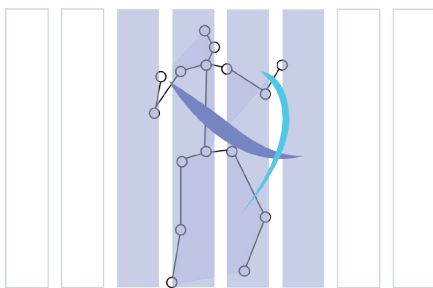


## 08 Rhapsody

**Rhapsody** [1] è un applicativo software volto a creare una rappresentazione audiovisiva dell'espressività gestuale di una persona. Pensato come una risorsa per potenziare uno spettacolo di danza dal vivo, consiste nell'utilizzare i movimenti del performer per generare in tempo reale tracce sonore e visive, facendo dialogare tecniche diverse quali l'analisi del movimento e la sonificazione interattiva [2].

**Siria Gasperoni**



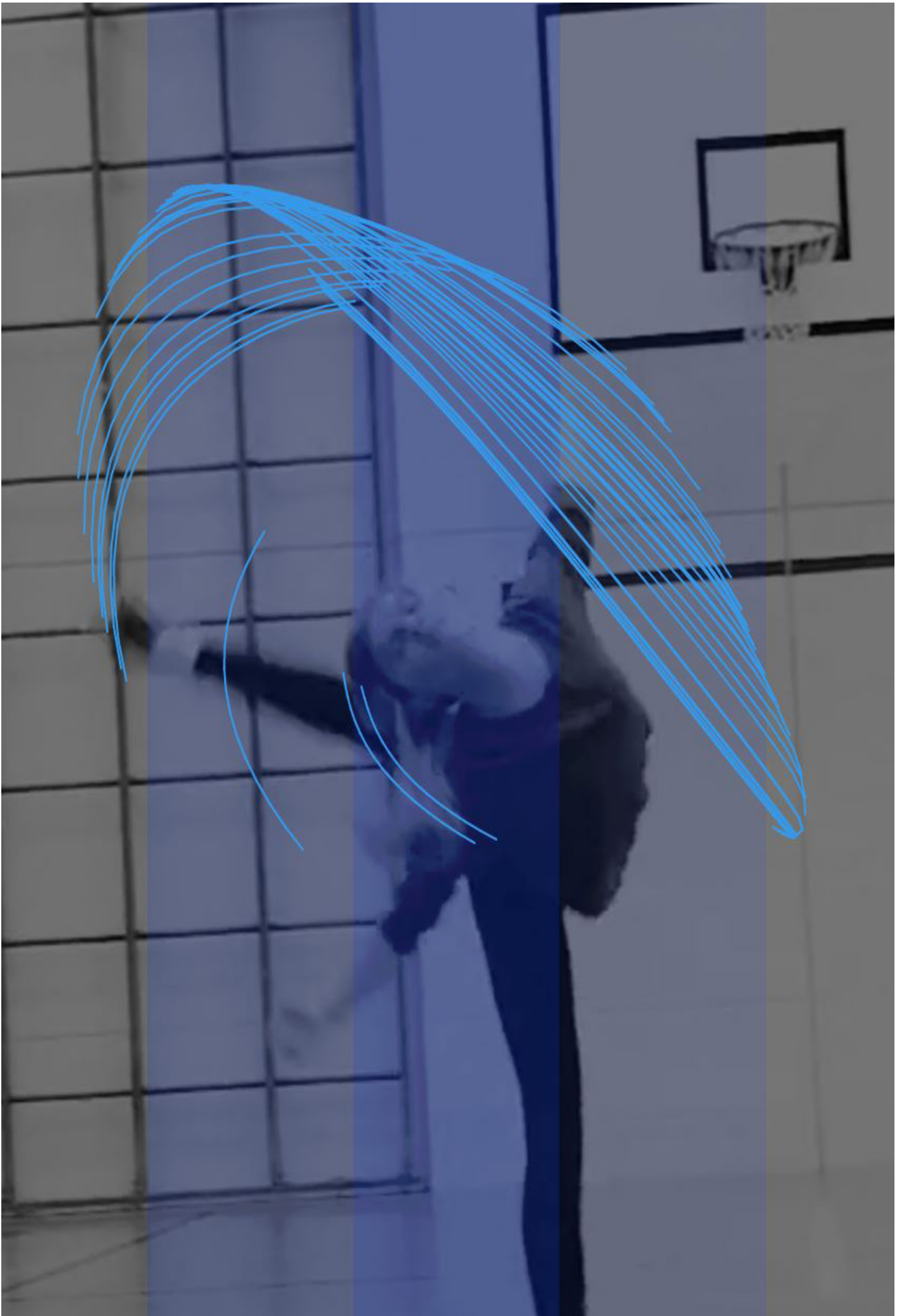
# Interactive sonification  
# PoseNet  
# Live performance  
# Tone.js  
# Musical human computer interaction

[github.com/ds-2021-unirms](https://github.com/ds-2021-unirms)  
[github.com/siriagas](https://github.com/siriagas)

[1] Una rapsodia è una composizione dalla struttura fluida, caratterizzata da una gamma di stati d'animo, colori e tonalità altamente contrastanti. Un'area di ispirazione spontanea e un senso di improvvisazione la rendono più libera nella forma di un insieme di variazioni.  
<http://www.carocci.it/lavia/capitoli.pdf>

[2] Il concetto di sonificazione, nato nell'ambito della HCI, si fonda sull'idea di utilizzare il suono non verbale come mezzo di rappresentazione dell'informazione.  
<https://conts.it/art/eggs-project/progetto-eggs-light-per-sito.pdf>

**a destra**  
Immagine elaborata dal prototipo



## Making visible

Come ascoltare e rendere visibili le idee che nascono dal corpo? Il linguaggio poetico del movimento trova la sua sintassi grammaticale nel corpo, nello spazio, nel ritmo, nelle qualità dinamiche, nelle azioni. Il corpo ha da sempre assunto un ruolo di primaria importanza nella conoscenza e nella comprensione del proprio essere, costituendo la base materiale e sociale della nostra esistenza e il luogo dove emerge la nostra soggettività. In psicologia, ormai da tempo si parla del fenomeno definito *embodiment* che vede il corpo come il luogo nel quale sono situati i pensieri, le emozioni e gli aspetti cognitivi di ogni individuo. [3] Recentemente, anche l'ambito della ricerca sembra declinare la questione della relazione tra mediazione tecnologica e produzione sonora a partire da tematiche legate alla percezione, al movimento ed alla corporeità. Esempio chiave è la *embodied music cognition theory* [4], secondo Marc Leman [5], l'esperienza musicale è innanzitutto un'esperienza incorporata, in quanto non è costituita unicamente da sensazioni auditive ma in egual misura da stimoli visivi, senso del movimento e sensazioni tattili. In quest'ottica il progetto intende approfondire conoscenza e rapporto con il suono, con l'obiettivo di valorizzare le potenzialità espressive del corpo e le sue gestualità.

## Il corpo e il suono nell'interazione uomo macchina

Il progetto Rhapsody nasce nel contesto di quel filone di ricerca che studia le relazioni tra gesto e suono alla luce delle nuove possibilità offerte dalle tecnologie digitali in ambito musicale. L'idea fondamentale è quella di riuscire a far percepire la tecnologia come un'estensione del proprio corpo al fine di poterla controllare in modo intuitivo così come avviene nell'ambito della tecnologia meccanica quando, per esempio, si guida un'automobile o si suona uno strumento musicale. Questo obiettivo presuppone uno sforzo di concezione, sviluppo e sperimentazione

[3] Embodiment  
<https://www.eduxo.it/2020/09/21/embodiment/>

[4] Questa ricerca sui gesti musicali è stata piuttosto influente in quanto pone maggiore enfasi sul feedback e sull'integrazione sensomotoria, nonché sull'accoppiamento tra percezione e azione.  
[https://en.wikipedia.org/wiki/Embodied\\_music\\_cognition](https://en.wikipedia.org/wiki/Embodied_music_cognition)

[5] Marc Leman è professore di ricerca di in Musicologia sistematica, la sua ricerca si basa sul campo dell'*embodied music cognition theory*.  
<https://research.flw.ugent.be/en/marc.leman>

**a destra**  
Future you, Motion-capture artwork reflecting your future self, London, 2019



di interfacce funzionali, ergonomiche ed intuitive, così come lo sono il volante e i pedali dell'acceleratore e del freno per la guida di un'automobile. Nel momento in cui il computer diventa strumento per la produzione di suoni, il problema diventa quello di progettare e realizzare delle interfacce efficaci per l'esecuzione musicale al pari di uno strumento tradizionale. [6]

In quest'ottica il rapporto tra il campo più ampio e generale dell'HCI [7] è il settore specifico del NIME [8]. Interessante è il punto di vista di un designer dell'interazione Bill Verplank, da anni attivo nella ricerca legata al NIME, per Verplank la musica e le nuove tecnologie costituiscono un luogo formidabile di invenzione e sperimentazione di nuove interfacce costruite ad hoc per la musica, che possono altresì fungere da ispirazione e metafora per lo sviluppo di interfacce da utilizzare in ambiti completamente diversi e non solo artistici. Il progetto Rhapsody, pur nascendo in un contesto di ricerca legato al NIME e pur essendo rivolto a fini artistici, si colloca in uno spazio intermedio tra musica, da una parte, e "sonificazione" applicata al movimento, dall'altra.

## Casi studio

Il progetto prende ispirazione da casi studio differenti ma che riflettono in particolare sull'importanza del beneficio dell'espressione personale attraverso pratiche artistiche come il disegno, il colore, la scrittura, la danza e che indagano il rapporto delle gesture e movimenti in relazione allo spazio e al processo di sonificazione.

### Messa di voce | Golan Levin

Messa di Voce è una performance che arricchisce il dialogo, le grida e le canzoni dei protagonisti con visualizzazioni interattive in tempo reale. Alla base del progetto si trovano temi di comunicazione astratta, relazioni sinestetiche, scritture puntiformi. Il software trasforma le sfumature vocali e sonore in grafiche sofisticate.

[6] Gestualità, Suono, Rappresentazione. Ricerca e sviluppo di risorse per le multimedialità in ambito artistico.  
<https://www.academia.edu/25233965/>

[7] L'interazione uomo-computer (HCI) è un campo di studio multidisciplinare incentrato sulla progettazione della tecnologia informatica e, in particolare, sull'interazione tra esseri umani (gli utenti) e computer.  
<https://www.interaction-design.org/literature/topics/human-computer-interaction>

[8] NIME, New Interfaces for Musical Expression -Nuove interfacce per l'espressionemusicale  
[www.nime.org](http://www.nime.org)

1

"Messa di Voce"

Golan Levin, Zachary Lieberman, Jaap Blonk, and Joan La Barbara  
<http://www.tmemo.org/messa/>  
2003.

2

"Body, Movement Language"  
Bill T. Jones, Google creative lab, Manifesto, consente all'utente di creare una scia di parole che rispondono ai loro movimenti in tempo reale.  
<https://www.billtjonesai.com/manifesto/2019>.

3

"Synesthesia "

Paulina Wyr

[https://www.](https://www.aleksandraslyz.com/)

[aleksandraslyz.com/](https://www.aleksandraslyz.com/)

1

2

3





Il suo obiettivo è quello di suscitare domande sul significato del parlato, degli atti e dell'ambiente. In particolare si pone la riflessione se potessimo vedere il nostro discorso, come potrebbe apparire?[9].

## Body, Movement Language | Bill T. Jones

Body, Movement Language è una raccolta di esperimenti che è il risultato della collaborazione tra il pluripremiato coreografo Bill T. Jones e Google Creative Lab [10]. Questi esperimenti utilizzano l'intelligenza artificiale per invitare gli utenti a esplorare le possibilità creative del linguaggio e del movimento e creare nuove connessioni con il lavoro iconico di Jones, utilizzando nient'altro che un laptop.

## Synesthesia | Paulina Wyr

Synesthesia è una performance interattiva in tempo reale creata dalla sonificazione del movimento, in questo caso gli artisti utilizzano sensori di movimento. I dispositivi consentono agli utenti di acquisire tutti i cambiamenti di posizione del braccio, velocità, rotazione e tensione muscolare. Tali informazioni vengono trasmesse in tempo reale direttamente al software che converte i dati in valori sonori. Questa installazione performativa è completata da una proiezione che riflette la traiettoria del movimento nello spazio.

## Prototipo | Rhapsody

Partendo dal presupposto che non ci sono relazioni necessarie tra gesto e suono, ho sperimentato la possibilità di costruire tra essi rapporti nuovi e apportatori di significato. L'idea fondamentale è stata quella di definire un panorama audio, costituito da diversi segmenti che simulano una "tastiera virtuale", in base ai quali è stato possibile suddividere il movimento del corpo nello spazio, e assegnare a ciascun movimento uno specifico suono modulato a seconda dei parametri dei singoli segmenti sonori [11]. Per il riconoscimento delle parti del corpo ho utilizzato il modello di machine learning PoseNet [12], tale modello mi ha

[9] paper progettuale su:  
[http://tmema.org/messa/npar/messa\\_NPAR\\_2004\\_150dpi.pdf](http://tmema.org/messa/npar/messa_NPAR_2004_150dpi.pdf)

[10] Google Creative Lab è una raccolta di esperimenti che utilizzano Chrome, Android, AI, AR e altro ancora.

[11] Ogni segmento suonerà un intervallo diverso di un accordo in scala musicale, se viene rilevata una parte del corpo in quel particolare segmento, consentendo così agli utenti di creare melodie musicali nuove esplorando un set di dati derivato dalla sonificazione dei movimenti.

[12] PoseNet ci fornisce modelli pre-addestrati necessari per rilevare i gesti dell'utente. Questi modelli pre-addestrati vengono eseguiti nel nostro browser.

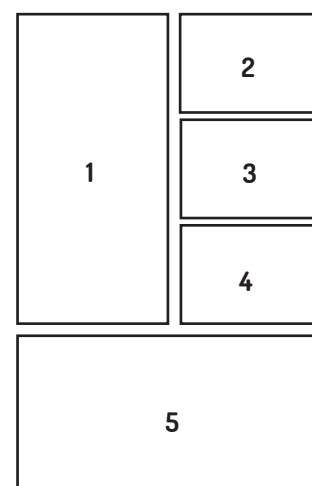
1  
Schema di funzionamento del prototipo.

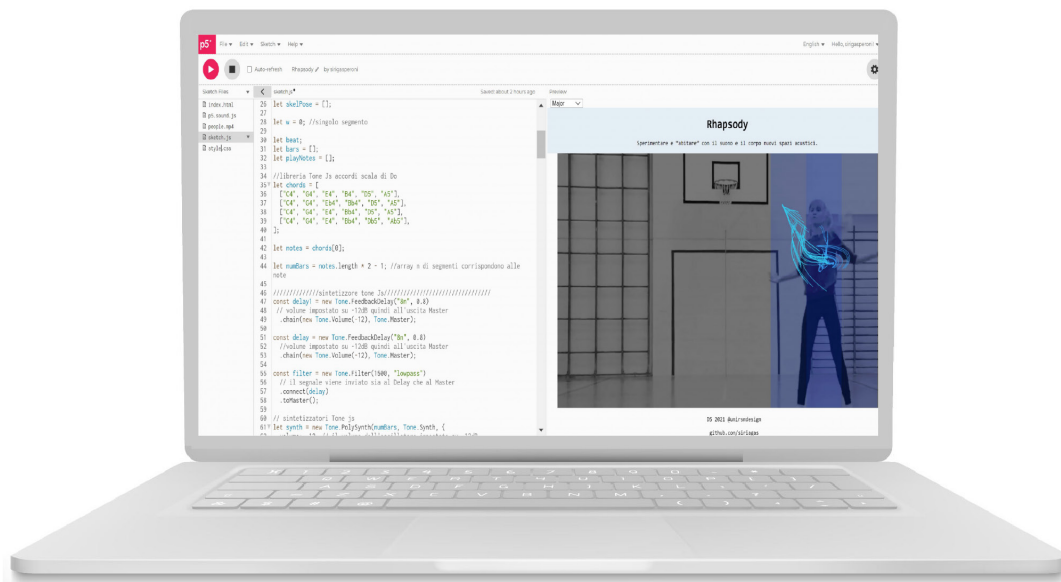
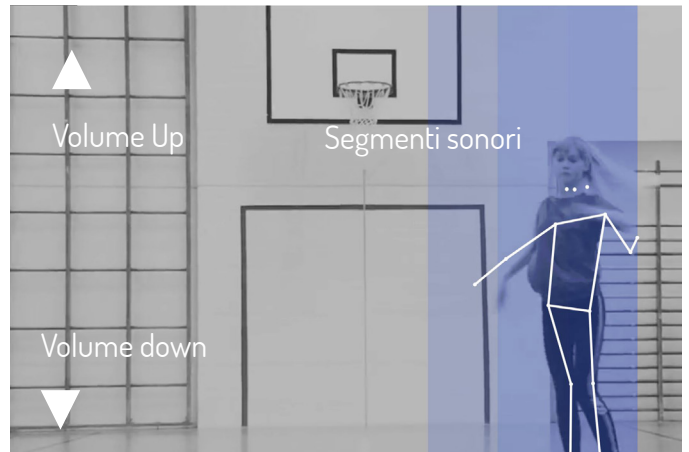
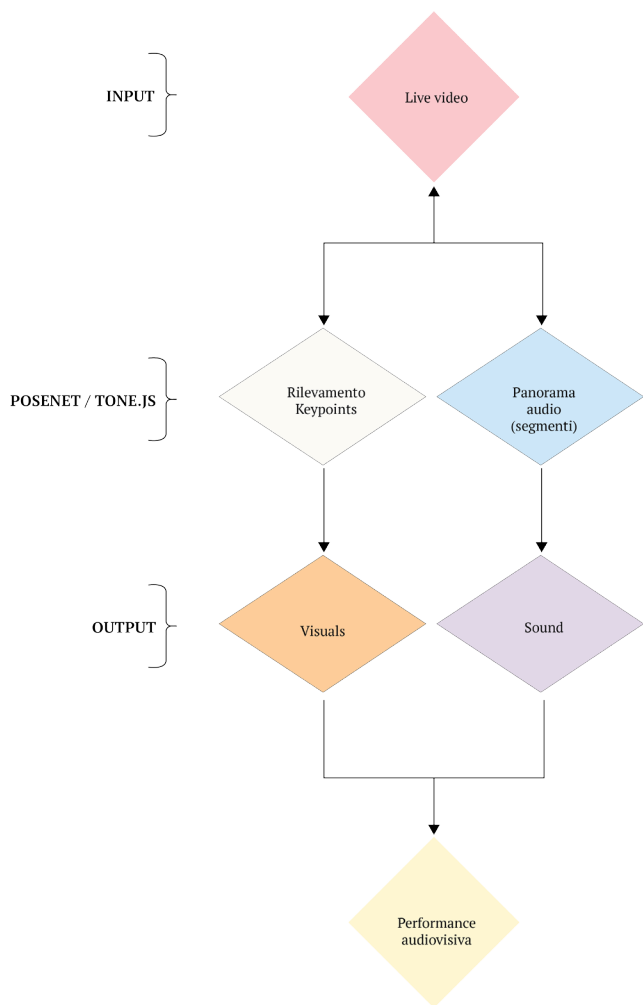
2  
Rilevamento keypoints dell'utente, utilizzando PoseNet.

3  
Attivazione dei segmenti sonori.

4  
Implementazione grafica del movimento.

5  
Focus sul codice del progetto sviluppato sull'editor p5js.







permesso di programmare utilizzando diversi punti del corpo che sono stati riconosciuti tramite un ingresso video della webcam, mentre l'output sonoro dei segmenti è generato facendo riferimento alla libreria Tone.js, un framework Web Audio per la creazione di musica interattiva nel browser. Unendo le due parti realizzate ho ottenuto il prototipo di base che permette di stimolare il performer (o utilizzatore) a controllare, adattare e esplorare la propria gestualità in base alla risposta sonora continua ottenuta in conseguenza del proprio movimento. Un altro risultato del prototipo riguarda la rappresentazione visiva dei gesti e dei movimenti. Il risultato finale è composto da una partitura sonoro-coreografica, dando così al performer la possibilità di controllare la risposta sonora e di visualizzare in tempo reale l'astrazione grafica del proprio corpo in relazione alle gestualità e ai movimenti nello spazio. Rhapsody è stato pensato in modo da non richiedere particolari abilità ovvero in modo che qualsiasi gesto semplice produca una sonificazione soddisfacente. Tale aspetto è fondamentale sia nel caso di realizzazioni di installazioni interattive che per i successivi sviluppi in altri contesti.

### Sviluppi futuri | Rhapsody

Uno dei problemi concettuali che ho riscontrato nella sonificazione dei movimenti, riguarda la gamma delle scelte possibili e ammissibili dal punto di vista di una coerenza percettiva e cognitiva, in quanto è pressoché infinita. Per ora in questo prototipo sono implementate solo parti base di sonificazione, ma è chiaro che ci sono molte modalità in cui si potrebbe rendere il processo di sonificazione percettivamente più rilevante, una possibilità potrebbe essere quella di integrare l'algoritmo sviluppato sull'editor di p5.js [13] con la patch specifica per la modulazione sonora, Max/MSP/Jitter [14]. Un ulteriore miglioramento potrebbe essere quello di riprodurre il progetto esternamente all'editor p5js utilizzando la libreria TensorFlow [15] al

[13] P5.js è una piattaforma online che costituisce appunto uno sviluppo di Processing e che integra anche il linguaggio Javascript. <https://editor.p5js.org/>

[14] Max/MSP/Jitter, è un linguaggio di programmazione visuale per musica e multimedia .

1

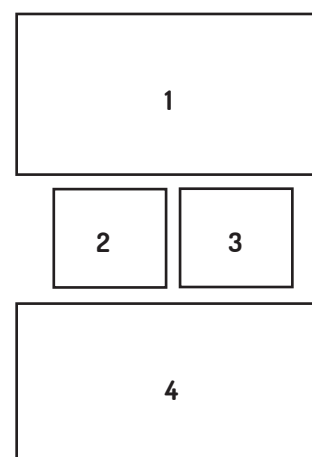
Visualizzazione a schermo intero prototipo.

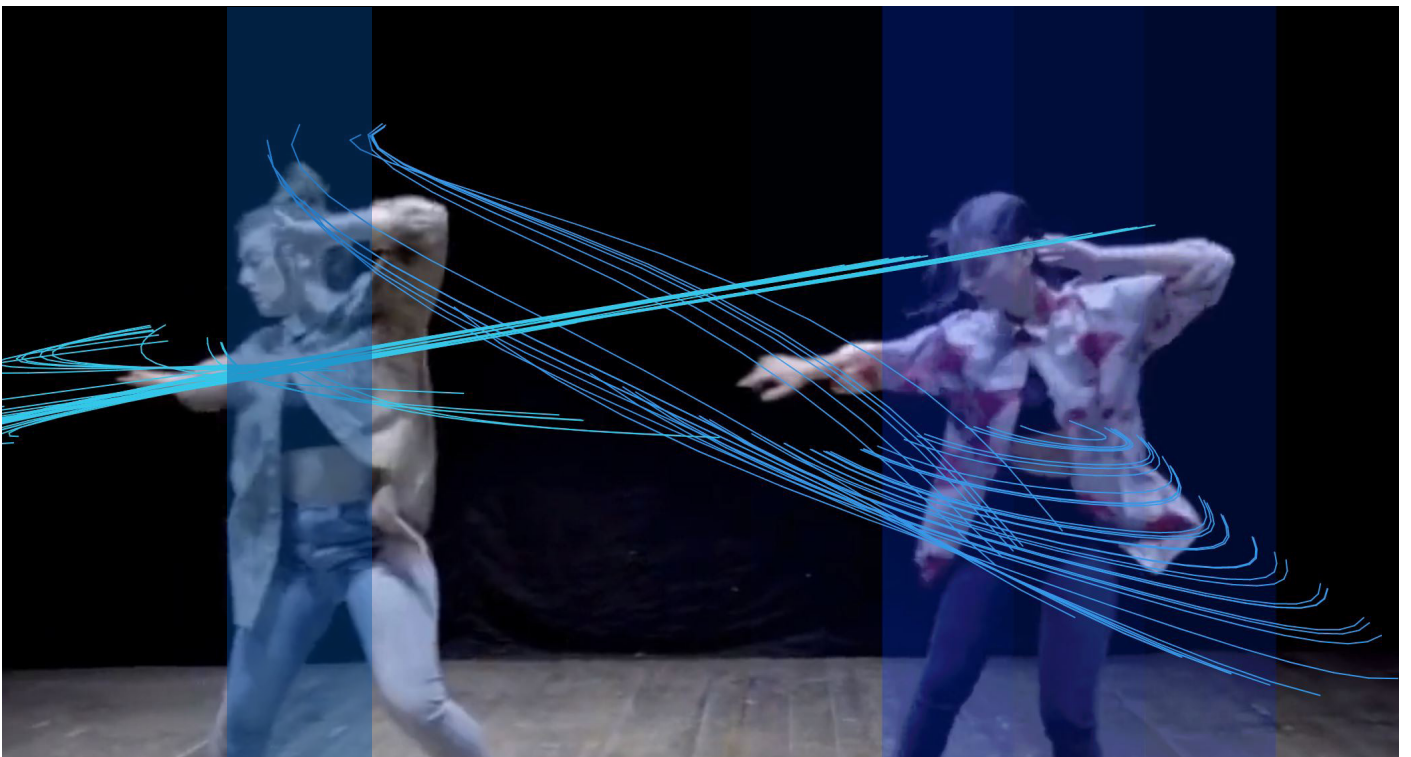
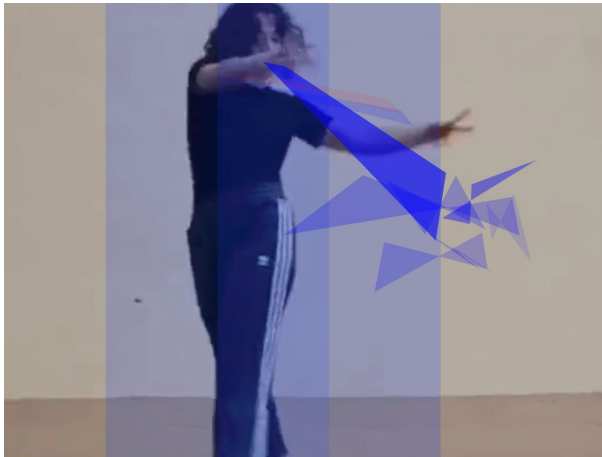
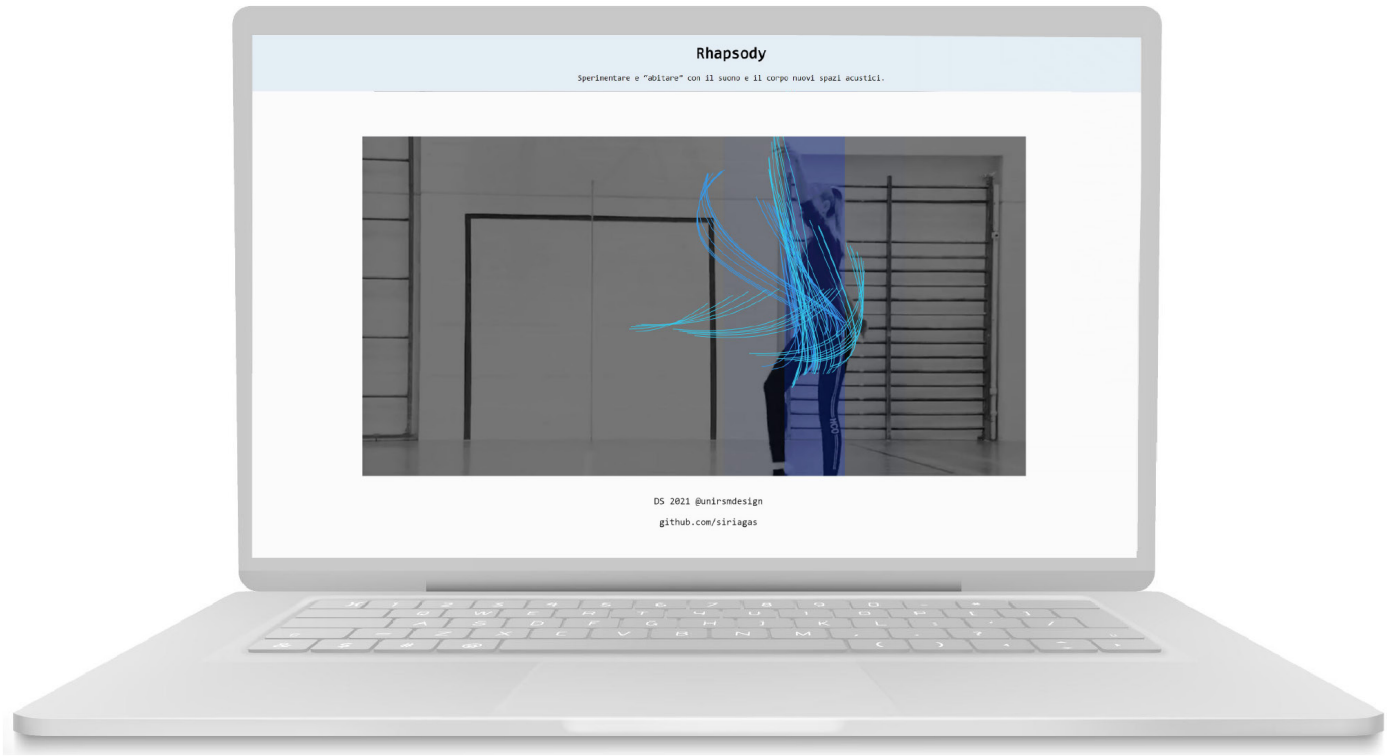
2-3

Prove utilizzando altre grafiche.

4

Prova di prototipazione utilizzando il multipose di PoseNet.





fine di ridurre la latenza nel riconoscimento dei movimenti, riscontrata soprattutto nel rilevamento di più utenti in contemporanea, in modo da rendere fluida l'esecuzione di più performer in tempo reale.

### **Sitografia**

-<http://motionsensorcomputing.com/index.php/motion-analysis/movement-sonification>

<https://medium.com/globant/posenet-your-gateway-to-gesture-detection-a15d0ed0ae40>

-[https://en.wikipedia.org/wiki/Sonic\\_interaction\\_design](https://en.wikipedia.org/wiki/Sonic_interaction_design)

-<https://blog.google/technology/ai/bill-t-jones-dance-art/>

-<https://link.springer.com/article/10.1007/s12193-020-00322-2>

-<https://experiments.withgoogle.com/billtjonesai>

-[https://www.researchgate.net/figure/Example-of-motion-descriptors-extracted-from-expressive-gestures-in-a-dance-performance\\_fig3\\_248904502](https://www.researchgate.net/figure/Example-of-motion-descriptors-extracted-from-expressive-gestures-in-a-dance-performance_fig3_248904502)

<https://www.taylorfrancis.com/chapters/edit/10.4324/9781315775845-18/music-perception-embodied-music-cognition-marc-leman-pieter-jan-maes>

<https://github.com/Tonejs>

### **Link protipo**

<https://editor.p5js.org/sirigasperoni/sketches/qAC-pf84->

<https://editor.p5js.org/sirigasperoni/sketches/ca1zbvF1E>

<https://editor.p5js.org/sirigasperoni/sketches/b6lmTdXBU>

[15] TensorFlow è una libreria software opensource. che fornisce moduli sperimentati e ottimizzati, utili nella realizzazione di algoritmi.

