

08 Spatial Being

Spatial Being è un ambiente virtuale collaborativo nel quale l'utente è svincolato dalle interfacce.

Spatial Being offre la possibilità a più persone di vedersi e interagire nello stesso spazio, tra loro e con gli oggetti, utilizzando modalità naturali di comunicazione come la voce, i gesti e il movimento del corpo.

Roberto Alesi



#humanbodyinterfaces
#virtualspaces
#naturaluserinterface
#humancomputerinteraction
#handtracking
#invisibleinterfaces

github.com/ds-2021-unirms
github.com/fupete
github.com/RobertoAlesi

a destra
render evocativo
del progetto



L'importanza dei gesti

Le condizioni di isolamento dovute alla pandemia globale [1], hanno reso gli **spazi virtuali** l'unico luogo d'incontro, connessione e condivisione.

Gli ambienti di co-progettazione che utilizziamo ogni giorno [2] sono ormai una nuova forma di interazione con le persone, tuttavia, non restituiscono la stessa esperienza.

Questo perchè usiamo sempre meno il nostro corpo per interfacciarci con le altre persone e con le cose che facciamo quotidianamente.

“Il vero problema di un'interfaccia è che è un'interfaccia.” [3]

I gesti fanno parte del linguaggio nativo degli esseri umani e, al giorno d'oggi, costituiscono una modalità sempre più importante per l'approccio uomo-computer.

Rispetto a molte interfacce esistenti, hanno il vantaggio di essere facili da usare, naturali e intuitivi. L'introduzione del linguaggio naturale [4] può creare un'esperienza utente più ricca che si sforza di comprendere il linguaggio umano e che potrebbe dare inizio a una rivoluzione tecnologica e culturale, verso una dimensione sempre più fluida e connessa agli altri.

Il progetto vuole indagare come le tecnologie, presenti nella gran parte dei nostri dispositivi, come webcam, microfoni e sensori spaziali, unite agli algoritmi di machine learning, come Eye tracking, Hand Pose, PoseNet, Speech Recognition, possano cambiare il modo in cui ci avviciniamo agli spazi virtuali e agli strumenti che utilizziamo ogni giorno, svincolati da qualsiasi interfaccia software e hardware.

La manipolazione diretta delle rappresentazioni grafiche sullo schermo, utilizzando il proprio corpo come dispositivo di controllo, porterà a una graduale smaterializzazione delle interfacce, a favore della valorizzazione dell'**esperienza**, della **presenza** e dell'**esistenza** all'interno degli spazi virtuali.

[1] L'emergenza sanitaria da COVID-19, ha avuto un enorme impatto sulla diffusione delle attività da remoto, portando a un incremento da 570 mila smart workers nel 2019 a ben 6,5 milioni nel 2020.

[2] Figma e Mirò, ad esempio, sono tra i software per eccellenza di co-design in tempo reale, utilizzabili sia da browser, sia tramite app.

[3] Donald Norman (1990)

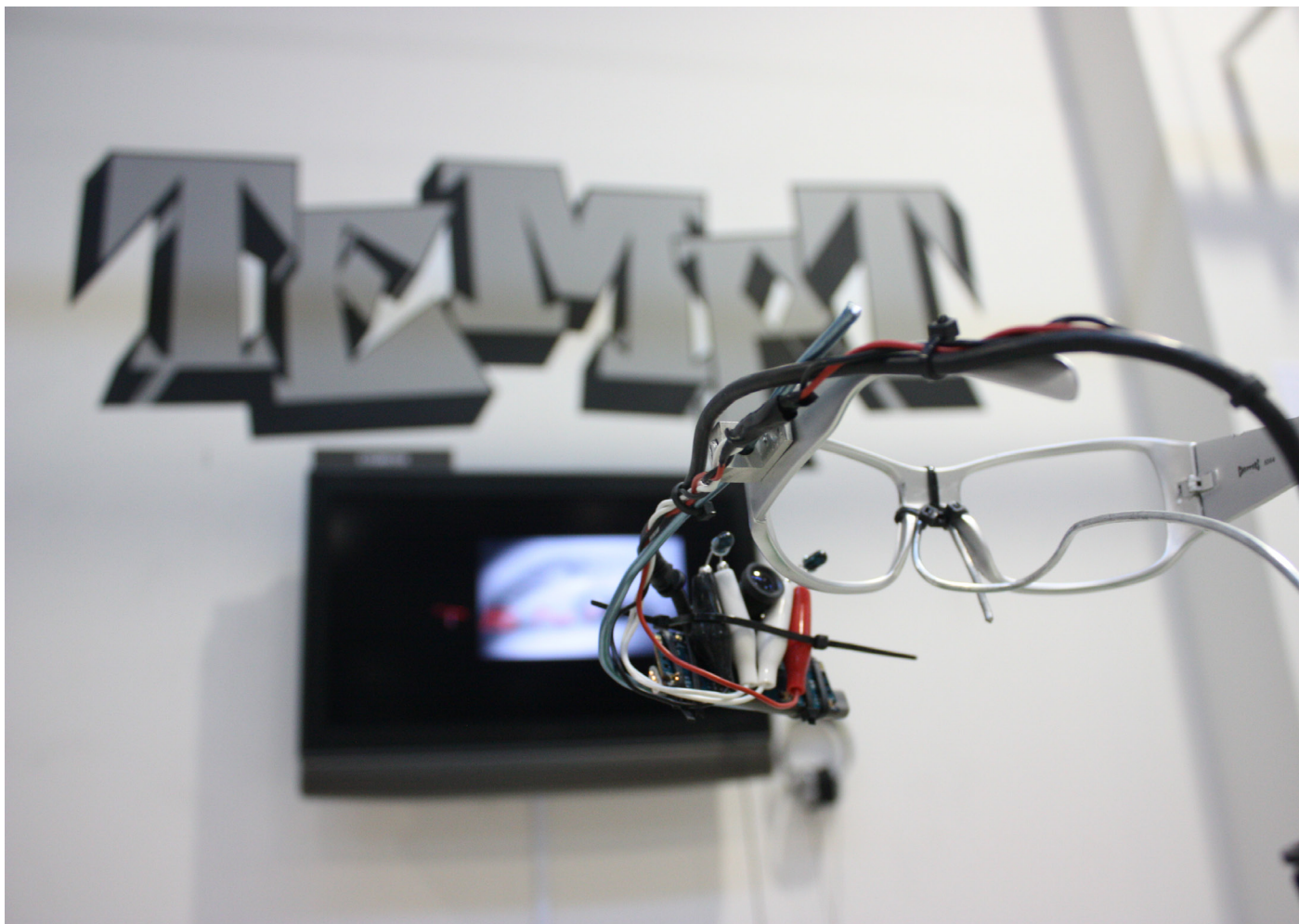
[4] Natural User Interface (NUI), è il termine utilizzato per fare riferimento a un'interfaccia utente basata su movimenti relativamente naturali, azioni e gesti.

in alto

Minority Report, dir. Steven Spielberg (2002). Il protagonista interagisce con i dispositivi utilizzando il corpo come interfaccia.

in basso

Eeyewriter è un progetto sperimentale sviluppato da un team internazionale, composto dai membri di Free Art and Technology (FAT), OpenFrameworks, Graffiti Research Lab e dalle comunità di Ebeling Group, sta lavorando per creare un sistema di tracciamento oculare open source e a basso costo, che consentirà ai pazienti con SLA di disegnare usando solo gli occhi.



Casi Studio

“Sixth Sense”, Fluid Interfaces Group, MIT, 2001 [5]

È un’interfaccia gestuale indossabile che aumenta il mondo fisico che ci circonda con informazioni digitali e ci consente di utilizzare i gesti delle mani per interagire con svariati servizi.

Il progetto, avviato nel 1994, ha portato alla creazione di un dispositivo in grado di raccogliere dati sull’ambiente circostante all’utente e di riconosce i movimenti delle mani tramite una videocamera.

Per lo sviluppo di questo progetto sono state studiate le gesture e le loro sfumature in base alla persona e alla cultura di provenienza, così da poter prevedere l’intenzione del gesto compiuto e ridurre il rischio di errori nell’esecuzione del comando.

“Hubs”, Mozilla, MIT, 2002 [6]

Hubs è una chat room VR open source, per ogni visore e browser, che esplora nuove modalità di comunicazione e di interazione.

È possibile aprire una stanza virtuale, condividerla tramite un URL e interagire tramite tool differenti con lo spazio circostante, in tempo reale.

Mozilla Hub offre nuove opportunità per imparare, creando un senso di umanità condivisa.

Prototipazione

Spatial Being permette a più utenti di incontrarsi dentro lo stesso spazio virtuale, 2D o 3D, per avviare processi di co-creazione e co-progettazione in tempo reale, utilizzando le componenti fondamentali del linguaggio naturale, i gesti e la voce.

Per sviluppare il progetto sono state individuate e analizzate le parti principali che lo costituiscono, in particolar modo: avviare una comunicazione bidirezionale server-client, riconoscere e usare i gesti delle mani e la voce per controllare gli oggetti all’interno di uno spazio virtuale.

Inizialmente, utilizzando “Socket.io” [7], è stato creato un server per consentire a più utenti di connettersi nello stesso spazio virtuale.

Il server è stato reso accessibile online grazie

[5] Nel 2012 Pranav Mistry ha indossato un device simile ma più avanzato che ha chiamato “WUW”, Wear your world.

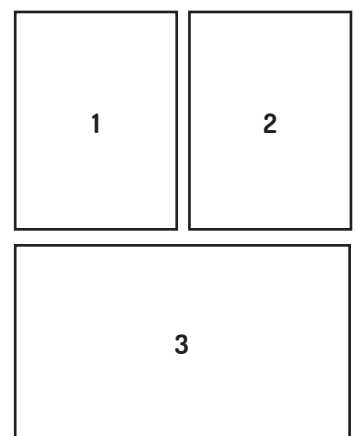
[6] Dal 2020, con Hubs Cloud, Mozilla offre la possibilità a chiunque di creare uno spazio personalizzarlo, con i propri contenuti e avatar, e di hostarlo su server esterni. (<https://hubs.mozilla.com/cloud>)

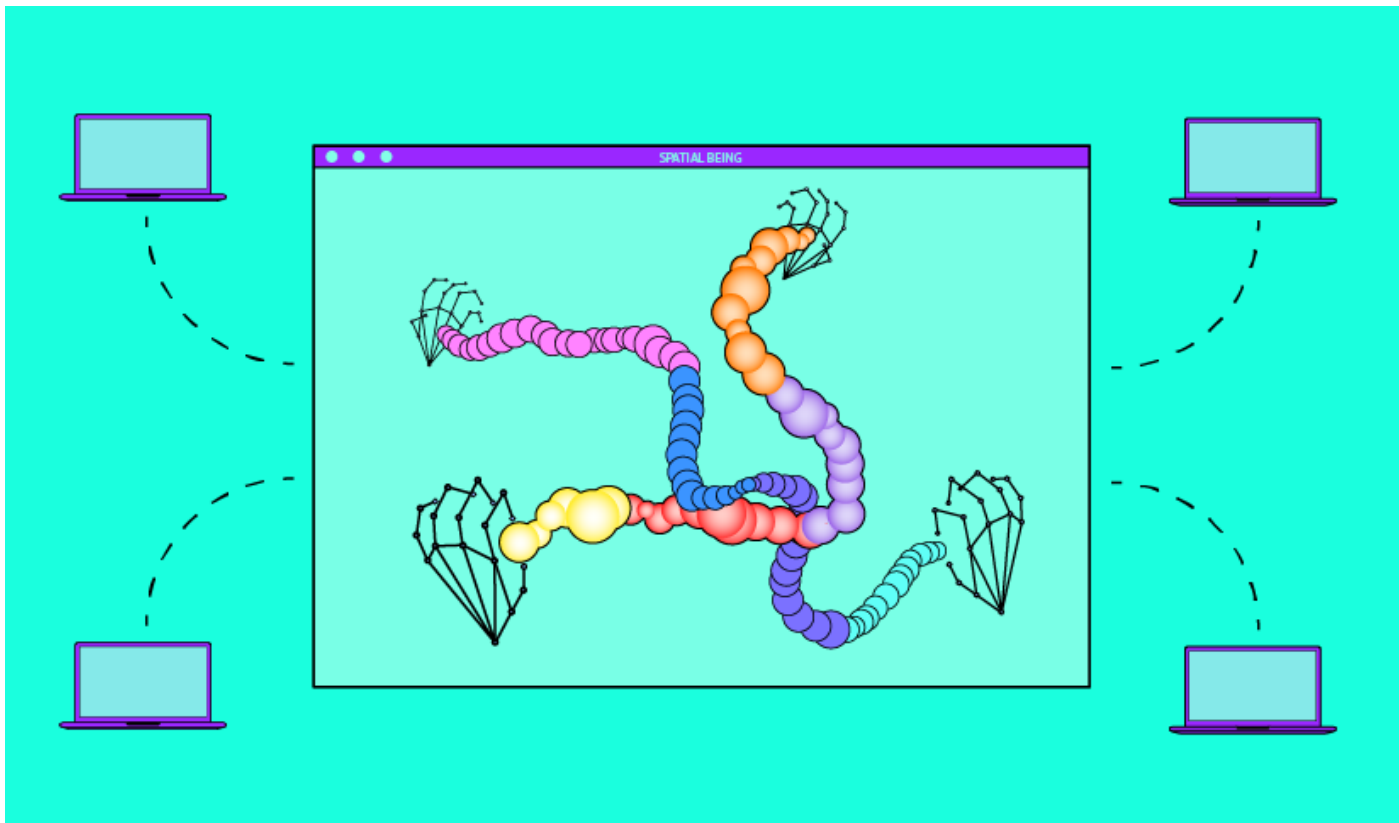
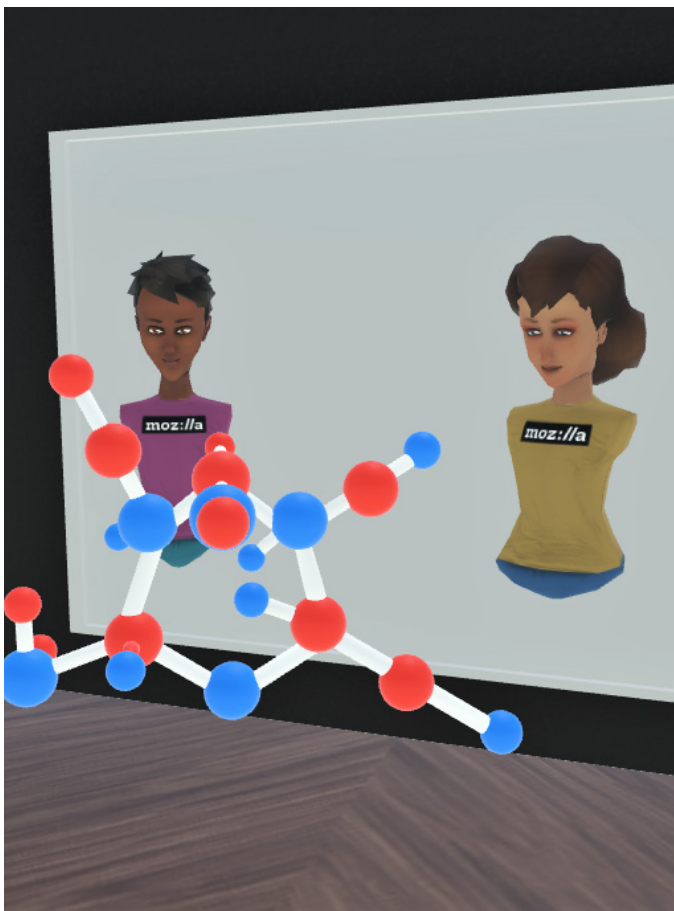
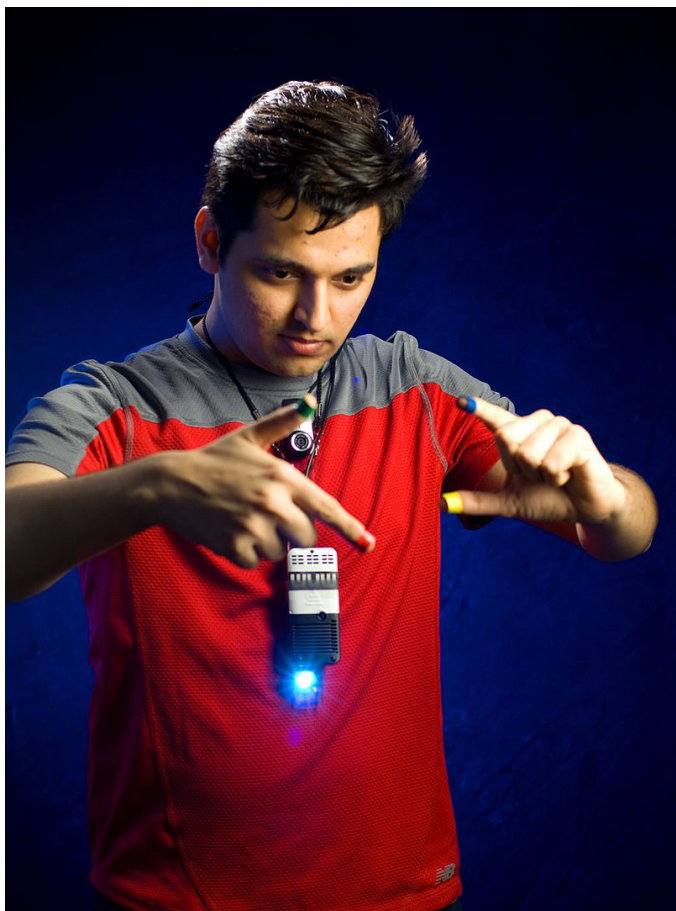
[7] Socket.io è una libreria Javascript per applicazioni web in tempo reale. Comprende una comunicazione bidirezionale realtime tra i web client e i server. (socket.io/docs/v4)

1
Pranav Mistry mentre indossa e utilizza Sixth Sense.

2
Link MozillaHubs (hubs.mozilla.com/).

3
schema del funzionamento di Spatial Being.





a un hosting sulla piattaforma “Glitch” [8], che gestisce la comunicazione dei vari Socket connessi. Il server permette, a un certo numero di client, di condividere i propri dati (posizione nello spazio, colore e dimensione), li organizza e li invia a tutti gli utenti per disegnare, in tempo reale, su tutti gli sketch.

In un secondo momento sono stati approfonditi gli ambiti della Natural User Interface e della Human Body Interaction [9] tramite algoritmi incentrati sul linguaggio naturale come “P5.speech” [10] e “HandPose” [11]. L’Hand Tracking ha permesso di utilizzare i gesti per interagire e manipolare oggetti come se fossero tra le proprie mani.

Un algoritmo di machine learning è stato allenato per essere in grado di eseguire compiti differenti in base alla *gesture* riconosciuta, utilizzando la webcam del computer.

Chiudendo la mano è possibile cambiare colore, con la mano aperta o semichiusa è possibile disegnare una sfera o un cubo e facendo pinch con indice e pollice è possibile ridimensionare l’elemento.

Tramite P5.Speech è stato invece possibile utilizzare la voce per muovere un oggetto tridimensionale nello spazio, cambiarne il colore e la dimensione.

Infine, dopo aver acquisito maggior consapevolezza e padronanza delle tecnologie utili, sono state unite le parti testate per giungere al prototipo di Spatial Being.

Il risultato finale permette, a più persone, di collegarsi su diversi sketch di P5.js e di interagire, tra loro e con lo spazio, utilizzando le mani per generare forme e la voce per cambiarne il colore o per dare comandi vocali come “Disegna”, “Stop” e “Salva”. Quando un nuovo utente esegue l’accesso su Spatial Being, il suo nome viene posizionato casualmente nello spazio e collegato a quello degli altri, generando una mappa tridimensionale della community di persone, che si incontrano per creare liberamente.

[8] Glitch è una piattaforma collaborativa per creare e hostare app e sitiweb. (glitch.com/)

[9] Human Body Interaction (HBI) è il termine utilizzato per far riferimento a quei tipi di interfacce che trasformano i segnali del corpo in comandi per controllare i device.

[10] Link P5.speech ([idmnyu.github.io/p5.js-speech/](https://github.com/idmnyu/p5.js-speech/))

[11] Link Hand Pose (<https://learn.ml5js.org/#/reference/handpose>)

1-2

Una volta eseguito l’accesso su Spatial Being vengono visualizzate le istruzioni per disegnare.

3

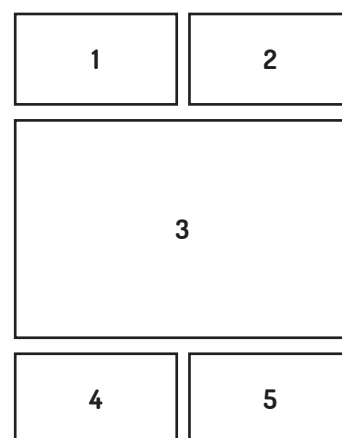
Artwork co-creato da più utenti su Spatial Being 2D. <https://editor.p5js.org/RobertoAlesi/full/0Tne-Js-i>

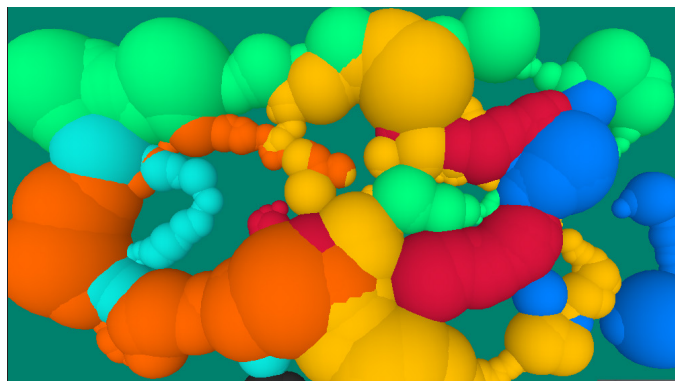
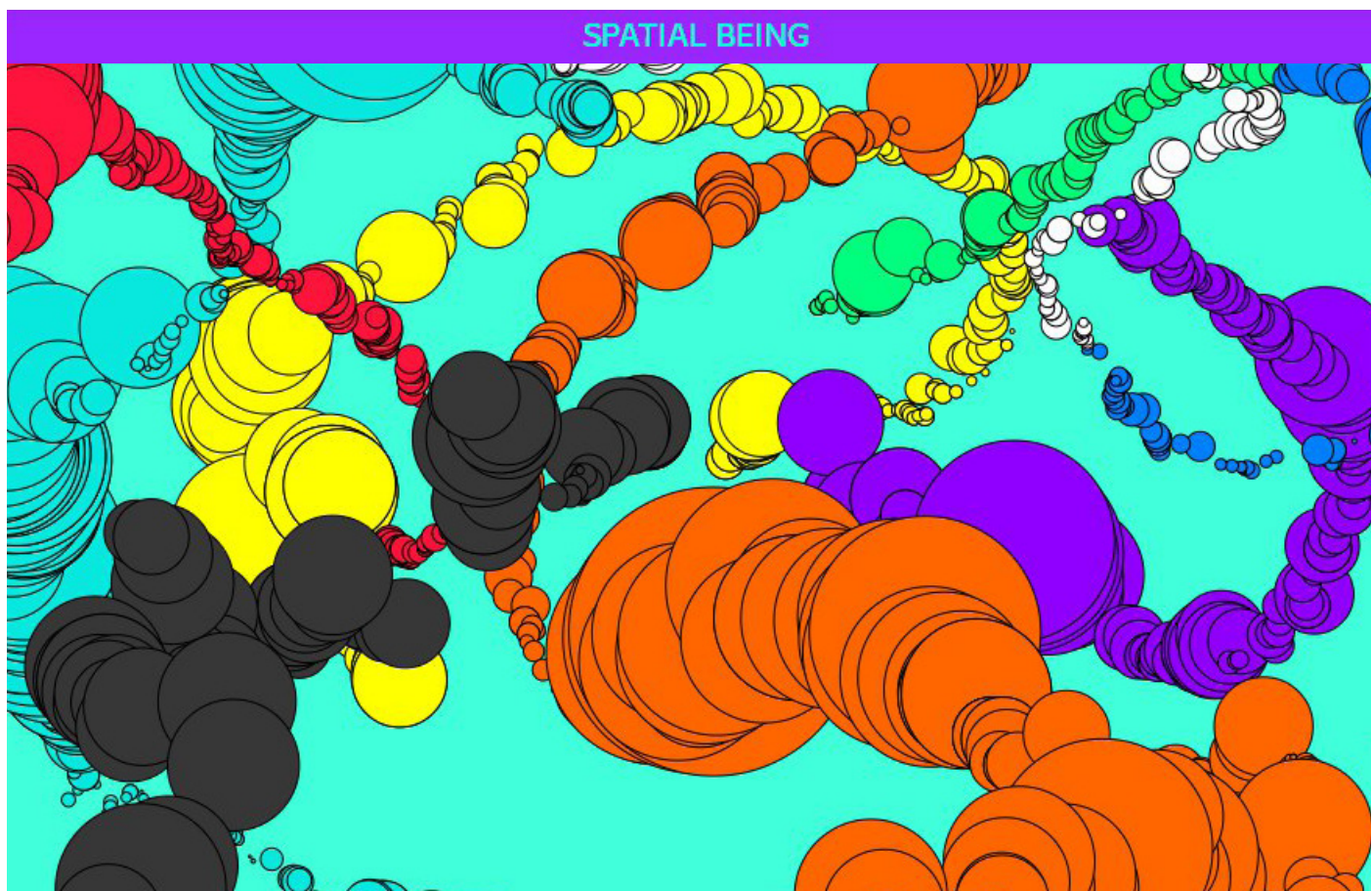
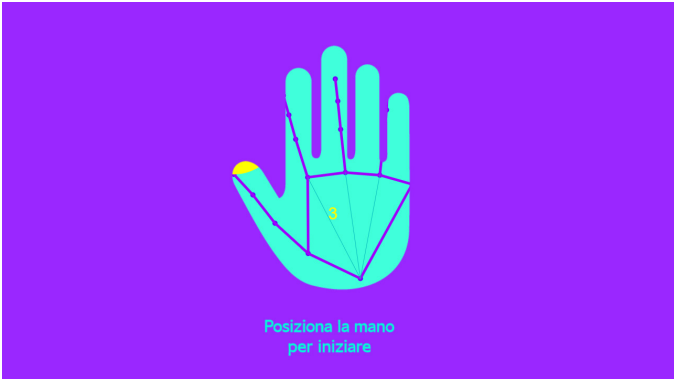
4

Artwork co-creato da più utenti su Spatial Being 3D. <https://editor.p5js.org/RobertoAlesi/full/0Tne-Js-i>

5

Mappa degli utenti connessi dentro Spatial Being https://editor.p5js.org/RobertoAlesi/full/8lwd_t86o





Sviluppi futuri

Durante lo sviluppo del prototipo sono state riscontrate delle criticità come la latenza della comunicazione, che rende poco fluida l'esperienza tra gli utenti connessi, e la difficoltà su alcuni dispositivi di mantenere sempre attivi microfono e webcam, creando problemi durante l'invio e la ricezione dei dati in tempo reale.

Sarebbe opportuno riprodurre Spatial Being fuori dall'editor di P5.js e WEBGL [12], utilizzando la libreria Three.js [13], in modo da rendere più preciso e affidabile l'*hand tracking* e più leggera la creazione dinamica di oggetti tridimensionali. Proprio per questo motivo, durante la fase di prototipazione, è stata esplorata la comunicazione WebSocket online utilizzando Three.js.

Il primo prototipo consente a più utenti di collegarsi all'interno di uno spazio 3D condiviso, di muovere il proprio cubo identificativo utilizzando le frecce della tastiera e di cambiare il colore di un elemento posto al centro della scena con un *click* del mouse. Con il secondo prototipo invece, è stato testato l'*hand tracking*, utilizzando la libreria javascript Tensorflow [14], consentendo riconoscimento più affidabile e lineare per disegnare elementi tridimensionali nello spazio. Apportando queste modifiche, Spatial Being, potrebbe diventare uno spazio di co-design accessibile da qualsiasi dispositivo, dotato di un browser e una fotocamera.

In futuro potrebbe essere implementato con una capacità di riconoscimento dei gesti più accurata e intelligente, in modo da riconoscere le sfumature di significato dei movimenti in base alla cultura di provenienza e ai contesti di utilizzo. Questo permetterebbe di espandere la fruizione dei servizi digitali e degli spazi virtuali a un numero di utenti più elevato, consentendone l'utilizzo anche alle persone diversamente abili, che al momento non vengono incluse nel mondo in cui ci muoviamo ormai da molto tempo.

[12] WWBGL è una libreria grafica 3D per il web (<https://get.webgl.org/>)

[13] Three.js è una libreria JavaScript cross-browser e un'interfaccia di programmazione utilizzata per creare e visualizzare grafica 3D (<https://threejs.org/>)

[14] Tensorflow è una libreria open-source per il machine learning (<https://www.tensorflow.org/>)

1

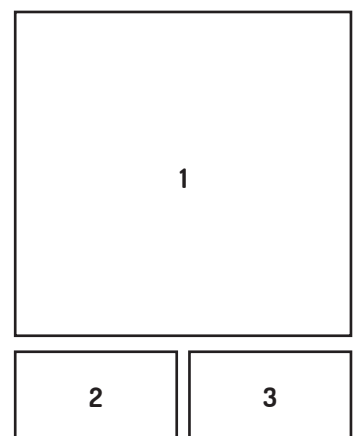
Il prototipo di Spatial Being permette di visualizzare la propria webcam in basso a destra. In futuro potrebbe essere implementato con lo streaming video, consentendo di effettuare videochiamate.

2

Primo prototipo realizzato utilizzando Socket.io e Three.js (<https://socket-three-due2.glitch.me>)

3

Secondo prototipo realizzato utilizzando Socket.io, Three.js, Tensorflow e HandPose (<https://hand-pose.glitch.me>)



ndP DEFDEF by RobertoAlesi

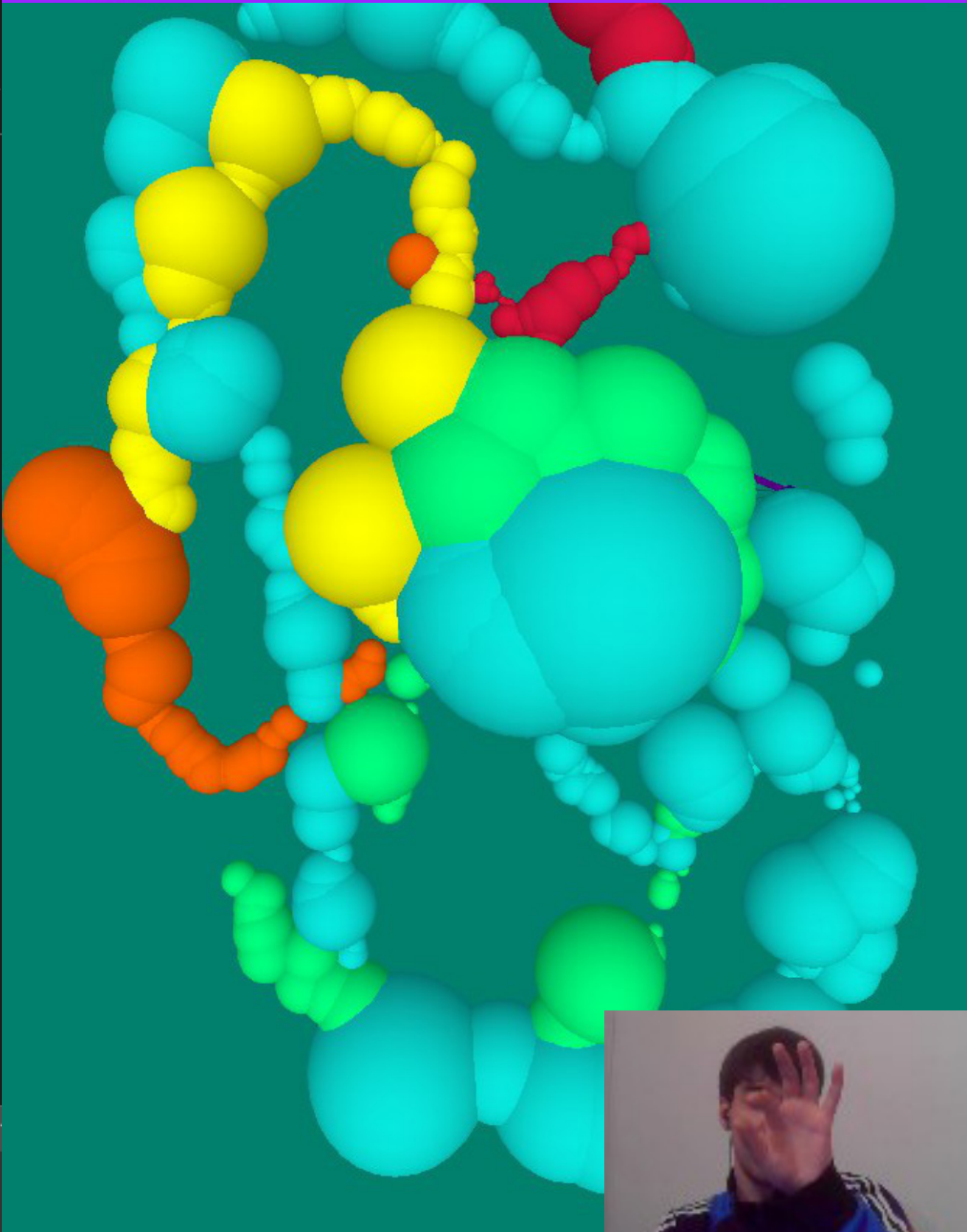
Saved: about 21 hours ago

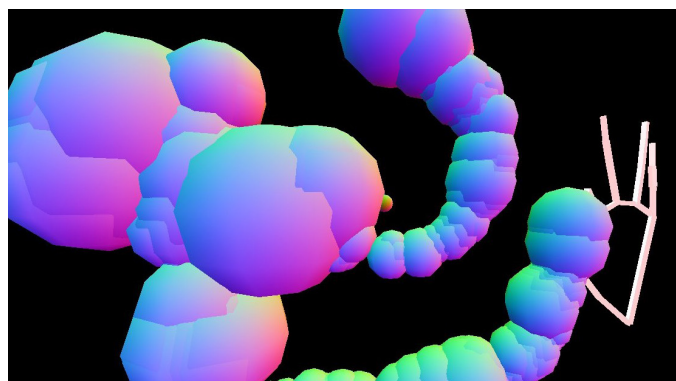
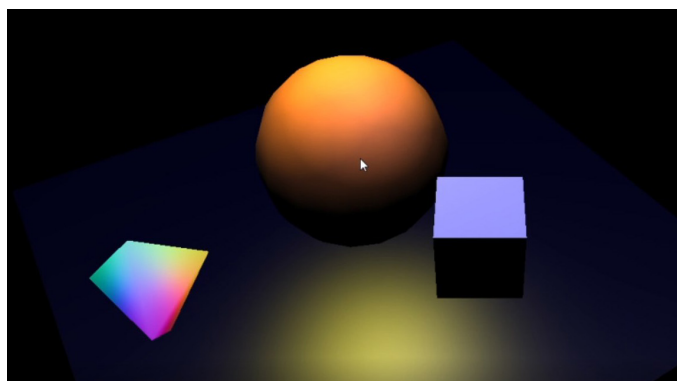
S

```
mb[3][0], 0, widthVideo,  
mb[3][1], 0, heightVideo,  
ky[3][0], 0, widthVideo,  
ky[3][1], 0, heightVideo,  
  
inger[3][0],  
  
ythumb-d0/2, col: colore,
```

Clear

SPATIAL BEING





Sitografia

- <https://arvrjourney.com/mozilla-hubs-bring-people-closer-from-distance-6345c1f399e0>
- <http://www.figlab.com/>
- <https://fingerspelling.xyz/>
- https://www.ted.com/talks/pranav_mistry_the_thrilling_potential_of_sixthsense_technology#t-547392
- <https://blog.scottlogic.com/2016/05/25/Body-as-Interface.html>
- <http://www.eyewriter.org/>
- <https://www.interaction-design.org/literature/article/natural-user-interfaces-what-are-they-and-how-do-you-design-user-interfaces-that-feel-natural>
- <https://www.sciencedirect.com/topics/engineering/gesture-interface>
- <https://uxdesign.cc/the-less-interface-the-better-c7e938cd5517>
- https://medium.com/@Mohan_Krishnaraj/designing-for-the-invisible-interface-mohan-krishnaraj-aaf40c95cc18
- <https://www.fastcompany.com/1150913/mits-sixth-sense-machine-makes-reality-better>
- <https://ml5js.org/>
- <https://idmnyu.github.io/p5.js-speech/>

Bibliografia

- Michele Zannoni, *Progetto e interazione, il design degli ecosistemi interattivi*, Quodlibet, 2018
- Edward De Bono, *Il pensiero laterale*, Univ. Rizzoli, 1967
- Dejan Chandra Gope, *Hand Gesture Interaction with Human-Computer*, Global Journals Inc., 2011
- Yong-Tian Wang, *Virtual Reality & Intelligent Hardware*, KeAi Communication, 2019

Filmografia

- Minority Report, dir. Steven Spielberg, 2002
- Iron Man, dir. Shane Black, Jon Favreau, 2008
- Tron: Legacy, dir. Joseph Kosinski, 2010

