

# MOTION CAPTURE IN VIDEOGAMES

Set virtuali e attori reali nell'animazione dei videogames al  
servizio della narrazione identitaria

## RELAZIONE TECNICA

a cura di Andrea Assorgia



**REGIONE AUTÒNOMA  
DE SARDIGNA**

**REGIONE AUTONOMA  
DELLA SARDEGNA**

Con il contributo di Regione Autonoma della  
Sardegna – Assessorato regionale della Pubblica  
istruzione, beni culturali, informazione, spettacolo e  
sport.

L.R. 20 settembre 2006, n. 15, art. 16, comma 3  
Studi e ricerche e progetti di ricerca e sperimentazione  
sui nuovi linguaggi e tecnologie audiovisive

## Introduzione

La presente sezione illustra l'insieme di procedure, sperimentazioni e ripetute iterazioni che hanno portato alla realizzazione del prototipo di videogame Gli Spiriti di Terra Mala.

Il gioco si inquadra nel genere **Walking Simulator**<sup>1</sup>, una categoria di videogame narrativi nella quale lo scopo di raccontare la storia prevale sulle altre meccaniche classiche. Ne "Gli spiriti di Terra Mala", come nella maggior parte dei walking simulator, non è possibile perdere o vincere, non si affrontano nemici, né è possibile far deviare la storia dal percorso previsto. Il giocatore deve "partecipare alla storia" interagendo con l'ambiente e con gli altri personaggi non giocanti. Tali interazioni, che all'apparenza servono solo a far andare avanti la storia con i tempi e la progressione decisa dal giocatore, celano in realtà lo scopo di amplificare l'immedesimazione del giocatore. Il meccanismo di azione e diretta reazione del sistema di gioco, incluso nei dialoghi e nelle altre meccaniche di gameplay, consentono infatti al giocatore di sentirsi parte della storia, in quanto gli avvenimenti hanno luogo come diretta risposta alle proprie azioni.

In particolare il videogioco oggetto dell'attività è stato ideato con lo scopo di richiedere l'utilizzo della tecnica del Motion Capture, oggetto del progetto di ricerca. Il primo intento dell'attività è stato infatti sperimentare la corretta procedura per catturare la performance recitativa di attori in carne ed ossa e trasferirla su personaggi sintetici. Modelli 3D realizzati con la possibilità di replicare i movimenti di corpo, mani e viso di un attore in carne ed ossa.

Pure essendo una tecnica in uso nei videogiochi da almeno un decennio nei progetti di tipo AAA<sup>2</sup> (produzioni multimilionarie) la tecnica del *motion capture* non è ancora molto affermata in ambito *Indie*<sup>3</sup> (piccoli studi di produzione che contano da 2 a 20 addetti). Le motivazioni sono da ricercarsi nell'elevato costo che le attrezzature per il mocap hanno avuto fino a tempi recenti e nella disponibilità di ampie librerie di animazioni già pronte per l'acquisto a una frazione del prezzo necessario per realizzarle in proprio. Per i due motivi appena citati le produzioni indipendenti di giochi prettamente narrativi hanno sempre privilegiato meccaniche di gioco nelle quali i personaggi non erano visibili (visuale in prima persona, ambienti disabitati, comunicazioni con personaggi non giocanti solo via voce)<sup>4</sup>

Lo scopo dell'attività è stato quindi di sperimentare le corrette procedure per l'utilizzo delle tecniche di motion capture all'interno di produzioni a basso budget, definendo una *pipeline* di produzione che partisse dalla creazione di modelli di personaggi sintetici, passasse dalla registrazione delle performance degli attori in carne e ossa e unisse le due cose all'interno di un motore di gioco. Per ciascuno di questi passaggi è stata effettuata una ricognizione su caratteristiche, prezzo e compatibilità con le altre componenti, fino a giungere alla definizione della pipeline "Creazione dei personaggi → MoCap → Game Engine", che illustreremo nelle prossime pagine.

**Il videogioco prodotto è disponibile sul sito web: <https://net-press.itch.io/spiriti-di-terramala>**

1 Per il genere Walking Simulator si vedano [Gone Home](#), [Dear Esther](#) e [The vanishing of Ethan Carter](#)

2 [Studi di produzione AAA](#) su Wikipedia: [https://it.wikipedia.org/wiki/AAA\\_\(videogiochi\)](https://it.wikipedia.org/wiki/AAA_(videogiochi))

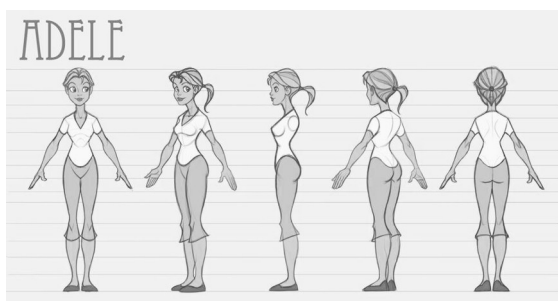
3 [Studi di produzione Indie](#) su Wikipedia: [https://it.wikipedia.org/wiki/Videogioco\\_indipendente](https://it.wikipedia.org/wiki/Videogioco_indipendente)

4 Si vedano a tale proposito i gameplay di [Gone Home](#), [Dear Esther](#) e il recente [The suicide of Rachel Foster](#)

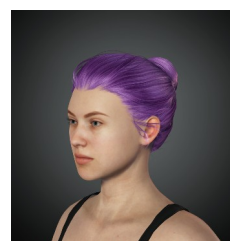
## Creazione dei personaggi

Il primo passo nella pipeline di produzione riguarda la creazione dei personaggi 3D. Il processo prevede che l'artista riceva informazioni dettagliate sul personaggio da realizzare, non solo fisiche ma anche caratteriali. Il processo creativo inizia quindi con la scrittura da parte dello sceneggiatore, che consegna a tutte le professionalità coinvolte nella produzione una *Story Bible* nella quale siano descritti i personaggi sia nell'aspetto esteriore, sia nelle back-stories, che aiutino i professionisti a inquadrare meglio il personaggio che devono realizzare.

Di norma l'artista, dopo aver esaminato il documento, crea dei bozzetti del personaggio, da presentare al producer. Se però nella creazione dei modelli vengono utilizzati dei programmi che automatizzano il processo è possibile che i bozzetti abbiano già a forma di modelli 3D.



2D Concept Character Design



3D Concept Character Design

Come appena accennato, la modellazione dei personaggi 3D rientra tra i compiti del 3D Artist. Il processo può avvenire in due modi

- **Creazione da zero con software di modellazione 3D**

Questo processo, in uso fino a poco tempo fa, è oggi prerogativa soprattutto delle grosse produzioni che vogliono garantirsi il totale controllo su quanto creato e utilizzato nella propria produzione. Prevede che l'artista 3D utilizzi un software di modellazione (Maya, Blender...) per dare volume alla reference 2D disegnata dal *Character Concept Creator* spesso una figura separata dal modellatore 3D. Dopo la modellazione dei volumi l'artista procede alla colorazione (texturing) dei volumi, alla creazione dello scheletro che dovrà muovere il personaggio e alla definizione dei morph facciali per l'animazione delle espressioni del viso. Come facilmente intuibile questo processo, che va ripetuto dall'inizio per ogni singolo personaggio, garantisce il totale controllo sul risultato finale, ma risulta molto lungo e dispendioso, motivo per il quale le produzioni con budget più limitato o che non hanno grosse esigenze di discostarsi dalla norma, si rivolgono a software più automatizzati.

- **Creazione di personaggi umanoidi attraverso sistemi automatizzati**

Nati già nel corso degli anni 2000 questi software si sono evoluti enormemente nel decennio successivo e soprattutto negli ultimi anni, fino a diventare una valida alternativa alla creazione "da zero", soprattutto nel costo e nella velocità di produzione.

Questi software offrono la possibilità di partire da un modello base (maschile o femminile) e di modificare attraverso delle manopole età, etnia, conformazione del fisico, forma e connotati del viso, fino a ottenere un personaggio umanoide totalmente personalizzato secondo le proprie esigenze. Le versioni più avanzate permettono perfino di emulare l'aspetto di persone reali, partendo da una fotografia, che viene analizzata da un software di intelligenza artificiale e trasformata in volumi 3D. Una volta terminata la modellazione del corpo è possibile aggiungere capelli, vestiti e accessori, attingendo a una vasta libreria di modelli già disponibili, oppure importando i propri.

### L'importanza degli standard

A prescindere dalla tecnica prescelta è di massima importanza che i modelli realizzati si attengano comunque a determinati standard dell'industria, per garantire l'interoperabilità tra i diversi software coinvolti nella produzione del videogioco. In particolare

- **il formato di esportazione**

Deve avvenire in in uno dei formati standard, maggiormente accettati dagli altri software. Attualmente i più diffusi sono OBJ<sup>5</sup> e FBX<sup>6</sup>.

- **Lo scheletro (rig)<sup>7</sup>**

lo scheletro del personaggio deve avere una gerarchia ben definita che parte dal bacino (hip) e si dirama verso l'alto in spina dorsale, braccio, avambraccio, mano e dita e verso il basso in coscia, gamba, piede e punta del piede. Avere un rig standard è fondamentale perché ciò che viene definito in sede di modellazione deve essere utilizzato poi dai software di MoCap e dal game engine. Attualmente quasi tutti i software convergono su uno standard con gerarchia come descritta.

- **Le espressioni del viso (blendshapes)**

Se il gioco prevede la necessità di animare anche il viso è necessario che il personaggio sia dotato preventivamente di espressioni facciali. E' possibile ottenere tale risultato attraverso una ossatura (rig) dedicata, oppure, più comunemente ed efficacemente, prevedendo delle deformazioni facciali (blendshapes<sup>8</sup>) che permettano di muovere le varie parti del viso, fino ad ottenere delle espressioni credibili. Tali deformazioni, che nella realtà prevedono l'utilizzo di svariati muscoli facciali, vengono semplificate in singole entità (es: sopracciglio sx in alto) e ricombinate dai software di MoCap per riprodurre la prova attoriale. A differenza dei punti precedenti (formato di esportazione e rig) nel caso delle blendshapes non si è ancora affermato un standard de-facto, ed ogni produttore di software implementa un proprio set di espressioni. Nei tempi recentissimi tuttavia Apple ha introdotto **Apple ARKit52**, una serie di 52 blendshapes<sup>9</sup> che iniziano ad essere riconosciute anche da altri software.

5 Il formato OBJ: [https://it.wikipedia.org/wiki/Formato\\_Wavefront\\_.OBJ](https://it.wikipedia.org/wiki/Formato_Wavefront_.OBJ)

6 Il formato FBX: <https://it.wikipedia.org/wiki/FBX>

7 Skeletal animation: [https://en.wikipedia.org/wiki/Skeletal\\_animation](https://en.wikipedia.org/wiki/Skeletal_animation)

8 Unity Blendshapes: <https://docs.unity3d.com/Manual/BlendShapes.html>

9 Apple ARKit52 blendshapes: <https://developer.apple.com/documentation/arkit/arfaceanchor/blendshapelocation>

## La scelta di un software di creazione personaggi automatizzato

Come già detto il presente progetto si prefigge la definizione di una pipeline dedicata a produzioni a basso budget, nella quale i componenti siano interpolabili. L'attenzione è quindi stata rivolta ai software di generazione di personaggi automatizzata. Una attenta ricognizione ha identificato diversi prodotti, che sono stati classificati per caratteristiche di costo, versatilità, aderenza agli standard, facilità di uso. I software esaminati sono stati i seguenti:

- **Adobe Fuse**<sup>10</sup>  
lanciato nel 2014 consente di scegliere le proporzioni di viso e corpo, partendo da standard differenziati in base a età ed etnia. Il progetto è stato acquisito da Adobe che ne ha definitivamente bloccato lo sviluppo nel 2020. E' ancora disponibile gratuitamente sulla piattaforma Steam, ma risente ormai degli anni e del mancato aggiornamento. I personaggi non sono dotati di blendshapes facciali.
- **MakeHuman**<sup>11</sup>  
Unico progetto Open Source della lista è disponibile gratuitamente sotto licenza AGPL. Dispone, tra gli altri, di morph in grado di cambiare le proporzioni del corpo anche in base a genere ed età, permettendo così di ottenere personaggi di tutte le sfumature. E' curato da una nutrita community di appassionati che mettono a disposizione numerosi accessori quali abiti, capigliature e morph speculari. Risente comunque della mancanza di scopo professionale, che ne limita disponibilità di scelta in termini di accessori. I personaggi non sono dotati di blendshapes facciali.
- **Poser**<sup>12</sup>  
Lanciato nel 1995 è il più antico progetto della lista. Permette non solo la creazione di personaggi, ma anche la possibilità di costruire set virtuali, di illuminarli e di mettere i personaggi in posa come in un set fotografico. Dispone di un buon marketplace dove è possibile acquistare capigliature, abbigliamento e accessori per i propri personaggi, ma mantiene comunque un costo abbastanza elevato per l'utilizzo del software base.
- **Daz 3D**<sup>13</sup>  
Considerato il diretto concorrente di Poser, ma nato come spin-off, soltanto cinque anni dopo ne conserva molte delle caratteristiche, compresa la possibilità di creare set virtuali che lo rendono di fatto un set fotografico virtuale. I personaggi si sono evoluti nel tempo fino alla versione Genesis 8 che garantisce la presenza di rig umanoide standard e espressioni facciali, ma non nello standard Apple Arkit52. Dispone di un grande marketplace di capigliature, abiti e accessori, ma a differenza di Poser è disponibile gratuitamente, consentendo all'utente di spendere solo per l'acquisto all'interno del marketplace.
- **Character Creator 3**<sup>14</sup>  
E' parte di una suite di programmi più orientata verso l'animazione e il motion capture che alla realizzazione di set fotografici. La creazione dei personaggi avviene come di consueto modificando le proporzioni di viso e corpo. Dispone anche esso di un vasto marketplace dove è possibile acquistare capigliature e accessori. E' l'unico della lista a disporre di blendshapes nello standard Apple Arkit52, fondamentale nella pipeline del presente progetto. Questa caratteristica, assieme all'interoperabilità con iClone ne fanno propendere la scelta per il

10 Adobe Fuse: [https://en.wikipedia.org/wiki/Adobe\\_Fuse\\_CC](https://en.wikipedia.org/wiki/Adobe_Fuse_CC)

11 MakeHuman: <https://en.wikipedia.org/wiki/MakeHuman>

12 Poser: [https://en.wikipedia.org/wiki/Poser\\_\(software\)](https://en.wikipedia.org/wiki/Poser_(software))

13 DAZ 3D: [https://en.wikipedia.org/wiki/Daz\\_3D](https://en.wikipedia.org/wiki/Daz_3D)

14 Character Creator 3: <https://en.wikipedia.org/wiki/Reallusion>

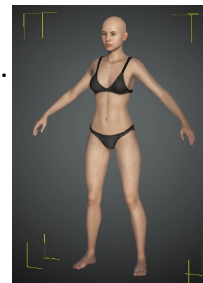
presente progetto.

### Creare personaggi con Character Creator 3

In questa sezione vediamo brevemente come creare un personaggio utilizzando Character Creator

#### 1. Importazione del personaggio base

Tutti i personaggi che nascono all'interno del software partono da un avatar. E' possibile scegliere l'avatar in base ad alcune caratteristiche tecniche, come ad esempio la presenza di blendshapes. Il personaggio base compare inizialmente senza capelli ne vestiti.



#### 2. Modificare i connotati

Attraverso i morphs disponibili sul programma è possibile modificare le proporzioni del corpo e i connotati del viso. Nell'esempio a destra una estremizzazione delle modifiche che è possibile effettuare sul modello.



#### 3. I capelli

E' possibile scegliere tra svariate capigliature, combinarne diverse per creare uno stile unico e successivamente colorare i capelli a piacimento



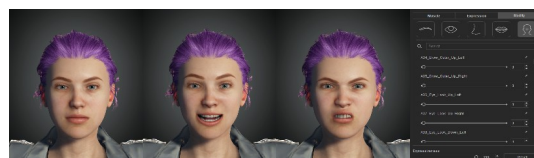
#### 4. Abbigliamento

Allo stesso modo è possibile scegliere l'abbigliamento tra una vasta libreria di capi e modificarne colorazione e motivo stampato



#### 5. Espressioni

E' necessario assicurarsi che il modello abbia impostate le espressioni facciali Apple Arkit52 e modificarle secondo il piacimento



#### 6. Esportazione

Prima di esportare è necessario impostare il modello nella T-Pose standard, in modo che la posizione del modello e dello scheletro (rig) coincidano con il rig umanoide degli altri programmi coinvolti.

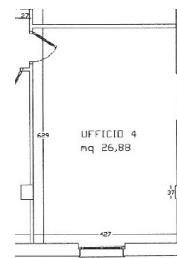


## Allestimento di una sala

anche se le tute inerziali possono essere utilizzate virtualmente ovunque si possa installare una rete WiFi è consigliabile avere una sala attrezzata per registrare le sessioni con efficienza e comodità. Un locale adatto a ospitare sessioni di MotionCapture deve avere determinate caratteristiche:

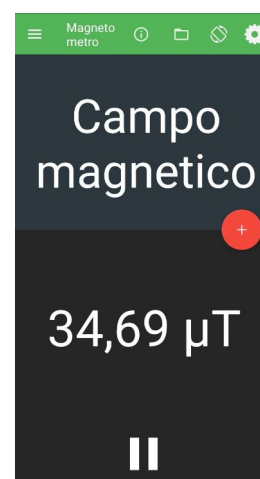
- **Essere di dimensioni medio-grandi**

i recitativi dei videogiochi narrativi prevedono infatti la presenza di almeno due attori che dialogano. Anche se nel caso del motion capture è assente la scenografia lo spazio richiesto è più o meno lo stesso di un set cinematografico. Per evitare scene troppo fisse la sala di posa deve consentire agli attori di muoversi liberamente, camminare per qualche passo e anche eseguire dei semplici stunt (cadute, salti, combattimenti ecc...).



- **Non soffrire di interferenze elettromagnetiche**

Come già accennato, la tecnologia delle tute inerziali si basa sui dati raccolti da un insieme di sensori che sfruttano il campo elettromagnetico terrestre per orientarsi nello spazio. Se il locale nel quale si intende registrare le performance soffre di eccessive interferenze elettromagnetiche, causate da un eccesso di armatura ferrosa o da apparecchiature elettriche, la qualità dei dati sarà molto povera o addirittura nulla. Per questo motivo è necessario sincerarsi prima che il locale sia adatto. Per questo progetto è stata utilizzata una app per cellulare<sup>15</sup> che fornisce in tempo reale una lettura dei microtesla rilevati in. Di norma, valori inferiori ai 40 garantiscono una ottima qualità dei dati e nessun disturbo ai sensori delle tute.



- **Essere adeguatamente insonorizzato**

Se le scene da registrare con il motion capture prevedono dei dialoghi è buona norma che vengano registrati in presa diretta, in modo da avere la massima sincronia con i movimenti del viso e una maggiore naturalezza della performance. E' tuttavia necessario che l'ambiente sia adeguatamente insonorizzato e privo di riverberi, soprattutto nelle basse frequenze nelle quali si trova la voce umana. Abbiamo svolto tali analisi attraverso una app<sup>16</sup> che ci ha fornito i dati utili a capire quali interventi erano necessari.

| RIASSUNTO DELLA MISURAZIONE |         |      |      |      |      |
|-----------------------------|---------|------|------|------|------|
| Fc [Hz]                     | 250     | 500  | 1k   | 2k   | 4k   |
| RT [s]                      | 0.11    | 0.12 | 0.12 | 0.08 | 0.09 |
| RTimed [s]                  | 0.12    |      |      |      |      |
| INFORMAZIONI STANZA         |         |      |      |      |      |
| Altezza:                    | 3.1 m   |      |      |      |      |
| Superficie pavimento:       | 27 m²   |      |      |      |      |
| Volume stimato:             | 83.7 m³ |      |      |      |      |

La sala allestita per il progetto è stata allestita con un tappeto in schiuma di neoprene dello spessore di 1cm (utile anche per gli stunts), pannelli fono-assorbenti e bass-traps e teli pesanti gree e blue screen, utili anche per sessioni di virtual production.

15 Physics Toolbox Sensor Suite : <https://play.google.com/store/apps/details?id=com.chrystianvieyra.physicstoolboxsuite>

16 APM Tool Lite: <https://play.google.com/store/apps/details?id=it.suonoevita.apmtoolite&gl=IT>

## Il Motion Capture

### I diversi tipi di Motion Capture

Come già ampiamente esposto nelle sezioni precedenti esistono diverse tecnologie per catturare i movimenti del corpo e riprodurre questi movimenti informazioni in forma digitale applicati a un modello 3D. Ciascuna tecnologia offre dei vantaggi e degli svantaggi e ovviamente ricopre una fascia di costo differente.

- **Sensori di profondità ad infrarosso**

Si tratta di una soluzione molto economica e che rappresenta l'entry level nell'industria del mocap. I dispositivi utilizzati sono in grado di percepire la luce infrarossa generata dagli stessi sensori e calcolare distanza e angolazione degli oggetti inquadrati. Di norma questi dispositivi sono progettati per scopi di intrattenimento (es: Microsoft Kinect o HTC Vive VR) e vengono impiegati nel mocap per progetti amatoriali. La qualità dei dati catturati è infatti povera e tende a precipitare verticalmente con l'introduzione di più di un attore in scena.

- **Inerziale**

Si tratta di sistemi che sfruttano una serie di sensori di movimento inerziali (IMU), simili a quelli dei telefoni cellulari, disposti strategicamente sulle varie parti del corpo. Questi sistemi combinano una buona qualità dei dati catturati a un prezzo relativamente ridotto e accessibile per produzioni indipendenti. Rispetto ai più costosi sistemi volumetrici hanno anche il vantaggio di poter essere trasportati e impiegati in qualsiasi ambiente, anche all'aperto. Di contro, in ambienti chiusi soffrono delle interferenze magnetiche generate da componenti metalliche e apparecchiature elettriche.

- **Volumetrico**

Questa tecnologia rappresenta lo stato dell'arte del MoCap. Prevede di coprire il perimetro di una stanza (il volume) con telecamere ad alta definizione e ad alta velocità di scansione. Il numero di telecamere e le loro caratteristiche può variare da 8 a 24 e anche oltre, in base alla dimensione del volume. Il mocap volumetrico non risente di interferenze elettromagnetiche, è molto preciso ed è applicabile anche a personaggi non umani (animali) o ad oggetti di scena in movimento (props). Di contro è molto più costoso dei sistemi inerziali e non è possibile trasferirlo facilmente. Per catturare i movimenti di uno scalatore o di un cavallerizzo è necessario perciò costruire una parete da arrampicata all'interno del volume o portare un cavallo all'interno del volume.

A seguito di questa prima analisi si è deciso di concentrare il resto della ricognizione sui sistemi inerziali, che offrono, lo ripetiamo, un minor costo iniziale a fronte di una maggiore versatilità e una qualità dei dati di buon livello.

### Ricognizione tra i sistemi di Motion Capture inerziali

Abbiamo quindi passato in rassegna le diverse soluzioni di MoCap inerziale, concentrandoci sui tre maggiori produttori. Anche in questo caso il mercato offre soluzioni entry-level e soluzioni al top di gamma. Su ciascuno di essi abbiamo indagato analizzando i dati forniti dai produttori e il livello di soddisfazione dei possessori anche contattandoli in privato. Il risultato è riassunto nella tabella successiva.

| Sistema IMU       | Prezzo iniziale | Prezzo annuo | Face capture | Hands capture            | Soddisfazione clienti |
|-------------------|-----------------|--------------|--------------|--------------------------|-----------------------|
| Noitom perception | € 1.300         | 0            | no           | Si, a parte<br>+\$600,00 | *                     |
| Neuron v32        |                 |              |              |                          |                       |
| Noitom PN Pro     | € 3.000,0       | 0            | No           | No                       | ***                   |
| Rokoko Smartsuit  | € 4.000,00      | € 600,00     | Si           | Si, compreso             | ***                   |
| Noitom PN Studio  | € 5.200,00      | 0            | No           | Si                       | ****                  |
| Xsense Awinda     | € 6.000,00      | € 2.300,00   | No           | No                       | *****                 |

### Rokoko Smartsuit

La scelta è ricaduta su un sistema Rokoko Smartsuit, che come si evince dalla tabella, si colloca tra le soluzioni nella fascia medio-alte e presenta numerosi punti di vantaggio rispetto alle soluzioni di costo superiore.

- un unico sistema per la cattura dei movimenti corpo, mani e viso body+hands+face
- Compatibilità con Apple ARKit52 per il facial motion capture
- Possibilità di registrare il video della sessione di mocap, al fine di avere un riferimento visuale

*E' da notare che l'azienda, pur dimostrandosi molto professionale nel servizio post-vendita, ha consegnato le due tute con estremo ritardo, soltanto tre mesi dopo l'acquisto, il 9 di giugno.*

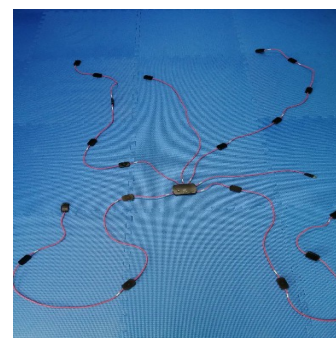


*Le due Rokoko Smartsuit acquistate*

Il prodotto consiste in una tuta di tessuto elastico e traspirante dotata di numerose tasche interne, accessibili tramite zip disposte longitudinalmente lungo gli arti e sulla schiena. Queste tasche custodiscono il sistema di sensori IMU che sono tutti connessi a un hub che funziona da centro di controllo e tramette le informazioni sui movimenti dell'attore tramite WiFi. Il reticolo dei sensori può essere estratto per lavare la tuta o per montarlo su una tuta di taglia diversa. Sulla tuta sono presenti inoltre delle fasce velcro esterne per fissare i sensori in posizione e impedire che si muovano durante la performance



*Lo schema di montaggio*



*Il reticolo dei sensori IMU*

## Il MoCap facciale

Il software Rokoko Studio consente di registrare la mimica del viso connettendosi a un dispositivo Apple dotato di riconoscimento facciale. Tale caratteristica è disponibile tra gli smartphone iPhone e iPad a partire iPhone X.

Catturare le espressioni del viso prevede pertanto la necessità di tenere il dispositivo di fronte al viso durante l'intera performance. La procedura può sembrare facile quando si recita in poltrona seduti davanti a una webcam (MoCap solo viso), ma quando si intende integrare la mimica corporale con quella del viso, il processo diventa molto più complesso.

E' utile segnalare che i movimenti del capo e quelli del viso vengono registrati da Rokoko Studio su due sistemi differenti. I primi, registrati da un sensore da piazzare sul retro della testa attraverso una fascia elastica, fanno parte dell'insieme dei movimenti corporei; i secondi vengono registrati dal dispositivo Apple e inseriti su un file di animazione separato.

Mentre il sistema di Rokoko risulta perfettamente funzionante e richiede pochissimi accorgimenti dal momento in cui si indossa la tuta a quello in cui si inizia a registrare, il fissaggio in posizione del dispositivo iPhone risulta in capo all'ingegno e alla creatività dell'utente. Durante il progetto, prima di trovare quella ottimale, sono state studiate diverse soluzioni.

**Soluzione 1:** Rokoko fornisce come optional un sostegno per il dispositivo di cattura facciale. Consiste in una pettorina da indossare come uno zaino portata sul davanti. La stecca di sostegno posiziona la videocamera del telefono proprio di fronte al viso. Il problema è dato dal fatto che essendo questo sostegno collegato al petto, il telefono non ruota assieme al collo, portando così alla cattura di soltanto metà volto nel caso in cui l'attore si volti a destra o a sinistra.



**Soluzione 2:** la maggior parte degli utenti di sistemi di Motion Capture risolve il problema adattando un casco leggero (sono molto indicati quelli da skate, perchè tengono libera la fronte). Al casco viene applicato un bastone da selfie in grado di reggere il telefono. Anche questa soluzione è stata scartata in quanto il telefono attaccato a un bastone così lungo crea un peso che alla lunga diventa fastidioso per l'attore non allenato.



**Soluzione 3:** la soluzione ottimale è stata raggiunta utilizzando un sostegno per GoPro, corto ma ricurvo in modo da risultare di fronte al viso creando una leva ricurva minore che viene comunque compensata dal peso della batteria di alimentazione degli smartgloves posizionata sul retro della testa.



## Le sessioni di registrazione

Il software Rokoko studio non presenta particolari difficoltà nell'utilizzo. E' necessario seguire alcuni step

### Operazioni preliminari

1. Approntare una rete wifi dedicata, che non venga utilizzata da altri dispositivi e possibilmente che non abbia accesso alla rete internet
2. Configurare la tuta collegando l'hub via USB e inserendo l'indirizzo del computer sul quale gira Rokoko Studio
3. Configurare allo stesso modo iPhone e smartgloves
4. Creare un nuovo profilo attore inserendo l'altezza e (se fuori standard) anche le misure degli arti. Questi dati opzionali sono necessari solo quando si vogliono simulare creature fantastiche o animali antropomorfi (es orango, gorilla, scimpanzè), utilizzando protesi modificano le normali proporzioni umane.



*Il software Rokoko Studio in fase di registrazione*

### Per ogni take

5. Verificare che la telecamera sia attiva per registrare la reference video
6. Registrare contemporaneamente l'audio su un programma separato
7. Calibrare la posizione iniziale ad ogni take
8. verificare che i personaggi o parti di essi non abbiano subito drift esagerati, difficili da correggere in fase di cleanup



*Una sessione di MoCap*

Naturalmente sono occorse numerose sessioni di prova, durante le quali il team ha collezionato una serie di *best practices* che possono arrivare solo dalla sperimentazione diretta.

Le animazioni de "Gli spiriti di Terra Mala" sono state registrate in due diverse sessioni: una per i recitativi (cutsscenes) e una per gli spezzoni che devono ripetersi a ciclo continuo fino alla successiva (idle, camminata, ecc...) le animazioni così ottenute sono state passate al software iClone per un breve cleanup, riguardante soprattutto la correzione di piccoli drift (spostamenti non voluti di) e il riposizionamento di mani e piedi in relazione agli oggetti o alle parti del corpo toccate.

Successivamente è avvenuta l'esportazione finale in formato FBX e l'importazione all'interno del Game Engine Unity.

**L'intera pipeline di produzione è stata presentata agli allievi di Game Maker Academy ed è ora parte integrante del programma del secondo anno dei corsi per Game Developer e Game Artist.**

## Creazione del videogame



L'attività ha riguardato la creazione di un videogame che facesse uso di animazioni registrate con la tecnica del Motion Capture per la valorizzazione identitaria della Sardegna.

Com'è noto, il nuraghe Diana di Quartu Sant'elena<sup>17</sup> è oggetto di leggende tramandate dalla popolazione cittadina, che riguardano un tesoro custodito da alcuni spiriti. Secondo le voci circolanti, alcuni hanno addirittura perso la vita nel tentativo di mettersi in contatto con questi spiriti. Il videogioco realizzato per il progetto prende spunto da questa leggenda realmente circolante per raccontare una storia di fantasia.

Si è scelto quindi di realizzare un breve gioco del **genere Walking Simulator per PC Windows** nel quale due giovanissime ragazze compiono una bravata, rubando l'auto dai genitori per recarsi nottetempo al nuraghe Diana. Giunte sul posto raccolgono alcune pietre a formare un pentacolo, accendono delle candele e contattano gli spiriti, che magicamente rispondono.

Il gioco si usa in prima persona (il personaggio giocatore non è visibile). Ci si muove utilizzando i consueti **comandi WASD** per i movimenti nelle quattro direzioni, **più il mouse** per spostare lo sguardo, il pulsante sinistro per interagire e il destro per uscire dalla modalità cinematografica e riprendere il controllo del proprio personaggio.

Il gioco contiene tutti gli elementi necessari a sperimentare la pipeline oggetto della ricerca

- creazione di personaggi umani in 3D
- registrazione di performance mimica e voce
- assemblaggio in un videogioco
- trattazione di argomenti identitari quali un sito archeologico e una leggenda locale

## Preparazione della scenografia

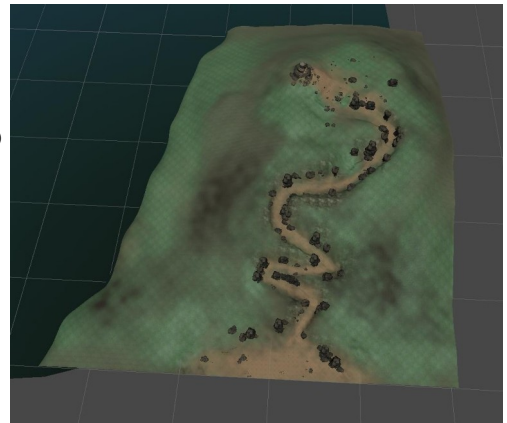
Il primo passo è stato la preparazione del terreno che potesse ospitare il videogame. Il gameplay prevede che i personaggi facciano un breve percorso lungo il sentiero prima di arrivare alla radura dove si terrà la seduta spiritica. Questo espediente, molto comune nei Walking Simulator, serve a introdurre i personaggi e fare la premessa alla storia che vedrà il suo vivo nella seduta spiritica.

17 Nuraghe Diana – Quartu Sant'Elena (CA) - [https://it.wikipedia.org/wiki/Nuraghe\\_Diana](https://it.wikipedia.org/wiki/Nuraghe_Diana)

## Attività svolte

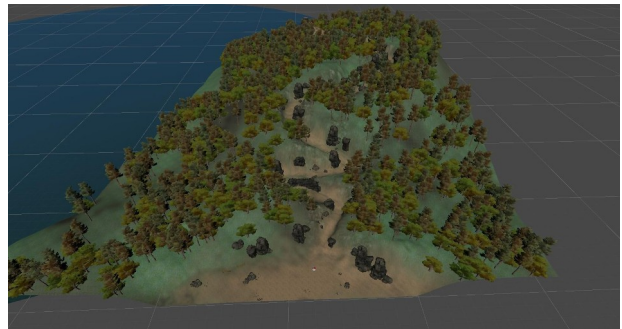
- **Preparazione del terreno su Unity**

E' stato realizzato un terreno di 500m per lato, modellato come una collina, in modo che i personaggi possano compiere l'ascesa verso la radura. Si noti il sentiero che parte dalla parte inferiore dell'immagine e che i personaggi devono fino ad arrivare alla radura più in alto dove si trova il nuraghe.



- **Aggiunta vegetazione**

Il terreno è stato successivamente popolato di vegetazione, alberi e cespugli, al fine di arricchire la scena, limitare i movimenti del giocatore e occludere lo sguardo, celando quindi la ristretta dimensione del terreno.



- **Modellazione del nuraghe Diana**

A parte, utilizzando un software di modellazione 3D è stata realizzata una ricostruzione del Nuraghe Diana, successivamente collocata nel terreno finora preparato, esattamente sulla sommità della collina.



- **Illuminazione ed effetti**

Gli ultimi tocchi alla fotografia sono stati dedicati all'inserimento di un cielo notturno (skybox), all'inserimento della corretta illuminazione e al bake delle luci, procedura che consente di calcolare in precedenza l'illuminazione della scena, evitando di doverlo fare in tempo reale durante l'esecuzione del gioco.



## Creazione delle meccaniche base dentro Unity

Dopo la preparazione del terreno di gioco il gruppo di lavoro ha proceduto a implementare le meccaniche base

- **Importazione dei personaggi non giocanti**

I personaggi creati durante la fase precedente sono stati finalmente importati all'interno di Unity.

Si noti che a causa di un conflitto tra il rig umanoide standard di unity e i blendshapes di Character Creator è necessario creare una *avatar mask*<sup>18</sup> per disattivare l'ossatura della mascella, in modo che possa essere comandata dai blendshapes.



- **AI dei personaggi non giocanti**

Sono state implementate le logiche di comportamento per i personaggi non giocanti.

In particolare è stato creato il **nav-mesh**<sup>19</sup>, la porzione di terreno calpestabile che deve essere considerata per il calcolo del percorso più breve. In questo modo si può impostare una origine, una destinazione e una velocità, per farli che i personaggi camminino in modo naturale verso una meta, evitando ostacoli e compiendo le deviazioni necessarie.

Sono stati inoltre creati dei **trigger-volume**: dei volumi sferici invisibili durante il runtime che percepiscono quando un personaggio (giocante o non giocante) entra nella sfera di interesse e innescano gli eventi programmati.



*Il nav-mesh (il sentiero serpeggiante) e i trigger volumes sotto forma di sfere*

- **Inserimento di un FPS controller**

per compeltare le meccaniche di gioco base (world-rules) è stato inserito un *fps-controller*<sup>20</sup>. Un personaggio invisibile che rappresenta gli occhi del giocatore e che si muove secondo i consueti comandi WASD+Mouse.

## Implementazione delle meccaniche di gioco narrative

Da ultimo sono state implementate le meccaniche di gioco (Gameplay Rules), quelle che decidono come si avanza nel gioco e come si completa, anche quando non è prevista una vittoria o una sconfitta.

- **interazione con i bot**

18 Unity Avatar Mask: <https://docs.unity3d.com/Manual/class-AvatarMask.html>

19 Unity Nav-Mesh: <https://docs.unity3d.com/Manual/nav-BuildingNavMesh.html>

20 FPS Controller: [https://it.wikipedia.org/wiki/Sparatutto\\_in\\_prima\\_persona](https://it.wikipedia.org/wiki/Sparatutto_in_prima_persona)

E' stata programmata una interazione con il personaggio non giocante di Gabry, la ragazza che ci accompagna nell'avventura spiritica. Avviene inquadrando il personaggio al centro dello schermo e utilizzando il pulsante sx del mouse. Per suggerire questa azione (peraltro molto comune nei walking simulator e più in generale in tutti i giochi controllati in prima persona (fps-controller), è stato inserito un puntino al centro dello schermo che cambia aspetto quando viene puntato un oggetto col quale si può interagire.

- **I dialoghi**

Interagendo con il personaggio non giocante si innescano sessioni di dialogo (cutscenes) che interrompono temporaneamente il gameplay tradizionale.

- I comandi di movimento sono temporaneamente sospesi;
- cambia la visuale di gioco includendo anche il personaggio del giocatore;
- Viene offerta al giocatore la possibilità di selezionare il testo della prossima battuta

I dialoghi sono stati implementati attraverso la libreria Fungus<sup>21</sup>, che facilita la visualizzazione di testi in modalità ticker e hanno la possibilità di ramificare gli esiti dei dialoghi attraverso una interfaccia visuale.

- **Missioni**

Dai dialoghi con l'unico personaggio scaturiscono delle piccole missioni che il giocatore deve compiere per far andare avanti la storia. Ad esempio deve aiutare il personaggio a raccogliere delle pietre e disporle su una forma di pentacolo e successivamente posizionare delle candele e accenderle. Solo quando queste piccole missioni vengono portate a termine la storia va avanti al capitolo successivo. Anche in questo caso la logica di avanzamento è stata implementata attraverso Fungus.

- **Identificazione dei capitoli e relativi trigger**

Il completamento delle singole missioni porta avanti la storia muovendola verso un nuovo capitolo. La storia che è stata implementata è di tipo lineare, nella quale i capitoli si succedono l'uno a l'altro nell'ordine prestabilito. Anche in questo caso è stata utilizzata una Flowchart di Fungus per definire e memorizzare lo stato di avanzamento della storia

---

21 Fungus un plugin per Unity: <https://fungusgames.com/>

## Sviluppi futuri

Non è da sottovalutare inoltre la possibilità di utilizzare la pipeline sperimentata durante questo progetto anche in altri ambiti affini all'industria videoludica ma alle quali manca la caratteristica dell'interattività.

- **Produzioni TV con set virtuali e personaggi 3D**

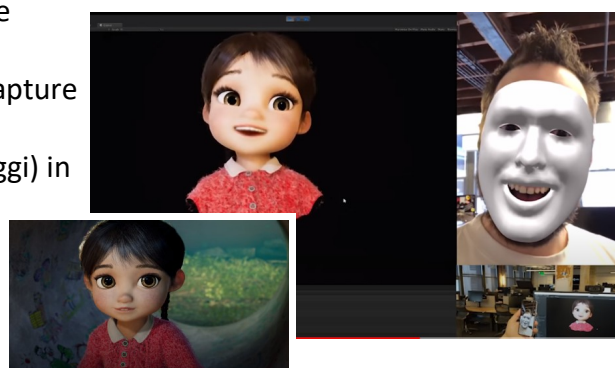
L'immagine qui illustrata mostra un momento di produzione di un programma TV per bambini, nel quale la presentatrice (a destra) è accompagnata da un simpatico coccodrillo. Quest'ultimo in realtà è animato in tempo reale da un secondo attore che indossa una tuta Rokoko. In questo caso i dati rilevati dalla tuta inerziale vengono trasmessi in tempo reale al game engine Unity, che elabora i gesti del coccodrillo e restituisce l'immagine al software di regia televisiva, che infine provvede a "bucare" lo schermo verde, inserendo così gli attori reali e virtuali all'interno di una foresta lussureggiante.

Questa produzione utilizza la medesima pipeline sperimentata durante il nostro progetto, modificandone solo l'utilizzo finale (produzione TV invece di videogiochi).

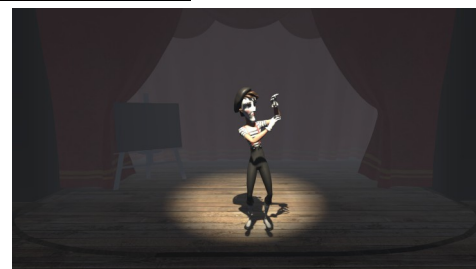


- **Cortometraggi in animazione 3D**

L'immagine qui illustrata mostra come sia possibile utilizzare il game Engine Unity3D per realizzare sequenze video di grande effetto. Con il Motion Capture è possibile inoltre sveltire la fase di animazione, realizzando cortometraggi (ma anche lungometraggi) in animazione 3D per una frazione del costo solitamente necessario per le produzioni realizzate con animazione tradizionale.



**Durante le sperimentazioni del presente progetto Net Press ha realizzato un piccolo corto dimostrativo<sup>22</sup> di questa possibilità, presentandolo il 19 agosto scorso al Bisus Film Fest<sup>23</sup> di Teulada e illustrando il percorso che ha portato a definire la pipeline.**



*Dreaming mime: il corto dimostrativo prodotto da Net Press*

<sup>22</sup> Dreaming Mime corto dimostrativo: <https://www.facebook.com/gamemakeracademy/videos/561971211808147>

<sup>23</sup> Bisus Film Fest: <https://www.facebook.com/events/1223495944835833/>

## Conclusioni

### Risultati ottenuti

L'attività tecnica illustrata era rivolta alla sperimentazione del Motion Capture applicato ai giochi narrativi identitari. La prima sfida era perciò rivolta alla definizione di una pipeline a basso budget per la produzione di videogame narrativi. Come accennato infatti il motion capture e la realizzazione di cutscene recitative è rimasta per lungo tempo ad esclusivo appannaggio delle grandi produzioni AAA, a causa degli alti costi del MoCap volumetrico. Scopo dell'attività del nostro gruppo di lavoro tecnico è stato perciò la definizione una pipeline di produzione con budget accessibile agli studi indipendenti, come quelli che operano in Sardegna e nel resto d'Italia.

- **La pipeline Character Creator 3 → Rokoko → Unity**

Attraverso numerose sperimentazioni, tentativi e ripetuti perfezionamenti il nostro gruppo di lavoro ha definito, descrivendola in questo documento, una procedura per la creazione a basso budget di cutscenes recitative da inserire nei videogiochi. Partendo dalla **creazione di personaggi 3D** con Character Creator 3, animandoli con l'utilizzo dei **sistemi di Motion Capture** Rokoko e finalizzando le prime due fasi all'interno del **Game Engine** Unity è stato possibile creare cutscenes recitative con una piccola frazione del budget solitamente destinato a queste attività dalle grandi produzioni.

- **Una nuova via per i videogiochi identitari**

i videogame nati per la valorizzazione turistica portano tutti una grande necessità di contenuti narrativo, senza il quale non sarebbe possibile veicolare sufficienti informazioni. Se si osserva il panorama dei giochi prodotti con l'intento di valorizzare il territorio, si nota che la maggior parte delle produzioni, riguardano giochi narrativi in due dimensioni 2D. Ne sono un esempio il celebre "Father and Son" del museo Archeologico di Napoli<sup>24</sup>, ormai considerato un caso di studio, ma anche i più recenti Beyond Our Lives<sup>25</sup>, dedicato agli etruschi, ma perfino "The Secret island: a sardinina Journey"<sup>26</sup> pubblicato recentemente. Tutti questi esempi, ma ce ne sarebbero tanti altri, hanno in comune la caratteristica di essere videogiochi narrativi in 2D. La pipeline di produzione di animazioni 2D è infatti molto più consolidata, collaudata ed economica.

Di contro, grosse produzioni videoludiche quali ad esempio la serie di Tomb Raider<sup>27</sup> che ha per protagonista la famosa archeologa Lara Croft, ma soprattutto l'intera serie di Assassin's Creed<sup>28</sup>, dimostrano che videogiochi in 3D che ricostruiscono accuratamente le ambientazioni storiche, raccontandole nella trama attraverso numerose cinematiche, riescono ad attrarre attenzione e a veicolare il turismo, come nel caso di Monteriggioni<sup>29</sup>.

**Il presente progetto porta perciò nuova linfa al genere dei videogiochi identitari, consentendo di inserire all'interno delle opere videoludiche, non realizzate con budget milionari, cutscenes animate e cinematiche in 3D, realizzate con la tecnica del Motion Capture.**

24 Father And Son, Museo Archeologico di Napoli: <https://mannapoli.it/father-and-son-the-game/>

25 Beyond Our Lives: <https://www.tuomuseo.it/beyond-our-lives-disponibile-il-video-gioco-dedicato-agli-etruschi/>

26 The Secret Island... : <https://play.google.com/store/apps/details?id=it.tinkerstudio.thesecretisland&gl=IT>

27 Tomb Raider: [https://it.wikipedia.org/wiki/Tomb\\_Raider\\_\(serie\)](https://it.wikipedia.org/wiki/Tomb_Raider_(serie))

28 Assassin's Creed: [https://it.wikipedia.org/wiki/Assassin%27s\\_Creed\\_\(serie\)](https://it.wikipedia.org/wiki/Assassin%27s_Creed_(serie))

29 Monteriggioni e Assassin's Creed: <https://ivipro.it/it/speciale/monteriggioni-e-assassins-creed-ii-otto-anni-dopo/>