nell'ambiente virtuale ma lo costituisce, diventando la sua matrice percettiva e cognitiva.

Nel progetto SHARED VR, questa dinamica trova una traduzione esemplare: la presenza dell'utente è resa attraverso elementi essenziali, come le mani luminose che ne rappresentano l'identità nello spazio condiviso. Il corpo non è un semplice input per la macchina, ma il luogo in cui si incontrano percezione, relazione e memoria collettiva.

In questa prospettiva, il corpo si conferma come **medium essenziale** dell'esperienza, condizione primaria della percezione per la fenomenologia, fondamento della cognizione situata nelle scienze cognitive e interfaccia attiva nelle teorie dei media digitali. Nella VR è proprio il corpo, nella sua duplice dimensione fisica e virtuale, a garantire la possibilità di abitare uno spazio simulato e di viverlo come reale. Nel caso di SHARED VR, il corpo diventa anche il vettore della partecipazione collettiva, poiché ogni gesto individuale si iscrive in un ambiente condiviso, contribuendo a costruire la memoria sonora e visiva dell'opera.

4.2 Corpo, ambiente e interfacce digitali

Il rapporto tra il corpo e l'ambiente si configura come una relazione dinamica, in cui le possibilità d'azione offerte dallo spazio diventano parte integrante dell'esperienza. Come accennato nel capitolo precedente, James J. Gibson, in *The Ecological Approach to Visual Perception* (1979) [31], introduce il concetto di **affordance** per descrivere questo fenomeno. Le affordance non sono proprietà intrinseche degli oggetti né proiezioni soggettive, ma possibilità di azione che emergono dall'incontro tra le caratteristiche del corpo e quelle dell'ambiente. Gibson definisce il suo approccio ecologico perché la percezione non è un processo interno e isolato ma il risultato dell'interazione costante tra organismo e ambiente: l'esperienza percettiva si radica in un ecosistema di relazioni in cui il corpo coglie direttamente le opportunità d'azione offerte dal mondo circostante.

Questa visione ecologica mette in luce come la percezione non sia mai neutra, ma orientata all'azione. Gli spazi non si presentano come scenari astratti, ma come campi di possibilità, pronti a essere attivati dal corpo che li abita. Nell'ambito digitale, ciò significa che un ambiente virtuale non è soltanto una rappresentazione grafica, ma un contesto che suggerisce **gesti, movimenti e interazioni**. L'efficacia di una simulazione immersiva, quindi, dipende dalla sua capacità di evocare affordance riconoscibili, in continuità con le aspettative incarnate dell'utente.

Paul Dourish, in *Where the Action Is: The Foundations of Embodied Interaction* (2001) [40], ha elaborato questa intuizione sviluppando la nozione di *embodied interaction*. Secondo Dourish, l'interazione con la tecnologia non è mai un processo puramente simbolico o astratto, ma un atto situato che coinvolge il corpo. Le interfacce digitali non sono meri strumenti di input, ma ambienti in cui le azioni corporee acquistano significato. La comprensione passa attraverso l'atto stesso del fare, trasformando il gesto in una forma di pensiero incarnato. In questa prospettiva, la tecnologia non media soltanto l'accesso a un contenuto, ma diventa lo spazio in cui l'utente negozia la propria presenza.

Diversi studi sul design dell'interazione hanno sottolineato l'importanza del gesto come linguaggio naturale. Loke e Robertson (2013) [41] mostrano come il movimento corporeo non sia riducibile a un semplice comando impartito a un dispositivo, ma costituisca un atto creativo ed espressivo. In questa prospettiva, il gesto diventa un modo per **attribuire senso** all'interazione, generando significati che non sono predeterminati dal sistema ma emergono dall'esperienza situata del corpo nello spazio. Il movimento non serve solo a ottenere una risposta funzionale dal sistema, ma diventa materiale di progettazione e medium estetico: un gesto può aprire possibilità interpretative diverse, attivare emozioni, costruire un senso di partecipazione.

La centralità del corpo nell'interazione digitale mostra come l'efficacia di un ambiente immersivo non dipenda soltanto dalla qualità grafica o dal realismo visivo, ma dalla sua capacità di risuonare con gli schemi corporei dell'utente. Una superficie che invita a essere toccata, un oggetto che sembra rispondere a un gesto di avvicinamento, uno spazio che si apre al movimento: sono tutti esempi di affordance che trasformano la tecnologia in luogo di esperienza incarnata.

In questa prospettiva, progetti immersivi contemporanei dimostrano come la scelta di interfacce corporee non sia un dettaglio tecnico, ma una precisa opzione culturale ed estetica. Quando un'opera si affida ai gesti, non solo rende l'interazione più accessibile, ma valorizza la dimensione empatica del corpo come veicolo di relazione. Il gesto diventa così linguaggio universale, capace di superare le barriere cognitive e linguistiche, e di trasformare l'esperienza virtuale in un terreno di condivisione. È proprio in questa prospettiva che si comprendono le scelte progettuali di esperienze come SHARED VR, dove l'ambiente non viene "navigato" attraverso comandi artificiali, ma abitato tramite gesti semplici e immediati, che fanno del corpo il vero interprete e mediatore dell'opera.

4.3 Presenza e senso di "essere nel luogo"

Uno degli aspetti più discussi e centrali nella riflessione sulla realtà virtuale è il concetto di presenza, inteso come il senso soggettivo di trovarsi realmente all'interno di un ambiente che, in realtà, è simulato. Steuer, nel saggio *Defining Virtual Reality. Dimensions Determining Telepresence* (1992) [42], definisce la telepresenza come «*l'esperienza di essere presenti in un ambiente mediato piuttosto che nel luogo fisico in cui ci si trova*». La VR, dunque, non si limita a proporre immagini o suoni, ma crea le condizioni perché l'utente sospenda la consapevolezza della mediazione tecnologica e percepisca lo spazio virtuale come se fosse reale.

Lombard e Ditton (1997) [36] hanno ulteriormente chiarito questa dinamica, definendo la presenza come «*l'illusione per cui un'esperienza mediata non appare mediata*». In altre parole, l'utente non vive l'interazione con un medium, ma direttamente con l'ambiente rappresentato. Più l'interfaccia tecnologica riesce a rendersi trasparente, più l'individuo dimentica il mezzo e si concentra sull'esperienza. Questa **illusione di non mediazione** è una condizione psicologica che consente al soggetto di accogliere l'ambiente simulato come luogo abitabile.

Slater (2018) [43] ha contribuito a consolidare queste riflessioni, distinguendo tra immersione e presenza. L'**immersione** è la dimensione tecnica, determinata dal grado in cui un sistema VR riesce a coinvolgere i sensi dell'utente attraverso display stereoscopici, audio spazializzato, interfacce interattive. La **presenza**, invece, è la risposta psicologica soggettiva: il sentirsi "nel luogo". Slater individua due componenti complementari necessarie a rendere un'esperienza VR efficace: il **place Illusion**, che descrive la sensazione di essere fisicamente collocati nello spazio virtuale, e la **plausibility Illusion**, che riguarda la credibilità degli eventi che vi accadono.

La presenza non è quindi un dato oggettivo, ma un processo che nasce dall'incontro tra caratteristiche tecniche del sistema e predisposizioni cognitive dell'utente. L'attenzione selettiva, la capacità di sospendere l'incredulità, la disponibilità a lasciarsi assorbire dall'esperienza sono tutti fattori che concorrono a costruire il senso di "essere nel luogo". La VR riesce a favorire questa condizione proprio perché richiama la corporeità dell'utente: muoversi, guardarsi intorno, compiere gesti nello spazio rafforza la percezione di trovarsi realmente immersi.

Se la presenza individuale è un elemento cruciale, nelle esperienze collettive acquista rilievo anche la co-presenza, ossia la percezione di condividere lo stesso spazio con altri utenti. L'incontro con presenze virtuali degli altri utenti non ha solo una funzione cognitiva, ma rinforza la credibilità dell'ambiente come luogo sociale. La co-presenza moltiplica l'illusione di essere "dentro" lo spazio simulato, trasformando l'esperienza in un evento relazionale.

In questo senso, progetti immersivi come **SHARED VR** dimostrano come il senso di presenza non sia soltanto un fatto individuale, ma una costruzione condivisa. La percezione di agire insieme ad altri, di vedere le conseguenze dei gesti altrui nello spazio virtuale, di interagire con elementi comuni, rafforza la sensazione di trovarsi in un ambiente vivo e collettivo. La VR non si limita quindi a offrire un'illusione di luogo, ma diventa un terreno di incontro, in cui la presenza si trasforma in co-presenza e l'esperienza individuale si intreccia con quella degli altri partecipanti.

In sintesi, i concetti di **embodiment**, **affordance** e **presenza** mostrano come la realtà virtuale non sia soltanto un fenomeno tecnologico, ma un'esperienza percettiva e relazionale radicata nel corpo. Nel caso di SHARED VR, il valore dell'opera emerge proprio dalla capacità di trasformare principi teorici di corpo come medium, gesto come interfaccia, presenza come senso di "essere nel luogo", in scelte progettuali concrete. L'uso di mani luminose, di gesti semplici e di uno spazio partecipativo non rappresenta solo una soluzione tecnica, ma l'applicazione coerente di un impianto teorico che valorizza la centralità dell'esperienza incarnata.

Questa riflessione introduce direttamente la seconda parte della tesi, dedicata allo sviluppo tecnico del prototipo multiutente. I concetti di embodiment, affordance e presenza diventano criteri operativi per l'architettura del sistema: guidano le scelte di implementazione dell'hand tracking come interfaccia naturale, della struttura software e hardware che sostiene l'esperienza immersiva e del networking necessario a rendere possibile la co-presenza degli utenti.

PARTE 2

Questa sezione è dedicata alla descrizione delle scelte progettuali e delle soluzioni tecniche adottate per la realizzazione del prototipo di **SHARED VR**. Dopo aver delineato nella prima parte i fondamenti teorici, concettuali ed estetici del progetto, ora si analizzano gli aspetti ingegneristici e implementativi che hanno reso possibile tradurre i principi dell'opera in un ambiente immersivo multiutente.

Il percorso di sviluppo non è stato lineare: nel corso del lavoro, infatti, il progetto è stato progressivamente rimodellato per conciliare la visione artistica del Maestro Michelangelo Lupone con le possibilità e i vincoli offerti dal mezzo della realtà virtuale. Nella fase iniziale erano stati ipotizzati **tre livelli di movimento**: il riconoscimento di pose statiche delle mani, il tracciamento dei movimenti nello spazio sui tre assi e il riconoscimento di path gestuali tramite modelli di *machine learning*. Durante la prototipazione, i primi due livelli sono stati mantenuti e costituiscono ancora oggi lo scheletro dell'interazione, mentre il terzo è stato abbandonato per scelta progettuale: il riconoscimento di traiettorie specifiche avrebbe infatti implicato che l'utente fosse preventivamente istruito su gesti codificati, limitando così la spontaneità. L'obiettivo perseguito è stato invece quello di privilegiare un'esperienza più libera ed esplorativa, in cui le interazioni risultassero il più naturali possibile e coerenti con i principi di embodiment e partecipazione condivisa. Analogamente, la componente di **rete**, concepita per supportare la co-presenza di più utenti tramite un'architettura client-server, è attualmente ancora in fase di sviluppo. Sono stati tuttavia realizzati test preliminari che hanno consentito di definire l'approccio progettuale, i protocolli di sincronizzazione e le strategie per minimizzare latenza e conflitti. Questi risultati, che verranno illustrati nei capitoli successivi, rappresentano un tassello fondamentale in vista delle implementazioni future.

Pertanto, il prototipo descritto in questa tesi non deve essere inteso come un prodotto concluso, ma come una **tappa intermedia** di un processo di ricerca più ampio, in cui sperimentazione artistica e sperimentazione tecnica procedono di pari passo. Le semplificazioni adottate, dettate da vincoli di tempo ma anche dalla necessità di garantire stabilità e chiarezza all'esperienza, hanno permesso di mettere a fuoco le potenzialità specifiche della VR come strumento di **interazione partecipativa** e di **costruzione di una memoria condivisa**.

I capitoli che seguono entreranno nel dettaglio di questo percorso, descrivendo prima l'architettura del sistema e le sue componenti tecnologiche, a seguire lo sviluppo dell'hand tracking e la progettazione della rete multiutente, per arrivare infine alla fase di testing e validazione, che includerà le osservazioni raccolte durante la presentazione pubblica a Tashkent.

5 Architettura del sistema SHARED VR

L'obiettivo di questo capitolo è descrivere i requisiti funzionali e l'architettura software/hardware sviluppata per il prototipo multiutente di **SHARED VR**. La progettazione tecnica si è posta come scopo primario quello di tradurre i principi fondativi dell'opera in un sistema stabile, modulare e scalabile, adatto a essere presentato in contesti espositivi, rimanendo coerente con la visione artistica del progetto.

Se nelle fasi concettuali il progetto si caratterizza per l'attenzione alla partecipazione condivisa e alla memoria sonora collettiva, a livello implementativo esso si traduce nella necessità di garantire una sincronizzazione precisa fra più utenti, un'interazione gestuale priva di mediazioni strumentali e una pipeline audiovisiva coerente. Il capitolo che segue illustra i requisiti identificati, le soluzioni tecnologiche adottate e le motivazioni che ne hanno guidato la scelta, con l'obiettivo di mostrare come l'infrastruttura tecnica si ponga come mediatore indispensabile fra visione artistica e vincoli ingegneristici.

5.1 Requisiti del sistema e scelte tecnologiche

La definizione dei requisiti ha rappresentato il primo passo verso la progettazione del sistema. L'applicazione è stata concepita per ospitare fino a tre utenti connessi simultaneamente, in grado di interagire in tempo reale nello stesso ambiente virtuale. La dimensione multiutente pone tuttavia vincoli stringenti, infatti, per risultare credibile, la co-presenza richiede una sincronizzazione stabile degli stati dei client e latenze al di sotto della soglia percettiva, così che ogni gesto compiuto da un partecipante si rifletta in modo pressoché istantaneo anche sugli altri.

Un ulteriore requisito ha riguardato le modalità di interazione: l'esperienza è stata progettata per basarsi esclusivamente sui gesti delle mani, senza ricorrere a controller fisici. Questa scelta, che rende l'opera accessibile anche a utenti non esperti, assume un valore ancora più importante: fondare l'interazione sulla corporeità significa rendere la mano stessa il medium primario dell'azione, eliminando ogni strumento esterno che possa introdurre mediazioni artificiali. Dal punto di vista tecnico, questo ha comportato l'impiego del sistema di hand tracking integrato nei visori **Meta Quest 3**, che consente di riconoscere in tempo reale la posizione delle mani nello spazio e i movimenti delle dita.

Un aspetto altrettanto cruciale riguarda la componente sonora, che in SHARED non costituisce un semplice accompagnamento ma un vero e proprio archivio vivo dell'interazione collettiva. Il suono deve stratificarsi nel tempo, reagire dinamicamente ai gesti e restituire agli utenti un feedback in grado di testimoniare la partecipazione di ciascuno. Per ottenere questo risultato, si è resa necessaria l'integrazione di **Unity 6** con un motore esterno dedicato, capace di gestire la generazione e la spazializzazione audio in tempo reale ovvero **Max/MSP**, collegato a Unity tramite il protocollo **OSC (Open Sound Control)**, che garantisce uno scambio rapido e leggero di

parametri fra le due piattaforme. L'utilizzo del plugin SPAT ha reso possibile una spazializzazione tridimensionale accurata, elemento decisivo per collocare gli eventi sonori nello spazio virtuale e rafforzarne la credibilità immersiva.

In una fase preliminare era stata valutata anche l'ipotesi di sviluppare una versione standalone per Meta Quest 3, che avrebbe permesso di eliminare l'uso di computer esterni. Questa soluzione si sarebbe basata su RNBO, un compilatore di Cycling '74 progettato per esportare patch di Max su piattaforme embedded. Tuttavia, l'impossibilità di utilizzare alcune funzioni avanzate di Max, in particolare quelle relative alla spazializzazione, ha reso questa soluzione poco adeguata rispetto agli obiettivi del progetto.

Sulla base di questi requisiti, dunque, le tecnologie selezionate sono state: **Unity 6** come motore di sviluppo, scelto per la compatibilità con **OpenXR** e per la solidità del suo ecosistema; **XR Interaction Toolkit** combinato con **XR Hands** per la gestione dell'interazione gestuale; **Netcode for GameObjects** come infrastruttura di rete, adatta a supportare prototipi multiutente in configurazioni LAN; e l'**Universal Render Pipeline (URP)** per ottimizzare il bilanciamento tra qualità grafica e prestazioni.

5.2 Architettura software e hardware

L'architettura complessiva del sistema SHARED VR segue una logica **client-server** in configurazione **PC VR**, soluzione che si è rivelata la più adeguata a garantire stabilità, coerenza e fedeltà alla complessità sonora del progetto. Il sistema prevede tre visori **Meta Quest 3**, ognuno dei quali è collegato tramite **AirLink** a un PC Windows dedicato, che esegue un'istanza dell'applicazione Unity. I tre computer non operano in maniera indipendente, ma sono interconnessi tramite un router centrale che gestisce la rete locale (LAN) e coordina la comunicazione tra i vari computer. Questa infrastruttura, è stata progettata per assicurare basse latenze e per ridurre al minimo i rischi di instabilità o interferenze esterne, condizioni imprescindibili in un contesto installativo dove l'affidabilità del sistema deve prevalere sulla scalabilità o sulla connettività remota.

All'interno della rete, uno dei tre computer assume il ruolo di server autoritativo, incaricato di mantenere lo stato condiviso dell'ambiente virtuale e di sincronizzare le azioni degli altri due client. Questa impostazione risponde alla necessità di evitare incoerenze o conflitti nelle interazioni degli utenti, infatti un'architettura peer-to-peer, alternativa a quella adottata, avrebbe comportato una maggiore complessità nella gestione della coerenza fra le azioni degli utenti, con la possibilità di generare differenze nella percezione condivisa dell'ambiente virtuale. L'approccio client-server centralizzato garantisce invece che lo stato dell'oracolo e degli oggetti condivisi sia unico e costantemente aggiornato, permettendo a tutti i partecipanti di percepire trasformazioni coerenti e di vivere l'esperienza come realmente condivisa.

Il funzionamento del sistema può essere descritto come un ciclo continuo di **input**, **elaborazione** e **feedback audiovisivo**. I sensori integrati nei visori rilevano i movimenti delle mani degli utenti, che vengono interpretati tramite il pacchetto **XR Hands**, utilizzato per il riconoscimento e la classificazione delle pose fondamentali, come la mano aperta o il pugno chiuso. In parallelo, vengono letti i parametri più dettagliati relativi alla geometria della mano, come posizione e rotazione nello spazio, chiusura delle dita e coordinate dei singoli joint. Unity riceve e integra questi dati, utilizzandoli sia per aggiornare la scena visiva sia per alimentare la componente sonora attraverso un flusso di dati inviato a Max/MSP tramite protocollo OSC.

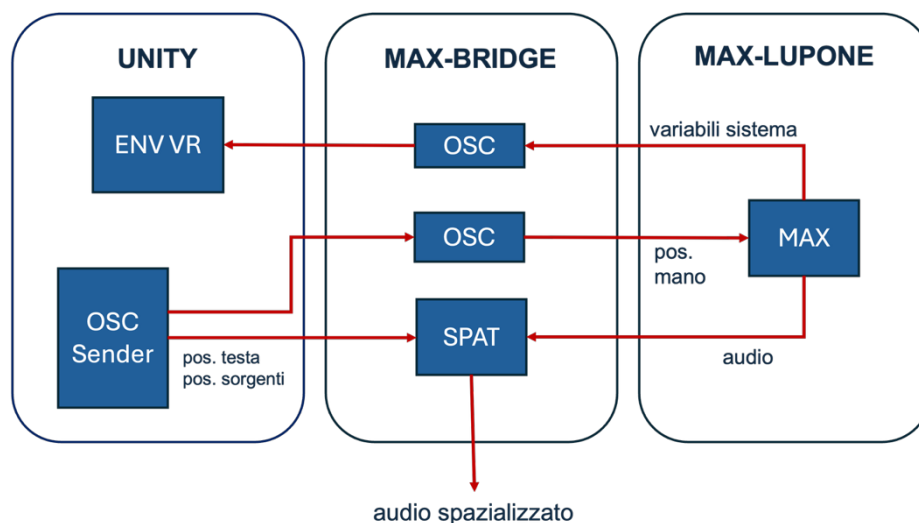


Figura 5.2.1: Schema funzionamento comunicazione OSC Unity-Max [F5.2.1]

Il **flusso dati** verso Max/MSP, illustrato in **figura 5.2.1**, non avviene in modo diretto ma è mediato da una patch dedicata che funge da ponte tra Unity e la patch principale di Michelangelo Lupone. I parametri dell'interazione come le coordinate spaziali della mano normalizzati tra 0 e 1, la posa, la mano che interagisce vengono raccolti su Unity inviati tramite protocollo OSC alla patch intermedia, che li riceve e li inoltra alla patch sviluppata da Michelangelo Lupone per la generazione dei suoni forma. Qui essi diventano materia sonora, modulando in tempo reale, attraverso l'uso di granulatori, la generazione dei cosiddetti **suoni forma**, che restituiscono acusticamente la traccia del gesto. Il processo è completato da un flusso di ritorno: la patch generativa produce l'output sonoro, che viene reindirizzato alla patch-ponte e quindi a Unity. Grazie all'integrazione con il plugin IRCAM **SPAT**, i suoni vengono poi collocati nello spazio tridimensionale in maniera precisa e immersiva, così che ciascun contributo sonoro risulti legato non solo al gesto che lo ha originato ma anche alla posizione dell'utente che lo ha compiuto. Il risultato è un ecosistema audiovisivo in cui la mano diventa il fulcro dell'interazione: ogni movimento produce effetti simultaneamente visivi e sonori, rafforzando la sensazione di **agency** e consolidando la credibilità della presenza condivisa.

Sebbene la versione standalone avrebbe semplificato l'allestimento, l'impossibilità di sfruttare pienamente le funzionalità avanzate di Max, in particolare la spazializzazione tridimensionale, avrebbe compromesso uno degli elementi più distintivi dell'opera. La decisione di adottare una configurazione PC VR con rete LAN locale, pur richiedendo un'infrastruttura hardware più articolata, ha permesso di salvaguardare la complessità artistica del progetto e di offrire un'esperienza immersiva che mantiene coerenza con la poetica originaria di SHARED.

Nei capitoli successivi, verranno analizzate la gestione dell'hand tracking e il sistema di networking, approfondendo come le librerie adottate siano state impiegate sia nelle fasi di sperimentazione preliminare sia nell'attuale prototipo. Questi due macroelementi rappresentano i cardini della pipeline interattiva del progetto, traducendo i principi concettuali di SHARED VR in pratiche concrete di sviluppo e determinando la qualità dell'esperienza immersiva.

6 Implementazione Hand Tracking

L'implementazione dell'hand tracking in SHARED VR si fonda sull'integrazione di tre strumenti messi a disposizione Unity: **OpenXR**, **XR Hands** e **XR Interaction Toolkit**. Ciascuno di questi componenti è necessario nel tradurre i movimenti dell'utente in dati interpretabili dal sistema virtuale.

OpenXR consente al motore Unity di comunicare direttamente con i sensori del Meta Quest 3, ricevendo in tempo reale i dati relativi alla posizione e all'orientamento della testa e delle mani.

XR Hands elabora queste informazioni e ricostruisce la mano dell'utente come una struttura scheletrica digitale, aggiornata costantemente in posizione e rotazione, consentendo di riconoscere configurazioni specifiche, come la mano aperta o il pugno chiuso.

Infine, **XR Interaction Toolkit** gestisce la struttura spaziale della scena XR e garantisce che i movimenti rilevati vengano rappresentati in modo coerente nell'ambiente virtuale.

La combinazione di questi tre strumenti ha reso possibile sviluppare un sistema di interazione privo di controller, in cui i gesti delle mani assumono il ruolo di unico canale di input. Su questa base sono stati realizzati diversi livelli di interazione: dal riconoscimento di pose statiche, all'utilizzo dei movimenti nello spazio e del grado di apertura delle dita per modulare parametri visivi e sonori, fino a sperimentazioni orientate al riconoscimento di traiettorie tridimensionali tramite tecniche di *machine learning*.

Il capitolo documenta l'evoluzione del sistema di hand tracking, a partire dalle funzionalità fornite da XR Hands e dalla loro integrazione con l'XR Interaction Toolkit. Verranno quindi presentate le diverse soluzioni sviluppate in Unity: dalle prime sperimentazioni basate su *prefab* geometrici e successivamente su sistemi particellari con VFX Graph, passando per i test sul riconoscimento di traiettorie tridimensionali tramite modelli ML, fino ad arrivare alla configurazione attuale che impiega i suoni forma e la cintura interattiva come elementi centrali dell'esperienza.

6.1 Librerie

Il sistema di tracciamento delle mani in **SHARED VR** si basa su una pipeline composta dalle librerie **OpenXR**, **XR Hands** e **XR Interaction Toolkit**, che collaborano per gestire il flusso dei dati e mantenere la coerenza tra il movimento reale dell'utente e la sua rappresentazione virtuale.

Ciascun livello della pipeline ha una funzione specifica: **OpenXR** raccoglie i dati dal visore, **XR Hands** li elabora e li interpreta, e lo **XR Interaction Toolkit** li integra nella logica della scena.

OpenXR rappresenta lo strato più basso di questa architettura, in quanto fornisce il punto d'ingresso ai dati provenienti dal visore Meta Quest 3. Attraverso questa tecnologia, Unity riceve in tempo reale le informazioni provenienti dai sensori del Meta Quest 3, che rilevano posizione, orientamento e movimento della testa e delle mani [44]. OpenXR funziona quindi da ponte tra hardware e software, raccogliendo i dati del tracciamento e rendendoli disponibili in un formato

stabile e coerente. Questo passaggio è essenziale per garantire che il sistema riceva continuamente informazioni aggiornate e precise, condizione necessaria per ottenere una risposta fluida e sincronizzata tra gesto reale e rappresentazione virtuale.

Su questo livello si innesta **XR Hands**, che rappresenta il cuore della logica di tracciamento. Introdotto da Unity per sfruttare le estensioni di OpenXR, questo pacchetto fornisce un accesso strutturato ai dati che descrivono la mano in tempo reale, permettendo di ricostruirne virtualmente la forma e il movimento [45]. A partire dalle informazioni fornite dal visore, XR Hands genera una struttura scheletrica digitale composta da articolazioni che corrispondono al polso, al palmo e alle dita. Ogni articolazione viene costantemente aggiornata con la propria posizione e rotazione, consentendo di seguire con precisione i movimenti della mano e di restituirne un comportamento coerente nello spazio virtuale.

Oltre a questa rappresentazione geometrica, XR Hands mette a disposizione strumenti per l'analisi della configurazione delle dita e il riconoscimento di pose specifiche. Attraverso il componente **Static Hand Gesture** [46], è possibile definire modelli di riferimento che descrivono pose particolari assunte dalla mano e stabilire le condizioni entro cui il gesto viene riconosciuto. Quando la configurazione reale della mano corrisponde a una di queste forme, il sistema genera un evento (**gesturePerformed** o **gestureEnded**) che può essere utilizzato per attivare azioni o modifiche all'interno della scena.

Nel contesto di **SHARED VR**, questa funzionalità è impiegata per distinguere in modo stabile e affidabile le due pose principali che regolano le interazioni con l'ambiente: la mano aperta, associata all'azione di rilascio o dispersione degli elementi sonori, e il pugno chiuso, che rappresenta l'interazione diretta con un suono forma.

Infine, **XR Interaction Toolkit** fornisce il framework di alto livello che gestisce la struttura della scena XR e garantisce la coerenza spaziale dell'esperienza [47]

Il suo componente principale, l'**XR Origin**, definisce la gerarchia degli oggetti XR e costituisce il punto di riferimento per la posizione della camera e delle mani nello spazio virtuale [48].

Questo sistema consente di mantenere un allineamento costante tra le coordinate fisiche e quelle digitali, adattando automaticamente parametri come l'altezza della camera rispetto al pavimento e l'orientamento dell'utente.

Nel progetto SHARED VR, il toolkit è servito come infrastruttura di supporto per integrare in modo coerente i moduli di hand tracking sviluppati, assicurando stabilità, modularità e una corretta sincronizzazione tra corpo, spazio e risposta dell'ambiente.

In sintesi, la pipeline segue un flusso continuo e coordinato: i sensori del visore rilevano i movimenti delle mani, OpenXR li raccoglie e li trasmette a Unity, XR Hands li elabora e li traduce in pose riconoscibili, e infine lo XR Interaction Toolkit li integra nella scena, generando la risposta interattiva.

Questo sistema modulare ha reso possibile un'esperienza senza controller, precisa e adattabile, in cui il gesto dell'utente diventa al tempo stesso strumento di interazione, atto performativo e parte integrante della composizione dell'opera.

6.2 Sperimentazione

La fase di sperimentazione ha avuto come obiettivo la costruzione progressiva dei blocchi funzionali che sarebbero poi confluiti nella versione definitiva del prototipo.

La scena XR è stata allestita in Unity a partire dal prefab **XR Origin (Hands)**, fornito dallo *XR Interaction Toolkit*, una variante dell'**XR Origin (XR Rig)** che, nella configurazione predefinita, include sia i controller sia le mani. Nel contesto di **SHARED VR**, i controller sono stati completamente rimossi, poiché l'interazione si basa esclusivamente sul tracciamento gestuale. L'**XR Origin (Hands)** è stato utilizzato come rig di riferimento per la camera e per le mani, assicurando l'allineamento costante tra i movimenti fisici dell'utente e la loro rappresentazione virtuale: infatti, questo componente consente di calibrare la posizione e le dimensioni del rig XR, mantenendo l'allineamento tra spazio fisico e virtuale e una percezione spaziale coerente.

OpenXR fornisce a Unity i dati dei joint della mano che vengono interpretati dal pacchetto **XR Hands**, che associa ciascuna articolazione a un identificatore univoco (`XRHandJointID`) e li espone tramite il sottosistema `XRHandSubsystem`, responsabile dell'aggiornamento frame-by-frame delle pose delle mani: in questo modo è possibile interrogare entrambe le mani e recuperare posizione, rotazione e stato di tracking dei joint principali.

Negli script sviluppati per il tracciamento, infatti, l'inizializzazione del subsystem e la lettura di joint fondamentali come il polso sono state essenziali; infatti, l'uso diretto dei joint come riferimenti spaziali, piuttosto che dell'intero oggetto mano, permette di ottenere una maggiore precisione e stabilità nel calcolo delle posizioni, evitando errori dovuti all'aggiornamento delle *mesh* o alla variabilità dell'orientamento complessivo. Ad esempio, il polso è stato utilizzato come riferimento per definire posizione e orientamento. In **HandMovementTracker.cs**, uno dei primi test di tracciamento, viene inizializzato il sottosistema di XR Hands all'interno della funzione **Start()**, in modo da poter accedere ai dati di entrambe le mani. *XRHandSubsystem*:

```
void Start()
{
    handSubsystem =
        XRGeneralSettings.Instance?.Manager?.activeLoader?.GetLoadedSubsystem<XRHandSubsystem>();
}
public void StartTracking()
{
    trackedHand = isLeftHand ? handSubsystem.leftHand : handSubsystem.rightHand;

    if (trackedHand.isTracked)
    {
        XRHandJoint wristJoint = trackedHand.GetJoint(XRHandJointID.Wrist);
        if (wristJoint.TryGetPose(out Pose wristPose))
        {
            startPosition = wristPose.position;
            isTracking = true;
        }
    }
}
```

In questo primo approccio al tracciamento, l'obiettivo era verificare la stabilità dei dati restituiti da *XRHandSubsystem* e sperimentare una prima forma di associazione diretta tra gesto e feedback visivo-sonoro. Una volta avviato il tracciamento, la posizione del polso veniva memorizzata come origine rispetto alla quale venivano calcolati gli spostamenti successivi della mano. I movimenti verticali e orizzontali erano quindi rimappati in un intervallo normalizzato compreso tra -1 e 1, e utilizzati per modificare la scala delle mesh associate alla posa della mano: lo spostamento verticale agiva sull'asse Y, mentre quello orizzontale agiva sull'asse X.

In questo modo ogni gesto produceva una variazione proporzionale nella dimensione dell'oggetto e nei parametri sonori collegati, costituendo il primo test di tracking-feedback per valutare la coerenza e la sensibilità del sistema.

Una volta avviato il tracciamento, nella funzione *Update* lo script calcolava lo spostamento rispetto alla posizione iniziale e lo rimappava su parametri di scala dell'oggetto e parametri audio:

```
XRHandJoint wristJoint = trackedHand.GetJoint(XRHandJointID.Wrist);
if (wristJoint.TryGetPose(out Pose wristPose)) trackedHand
{
    Vector3 delta = wristPose.position - startPosition;

    float targetScaleX = Remap(delta.x, minDelta, maxDelta, minScale, maxScale);
    float targetScaleY = Remap(delta.y, minDelta, maxDelta, minScale, maxScale);

    float targetPitch = Remap(delta.x, minDelta, maxDelta, minPitch, maxPitch);
    float targetVolume = Remap(delta.y, minDelta, maxDelta, minPitch, maxPitch);

    currentScaleX = Mathf.Lerp(currentScaleX, targetScaleX, Time.deltaTime * scaleSmoothSpeed);
    currentScaleY = Mathf.Lerp(currentScaleY, targetScaleY, Time.deltaTime * scaleSmoothSpeed);

    visualObject.localScale = new Vector3(currentScaleX, currentScaleY, currentScaleX);
    audioSource.pitch = Mathf.Lerp(currentPitch, targetPitch, Time.deltaTime * audioSmoothSpeed);
    audioSource.volume = Mathf.Lerp(currentVolume, targetVolume, Time.deltaTime * audioSmoothSpeed);
}
```

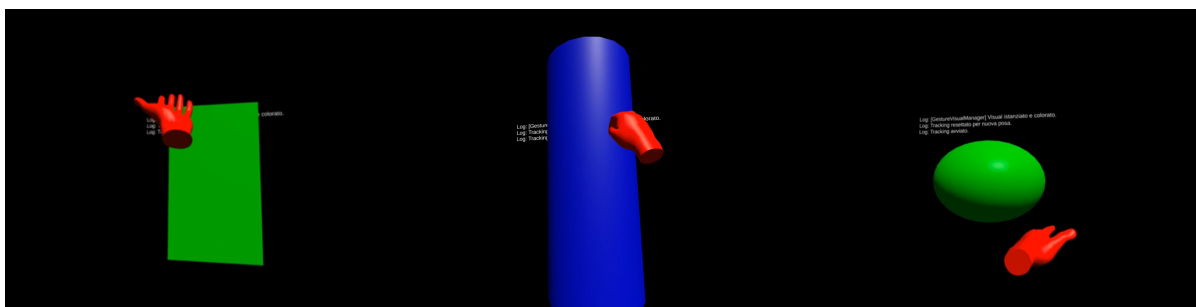


Figura 6.2.1: Screenshot scena *01HandTracking test*: tracking mano e riconoscimento pose[F6.2.1]

Parallelamente al tracciamento continuo, il secondo tassello della pipeline era il **riconoscimento di pose statiche**. Nel caso di questo primo prototipo, il tracciamento veniva avviato in corrispondenza del riconoscimento di una delle tre pose base tra Palm Down, Palm Up e Fist, che attivavano rispettivamente una sfera, un cubo e un cilindro come mesh di riferimento (**Figura 6.2.1**). Per l'implementazione di questa funzionalità, è stato utilizzato il componente *Static Hand*

Gesture del pacchetto XR Hands, introdotto da Unity per consentire il riconoscimento di configurazioni predefinite della mano.

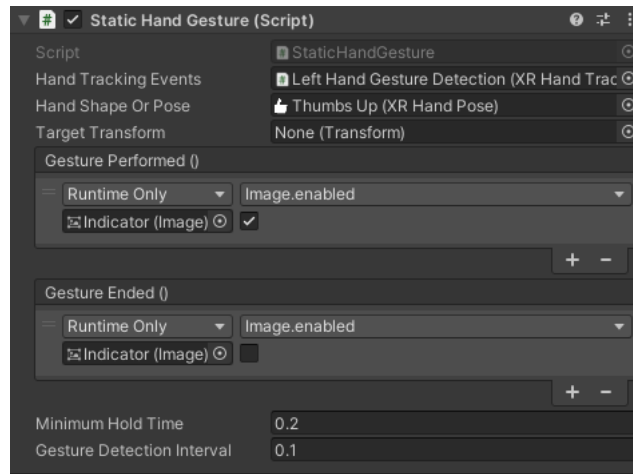


Figura 6.2.2: Componente Static Hand Gesture [F6.2.2]

Static Hand Gesture è uno script di XR Hands che riconosce la forma assunta da una mano, descritta attraverso due modalità diverse:

- Una **Hand Shape**, che definisce la configurazione relativa delle dita (livello di curl, apertura, spread, pinch) ed è indipendente dall'orientamento della mano;
- Una **Hand Pose**, che oltre alla configurazione delle dita tiene conto anche della posizione e della rotazione della mano nello spazio.

Ogni asset di tipo *Static Hand Gesture* può essere configurato dall'editor e include parametri come:

- la **soglia di confidenza** per il matching della gesture;
- il **tempo minimo di mantenimento** prima che l'evento venga considerato valido;
- l'**intervallo di rilevamento**, cioè la frequenza con cui il sistema confronta la mano reale con l'asset registrato.

Quando la configurazione della mano corrisponde ai valori salvati nell'asset, il componente emette un evento *gesturePerformed*, mentre al termine della corrispondenza emette un evento *gestureEnded*. Entrambi possono essere collegati a funzioni dello script tramite l'inspector o via codice.

Nel caso di SHARED VR sono state adottate entrambe le modalità: per *gesture* come **pugno chiuso** e **mano aperta** sono state utilizzate le Hand Shape, così da ignorare l'orientamento del palmo e rendere l'interazione più robusta; al contrario, nelle prime fasi di sperimentazione, per pose come **palmo in su** e **palmo in giù**, è stato necessario utilizzare Hand Pose, poiché l'orientamento nello spazio era parte integrante del gesto.

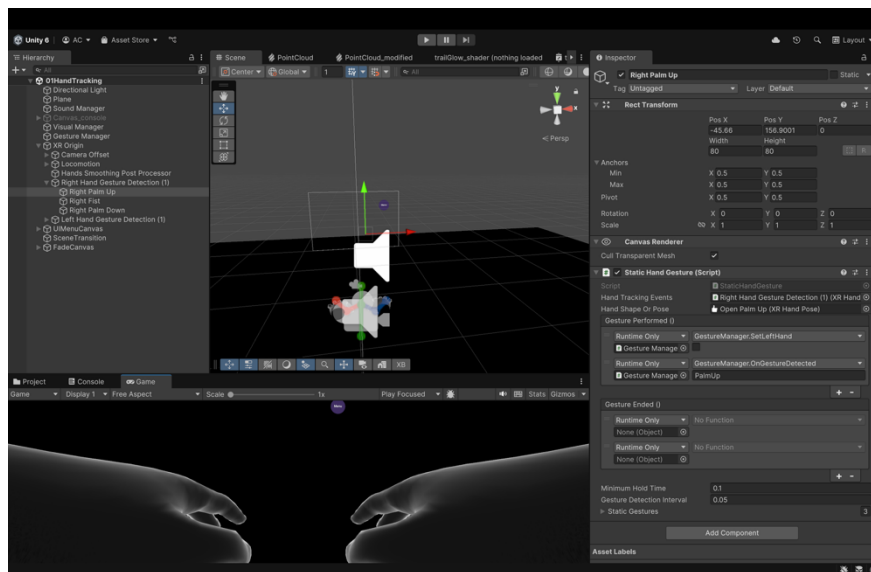


Figura 6.2.3: *01HandTracking* composizione scena Unity [F6.2.3]

Nella **figura 6.2.3** è mostrata una delle prime scene di test in Unity. Nella gerarchia della scena si distinguono i componenti principali:

- **XR Origin**, che include camera e tracking delle mani;
- i prefab **Right Hand Gesture Detection** e **Left Hand Gesture Detection**, ciascuno con gli asset per pugno, palmo su e palmo giù;
- i manager principali (**GestureManager**, **VisualManager**, **SoundManager**), che coordinavano la logica di feedback.

Nel pannello *Inspector* è visibile l'asset **Static Hand Gesture** configurato per la posa *Palm Up*: quando il sistema riconosce la mano destra con palmo verso l'alto, viene chiamata la funzione **OnGestureDetected** di **GestureManager.cs**, che a sua volta istanzia l'oggetto corrispondente e attiva il suono associato.

Gli script principali coinvolti in questa fase erano:

- **GestureManager**, responsabile del riconoscimento della gesture corrente e dell'invio degli eventi ai manager;
- **GestureVisualManager**, che istanziava e gestiva i prefab geometrici (cubo, sfera, cilindro);
- **GestureSoundManager**, che attivava i clip sonori e gestiva le transizioni con fade-in e fade-out;
- **HandMovementTracker**, che calcolava lo spostamento della mano rispetto a una posizione iniziale e lo rimappava in parametri di scala e suono.

In una fase successiva, la stessa logica è stata applicata a un sistema di **VFX Graph**. Invece di generare prefab statici, le gesture attivavano grafi particellari, e i movimenti della mano modulavano parametri come turbolenza, emissione o colore delle particelle. Dal punto di vista

del codice le differenze erano minime: gli stessi eventi *onGesturePerformed* venivano collegati a un componente *VisualEffect*, i cui parametri erano aggiornati runtime.

Il terzo livello di sperimentazione ha riguardato il **riconoscimento di traiettorie** tridimensionali compiute con la mano nello spazio. L'obiettivo era quello di distinguere tre gesti complessi: il simbolo dell'infinito, un movimento curvo concavo verso il basso e un movimento spigoloso con angolo verso il basso. A differenza delle pose statiche, questi gesti si sviluppano come sequenze temporali di punti 3D, ed è quindi necessario un approccio capace di catturare sia la dimensione spaziale che quella temporale.



Figura 6.2.4: Workflow path recognition [F6.2.4]

Per la raccolta dei dati è stata realizzata una scena dedicata, con lo script *XRHandTrajectoryRecorder.cs*. La registrazione partiva e terminava con il pinch, e durante l'intervallo venivano campionati i punti della traiettoria della punta dell'indice. Ogni gesto veniva normalizzato a 100 campioni uniformemente distribuiti nel tempo e salvato in formato JSON.

La scelta del modello è ricaduta su una **rete neurale convoluzionale (CNN)**. Le CNN, nate per il riconoscimento di immagini, sono particolarmente adatte a catturare pattern locali e invarianze in sequenze di dati, grazie all'uso di filtri che scorrono sugli input e individuano caratteristiche ricorrenti. In questo caso, la traiettoria viene rappresentata come una sequenza 1D di 100 punti, ciascuno con tre feature (coordinate X, Y, Z), dando origine a un tensore di input di dimensione (1, 100, 3). Il modello **GestureCNN**, sviluppato con **PyTorch**, elabora la sequenza dei punti della traiettoria attraverso tre strati di convoluzione 1D, che analizzano localmente le variazioni delle coordinate per estrarre i pattern caratteristici del gesto. Successivamente, un livello di **dropout** riduce il rischio di **overfitting** spegnendo casualmente alcune connessioni durante l'addestramento, mentre l'operazione di **pooling** (*AdaptiveAvgPool1d*) calcola la media delle attivazioni lungo tutta la sequenza, riassumendo il gesto in un unico vettore di caratteristiche. Infine, il livello **fully connected** trasforma questo vettore nei valori di output corrispondenti alle tre classi di gesto riconosciute dal sistema.

```
class GestureCNN(nn.Module):
    def __init__(self, num_classes):
        super(GestureCNN, self).__init__()
        self.conv1 = nn.Conv1d(3, 32, kernel_size=5, padding=2)
        self.conv2 = nn.Conv1d(32, 64, kernel_size=5, padding=2)
        self.conv3 = nn.Conv1d(64, 128, kernel_size=5, padding=2)
```

```

self.dropout = nn.Dropout(0.3)
self.pool = nn.AdaptiveAvgPool1d(output_size=1)
self.fc = nn.Linear(128, num_classes)

def forward(self, x):
    x = x.permute(0, 2, 1)
    x = F.relu(self.conv1(x))
    x = F.relu(self.conv2(x))
    x = F.relu(self.conv3(x))
    x = self.dropout(x)
    x = self.pool(x)
    x = x.squeeze(-1)
    x = self.fc(x)
    return x

```

Nella funzione **forward()**, la rete fa scorrere filtri lungo la sequenza dei punti per individuare pattern ricorrenti, attiva i risultati con funzioni **ReLU** e li sintetizza con un pooling adattivo. Il vettore finale, che rappresenta l'intero gesto, viene poi trasformato dal livello **fully connected** nei tre valori di output corrispondenti alle classi previste. Per migliorare la robustezza è stato applicato un processo di **data augmentation**: rotazioni casuali delle traiettorie, aggiunta di rumore gaussiano e variazioni di scala. Quest'ultimo aspetto è stato fondamentale, poiché una CNN è intrinsecamente robusta a traslazioni e, in parte, a rotazioni, ma non allo **scalamento**: senza augmentation, un gesto grande e uno piccolo venivano interpretati come diversi. Prima dell'inferenza, i dati raccolti in Unity venivano preprocessati nello script *PathRecognitionManager.cs* con una pipeline che comprendeva:

- **centrata** dei punti;
- **PCA** per ridurre la variabilità di orientamento, ordinando manualmente gli assi per autovalori decrescenti (per evitare inconsistenze tra Unity e Python);
- **normalizzazione max-norm** per riportare tutte le traiettorie alla stessa scala;
- **formattazione** in tensore $1 \times 100 \times 3$.

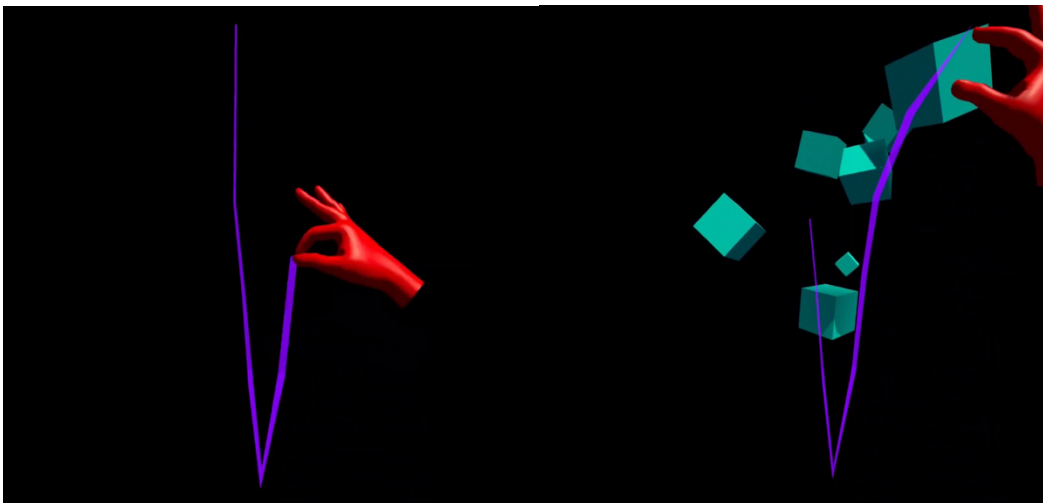


Figura 6.2.5: Riconoscimento path specifici con Unity Sentins e CNNN (movimento2) [F6.2.5]

Dopo l'addestramento, il modello PyTorch è stato esportato in formato **ONNX**, uno standard aperto che consente di trasferire modelli di machine learning in diversi ambienti di runtime. L'integrazione in Unity è stata possibile grazie al pacchetto **Sentis**, che permette di caricare e interrogare modelli ONNX direttamente all'interno del motore. Lo script *GestureClassifier.cs* creava un tensore a partire dalla traiettoria normalizzata, lo passava al modello e leggeva i logit di output per selezionare la classe più probabile:

```
public class GestureClassifier : MonoBehaviour
{
    public ModelAsset modelAsset;
    private Model model;
    private Worker worker;
    private readonly string[] classLabels = { "infinito", "movimento1", "movimento2" };

    void Start()
    {
        model = ModelLoader.Load(modelAsset);
        worker = new Worker(model, BackendType.GPUCompute);
    }

    public string ClassifyFromJsonData(GestureData jsonData)
    {
        var points = jsonData.positions.Select(p => new Vector3(p.posX, p.posY, p.posZ)).ToList();

        float[] inputData = new float[1 * 100 * 3];
        for (int i = 0; i < 100; i++)
        {
            inputData[i * 3 + 0] = points[i].x;
            inputData[i * 3 + 1] = points[i].y;
            inputData[i * 3 + 2] = points[i].z;
        }
        using var inputTensor = new Tensor<float>(new TensorShape(1, 100, 3), inputData);
        worker.Schedule(inputTensor);
        var outputTensor = worker.PeekOutput() as Tensor<float>;
        float[] logits = outputTensor.DownloadToArray();
        float[] probs = Softmax(logits);
        int bestIndex = System.Array.IndexOf(probs, probs.Max());
        string label = classLabels[bestIndex];
        float confidence = probs[bestIndex];
        return label;
    }
}
```

Il sistema era in grado di riconoscere correttamente i tre gesti, mostrando coerenza tra l'inferenza in Python e quella in Unity, verificata tramite script di confronto dei *logit*. Tuttavia, l'integrazione è rimasta confinata alla fase sperimentale poiché il riconoscimento di *path* avrebbe richiesto di **istruire l'utente** a riprodurre forme specifiche, introducendo una dimensione prescrittiva poco coerente con la filosofia del progetto SHARED, che privilegia la libertà esplorativa e la naturalezza del gesto. Nonostante ciò, la sperimentazione ha dimostrato la fattibilità tecnica e ha aperto prospettive interessanti per future evoluzioni, in cui modelli di machine learning potranno essere sfruttati per arricchire ulteriormente la varietà e la complessità delle interazioni.

Questa pipeline sperimentale, seppur semplice, ha permesso di validare i tre principi di base: l'affidabilità del riconoscimento di pose, la stabilità del tracciamento continuo della mano e la possibilità di estendere l'interazione a path più complessi. Le scene di test hanno rappresentato quindi il terreno su cui costruire gli sviluppi successivi, fino ad arrivare alla configurazione definitiva con SoundShape e cintura.

6.3 Sviluppo prototipo

La versione attuale del prototipo di **SHARED VR** integra i risultati delle sperimentazioni precedenti, realizzando un sistema stabile e utilizzabile in contesto espositivo. Questa fase del lavoro ha avuto come obiettivo principale l'unificazione dei diversi blocchi funzionali come tracciamento, interazione e feedback, in un unico flusso di esecuzione continuo, in cui ogni elemento risponde in tempo reale ai gesti dell'utente. L'architettura ottenuta è funzionalmente stabile, ma riflette ancora l'impostazione sperimentale delle fasi di sviluppo, caratterizzata da collegamenti diretti tra i moduli e da una gestione del flusso di dati non ancora pienamente ottimizzata. Questa configurazione costituisce la base su cui verrà eseguita la successiva riorganizzazione del sistema.

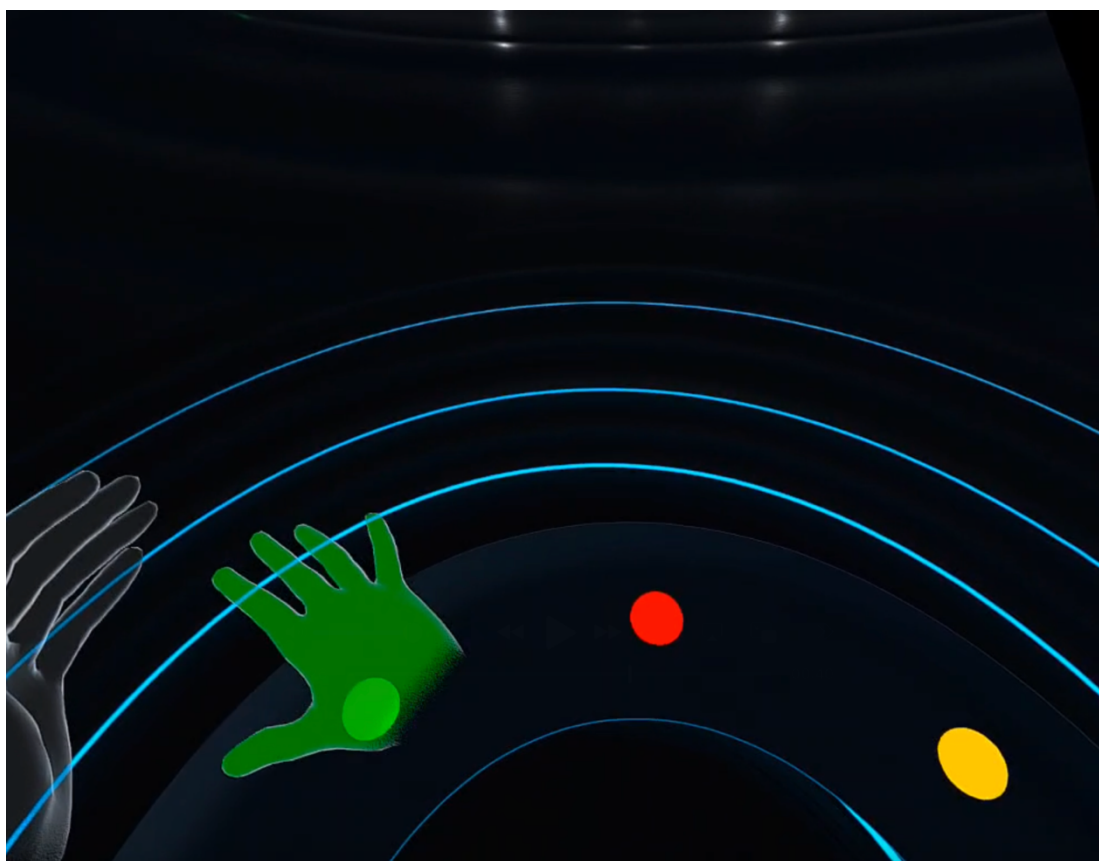


Figura 6.3.1: Screenshot scena *Uzbekistan-ArtIntegration*, cintura [F6.3.1]

Il sistema si articola attorno alla **cintura di soundshape**, una fascia virtuale posizionata davanti al bacino dell'utente che ospita tre sfere, ciascuna corrispondente a un diverso tipo di suono-forma. La cintura è gestita dal **BeltManager**, che ne controlla posizione, visibilità e stato generale. Ogni sfera possiede un proprio **SoundShapeTrigger**, responsabile di rilevare l'ingresso e l'uscita delle mani nel proprio collider e di comunicare con gli altri moduli. Il posizionamento della cintura è ancorato alla testa dell'utente tramite la XR Camera: a ogni frame, l'orientamento della testa viene proiettato sull'asse orizzontale e la cintura viene aggiornata in modo da restare sempre centrata davanti al corpo, indipendentemente da inclinazioni o movimenti verticali. Questo

accorgimento garantisce che le soundshape rimangano sempre accessibili e riduce la latenza percettiva tra movimento reale e aggiornamento visivo.

Ogni mano dell'utente è gestita da un'istanza separata dello script **HandTrackerSoundShape**. Questo componente costituisce il centro del sistema di interazione: raccoglie i dati di tracciamento forniti da **XR Hands**, calcola le coordinate normalizzate della mano all'interno del volume operativo e gestisce gli stati gestuali e visivi. I limiti del volume sono forniti dal **SoundModifierBoxController**, che definisce l'area di manipolazione, un box di riferimento posizionato davanti allo sguardo dell'utente. Ogni frame il tracker calcola la posizione del polso rispetto a questi limiti e la converte in valori compresi tra 0 e 1 per ciascun asse (X, Y, Z). Questi valori vengono poi inviati agli altri componenti per modulare feedback visivi e parametri sonori.

Oltre alle coordinate, il tracker calcola continuamente un valore di **openness**, ovvero la misura media di apertura delle dita, e determina la **posa corrente** della mano, ad esempio *openHand* o *fist*. Lo script mantiene in memoria anche la posa precedente, permettendo di riconoscere transizioni significative: un suono forma può essere afferrato solo se la mano entra nel collider con posa *openHand* e, una volta dentro, chiude in *fist*. Questa sequenza previene afferramenti accidentali e garantisce una dinamica naturale del gesto.

Il **SoundShapeTrigger**, assegnato a ogni sfera, si occupa di intercettare queste condizioni. Quando una mano entra nel suo collider, il trigger comunica con il relativo **HandTrackerSoundShape** per notificare lo stato di contatto e attivare il materiale di *collision*, che fornisce all'utente un segnale visivo di prossimità. Se, mentre la mano è all'interno, il tracker rileva la transizione da *openHand* a *fist*, il trigger attiva la procedura di presa: invia al **BeltManager** il comando di nascondere la cintura e istanzia un **SoundShapeController**, legato all'oggetto appena afferrato. In questa fase, il sistema assicura che **solo una mano alla volta possa trovarsi nello stato di manipolazione**, evitando conflitti di controllo.

Il **SoundShapeController** gestisce la fase di manipolazione vera e propria. In una prima versione, questo script si occupava anche di aggiornare i parametri del mixer audio in Unity; nella configurazione attuale, dove la gestione sonora è demandata a **Max/MSP** tramite OSC, la sua funzione principale è attivare ***persistent ghost sphere***: un oggetto sferico ancorato al palmo della mano che segue in tempo reale la sua posizione e simula la presenza fisica del suono-forma afferrato.

Durante la manipolazione, il **HandTrackerSoundShape** continua a fornire le coordinate normalizzate e il grado di apertura della mano. Questi valori sono letti dal **HandTrailManager**, che gestisce i trail assegnati a ciascuna mano. Il colore del trail varia in funzione della posizione nello spazio normalizzato lungo ogni asse, mentre lo spessore aumenta con l'apertura della mano, quindi, la scia è minima durante la presa e si allarga gradualmente con l'estensione delle dita, fornendo un riscontro visivo immediato sulla dinamica del gesto. Questo legame immediato tra

gesto e colore favorisce la comprensione intuitiva del mapping spaziale e rende visibili anche i movimenti più piccoli agli osservatori esterni.

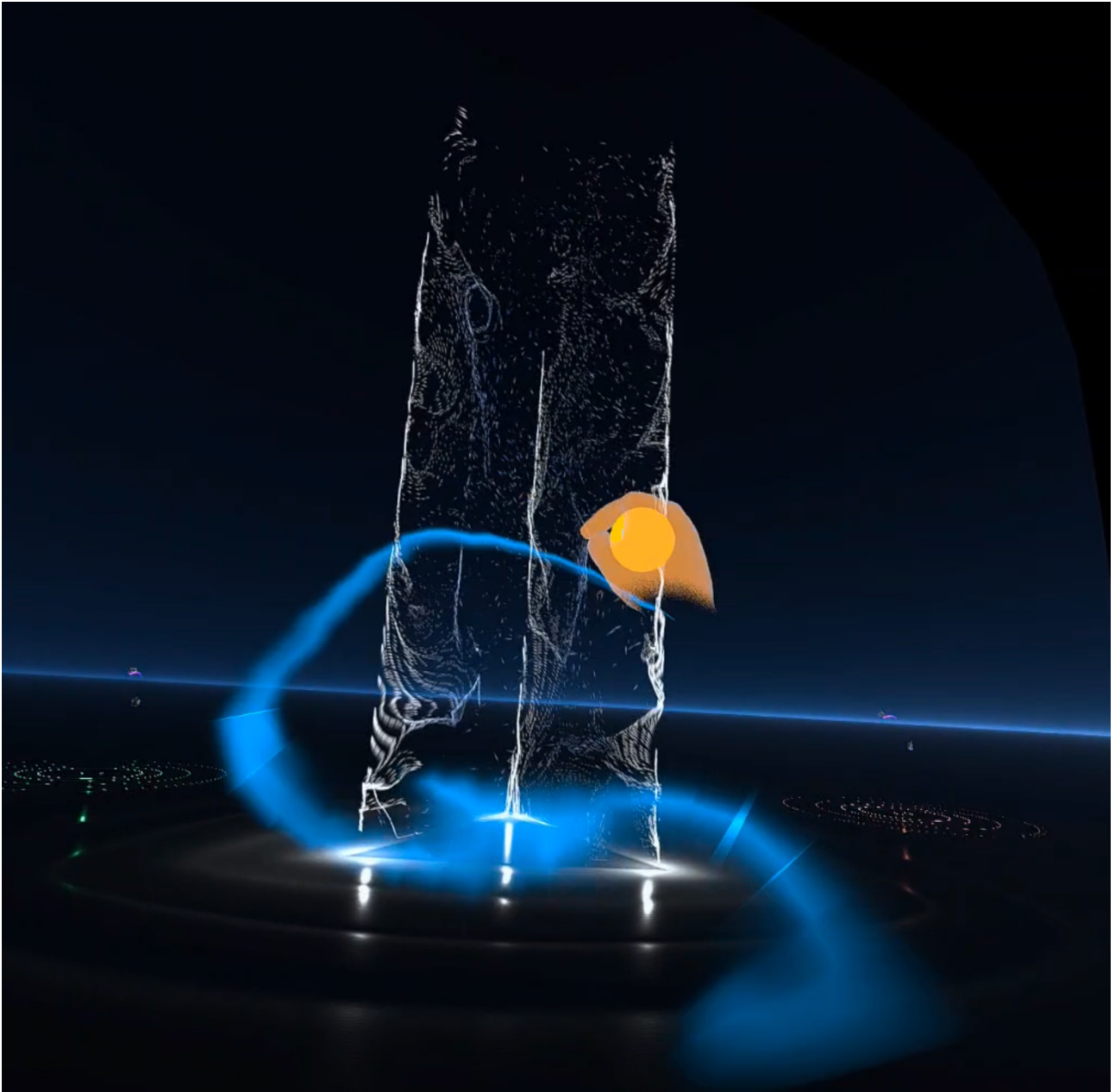


Figura 6.3.2: Screenshot scena *Uzbekistan-ArtIntegration*, interazione con il suono forma [F6.3.2]

Il ciclo di manipolazione si conclude quando la mano torna alla posa *openHand*. Il **SoundShapeController** avvia allora il **ReleaseSphereEffect**, che comunica con il **VFX Graph** per generare una scia di particelle luminose che confluisce all'interno dell'oracolo, simbolo della memoria sonora collettiva. Il **BeltManager** riceve l'informazione di fine manipolazione, riattiva la cintura e ripristina lo stato iniziale dei suoni forma, pronti per nuove interazioni.

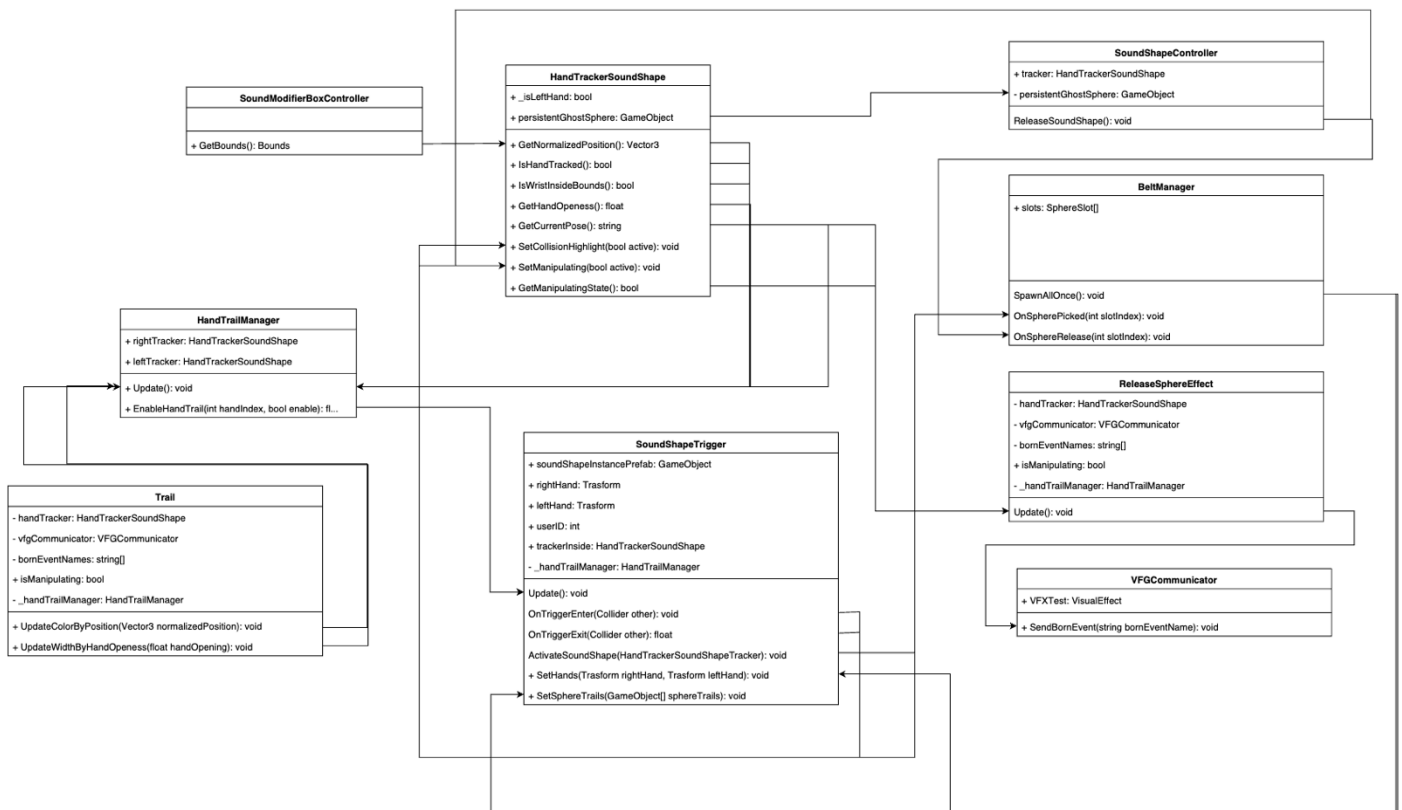


Figura 6.3.3: Schema UML del prototipo SHARED VR (versione Uzbekistan) [F6.3.3]

La **figura 6.3.1** illustra la struttura della versione attuale del prototipo. Ogni mano è gestita da un modulo **HandTrackerSoundShape**, che comunica con i **SoundShapeTrigger** e il **BeltManager**. I moduli **HandTrailManager** e **SoundModifierBoxController** forniscono rispettivamente il feedback visivo e le coordinate normalizzate. **SoundShapeController** e **ReleaseSphereEffect** regolano la logica di manipolazione e rilascio, mentre **VFGCommunicator** gestisce la comunicazione con il dominio grafico.

Lo schema evidenzia un'architettura di tipo incrementale, derivata da un processo di sviluppo iterativo basato su test continui e integrazione progressiva dei moduli, che ha portato alla configurazione stabile attualmente adottata.

Nonostante la solidità funzionale, l'analisi del codice ha evidenziato margini di miglioramento legati principalmente all'organizzazione interna e alla comunicazione tra moduli. Attualmente il passaggio dei dati avviene tramite riferimenti diretti e chiamate pubbliche, generando ridondanze e dipendenze che rendono difficile la manutenzione e l'estensione del sistema.

Il **refactoring** previsto ha l'obiettivo di razionalizzare la struttura attraverso la definizione di manager centralizzati, la semplificazione delle interfacce di comunicazione e l'introduzione di un sistema di eventi per la gestione dello stato e dei feedback. Questa revisione permetterà di ottenere un'architettura più pulita, leggibile e scalabile, che rappresenta il naturale completamento della fase di sviluppo descritta in questo capitolo.

7 Implementazione networking

La co-presenza multiutente è la condizione necessaria affinché SHARED VR esprima la propria natura partecipativa: uno spazio in cui i gesti di ciascun visitatore non restano privati, ma incidono in tempo reale sulla memoria sonora e visiva comune. L'obiettivo del networking consiste dunque nel **mettere in relazione** le azioni degli utenti dentro un sistema condiviso, mantenendo coerenza percettiva e **bassa latenza**.

Fin dalle prime fasi di progettazione, l'architettura del sistema è stata concepita come **client-server** in rete locale (**LAN**): un server dedicato che governa lo stato globale dell'ambiente e fino a tre client su **Meta Quest 3** che si connettono tramite indirizzo IP locale. La scelta di operare in rete locale da un lato riduce l'incertezza di rete e dall'altro permette di ottimizzare l'esperienza per contesti espositivi o installativi, dove la rete può essere configurata e gestita in modo stabile e controllato, garantendo connessioni dedicate solo all'applicazione.

La libreria individuata per la sincronizzazione dei dati è **Netcode for GameObjects (NGO)**, soluzione ufficiale di Unity che introduce concetti nativi come *NetworkManager*, *NetworkObject*, *NetworkBehaviour*, e integra **Unity Transport** su UDP⁵, adeguato a scenari LAN con frequenze di aggiornamento medio-alte e bassa latenza. [50]

Dal punto di vista architetturale, sono stati individuati tre vincoli:

- **Autorità server-side** in modo che solo il server determini lo stato effettivo degli oggetti condivisi, evitando conflitti tra client;
- **Replica continua delle mani** necessaria per mantenere la sensazione di embodiment e prossimità sociale all'interno dello spazio virtuale;
- **Canale di test osservabile**, una camera sul server per monitorare la scena.

Inoltre, per coerenza con la filosofia di SHARED, fondata su gestualità libera e partecipazione diretta, sono stati esclusi i controller XR e i servizi accessori forniti da Unity come *Vivox* e *Lobby* per il matchmaking. Questa scelta ha reso necessaria una riscrittura della pipeline di connessione in modo da adattarla a un'interazione *hands-only*.

In questo capitolo verranno descritte le scelte tecnologiche effettuate, l'architettura ipotizzata per il sistema multiutente e i test preliminari condotti. È importante sottolineare come, al momento della redazione della tesi, la priorità progettuale è stata la realizzazione e il testing della versione single-user del prototipo, presentata in occasione della trasferta a Tashkent per la validazione dell'esperienza immersiva. Conclusa questa fase, lo sviluppo si orienterà verso l'implementazione

⁵ UDP (*User Datagram Protocol*) è un protocollo di trasporto basato su pacchetti, utilizzato in combinazione con IP. È di tipo *connectionless*, cioè non stabilisce una connessione stabile tra mittente e destinatario e non garantisce né l'ordine né la ritrasmissione dei pacchetti persi. Questa struttura lo rende meno affidabile del TCP ma più efficiente e veloce, qualità che lo rendono particolarmente adatto ad applicazioni che richiedono comunicazioni in tempo reale, come videogiochi o ambienti VR. [49]

della componente di rete, già definita nei suoi principi architetturali e pronta per essere integrata a breve nella versione ufficiale di SHARED VR.

7.1 Netcode for Game Object (NGO)

Netcode for GameObjects (NGO) fornisce l'infrastruttura logica per la sincronizzazione in tempo reale tra più utenti. NGO è un framework ad alto livello che si appoggia a **Unity Transport** (UTP) per il trasporto dati e implementa un modello **client-server** in cui il server mantiene l'autorità sullo stato globale, propagando gli aggiornamenti ai client. Questo sistema è pensato per ridurre conflitti, garantire coerenza e bassa latenza, requisiti essenziali in un'installazione VR partecipativa.

A livello concettuale, NGO introduce gli elementi fondamentali della propria architettura:

- il **NetworkManager** che gestisce la sessione di rete e il trasporto dei pacchetti;
- i **NetworkObjects** che assegna a ogni entità un identificativo univoco e ne abilita la replica tra le istanze;
- i **NetworkBehaviour** che contiene la logica di rete e consente di definire Remote Procedure Calls (RPC) e NetworkVariable.

Il **NetworkObject** è il requisito minimo per qualunque entità che debba esistere in rete: fornisce un identificativo univoco sulla sessione, abilita operazioni di *spawn* e *despawn* gestite dal server e gestisce i parametri di visibilità e *ownership* tra le istanze connesse. Per impostazione predefinita il modello è **server-authoritative**: solo il server può gestire lo *spawn* dei NetworkObject e amministrare lo stato, mentre i client ne ricevono la replica secondo le regole di visibilità predefinite. Questo principio separa nettamente l'iniziativa del client, che produce eventi, dalla decisione del server, che stabilisce lo stato finale.

La logica di rete si sviluppa all'interno di script che si basano su **NetworkBehaviour**. Solo in questo contesto è possibile definire le **RPC** (*Remote Procedure Calls*) e accedere alle primitive di sincronizzazione. Il corretto funzionamento richiede il rispetto di tre condizioni principali:

- tutti le RPC devono essere dichiarate in una classe di tipo **NetworkBehaviour**;
- il GameObject che li ospita deve avere un **NetworkObject** attivo in scena;
- le **RPC** non devono essere invocate in *Awake()* o *Start()*, ma solo dopo che la connessione è inizializzata.

L'inizializzazione della logica di rete avviene in **OnNetworkSpawn()**, eseguito automaticamente quando l'oggetto viene registrato dal *NetworkManager*; in questa fase l'oggetto può accedere alle connessioni attive e inizializzare in modo sicuro la sincronizzazione, secondo il ciclo di vita previsto per i componenti di rete. Le invocazioni effettuate prima che l'oggetto sia registrato dal NetworkManager possono generare conflitti di temporizzazione (*race condition*) o errori di rete, poiché le connessioni non sono ancora inizializzate e i messaggi non possono essere instradati correttamente.

Le **RPC** rappresentano le chiamate di rete che consentono l'esecuzione di azioni a distanza tra client e server. NGO distingue chiaramente due tipi di RPC:

- **ServerRPC**, chiamati da un client ma eseguiti sul server, tipicamente utilizzati per notificare azioni locali come, ad esempio, l'ingresso di un player in un trigger;

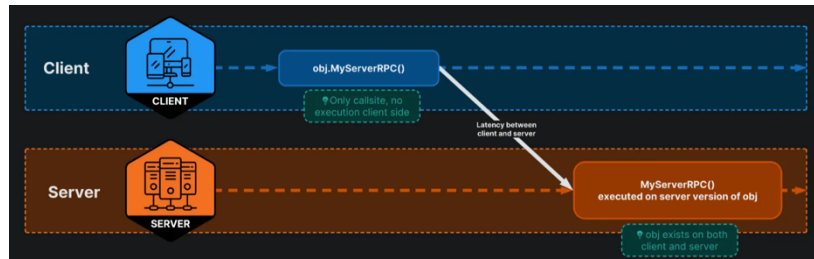


Figura 7.1.1: Funzionamento Server RPC, Unity Docs [F7.1.1]

- **ClientRPC**, chiamati dal server ed eseguiti su tutti i client connessi, utili per propagare uno stato globale aggiornato come il cambio di colore di un oggetto condiviso.

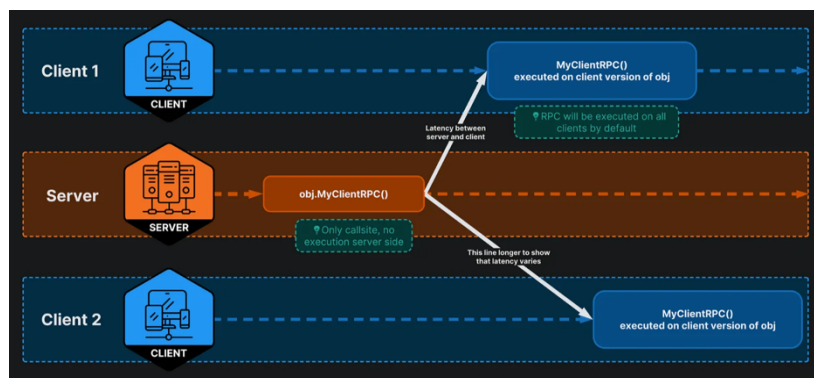


Figura 7.1.2: Funzionamento Client RPC, Unity Docs [F7.1.2]

Le RPC sono ideali per eventi puntuali quindi non persistenti, e si completano con lo strumento delle variabili di stato. La distinzione tra evento e stato è alla base del modello NGO: le RPC veicolano azioni momentanee, mentre la sincronizzazione automatica mantiene valori persistenti. A tal fine, oltre alle RPC, NGO fornisce le **NetworkVariable**, proprietà sincronizzate automaticamente dal server verso i client: si usano quando un valore deve rimanere **persistente e coerente** nel tempo. A differenza di una RPC, che è puntuale, una *NetworkVariable* rappresenta lo **stato** e può notificare cambiamenti tramite *on-value changed callbacks*. Strutture correlate, come *NetworkList<T>*, consentono inoltre di sincronizzare collezioni dinamiche, ad esempio gli array di rotazioni dei giunti della mano.

Alla base del sistema, opera **Unity Transport (UTP)**, il layer di rete **UDP** di Unity progettato per applicazioni real-time. UTP offre diverse modalità di trasmissione dei pacchetti, buffering e

strumenti di integrazione con i servizi **UGS**⁶ come *Relay* e *Lobby*⁷, ma nel contesto di SHARED VR, basato su LAN, si adotta UTP in modalità diretta, riducendo al minimo l'overhead e la latenza ed evitando i servizi intermedi. [51]

Per la fase di sperimentazione, è stato utilizzato il **Multiplayer Play Mode (MPPM)**, un'estensione ufficiale dell'Editor Unity pensata per il debugging multiplayer che permette di aprire più istanze in parallelo, simulando connessioni reali. Questo strumento ha ridotto drasticamente i tempi di test e debugging, permettendo di verificare la corretta sincronizzazione di *spawn*, RPC e *NetworkVariable*. [52]

L'analisi della **SampleScene** fornita dal **VR Multiplayer Template (VRMP)**, che mostra come integrare NGO in un ambiente XR, è stata fondamentale per lo sviluppo. Il template presenta tre elementi caratteristici:

- **XR Interaction Setup (MP Variant)**, variante del rig XR predisposta per il multiplayer, che funge da rig locale per il client e da base per le repliche remote degli altri giocatori;
- **NetworkManager_VR_Multiplayer_Template** gestisce la sessione di rete e definisce il *PlayerPrefab* da istanziare all'ingresso di un utente;
- **XRI Network Game Manager** coordina il flusso di gioco, la logica di *spawn* e *despawn* dei player e integrazione con UI di connessione.

Analizzando il template si nota però che la scena è **fortemente intrecciata** ai servizi Lobby e Vivox⁸, progettati per esperienze online complete. Nel flusso proposto da Unity la vita del player è gestita come **catena unica**: viene creata una lobby, si stabilisce connessione NGO e viene attivato il canale vocale. Nel contesto di SHARED VR, orientato a installazioni in LAN e privo di comunicazione vocale, questa architettura risulta eccessiva e poco flessibile. È stata dunque necessaria una **fase di test e refactoring** volta a separare la logica puramente NGO dalle dipendenze verso *Lobby* e *Vivox*, per ricomporre una scena meno complessa: connessione diretta via UTP, *spawn* controllato e un rig basato unicamente su **XR Hands**. Sulla base della scena di esempio fornita dal template, è stato sviluppato un ambiente di test personalizzato, in cui i blocchi non necessari sono stati sostituiti con implementazioni leggere più adatte al contesto installativo.

⁶ **UGS** (Unity Gaming Services) è la piattaforma cloud ufficiale di Unity che fornisce servizi backend, strumenti di networking e sistemi di analisi per esperienze multiplayer e connesse. [53]

⁷ **Relay** è un servizio incluso negli UGS che funge da server intermedio per instradare pacchetti di rete tra client che non si trovano sulla stessa rete locale, consentendo la connessione anche in assenza di IP pubblico o port forwarding. [54] **Lobby** è un servizio UGS che consente la creazione, la gestione e la scoperta di sessioni multiutente, fornendo meccanismi di matchmaking e connessione tra giocatori prima dell'avvio della comunicazione diretta tramite NGO o Relay. [55]

⁸ **Vivox** è il servizio UGS dedicato alla comunicazione vocale integrata, che fornisce canali audio in tempo reale per applicazioni e giochi multiplayer. Supporta chat di gruppo, comunicazioni private e gestione della prossimità spaziale del suono all'interno di ambienti 3D virtuali. [56]

7.2 Test preliminari

Per validare l'architettura multiutente è stata avviata una fase di sperimentazione mirata a isolare i nodi critici dell'integrazione di **Netcode for GameObjects** (NGO) con un progetto VR concepito per l'interazione *hands-only*. L'obiettivo di questi test è stato duplice: da un lato comprendere a fondo le dinamiche di connessione, *spawn* e sincronizzazione proposte da NGO; dall'altro verificare come adattarle a un contesto che non si appoggia ai servizi Unity di *Lobby* e *Vivox*, strettamente intrecciati nel template VRMP ufficiale ma non coerenti con la filosofia di SHARED VR. Si è dunque adottato un approccio volutamente minimale, costruendo uno script di avvio personalizzato in grado di avviare una sessione in modalità **server**, **host** o **client**, senza intermediari esterni.

Il codice riportato di seguito mostra la configurazione adottata per la connessione e la creazione della sessione:

```
public void StartServer() {
    var utp = NetworkManager.Singleton.GetComponent<UnityTransport>();
    utp.SetConnectionData(serverIPAddress.Trim(), ServerPort, "0.0.0.0");
    NetworkManager.Singleton.StartServer();
}

public void StartClient() {
    var utp = NetworkManager.Singleton.GetComponent<UnityTransport>();
    utp.SetConnectionData(serverIPAddress.Trim(), serverPort);
    NetworkManager.Singleton.StartClient();
}

public void StartHost() {
    var utp = NetworkManager.Singleton.GetComponent<UnityTransport>();
    utp.SetConnectionData(serverIPAddress.Trim(), serverPort, "0.0.0.0");
    NetworkManager.Singleton.StartHost();
}
```

Per quanto riguarda la configurazione del server, il metodo di avvio utilizza l'istruzione **SetConnectionData(serverIPAddress, ServerPort, "0.0.0.0")** che definisce l'indirizzo e la porta del server e specifica il *binding*⁹ su tutte le interfacce di rete disponibili, consentendo la connessione da qualsiasi client appartenente alla stessa LAN.

In questa configurazione, la variabile *serverIPAddress* rappresenta l'indirizzo IP della macchina che ospita il server: durante i test in locale è impostata su **127.0.0.1**, mentre nelle connessioni in rete locale corrisponde all'IP reale del dispositivo. L'indirizzo 127.0.0.1 è l'indirizzo di **loopback**, cioè un indirizzo riservato al dispositivo stesso che consente di stabilire una connessione di rete interna senza coinvolgere alcuna interfaccia fisica. La porta di comunicazione utilizzata è la **7777**, una scelta consolidata negli ambienti di sviluppo multiplayer, infatti, questo valore è comunemente impiegato nei framework di rete, inclusi *Unity Transport* e *Netcode for GameObjects*, in quanto

⁹ Nel contesto delle reti, il termine **binding** indica l'associazione di una porta e di un indirizzo IP a un socket, ossia al punto di ascolto utilizzato da un server per ricevere connessioni in ingresso. [57]

non riservato a servizi di sistema critici e riconosciuto come porta standard per applicazioni di gioco o test in LAN.

Durante i test è stata preferita la modalità **StartServer()**, che consente di mantenere la macchina del server priva di componenti XR e input, riservandola alle funzioni di regia e monitoraggio su desktop. I client su Meta Quest 3 si collegano con IP statico del server: non usando Lobby, il rilevamento della sessione avviene manualmente, in modo indipendente dai servizi UGS, garantendo controllo completo dell'infrastruttura e latenza minima.

Dopo aver definito la configurazione di rete e stabilito una connessione stabile tra server e client, il **primo passo del testing** riguarda la verifica dei componenti fondamentali di comunicazione, in particolare le **Remote Procedure Call (RPC)** e le **NetworkVariable**, utilizzate per la trasmissione dei messaggi e la sincronizzazione degli stati tra le istanze.

Per questa fase è stata realizzata una **scena di test dedicata**, un ambiente sperimentale costruito intorno a un **Totem centrale**, elemento condiviso da tutti gli utenti connessi. Il Totem funge da riferimento comune e da indicatore visivo dell'andamento della rete: serve a monitorare la posizione media dei giocatori nella scena e a visualizzare, in tempo reale, l'efficacia della sincronizzazione.

Quando i giocatori si avvicinano, il Totem cresce in altezza; se restano fermi per alcuni secondi, comincia a lampeggiare; e quando tutti entrano nell'area circolare che lo circonda, cambia colore diventando verde, segnalando il raggiungimento della condizione collettiva di prossimità.

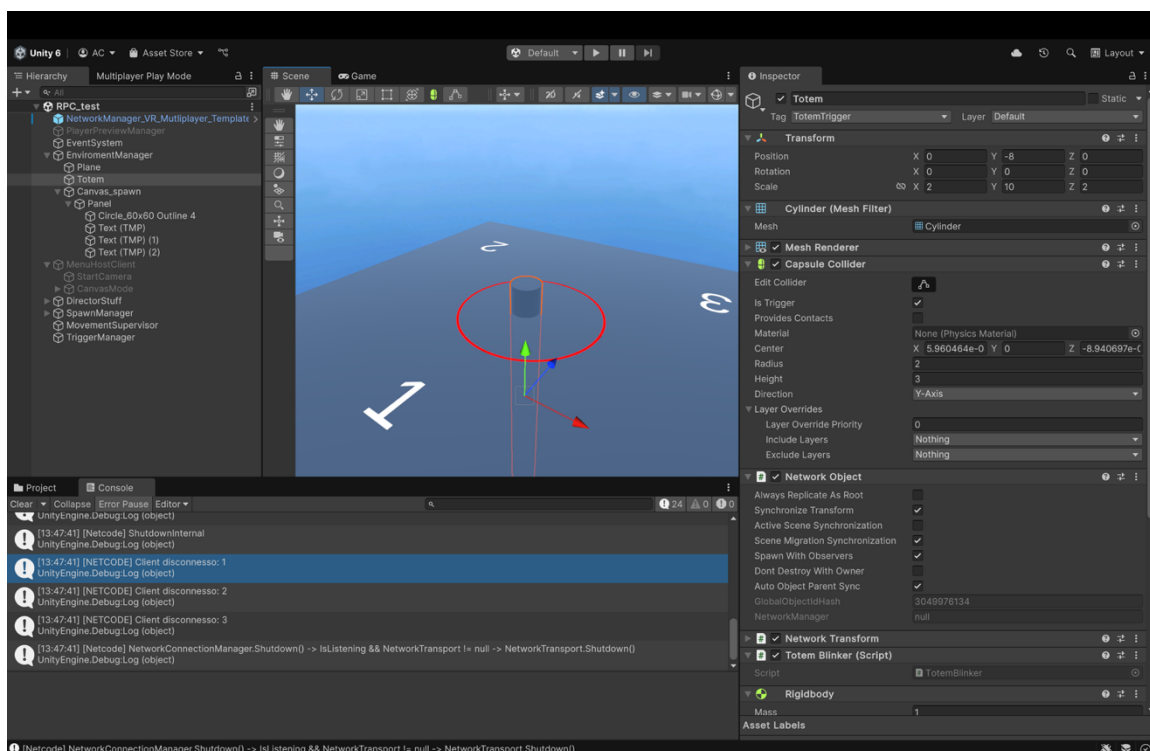


Figura 7.2.1: Screenshot scena RPC_test [F7.2.1]

Il funzionamento del Totem è gestito interamente dal server attraverso lo script **MovementMonitor.cs**, che calcola la distanza media tra i giocatori e determinava lo stato visivo corrispondente. La logica opera tramite una coroutine che, a intervalli regolari, sincronizza le posizioni dei client aggiornando lo stato condiviso. La funzione *EvaluateBlinkAndRaise()* controlla se tutti i client sono presenti e fermi oltre un certo intervallo temporale; in caso positivo, attiva il lampeggio del Totem, sincronizzato automaticamente su tutte le istanze tramite *NetworkVariable*.

Quando tutti i giocatori entrano nell'area del trigger, un messaggio *ClientRpc* tramite lo script **TriggerManager.cs**, aggiorna lo stato visivo su tutti i client, impostando il colore verde e disattivando il lampeggio:

```
[ClientRpc]
void SetGreenClientRpc()
{
    if (totem == null) totem = EnvironmentManager.Instance?.Totem;

    var r = totem?.GetComponent<Renderer>();
    if (r) r.material.color = Color.green;

    var b = totem?.GetComponent<TotemBlinker>();
    if (b) b.SetBlink(false);
}
```

La gestione coordinata di questi stati ha dimostrato la corretta replicazione di eventi e variabili di rete tra server e client. I test in LAN hanno confermato che il modello *server-authoritative* garantisce coerenza e latenza minima, requisito essenziale per un'esperienza multiutente sincrona.

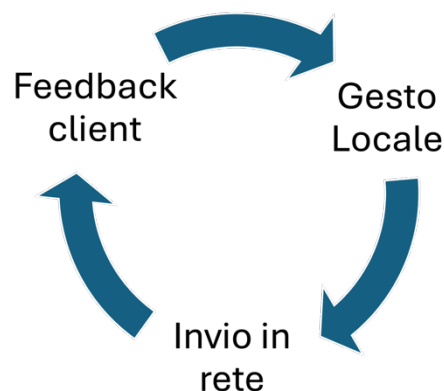


Figura 7.2.2: Ciclo di interazione [F7.2.2]

Alla base del sistema c'è un ciclo continuo di interazione: l'utente compie un **gesto locale**, che viene interpretato in tempo reale dal sistema e trasmesso in rete tramite i meccanismi di comunicazione di NGO. Il **server** elabora l'evento, aggiorna lo stato condiviso e invia ai client le modifiche corrispondenti, generando così un **feedback** coerente per tutti i partecipanti. Questo

processo ciclico è alla base della logica di rete co-presenza multiutente, e consente a ogni azione individuale di trasformarsi immediatamente in una variazione collettiva dell'ambiente virtuale.

Per minimizzare la latenza e i conflitti si prevedono alcune strategie:

- **Autorità server-side:** il server è l'entità che decide lo stato finale degli oggetti, evitando conflitti tra client.
- **Compressione dati:** invio solo dei parametri essenziali (posizione, rotazione, azione).

Completata la validazione del Totem, il passo successivo del testing si concentra sulla sincronizzazione del player XR. La **sample scene** di VR Multiplayer Template fornisce un Player Prefab già predisposto alla rete: il prefab include il *NetworkObject*, i componenti di sincronizzazione delle trasformazioni per testa e mani, uno script di lancio che istanzia il player al momento della connessione e, soprattutto, un sistema di **replica** delle mani che legge i dati locali e li riproduce sulle istanze remote tramite NGO.

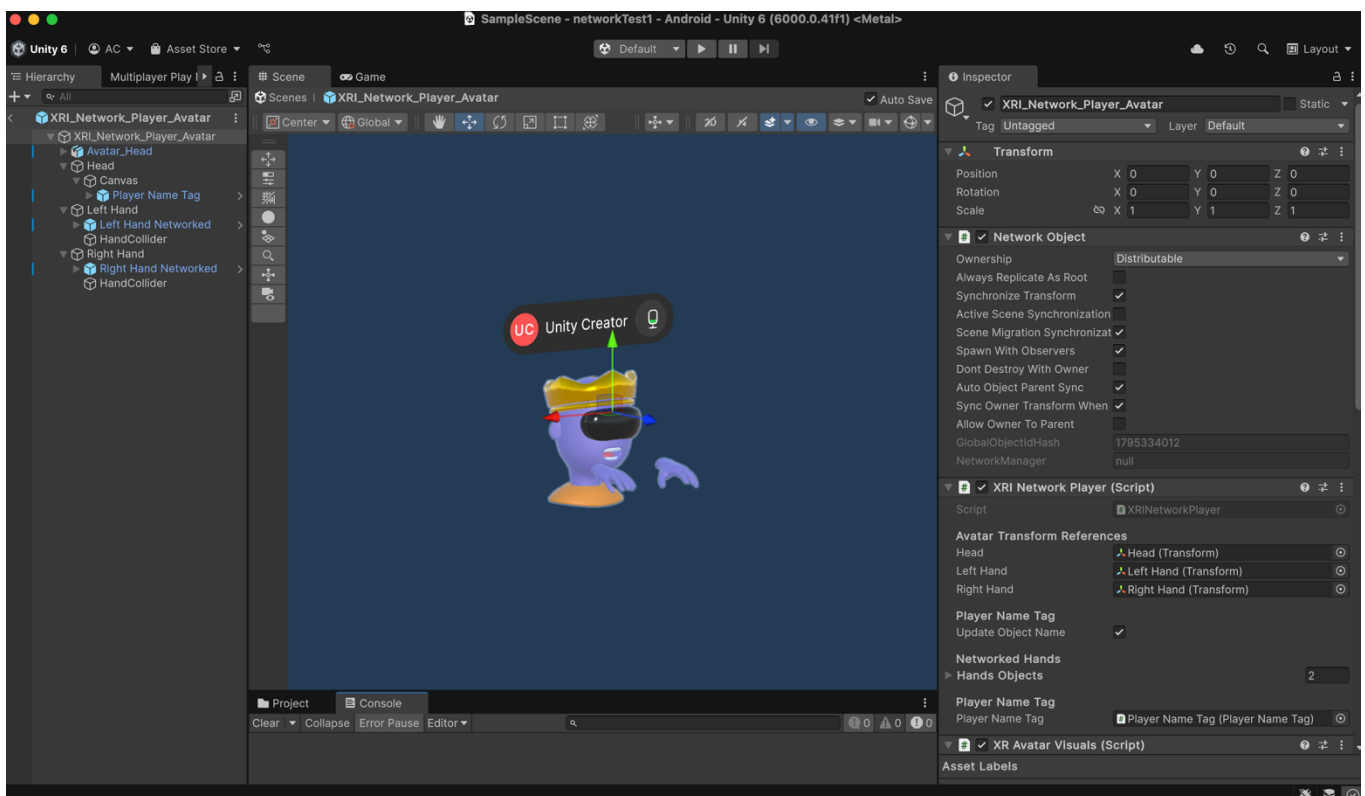


Figura 7.2.3: Screenshot XRI_Network_Player_Avatar VRMPAsset [F7.2.3]

Tuttavia, nel template originale, lo script **XR HandposeReplicator** associato al prefab del player remoto e responsabile della replica dei movimenti dell'utente locale, include riferimenti ai controller XR basati su azioni o input fisici. Lo script prevede infatti il seguente metodo:

```

void SetupLocalHands()

{
    m_LeftControllerTransformReference = m_XRModalityManager.leftController.transform;
    m_RightControllerTransformReference = m_XRModalityManager.rightController.transform;

    if (m_XRModalityManager.leftHand == null)    //Rig doesn't have hands setup
    {
        m_LeftHandTransformReference = m_XRModalityManager.leftController.transform;
        m_RightHandTransformReference = m_XRModalityManager.rightController.transform;
        SetTrackingType(XRInputModalityManager.InputMode.MotionController);
    }
    else    //Setup Hands and modality change listeners
    {
        m_LeftHandTransformReference =
        m_XRModalityManager.leftHand.GetComponentInChildren<XRHandSkeletonDriver>()
        .rootTransform;
        m_RightHandTransformReference =
        m_XRModalityManager.rightHand.GetComponentInChildren<XRHandSkeletonDriver>()
        .rootTransform;

        SetTrackingType(XRInputModalityManager.currentInputMode.Value);

        m_XRModalityManager.trackedHandModeStarted.AddListener(SwapToHands);
        m_XRModalityManager.motionControllerModeStarted.AddListener(SwapToControllers);
    }
}

```

Questa funzione rileva la modalità di input attiva e associa dinamicamente i riferimenti per la sincronizzazione della mano sinistra e destra. Tuttavia, lo script mantiene le dipendenze dai controller fisici, utilizzati come sorgente primaria per i dati di posizione e rotazione.

Nel contesto di SHARED VR, invece, la pipeline è interamente *hands-only*, basata sul tracciamento delle mani tramite XR Hands. L'assenza di controller rende questi riferimenti non validi, generando errori di inizializzazione e impedendo la trasmissione dei dati di movimento in rete. Per risolvere questo problema, è stato quindi sviluppato uno script dedicato, **XRHandPoseReplicator_HandsOnly**, che mantiene la logica di replicazione originale ma elimina completamente i riferimenti ai controller, semplificando la struttura.

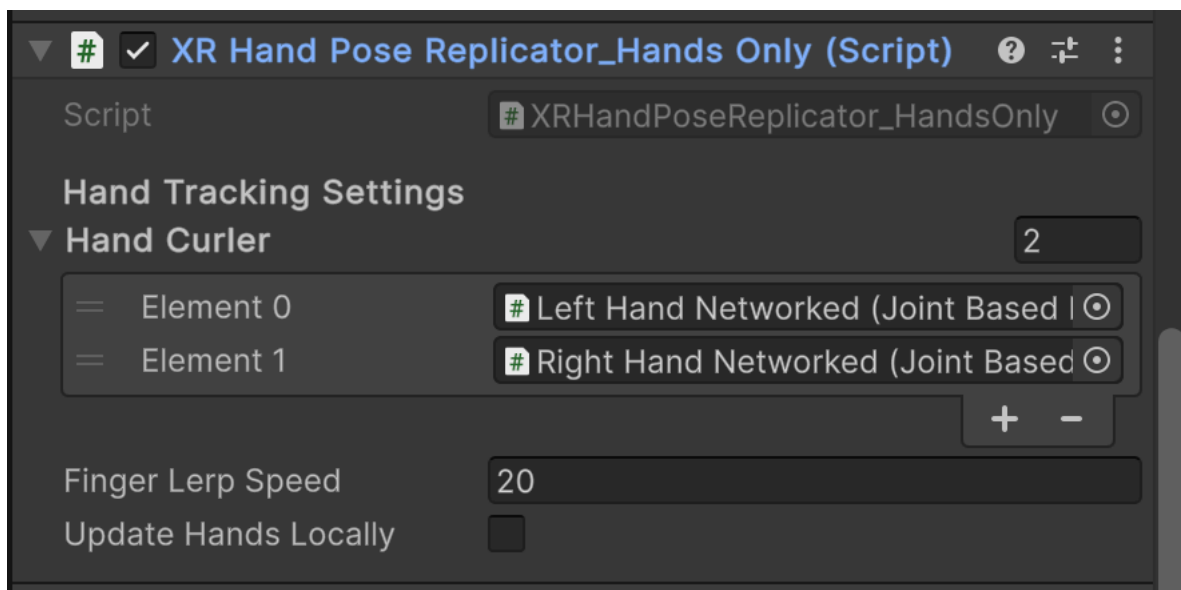


Figura 7.2.4: Screenshot Inspector XRHandPoseReplicator_HandsOnly [F7.2.4]

In sintesi, il **player owner** legge ogni frame le rotazioni dei giunti da XR Hands e le pubblica in rete tramite `NetworkList<Quaternion>`, mentre le **repliche remote** applicano i valori con interpolazione. In parallelo, un altro elemento, **XRI_NetworkPlayer_NonVoice** aggiorna in `LateUpdate()` la posizione della testa e delle mani, disattiva localmente i renderer del proprio avatar per evitare duplicazioni e applica gli aggiornamenti solo all'istanza di proprietà. Le copie remote rimangono sincronizzate in sola lettura.

Grazie a questa architettura, il sistema di rete replica fedelmente i movimenti di testa e mani in LAN, senza dipendenze da controller, *Lobby* o *Vivox*, eliminando i warning generati dai pacchetti non utilizzati del template originale.

Le modifiche apportate, che consistono nella riscrittura di alcuni componenti e la rimozione di riferimenti superflui, hanno permesso di ottenere un'infrastruttura stabile, leggibile e perfettamente compatibile con l'obiettivo di SHARED VR.

Le sperimentazioni condotte finora, tuttavia, hanno confermato la stabilità della replica delle mani e la corretta propagazione dei messaggi di rete tramite RPC e `NetworkVariable`, fornendo la conferma tecnica della solidità del sistema e una base pronta per la futura estensione multiutente.

8 Testing e validazione

La **fase di testing** si è svolta durante la tappa del programma ItARTS a **Tashkent**, dove è stato presentato al pubblico il prototipo single-user di SHARED VR. L'obiettivo era quello di osservare come i visitatori interagissero con l'ambiente virtuale e raccogliere le loro impressioni sull'usabilità del sistema e sulla qualità dell'esperienza immersiva.

È importante precisare che la sperimentazione non ha riguardato la versione definitiva dell'opera, ma un prototipo **single-user** sviluppato per verificare il corretto funzionamento della pipeline di interazione gestuale e la coerenza dei feedback audiovisivi. L'esperienza proposta ai visitatori aveva quindi un carattere puramente prototipale e valutativo, concepito per raccogliere osservazioni utili all'ottimizzazione successiva del sistema.

L'allestimento è stato ospitato nella **Sala dell'Organo del Conservatorio di Tashkent**, dove sono state predisposte **tre postazioni indipendenti**, ciascuna composta da un PC desktop e un visore **Meta Quest 3** collegato tramite **Oculus Link**.

A causa dell'instabilità della rete locale, non è stato possibile utilizzare la connessione **AirLink** wireless: è stata dunque adottata una **configurazione cablata**, che garantiva maggiore stabilità e controllo tecnico. Il progetto è stato eseguito direttamente dall'editor di Unity, in modo da poter monitorare in tempo reale il flusso dei dati verso Max/MSP e intervenire rapidamente sulla scena in caso di necessità.

Ogni postazione permetteva all'utente di interagire con l'ambiente sonoro e visivo, utilizzando esclusivamente le mani come strumento d'interazione. Il test mirava quindi a verificare la fluidità del tracciamento gestuale, la coerenza tra gesto e feedback audiovisivo e il livello di coinvolgimento individuale percepito.

8.1 Metodologia e strumenti di valutazione

Le sessioni di testing si sono svolte nell'arco di due giornate, coinvolgendo un totale di **22 partecipanti** con età compresa tra 18 e 68 anni. Il campione presentava profili differenti per formazione e livello di familiarità con le tecnologie interattive: la maggioranza non aveva esperienze precedenti in realtà virtuale, mentre alcuni hanno dichiarato un utilizzo occasionale di videogiochi o applicazioni immersive. Questa eterogeneità ha consentito di osservare come utenti con competenze diverse si adattassero all'interazione gestuale in un contesto di fruizione libera e non guidata.

Ogni partecipante ha affrontato un'esperienza immersiva della durata di circa **cinque minuti**, preceduta da una breve introduzione in cui venivano illustrate le modalità d'interazione e gli obiettivi della ricerca. Durante la sessione, l'utente era libero di esplorare e sperimentare le reazioni luminose e sonore generate dai propri gesti, scoprendo progressivamente la logica

dell'interazione attraverso un processo di **apprendimento implicito**. Al termine dell'esperienza, ogni partecipante ha compilato un **questionario anonimo** volto a raccogliere valutazioni sull'usabilità e sulla qualità percettiva della realtà virtuale.

La scena utilizzata per il testing corrispondeva alla **demo della cintura di soundshape**, realizzata per verificare il corretto funzionamento della pipeline di tracciamento e interazione gestuale.

Pur non rappresentando la versione definitiva dell'ambiente virtuale, la demo riproduceva la struttura spaziale e concettuale dell'opera, dal punto di vista visivo, la scena era ancora in fase di lavorazione, ma già rispondeva all'impostazione poetica di SHARED VR, basata su un'estetica minimale e sulla centralità della risposta audiovisiva alle azioni corporee dell'utente.

Durante il testing sono state proposte due varianti visive dell'effetto associato al gesto di rilascio, elemento cardine dell'interazione. Nella **prima versione** i trail luminosi erano costantemente attivi per entrambe le mani e, al momento del rilascio, l'oggetto nella mano dell'utente veniva attratto e assorbito dall'oracolo. Nella **seconda versione**, invece, il trail risultava attivo solo per la mano che afferrava l'oggetto, e al rilascio veniva generata una scia luminosa dal palmo che si riversava all'interno dell'oracolo, dando origine a una particella colorata legata all'identità dell'utente.

I questionari erano articolati in tre sezioni:

- dati anagrafici e grado di familiarità con la realtà virtuale;
- valutazione dell'usabilità basata **sul System Usability Scale (SUS)**;
- valutazione della presenza basata sull'**Igroup Presence Questionnaire (IPQ)**.

Le risposte sono state elaborate per ottenere indicatori statistici e per confrontare qualitativamente le percezioni emerse durante l'esperienza. L'analisi complessiva ha permesso di valutare la comprensibilità del gesto, la coerenza dei feedback e il livello di immersione percepito, costituendo una base di validazione utile alla successiva evoluzione del sistema verso la modalità multiutente.

La scelta di utilizzare la **System Usability Scale (SUS)** risponde alla necessità di disporre di uno strumento affidabile per la valutazione dell'usabilità percepita. Ideata da John Brooke (1986) [58], la SUS è composta da dieci affermazioni valutate su scala Likert da uno a cinque. Il punteggio complessivo, normalizzato su una scala da 0 a 100, fornisce una misura comparabile tra diversi sistemi interattivi: secondo le soglie individuate da Bangor, Kortum e Miller (2008), valori superiori a 68 indicano un livello di usabilità buono o accettabile. [59]

Parallelamente, l'**Igroup Presence Questionnaire (IPQ)**, elaborato da Schubert, Friedmann e Regenbrecht (2001) [60], è stato utilizzato per misurare la percezione di presenza all'interno dell'ambiente virtuale. Questo strumento si articola in tre dimensioni principali: la **presenza**, intesa come sensazione di "essere nel luogo" virtuale; il **coinvolgimento**, cioè il grado di attenzione e

assorbimento nella scena; e il **realismo**, ovvero la coerenza percettiva e cognitiva dell'esperienza simulata.

L'uso congiunto dei due questionari ha permesso di integrare misure oggettive e soggettive, confrontando i dati quantitativi con le osservazioni qualitative raccolte sul campo. Questo approccio ha fornito una visione coerente dell'esperienza utente, utile non solo per validare l'efficacia del prototipo, ma anche per interpretare in chiave critica le relazioni tra tecnologia, percezione e partecipazione che costituiscono il nucleo concettuale dell'opera.

8.2 Analisi dei risultati

Il campione mostra un'età media di **34.6 anni**, con una distribuzione bilanciata tra i generi. La maggior parte dei partecipanti non aveva precedenti esperienze di realtà virtuale immersiva, ma ha riferito un apprendimento rapido e una buona familiarità con l'ambiente dopo pochi secondi di interazione.

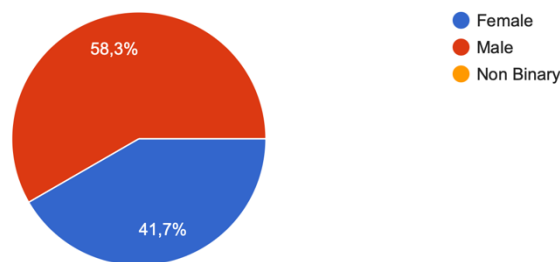


Figura 8.2.1 – Grafico distribuzione di genere [F8.2.1]

How often do you use interactive applications on your computer or mobile device (for example, video games)?

22 risposte

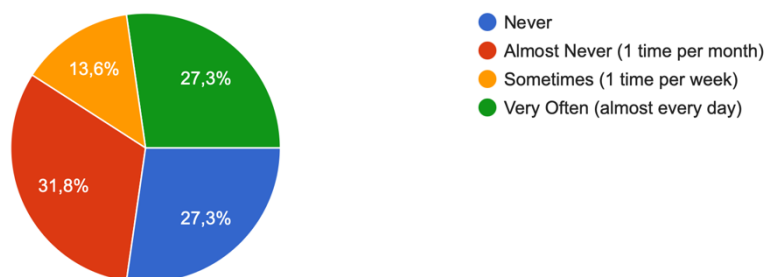


Figura 8.2.2 – Grafico usabilità *interactive application* [F8.2.2]

How often do you use immersive virtual reality applications (using a headset)?

22 risposte

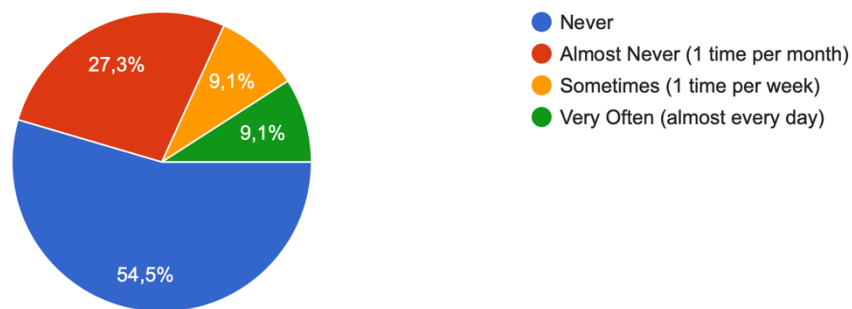


Figura 8.2.3 – Grafico usabilità VR application [F8.2.3]

L'analisi dei punteggi relativi alle domande del SUS evidenzia una **valutazione positiva dell'usabilità complessiva** del sistema. Le affermazioni riguardanti la **semplicità d'uso**, la **rapidità di apprendimento** e la **sensazione di controllo durante l'interazione** hanno ottenuto una media tra **3.6 e 3.8 su 5**, confermando che il sistema è percepito come intuitivo e facilmente gestibile anche da utenti privi di esperienze precedenti in realtà virtuale.

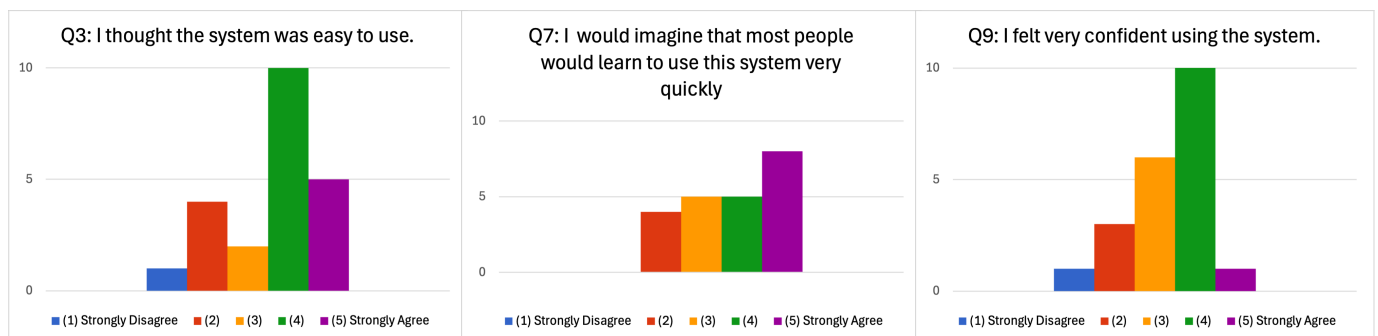


Figura 8.2.4 – Distribuzione delle risposte relative all'usabilità del sistema VR (Q3, Q7, Q9) [F8.2.4]

Al contrario, le domande relative alla **necessità di supporto tecnico** o alla **percezione di complessità** (Q2–Q4) mostrano valori inferiori con una media tra **2.1 e 2.5**, indicando che l'esperienza risulta accessibile e autonoma.

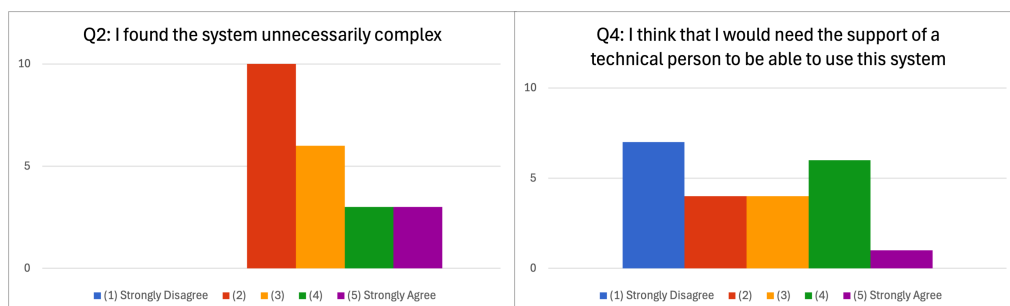


Figura 8.2.5 – Distribuzione delle risposte relative all'usabilità del sistema VR (Q2, Q4) [F8.2.5]

Il punteggio medio complessivo, calcolato secondo la formula standard del SUS, è pari a **67.9**, un valore che si colloca **intorno alla soglia di accettabilità** e **vicino alla fascia “Good”** individuata da *Bangor et al. (2008)*, confermando la buona percezione di usabilità del sistema anche nella sua fase prototipale. Ciò suggerisce che il sistema, pur essendo un prototipo, garantisce **un’esperienza intuitiva e coerente** anche per utenti non esperti.

Le domande dell’**Igroup Presence Questionnaire (IPQ)** hanno restituito valori medi compresi tra **4.6 e 5.2 su 7** nelle dimensioni di *presence*, *realismo* e *coinvolgimento*.

I punteggi più elevati riguardano la **sensazione di essere presenti nello spazio virtuale** e la **percezione di agire direttamente all’interno dell’ambiente**, con medie rispettivamente di **5.2** e **4.8**, ciò indica una forte sensazione di **immersione e agency**, elementi centrali per la fruizione artistica di *SHARED VR*.

I felt present in the virtual space.

22 risposte

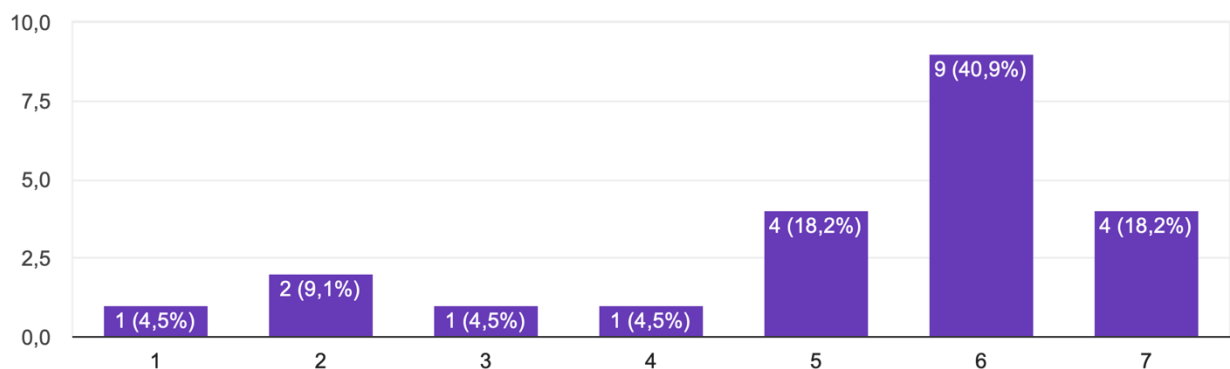


Figura 8.2.6 – Distribuzione delle risposte relative alla percezione di presenza nello spazio virtuale [F8.2.6]

I had a sense of acting in the virtual space, rather than operating something from outside.

22 risposte

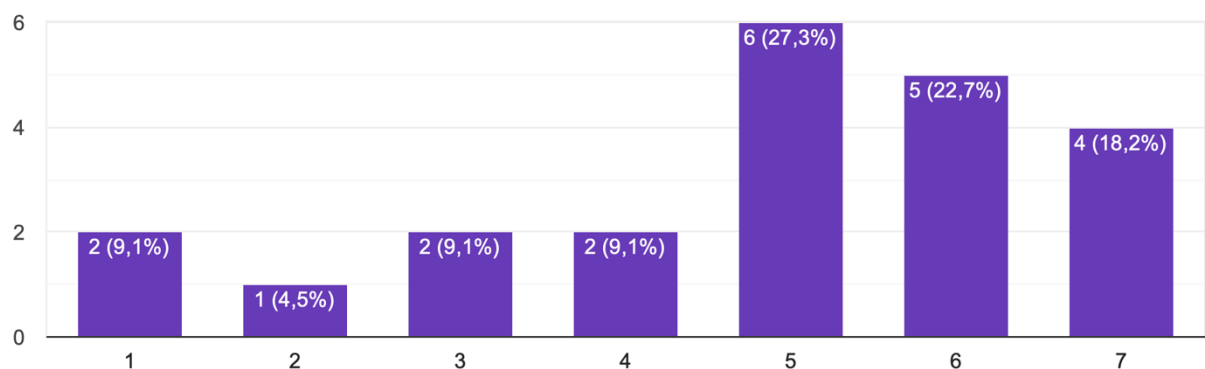
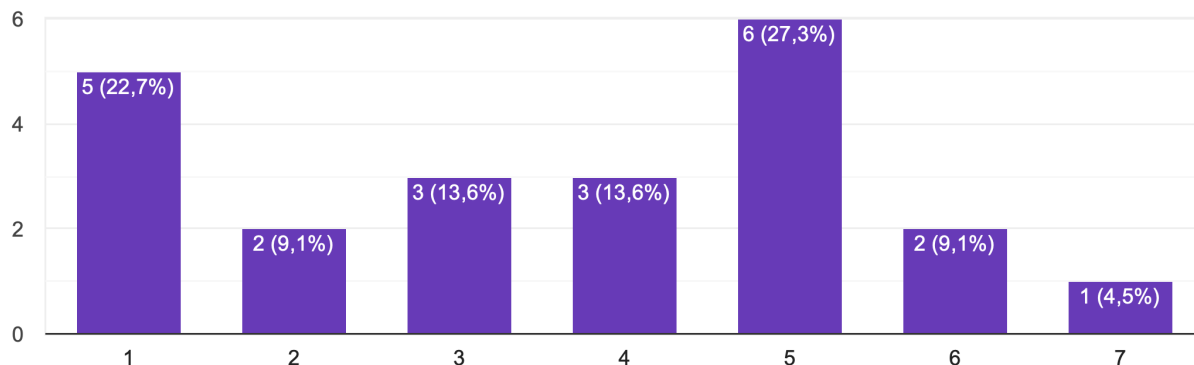


Figura 8.2.7 – Distribuzione delle risposte relative alla percezione di *agency* [F8.2.7]

Le voci relative all'**attenzione verso l'ambiente reale** hanno ottenuto valori medi inferiori, compresi tra **3.5 e 3.9**, coerenti con una **buona captazione sensoriale** e con la **temporanea sospensione della consapevolezza del contesto fisico** tipica delle esperienze immersive.

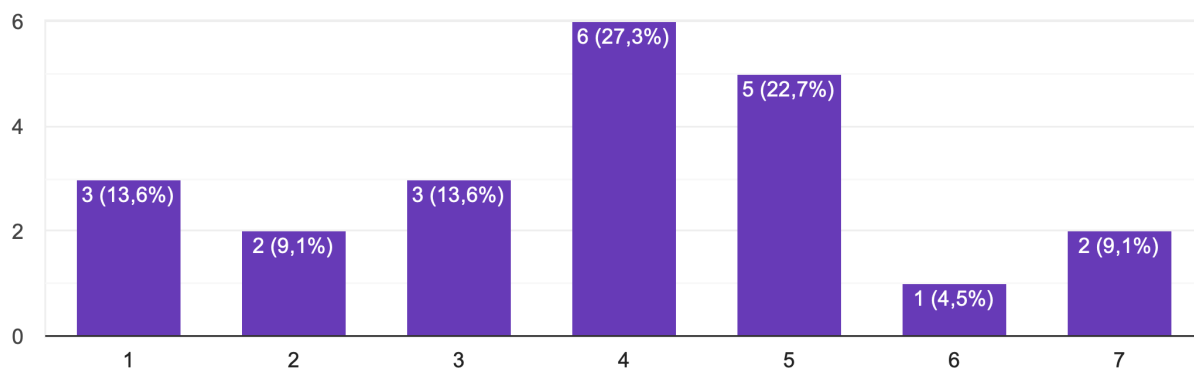
How aware were you of the real world surrounding while navigating in the virtual world? (i.e. sounds, room temperature, other people, etc.)?

22 risposte



I was not aware of my real environment.

22 risposte



I still paid attention to the real environment.

22 risposte

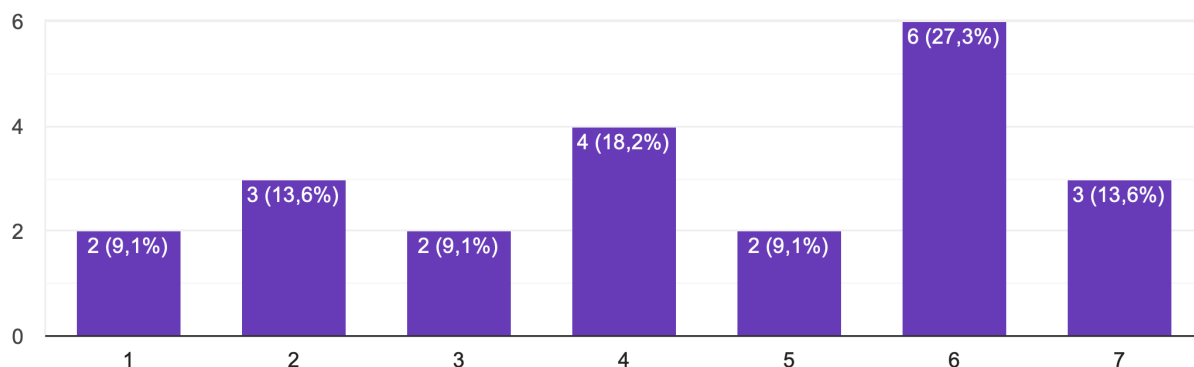


Figura 8.2.8– Distribuzione delle risposte relative alla dimensione di *awareness* del mondo reale [F8.2.8]

Nel complesso, il questionario conferma un **alto livello di presenza soggettiva**, una **buona credibilità percettiva** e un **coinvolgimento coerente con la poetica dell'opera**, in cui il gesto corporeo diventa il principale mediatore tra spazio virtuale e costruzione sonora condivisa.

La **Tabella 8.1** riassume i valori medi e le deviazioni standard dei punteggi SUS e IPQ, evidenziando la coerenza tra le misure di usabilità e la qualità esperienziale percepita.

Dimensione	Scala	Media	Dev. Std	Interpretazione
SUS – Usabilità complessiva	0–100	67.9	11.1	Buona usabilità, interazione intuitiva
IPQ – Presenza	1–7	5.18	1.74	Forte senso di “essere nel luogo”
IPQ – Agency	1–7	4.82	1.84	Sensazione di agire nello spazio virtuale
IPQ – Realismo	1–7	4.73	1.70	Elevata credibilità percettiva
IPQ – Coinvolgimento	1–7	4.64	1.68	Coinvolgimento e attenzione elevati
IPQ – Awareness reale	1–7	3.59	1.89	Bassa attenzione all’ambiente esterno

Tabella 8.2.1 – Sintesi dei risultati quantitativi

8.3 Discussione e osservazioni qualitative

L’analisi complessiva dei dati, integrata con le osservazioni dirette raccolte durante le sessioni, evidenzia un livello di usabilità e di coinvolgimento soddisfacente per la fase di sviluppo in cui si trova il progetto, ma anche alcune criticità. I risultati del questionario **SUS**, con un punteggio medio pari a **67.9**, confermano che il sistema è percepito come stabile e intuitivo, sebbene non ancora pienamente ottimizzato.

L’Interazione basata sul movimento naturale delle mani è risultata comprensibile per la maggior parte dei partecipanti, ma non sempre pienamente fluida: in diversi casi si sono osservati piccoli ritardi nel riconoscimento gestuale soprattutto nelle prime fasi dell’esperienza.

L’analisi dell’**IPQ** evidenzia un alto livello di presenza percepita, infatti gli utenti hanno percepito un efficace senso di **immersione** e di **agency**, interpretando le proprie azioni come parte integrante del mondo virtuale. Le osservazioni in sede di testing confermano questo comportamento: i partecipanti tendevano a **muovere le mani con naturalezza crescente** nel corso della prova, adattandosi progressivamente al sistema di tracciamento e mostrando **curiosità esplorativa** verso le reazioni visive e sonore dell’ambiente. Il valore relativamente più basso della dimensione awareness reale indica che la maggior parte dei partecipanti ha temporaneamente ridotto l’attenzione verso l’ambiente fisico, a favore di un assorbimento

sensoriale completo nell'esperienza. Il risultato è coerente con l'obiettivo del progetto SHARED, che ricerca uno stato di partecipazione immersiva in cui il gesto non è più semplice controllo tecnico, ma atto estetico e relazionale.

Il confronto tra le **due varianti visive** del rilascio ha offerto ulteriori spunti interpretativi. La versione con trail costantemente attivi e oggetto fisico assorbito dall'oracolo è stata percepita come più spettacolare e immediata, ma la seconda versione è risultata più leggibile e coerente con la poetica del gesto generativo. Diversi partecipanti hanno riferito che questa modalità trasmetteva una maggiore **sensazione di legame** con l'ambiente virtuale, poiché ogni gesto di rilascio produceva una scia luminosa che rimaneva visibile nello spazio, suggerendo una forma di continuità tra l'azione e la scena.

Sebbene la meccanica fosse implementata solo a livello visivo, questa risposta percettiva ha contribuito a rafforzare il **coinvolgimento emotivo** e la **percezione di connessione** con l'opera, in linea con il concetto di memoria condivisa che ne guida la progettazione.

Nel complesso, i risultati confermano che il prototipo di SHARED VR soddisfa i requisiti di **stabilità, usabilità e coinvolgimento** propri di una piattaforma interattiva destinata a contesti artistici e installativi. L'esperienza si distingue per la sua capacità di integrare il gesto corporeo con la risposta sonora e visiva, offrendo all'utente una percezione attiva del proprio ruolo nella costruzione dell'opera. I dati raccolti forniscono quindi indicazioni concrete per la prosecuzione del progetto: ottimizzare il sistema di tracciamento e la resa visiva del rilascio, potenziare la componente audio generativa e ampliare le possibilità di interazione collaborativa, mantenendo come riferimento la poetica della memoria condivisa che costituisce l'essenza di SHARED-ItARTS.

9 Conclusioni e Future Works

Il lavoro svolto in questa tesi rappresenta **una fase centrale nello sviluppo** dell'opera **SHARED VR**, in cui la ricerca teorica, la progettazione tecnica e la sperimentazione artistica si sono intrecciate in modo concreto.

Questa fase di ricerca ha consentito di **trasporre in forma applicativa** i principi fondativi del progetto, ovvero gesto, partecipazione e memoria condivisa, verificandone la coerenza all'interno di un ambiente virtuale interattivo.

L'attività svolta ha portato alla definizione di un'architettura integrata basata su **Unity, XR Hands, VRMP e Max/MSP**, in grado di mettere in relazione la dimensione visiva, sonora e gestuale in tempo reale.

La struttura implementata costituisce una **base solida per le successive fasi di sviluppo**, sia in termini di stabilità e scalabilità del sistema, sia per quanto riguarda l'allineamento tra componente tecnica e visione artistica.

Le prove sperimentali hanno evidenziato la **credibilità dell'esperienza immersiva** e la **chiarezza dell'interazione gestuale**, confermando la validità delle scelte progettuali, pur segnalando margini di miglioramento in relazione alla precisione del tracciamento e all'ottimizzazione della rete locale.

9.1 Valutazione critica e contributo alla ricerca artistica

Dal punto di vista tecnico, lo sviluppo del prototipo ha permesso di consolidare una pipeline di lavoro stabile, scalabile e coerente con la visione artistica originaria. L'integrazione tra Unity e Max/MSP si è rivelata efficace nel garantire un **dialogo continuo tra spazio visivo e spazio sonoro**, elemento cardine nella visione di Lupone.

Sul piano concettuale e artistico, SHARED VR rappresenta un contributo significativo alla **ricerca sulle forme di interazione partecipativa** nella realtà virtuale, in quanto l'opera si colloca nel solco delle esperienze di **arte immersiva e generativa**, ma ne rinnova la prospettiva attraverso l'attenzione al gesto come atto collettivo di costruzione della memoria.

Da un punto di vista teorico, dunque, SHARED VR si configura come un **dispositivo estetico e relazionale**, in cui il gesto individuale diventa un frammento di una costruzione collettiva. Ogni azione compiuta dagli utenti lascia traccia nella memoria sonora dell'opera, contribuendo alla sua continua trasformazione. In questo modo, il progetto concretizza, nel linguaggio della VR, la tensione tra individuale e collettivo, tra presenza e memoria, tra materia tecnologica e intenzione artistica.

L'attuale versione del prototipo costituisce pertanto una fase di validazione e di crescita del progetto, destinata a evolversi ulteriormente nelle successive sperimentazioni.

9.2 Prospettive e sviluppi futuri

I risultati ottenuti costituiscono una base solida per la prosecuzione del lavoro e delineano **prospettive di sviluppo concrete**, mirate a consolidare il prototipo e a portarlo verso una versione più completa e stabile.

Le prossime fasi di progetto potranno concentrarsi su tre direzioni principali:

- **Implementazione della modalità multiplayer**, con l'obiettivo di estendere l'esperienza alla dimensione multiutente e di realizzare così la co-presenza all'interno dell'ambiente virtuale;
- **Revisione e potenziamento dell'apparato visivo**, per ottenere una rappresentazione più coerente con la poetica di *SHARED*, migliorando la resa estetica e la leggibilità delle reazioni luminose legate al gesto tuttora in fase di definizione;
- **Refactoring del codice e ottimizzazione dei moduli esistenti**, in particolare per quanto riguarda il sistema di hand tracking, così da garantire maggiore stabilità, precisione e scalabilità in contesto installativo;

Questi interventi rappresentano **il naturale proseguimento del percorso di sviluppo** avviato con questa tesi e sono finalizzati a portare *SHARED VR* verso un livello di maturità tecnica e artistica più avanzato.

9.3 Considerazioni finali

Lo sviluppo di *SHARED-ItARTS VR* ha rappresentato un percorso di sintesi tra **ricerca artistica**, **progettazione tecnica** e **riflessione teorica**, dove la tecnologia non è stata impiegata solo come un mezzo, ma è diventata materia viva di relazione, capace di generare esperienze condivise e significative.

Il progetto dimostra come la realtà virtuale possa essere percepita non semplicemente come uno spazio di simulazione, ma anche come un luogo di incontro e partecipazione, in cui il gesto corporeo diventa principio generativo di senso, memoria e presenza.

In questa prospettiva, *SHARED VR* non è soltanto un esperimento tecnico o un prototipo artistico, ma un atto di continuità con la visione di Michelangelo Lupone: un invito a vivere l'arte come processo partecipativo, aperto e dinamico. Le prospettive future avranno l'obiettivo di rafforzare la **dimensione collaborativa** e **performativa** del sistema, completandone la struttura tecnica e approfondendo le possibilità espressive legate alla co-presenza multiutente.

L'opera prosegue così il proprio percorso di evoluzione, mantenendo come fulcro il dialogo tra **corpo**, **suono** e **spazio virtuale**, elementi che ne definiscono l'identità poetica e concettuale.

Ringraziamenti

Al nonno Paolo, il nonno più bello del mondo. Avrei voluto che fossi qui più di ogni altra cosa. A te, che sei stato il primo ad aver creduto in questo percorso chiamandomi sin da subito *"ingegnere"*. Grazie per i tuoi sacrifici e per i tuoi insegnamenti, per tutte le volte che mi hai difeso a spada tratta da ogni bocciatura, da ogni 18. Se oggi il tuo ingegnere ce l'ha fatta è solo merito tuo.

Alla nonna Ada, che ogni sera in videochiamata ha ascoltato con pazienza ogni mio delirio, dalla chimica alla tesi, e non ha mai smesso di chiedermi *"Come è andata oggi? Che hai fatto?"*. Grazie perchè, pur non conoscendo VR e informatica, ti sei sforzata di capire ogni errore di Unity che ti ho raccontato. Le mani, nonna, adesso funzionano! Grazie per aver sempre una risposta per tutto.

Alla nonna Tilde che ogni sera mi ha tenuto informata sul meteo di casa e mi ha sempre fatto trovare i tenerumi al mio ritorno. Grazie per la chiamata della sera e darmi spunti per la cena.

Alla mia famiglia, un bellissimo **Ca(va)sino**.

A mamma e papà e al loro instancabile sostegno.

Papà, che ha sempre la valigia pronta, o meglio lo zaino, per accompagnarmi in ogni nuova città. Se domani mi trasferissi in capo al mondo, so che verresti comunque, a riempirmi casa di casse d'acqua come se dovessi affrontare una quarantena.

Mamma, che con i suoi silenzi, dice più di mille parole. Grazie per gli abbracci che sanno di casa, la pasta con le zucchine lunghe e per ascoltare tutte le mie spiegazioni scientifiche non richieste. Senza di voi questo non sarebbe stato possibile, grazie per aver sempre provato a capire questa figlia strana, nomade e imprevedibile.

A Sarah e Ada, le mie sorelle simpatiche. Grazie per aver sopportato tutti i miei *"non posso giocare a the Sim, devo studiare"* in questi 5 anni. Ada, la mia *sorellonaonaona*, che è stata la nostra madre fuorisede, grazie per aver letto questa tesi 179 volte e per i deliveroo da Novara. Sarah, **LA** sorella simpatica, la duchessa, che è sempre l'uragano di casa, grazie per i pacchetti di the Sim, i concerti improvvisati e le risate sincronizzate.

A Marlena e Nebula, le mie sorelle pelose. Marlena, la biondona spelacchiata di casa, la mia coinquilina con eccellenza, che mi è sempre stata accanto, sin dalle lezioni in DAD di Analisi 1. Anche se nell'ultimo anno non ho potuto tenerti con me, grazie per la tua intelligenza emotiva, la tua empatia, le tue coccole, la tua tontaggine, la tua zampetta pietosa che mi ricorda che è ora di chiudere il computer perché devi mangiare giustamente. Nebula, l'ultima arrivata di casa, grazie per gli agguati, per tutte le volte che ti fai strapazzare, per i tuoi bacetti e per le tue cadute da letti e divani.

Ad Alex, mio fratello. Ale come hai detto anche tu, forse 6 pagine di ringraziamenti sono troppe anche se te le meriteresti tutte. Il destino a volte è strano, anche se siamo legati da 3 anni è come se lo fossimo stati da tutta la vita. Grazie per aver reso stupendo questo percorso. E, anche se sei scappato in Francia pur di non laurearti con me, questo traguardo è anche un po' tuo. Perché senza le risate, gli abbracci, le discussioni perché "*nono non è così... ah sì*" e "*sisisi esatto... ah no aspetta*", l'ansia preesame, il vino del pam, i vini più costosi, l'hummus, il pesto di rucola, le cene di fine esame, le canzoni cantate a squarciagola a casa e in macchina, "*Massimo Ranieri*", niente sarebbe stato ugualmente memorabile. Quindi questo traguardo lo dedico a noi, scemi ambiziosi autocritici, e ai nostri futuri traguardi insieme. Grazie per essere stato ed essere ancora parte della mia vita.

A Giorgia, *the moon of my life*, che da 10 anni mi supporta e mi sopporta. Non avrò imparato a tradurre dal greco (*le sirene avevano la forma di cosce di pollo*), ma grazie a te ho imparato che la distanza non conta quando un legame è vero. Grazie per non aver mai smesso di credere in me, per dirmi quando sono *tutta cretina*, per supportare i miei sicilianismi e per tutti i momenti insieme. Sono grata di poterti chiamare amica. So che ovunque andrò tu sarai sempre presente, anche a chilometri di distanza.

Alla famiglia di SHARED, una *famiglia di disperati* ma pur sempre una famiglia.

A Carlo, bandito gitano, che è stato e continua ad essere un punto di riferimento. Da te ho imparato e continuo a imparare molto. Grazie per tutti i consigli fraterni e per tutte le volte che mi hai aiutato, per la sincerità e la complicità.

A Federico e Daniel, gli amici abruzzesi. Daniel, che riesce a darmi conforto con la sua calma assoluta; Federico, che con la sua follia riesce sempre a dispensare saggezza. Grazie per tutti i passaggi e per le chiacchierate vere e profonde, per i vostri consigli e per avermi fatto sempre sentire a casa.

A Francesco che mi ha supportato sin dall'inizio di questa avventura. Grazie per la pazienza, l'ascolto, i rimproveri e tutti i tuoi insegnamenti. Lavorare con te mi ha aiutato a crescere tanto professionalmente e non solo.

Alla professoressa Mazali, una professoressa rara. Grazie per la fiducia, il sostegno e tutte le opportunità che mi ha dato, ma soprattutto per aver creduto in me sin dal primo momento. La sua guida è stata fondamentale.

Grazie, infine, a tutte le persone che ho avuto modo di conoscere in questo percorso, che hanno potuto influire sulla mia vita in modo positivo o negativo con un'occhiataccia, un saluto o un tenero sorriso. Ognuno, a modo suo, mi ha lasciato qualcosa: un ricordo, una lezione, una spinta a fare meglio. A tutti, grazie: perché anche da ciò che non resta, s'impara sempre.

Elenco Figure

[F3.2.1] Bonnie Marranca (2023, *"Telling Stories in Virtual Reality: In Conversation With Laurie Anderson"*, The MIT Press Reader. Available at: <https://thereader.mitpress.mit.edu/telling-stories-in-virtual-reality-in-conversation-with-laurie-anderson-2018/>

[F3.2.2] Duffy, C. (2023) Salon 009: *We live in an ocean of air* " Saatchi Gallery, Saatchi Gallery ". Available at: https://www.saatchigallery.com/exhibition/salon_009_we_live_in_an_ocean_of_air

[F3.2.3] teamlab borderless Tokyo, *Azabudai Hills teamLab*. Available at: <https://www.teamlab.art/e/tokyo/#modal-6>

[F5.2.1] Schema funzionamento comunicazione OSC Unity-Max.

[F6.2.1] Screenshot scena *01HandTracking* test: tracking mano e riconoscimento pose.

[F6.2.2] Componente Static Hand Gesture, XR Hand.

[F6.2.3] Screenshot *01HandTracking* composizione scena Unity.

[F6.2.4] Workflow path recognition.

[F6.2.5] Riconoscimento path specifici con Unity Sentins e CNN (movimento 2).

[F6.3.1] Screenshot scena *Uzbekistan-ArtIntegration*, cintura.

[F6.3.2] Screenshot scena *Uzbekistan-ArtIntegration*, interazione con il suono forma.

[F6.3.3] Schema UML del prototipo SHARED VR (versione Uzbekistan).

[F7.1.1] Funzionamento Server RPC, Unity Docs.

[F7.1.2] Funzionamento Client RPC, Unity Docs.

[F7.2.1] Screenshot scena *RPC_test*.

[F7.2.2] Ciclo di interazione.

[F7.2.3] Screenshot *XRI_Network_Player_Avatar VRMPAsset*.

[F7.2.4] Screenshot Inspector *XRHandPoseReplicator_HandsOnly*.

[F8.2.1] Grafico distribuzione di genere.

[F8.2.2] Grafico usabilità *interactive application*.

[F8.2.3] Grafico usabilità *VR application*.

[F8.2.4] Distribuzione delle risposte relative all'usabilità del sistema VR (Q3, Q7, Q9).

[F8.2.5] Distribuzione delle risposte relative all'usabilità del sistema VR (Q2, Q6).

[F8.2.6] Distribuzione delle risposte relative alla percezione di presenza nello spazio virtuale.

[F8.2.7] Distribuzione delle risposte relative alla percezione di *agency*.

[F8.2.8] Distribuzione delle risposte relative alla dimensione di awareness del mondo reale.

Bibliografia

- [1] ItARTS -Italy for Arts: Resources, Time and Space. Available at: https://www.mur.gov.it/sites/default/files/2024-06/Scheda%20progetto_INTAFAM00079.pdf
- [2] Ministero dell'Università e della Ricerca. (2025). Afam | Ministero dell'Università e della Ricerca. Available at: <https://www.mur.gov.it/it/aree-tematiche/afam>
- [3] Yakman, G. (2008). (PDF) STEAM Education: an overview of creating a model of integrative education. [online] ResearchGate. Available at: https://www.researchgate.net/publication/327351326_STEAM_Education_an_overview_of_creating_a_model_of_integrative_education
- [4] Maeda, J. (2013) "STEM + Art = STEAM," The STEAM+ Journal: Vol. 1: Iss. 1, Article 34. DOI: 10.5642/steam.201301.34; Available at: <https://scholarship.claremont.edu/steam/vol1/iss1/34>
- [5] Gabriele, A. (2025). CRM. [online] Crm-music.it. Available at: <https://crm-music.it/index.php/crm/chi-siamo>
- [6] Gioia, C.; Strada, F.; Cavasino, A.; Sgarzi, F.; Bottino, A.; Mazali, T. (2025). *Beyond Mimesis: Extensive VR Design in the SHARED Case*. Intervento presentato al convegno 14th EAI International Conference: ArtsIT, Interactivity & Game Creation tenutosi a Dubai (Emirati Arabi Uniti) nel 7-9 Novembre 2025. <https://hdl.handle.net/11583/3003242>
- [7] Borgdorff, H.: *The Conflict of the Faculties: Perspectives on Artistic Research and Academia*. Leiden University Press, Leiden (2012).
- [8] Nelson, R.: *Practice as Research in the Arts: Principles, Protocols, Pedagogies, Resistances*. Palgrave Macmillan, London (2013).
- [9] GDD Shared VR.
- [10] Bekele, M. K., Pierdicca, R., Frontoni, E., Malinverni, E. S., & Gain, J. (2018). A survey of augmented, virtual, and mixed Journal on Computing and reality for cultural heritage. *Cultural Heritage*, 11(2), Article 7. DOI: <https://doi.org/10.1145/3145534>
- [11] Carrozzino, M., & Bergamasco, M. (2010). Beyond virtual museums: Experiencing immersive virtual reality in real museums. *Journal of Cultural Heritage*, 11(4), 452-458. DOI: <https://doi.org/10.1016/j.culher.2010.04.001>
- [12] Bolter, J. D., & Grusin, R. (1999). *Remediation: Understanding new media*. MIT Press. <https://mitpress.mit.edu/9780262024525/>
- [13] L. Manovich, *The Language of New Media* (2001). MIT Press. <https://mitpress.mit.edu/9780262632553/>

- [14] Addison, A. C. (2001). Virtual heritage: Technology in the service of culture. In D. Arnold, A. Chalmers, & D. Fellner (Eds.), *VAST 2001 – Virtual Reality, Archaeology and Cultural Heritage* (pp. 343–354). ACM. DOI: <https://doi.org/10.1145/584993.585055>
- [15] Uyan Dur, B. I. (2021), Virtual reality art and immersive experimental typography. *Electronic Journal of New Media* 5(3):219–233. DOI: https://doi.org/10.17932/IAU.EJNM.25480200.2021/ejnm_v5i3002
- [16] Laurel, B. (1995). Virtual Reality. *Scientific American*, 273(3), 90–90. Retrieved March 8, 2021, from <http://www.jstor.org/stable/24981732>
- [17] McLuhan, M. (1964). *Understanding media: The extensions of man*. McGraw-Hill. New York: McGraw Hill.
- [18] Lanier, J. (2017). *Dawn of the new everything: encounters with reality and virtual reality*. New York, NY: Henry Holt and Company.
- [19] Anderson, L., & Huang, H.-C. (2017). *Chalkroom* [VR artwork]. HTC Vive Arts. <https://vivearts.com/works/chalkroom>
- [20] Ambrose, G. and Harris, P. (2015). *Typography*. Lausanne, Switzerland: AVA Publishing SA.
- [21] Anderson, L. (2017) Interview: A Virtual Reality of Stories [Video file]. Louisiana Channel. <https://www.youtube.com/watch?v=zHT016FbR30>
- [22] Marshmallow Laser Feast. (2018). *We Live in an Ocean of Air* [VR installation]. Saatchi Gallery, London. https://www.saatchigallery.com/exhibition/salon_009_we_live_in_an_ocean_of_air
- [23] TeamLab Borderless TOKYO, Azabudai Hills. teamLab. <https://www.teamlab.art/it/e/tokyo>
- [24] Benford, S., Greenhalgh, C., Reynard, G., et al. (2001). *Understanding and constructing shared spaces with mixed-reality boundaries*. *ACM Transactions on Computer-Human Interaction*, 5(3), 185–223. DOI: <https://doi.org/10.1145/502432.502436>
- [25] Benford S., Bowers J., Fahlen L. E., Mariani J., and Rodden T. (1994). *Supporting co-operative work in virtual environments*. *Comput.*
- [26] Milgram P. and Kishino, F. (1994). *A taxonomy of mixed reality visual displays*. *IEICE*
- [27] Manghisi, V. M., Uva, A. E., Fiorentino, M., Gattullo, M., Boccaccio, A., & Monno, G. (2018). *Enhancing user engagement through the user-centric design of a mid-air gesture-based interface* *Journal of Cultural Heritage*, 32, for the navigation of virtual tours in cultural heritage expositions. 186–197. DOI: <https://doi.org/10.1016/j.culher.2017.10.007>

- [28] De Boer, B. (2021). *Explaining multistability. Postphenomenology and affordances of technologies*. *AI & Society*, 38, 2267–2277. DOI: <https://doi.org/10.1007/s00146-021-01272-3>
- [29] Rosenberger R (2020a) “But, that’s not phenomenology!”: a phenomenology of discriminatory technologies. *Techné* 24:83–113. <https://doi.org/10.5840/techné2020210117>
- [30] Whyte KP (2015) *What is multistability? A theory of the keystone concept of postphenomenological research*. In: Crease RP, Friis JKBO (eds) *Technoscience and postphenomenology: The Manhattan papers*. Lexington Books, New York, pp 69–81
- [31] Gibson J (2015) [1979] *The ecological approach to visual perception*. Psychology Press, New York. Ristampa Routledge 2015: <https://www.routledge.com/9781138942530>
- [32] Rietveld E, Kiverstein J (2014) *A rich landscape of affordances*. *Ecolog Psychol* 26:325–352. <https://doi.org/10.1080/10407413.2014.958035>
- [33] Heft H (2001) *Ecological psychology in context: James Gibson, Roger Barker, and the legacy of William James’s radical empiricism*. Lawrence Publishers, London
- [34] Tolle F (2021) *Artifacts and affordances: from designed properties to possibilities for action*. *AI & Soc.* <https://doi.org/10.1007/s00146-021-01155-7>
- [35] Verbeek PP (2005) *What things do: philosophical reflections on technology, agency, and design*. University of Pennsylvania Press, Pennsylvania
- [36] Lombard, M., & Ditton, T. (1997). *At the heart of it all: The concept of presence*. *Journal of Computer-Mediated Communication*, 3(2). DOI: <https://doi.org/10.1111/j.1083-6101.1997.tb00072.x>
- [37] Merleau-Ponty, M. (1945/2012). *Phenomenology of perception* (D.A. Landes, Trans.). Routledge. DOI: <https://doi.org/10.4324/9780203720714>
- [38] Varela, F. J., Thompson, E. T., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. MIT Press. <https://mitpress.mit.edu/9780262720212/>
- [39] Hansen, M. B. N. (2006). *Bodies in code: Interfaces with digital media*. Routledge. <https://www.routledge.com/9780415970167>
- [40] Dourish, P. (2001). *Where the action is: The foundations of embodied interaction*. MIT Press. <https://mitpress.mit.edu/9780262541787/>
- [41] Loke L, Robertson T. (2013). *Moving and making strange: An embodied approach to movement-based interaction design*. *ACM Trans. Comput.-Hum. Interact.* 20, 1, Article 7. DOI: <https://doi.org/10.1145/2442106.2442113>

- [42] Steuer, J. (1992). *Defining virtual reality: Dimensions determining telepresence*. DOI:????
- [43] Slater, M. (2018). Immersion and the illusion of presence in virtual reality. *British Journal of Psychology*, 109(3), 431–433. DOI: <https://doi.org/10.1111/bjop.12305>
- [44] Unity. (2025). *OpenXR Plugin Manual*. Unity Documentation.
<https://docs.unity3d.com/Packages/com.unity.xr.openxr@1.11/manual/index.html>
- [45] Unity. (2025). *XR Hands Manual*. Unity Documentation.
<https://docs.unity3d.com/Packages/com.unity.xr.hands@1.4/manual/index.html>
- [46] Unity. (2025). *Static Hand Gesture Component*. Unity Documentation.
<https://docs.unity3d.com/Packages/com.unity.xr.hands@1.4/manual/static-hand-gestures.html>
- [47] Unity. (2025). *XR Interaction Toolkit Manual*. Unity Documentation.
<https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@3.0/manual/index.html>
- [48] Unity. (2025). *XR Origin Overview*. Unity Documentation.
<https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@3.0/manual/xr-origin.html>
- [49] Wikipedia. User Datagram Protocol. In Wikipedia, The Free Encyclopedia. Retrieved from https://it.wikipedia.org/wiki/User_Datagram_Protocol
- [50] Unity. (2025). NGO.
<https://docs.unity3d.com/Packages/com.unity.netcode.gameobjects@2.5>
- [51] Unity. (2025). Unity Transport (UTP).
<https://docs.unity3d.com/Packages/com.unity.transport@2.6>
- [52] Unity. (2025). Multiplayer Play Mode (MPPM).
<https://docs.unity3d.com/Packages/com.unity.multiplayer.playmode@0.1>
- [53] Unity. (2025). Unity Gaming Services Overview. Unity Technologies.
<https://unity.com/products/unity-gaming-services>
- [54] Unity. (2025). Relay Overview. Unity Documentation.
<https://docs.unity.com/relay/en/manual/relay-overview>
- [55] Unity. (2025). Vivox Voice Chat Overview. Unity Documentation.
<https://docs.unity.com/vivox/en/manual/vivox-overview>

- [56] Microsoft. (2025). Socket.Bind Method (System.Net.Sockets). Microsoft Learn.
<https://learn.microsoft.com/en-us/dotnet/api/system.net.sockets.socket.bind>
- [57] Brooke, J. (1986). System Usability Scale (SUS): A Quick and Dirty Usability Scale. Digital Equipment Co. Ltd, Reading, UK.
- [58] Bangor, A., Kortum, P., & Miller, J. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human–Computer Interaction*, 24(6), 574–594.
- [59] Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments*, 10(3), 266–281.