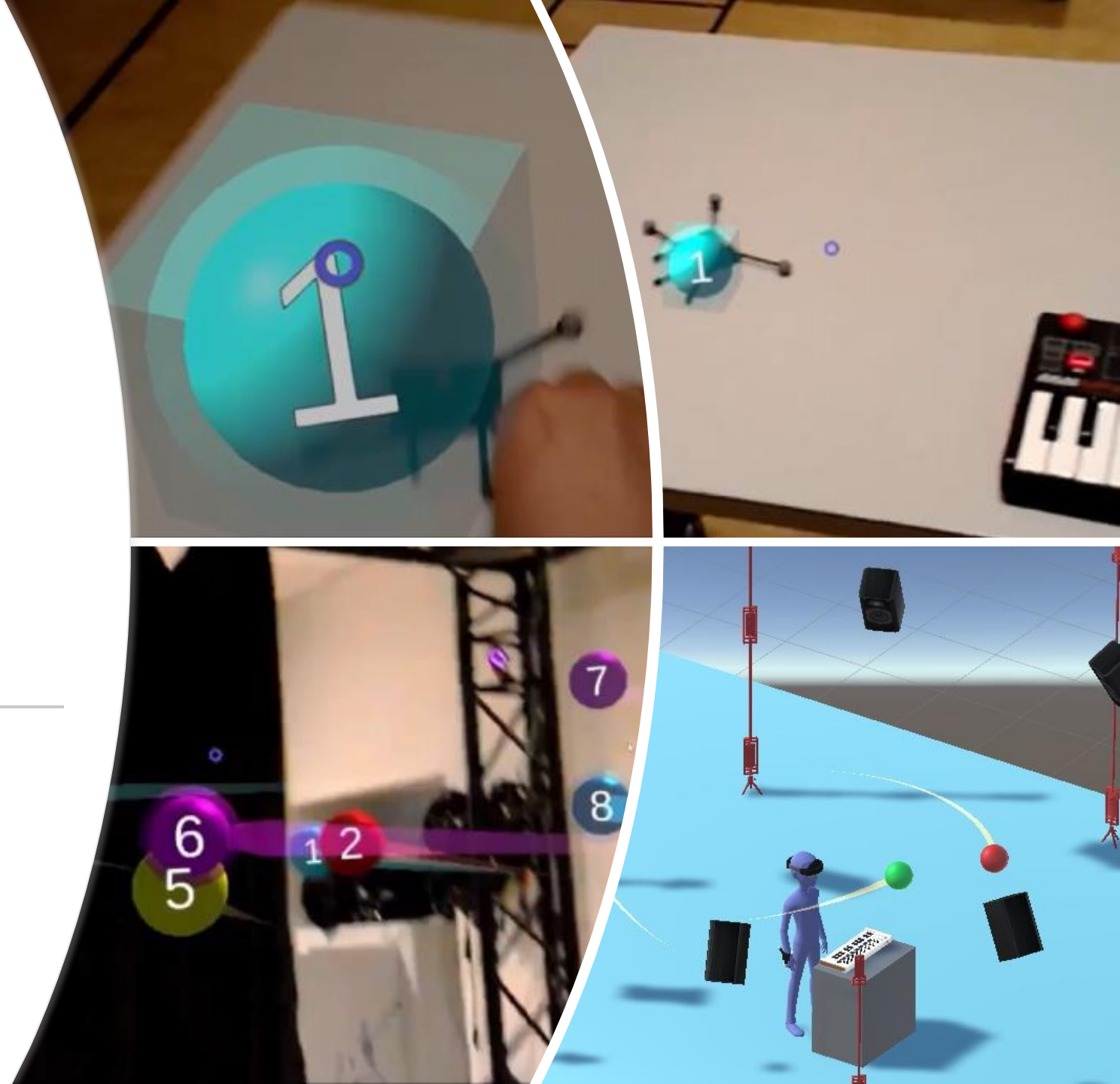
A Human-Agents Music Performance System in an Extended Reality Environment

Pedro Lucas

Stefano Fasciani



Agenda

- **Introduction**
- System Design
- Implementation
- Autonomous Behavior
- Results
- Conclusions

Introduction

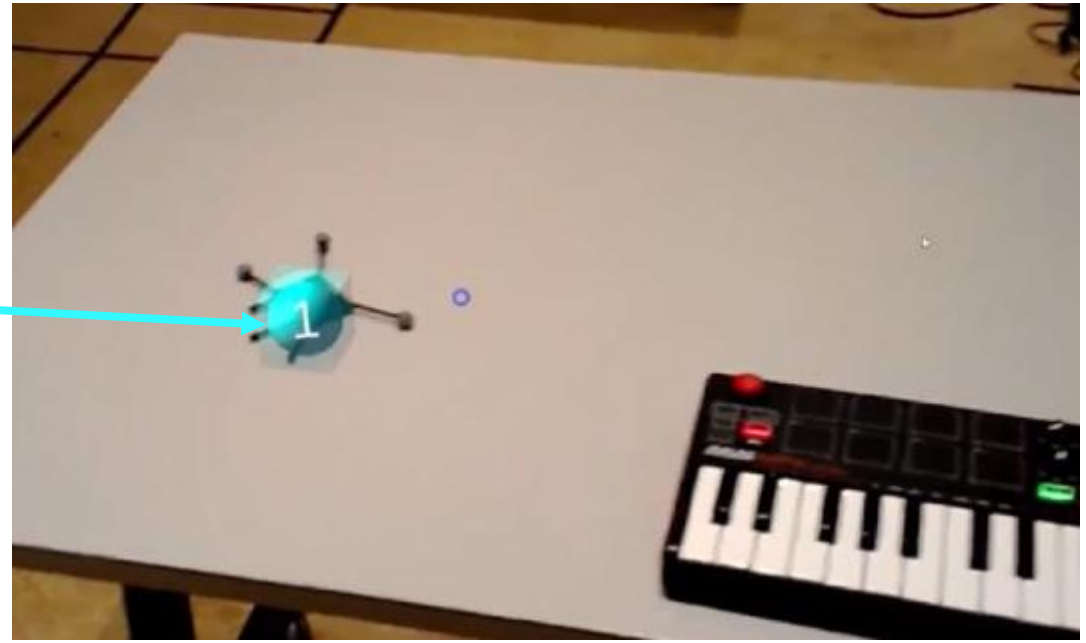
- **Research Question**
  - *How can we design and implement a system for human-machine live music performances in a multimodal environment?*

- **Key Aspects:**
  - **Human-machine** -> Human Musician + Artificial Musicians (Agents)
  - **Live music performance**-> Improvisation - Multi-track looping session (Like in Ableton Live)
  - **Multimodal environment** -> Sound embodiment into a 3D environment

Introduction

- **Basic elements (Physical-Virtual)**
  - ***A loop track*** *-> Musical material (through a physical keyboard), Musical Agent (Sound source – 3D Sphere)*
  - ***Agent Manual Movement*** *-> Physical object (Spatial Positioner)*

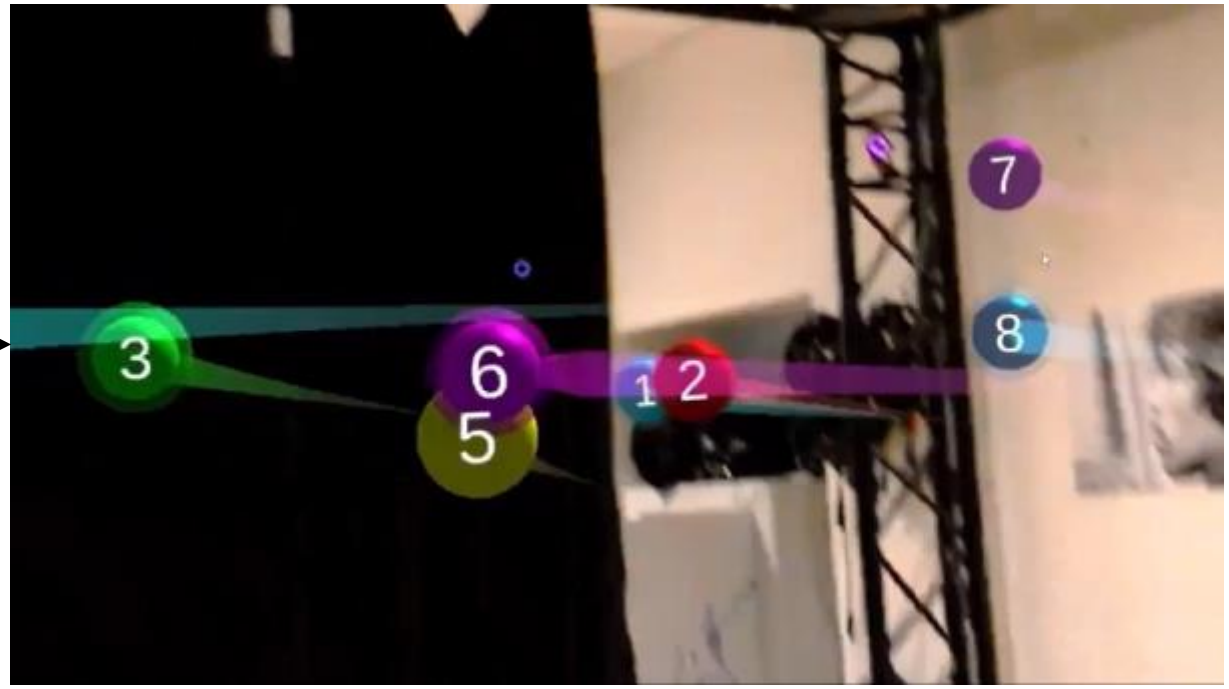
Agent (Musical Loop)
-Virtual object that can be heard and seen in the 3D space



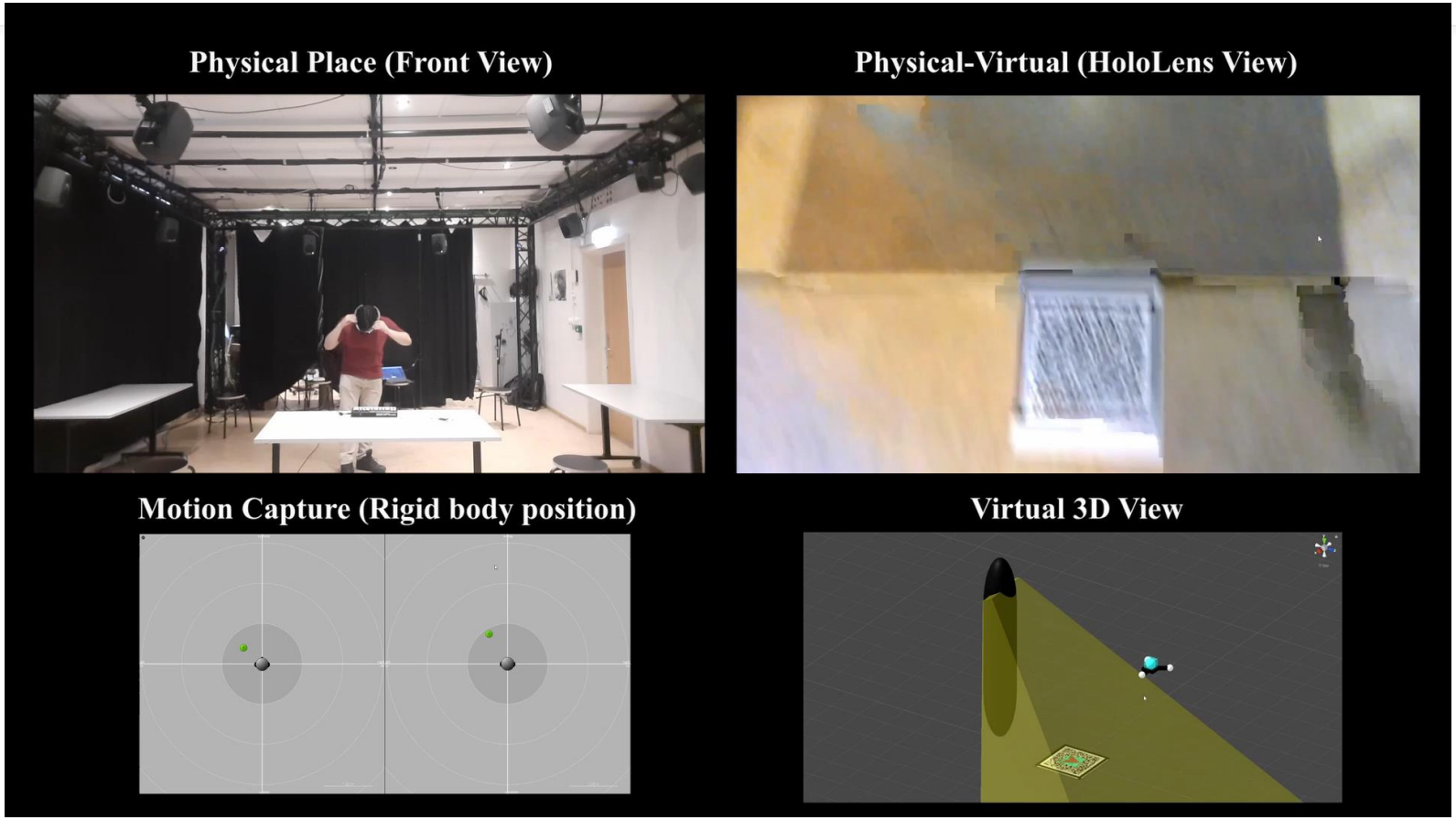
Introduction

- **Basic elements (Physical-Virtual)**
  - ***Autonomous Behavior***-> *Music Generation, Autonomous Movement*

Swarm of Musical Agents →



System Operation

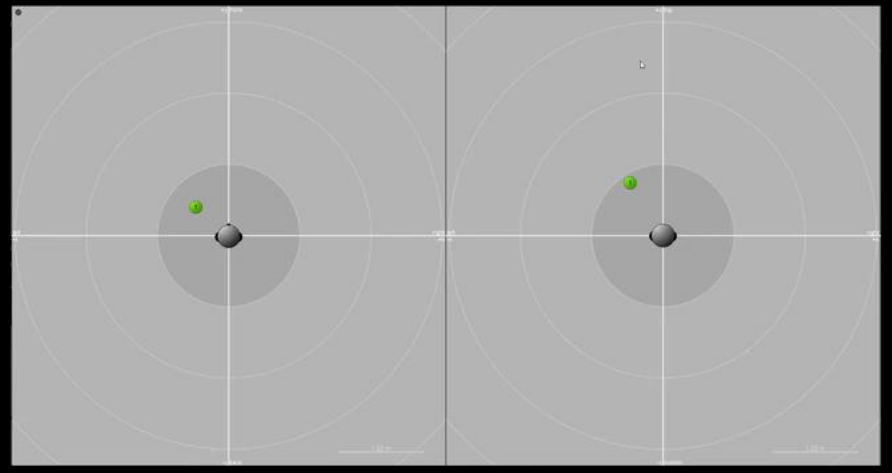


Demonstration Video

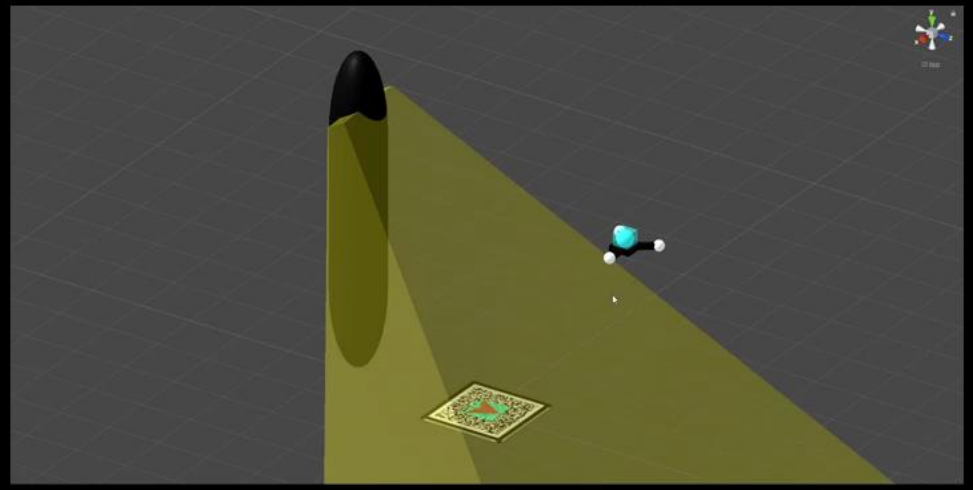
Physical Place (Front View)

Physical-Virtual (HoloLens View)

Motion Capture (Rigid body position)



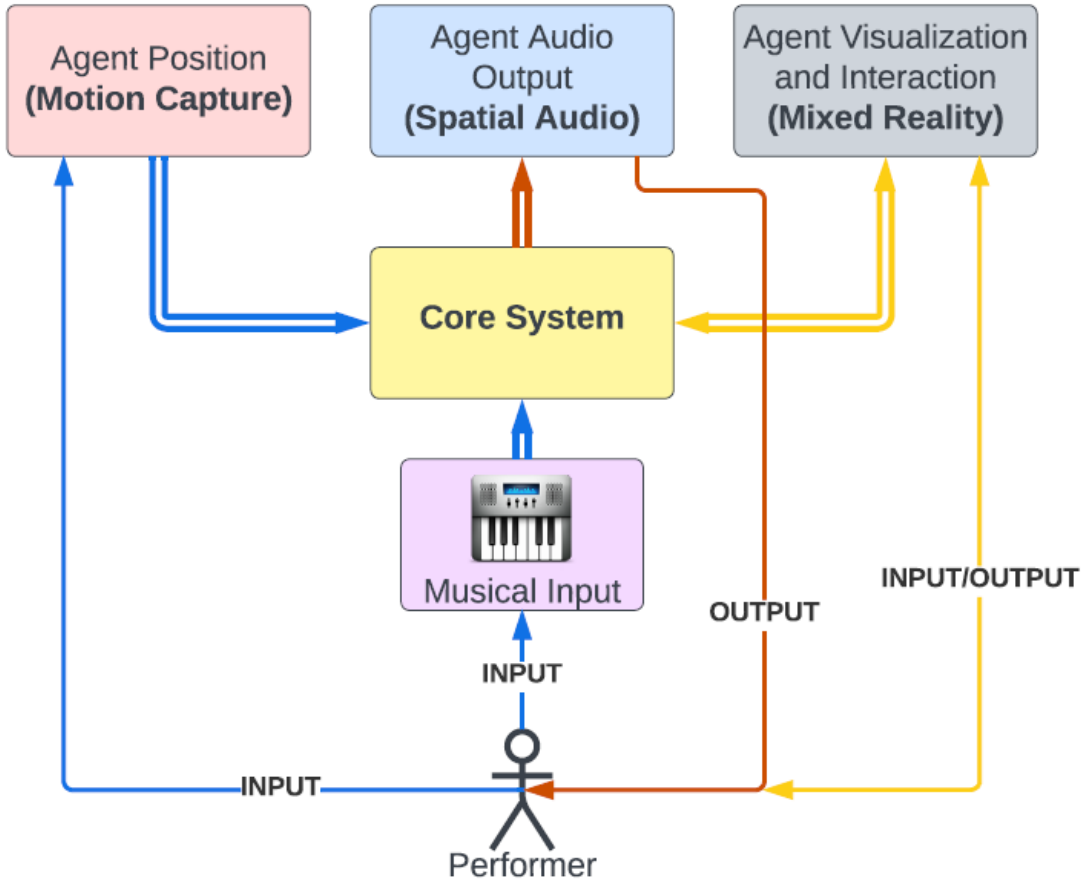
Virtual 3D View

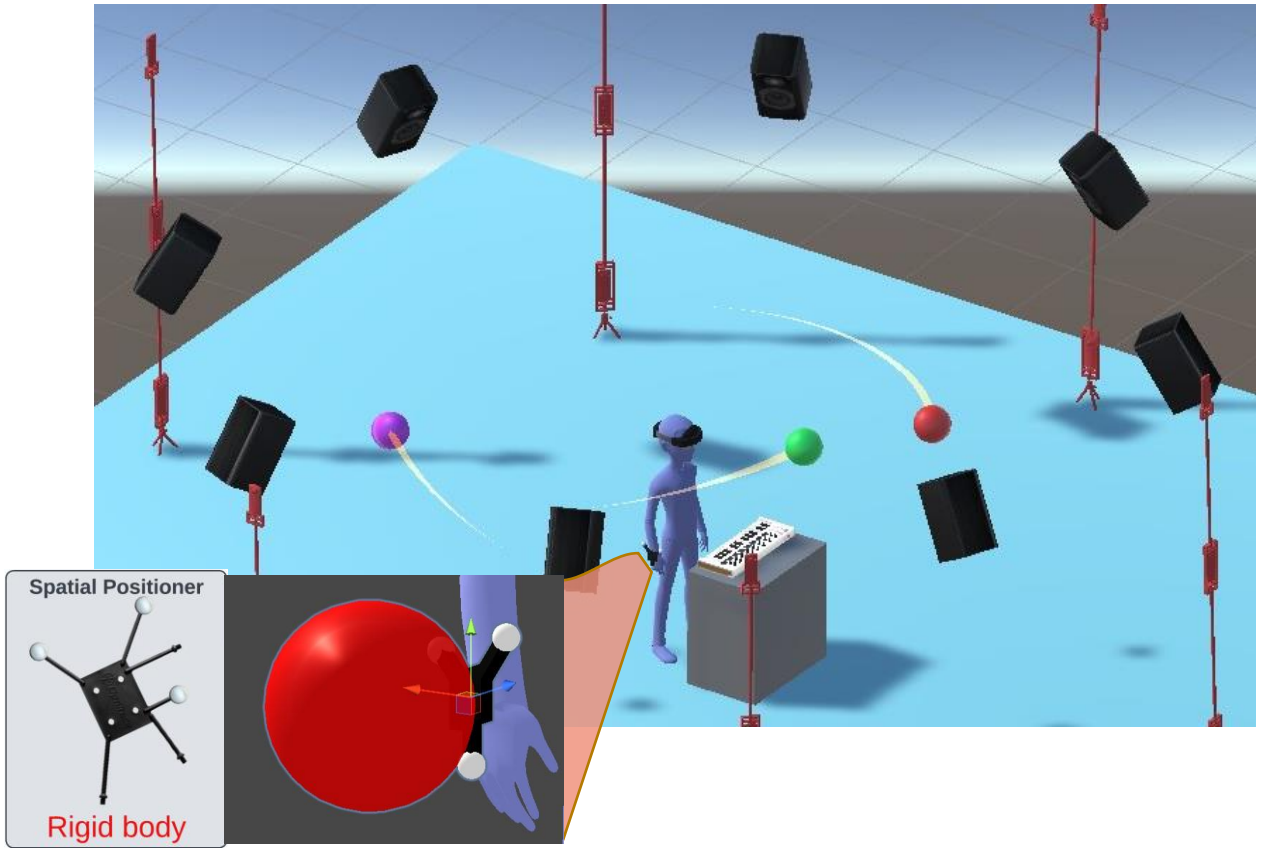


Agenda

- Introduction
- **System Design**
- Implementation
- Autonomous Behavior
- Results
- Conclusions

System Design





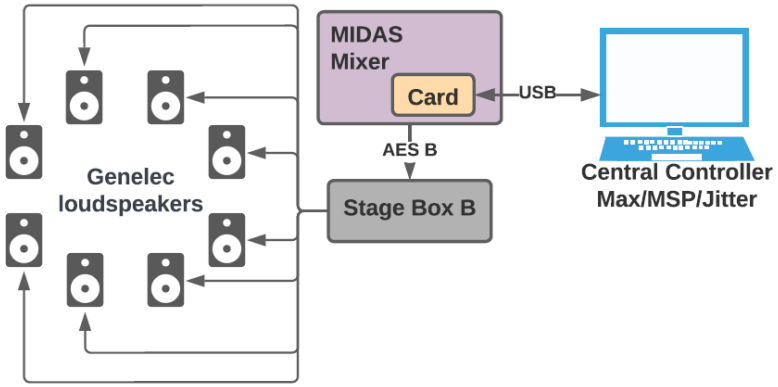
System Architecture and Operation

Agenda

- Introduction
- System Design
- **Implementation**
- Autonomous Behavior
- Results
- Conclusions

Implementation

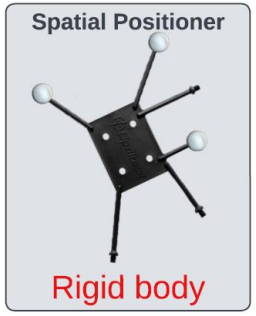
MCT Portal

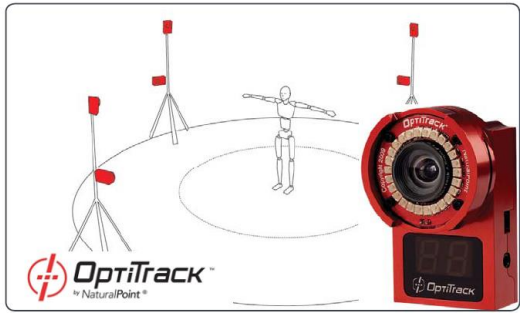


Spatial Audio System

Microsoft HoloLens 1

MIDI Controller





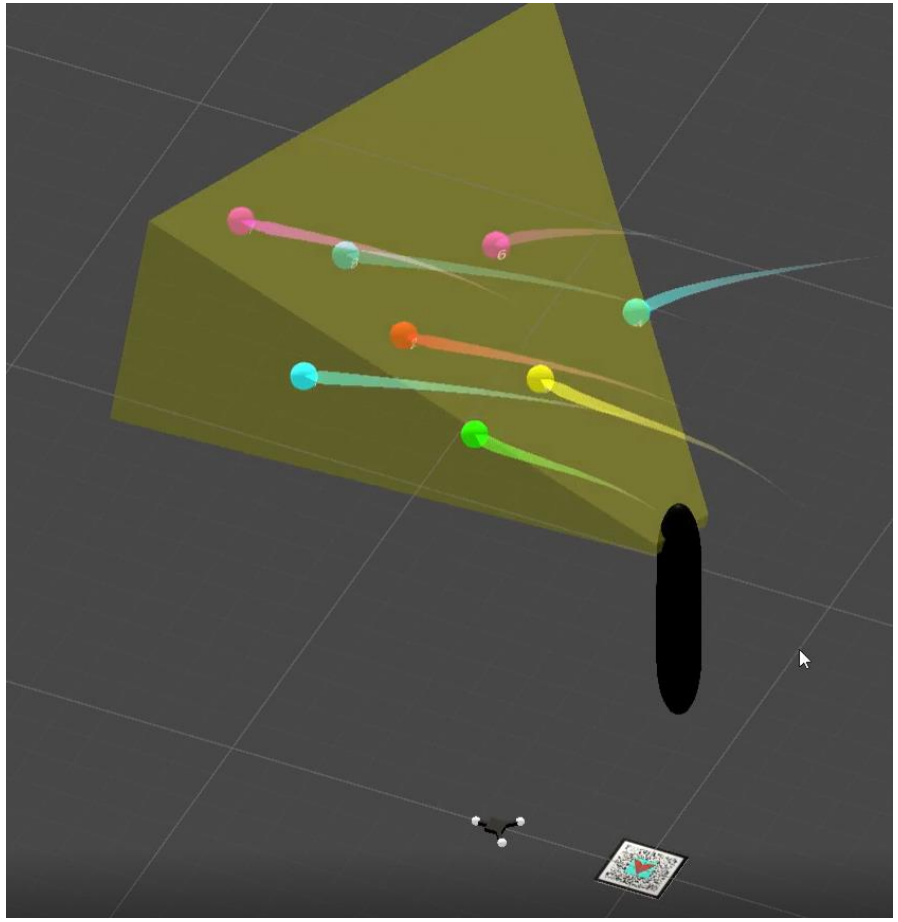
MOCAP System

MAX

Implementation

Playback System from a Music Session

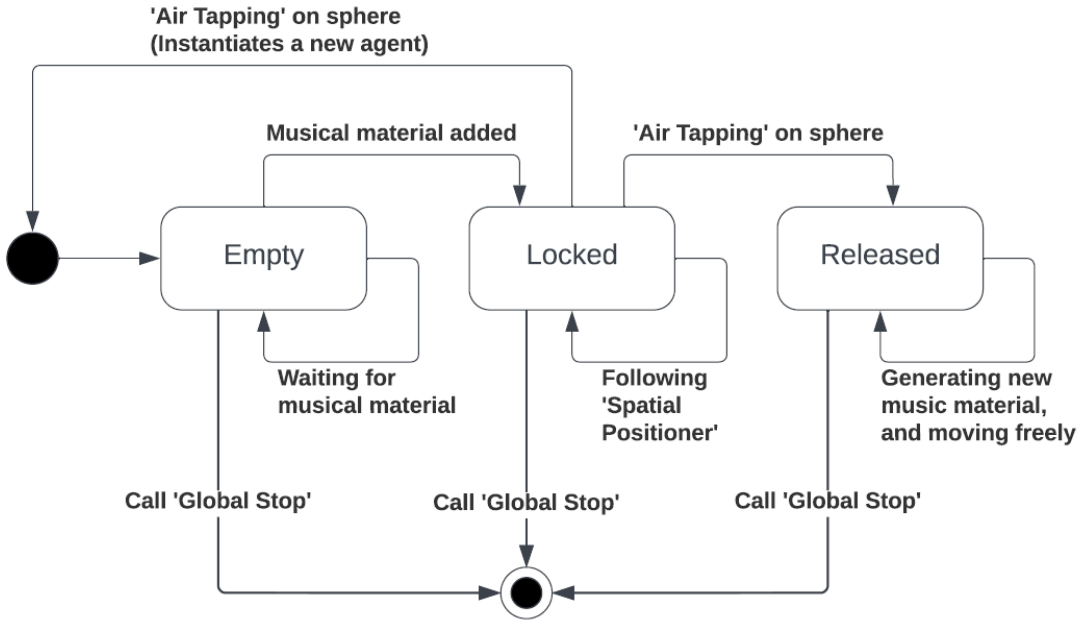
Unity



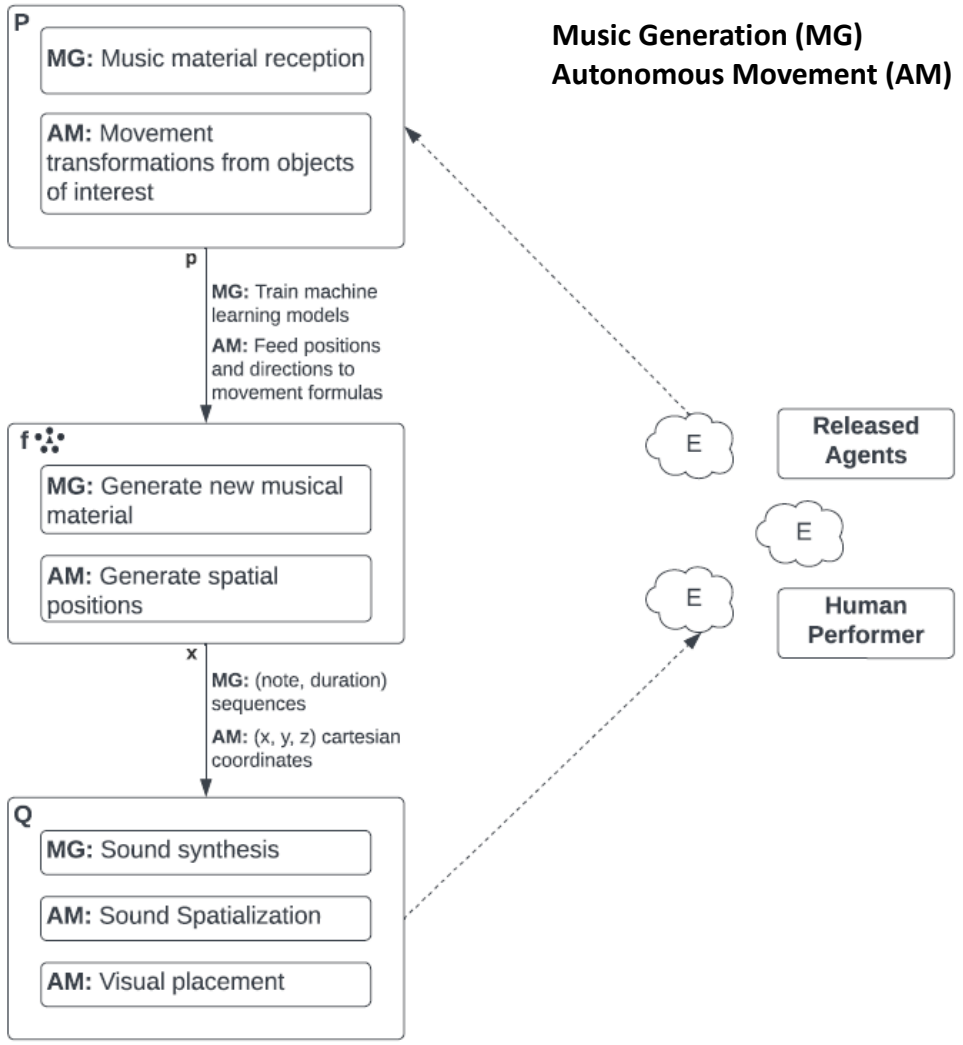
Agenda

- Introduction
- System Design
- Implementation
- **Autonomous Behavior**
- Results
- Conclusions

Autonomous Behavior



Agent Behavior (FSM)

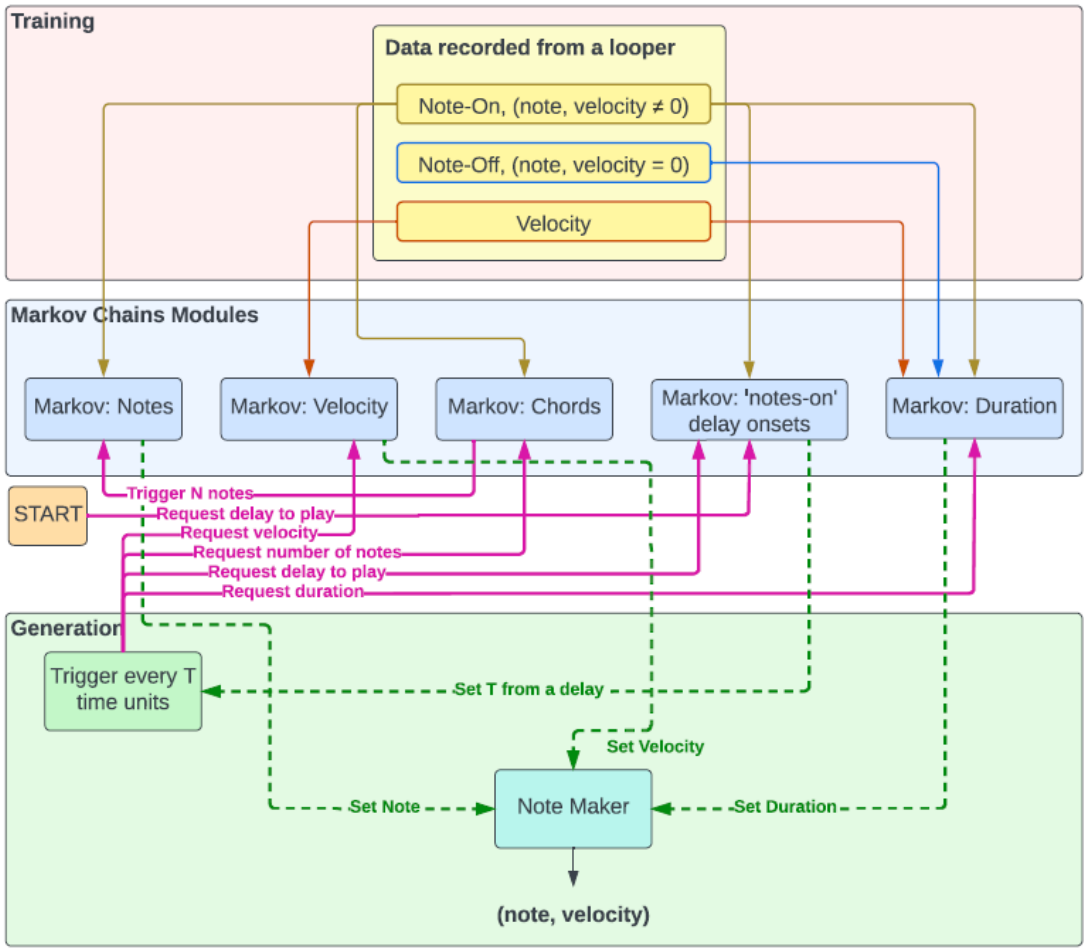


Swarm Behavior (PQf architecture – Live Algorithm)

T. Blackwell, O. Bown, and M. Young. Live Algorithms: Towards Autonomous Computer Improvisers

Autonomous Behavior

MUSIC GENERATION



Markov Chains System (https://spearced.com/algorithmic-process-ai/***)***

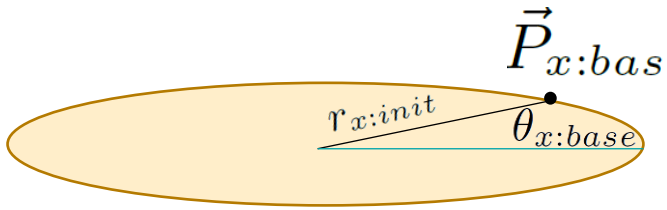
Autonomous Behavior

AUTONOMOUS MOVEMENT

Goal: Provide spatial awareness of the performer and other agents.

Spatial Sources for an Agent X:

1. Base circular constant movement



2. Position to be equally spread around the performer *h*.

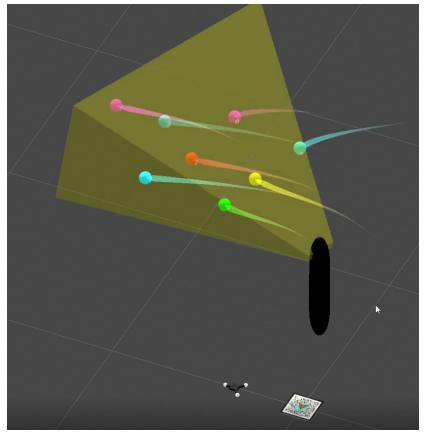
$$\vec{P}_{x:swarm} = \vec{P}_h - \sum_{i=1, i\neq x}^{N} \vec{P}_i : \vec{P}_h, \vec{P}_i \in \mathbb{R}^3; i, x, N \in \mathbb{N}^+$$

3. Performer's gaze direction $\hat{\mathbf{dir}}_h$

$$\vec{P}_x = \alpha(\vec{P}_{x:base} + \vec{P}_{x:swarm} + \hat{\mathbf{dir}}_h)$$

$$: \vec{P}_{x:base}, \vec{P}_{x:swarm}, \hat{\mathbf{dir}}_h \in \mathbb{R}^3; \alpha \in \mathbb{R}$$

α = 1/2 when the agent size is 1, otherwise α = 1/3



Agenda

- Introduction
- System Design
- Implementation
- Autonomous Behavior
- **Results**
- Conclusions

Results

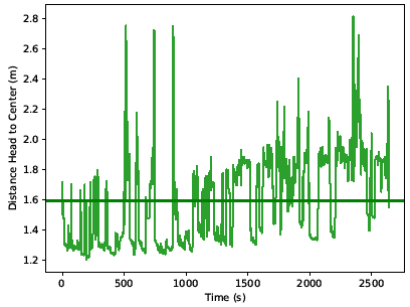
System Measurements (for 8 agents)

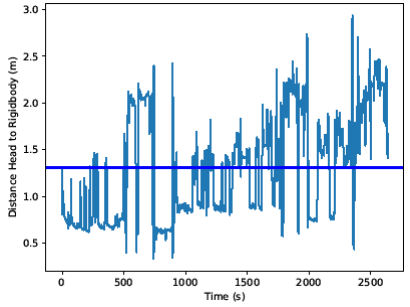
Parameter	Agents' Group Size							
	1	2	3	4	5	6	7	8
Physical keyboard to Sound Output Latency (ms)	40.56	38.43	36.20	39.62	37.33	38.93	35.83	39.07
Physical keyboard to Sound Output Jitter (ms)	6.00	6.45	6.78	8.04	6.81	6.73	7.83	7.05
Spatial Audio Placement Latency (ms) (Between the rigid body movement and the sound output panning from the loudspeaker array)	118.23	112.67	121.00	129.34	123.78	129.34	126.56	134.89
Sound to Visualization Latency (ms) (Between hearing a sound source - agent - from a point in the loudspeaker array and visualizing it through the HoloLens)	3.25	3.03	4.40	4.39	5.36	5.37	4.62	6.75
Sound to Visualization Jitter (ms)	8.62	9.81	12.30	12.81	12.89	15.16	11.66	13.44
Packet Loss Core-to-HoloLens (%)	19.08	20.80	20.33	20.71	21.72	24.14	23.56	23.89

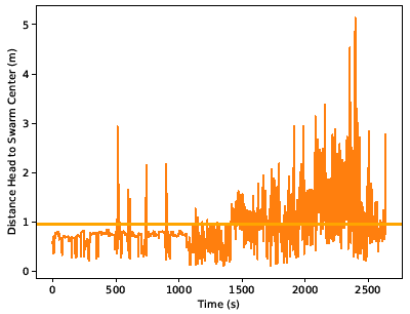
Summary of system measurements for agents' size 1 to 8

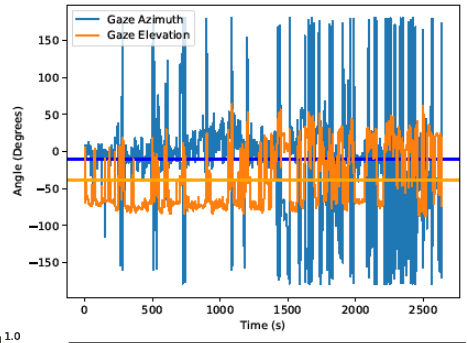
Results

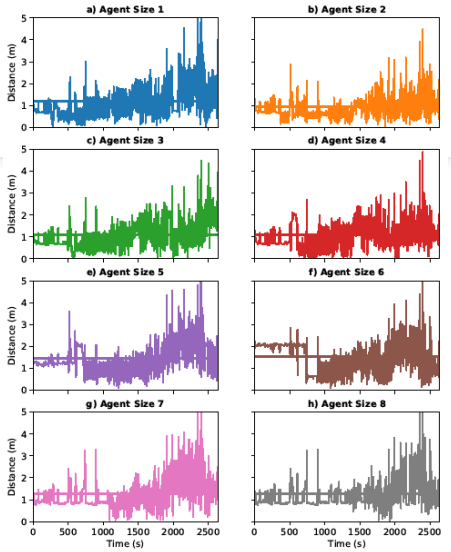
User Evaluation (7 musicians): Captured Data

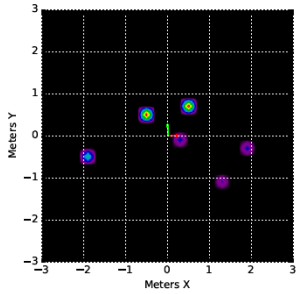


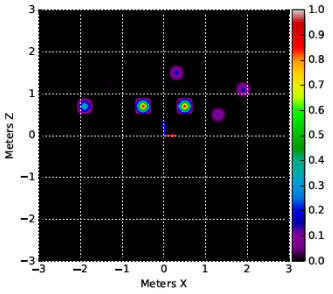


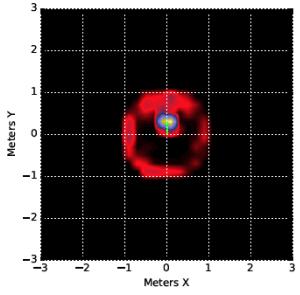


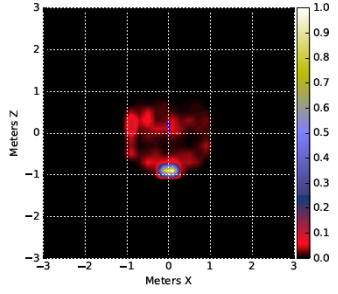


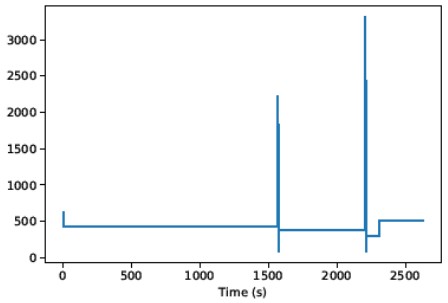


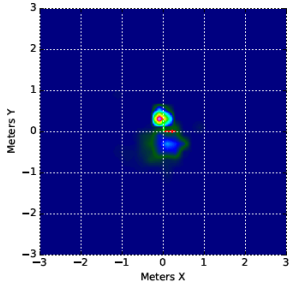


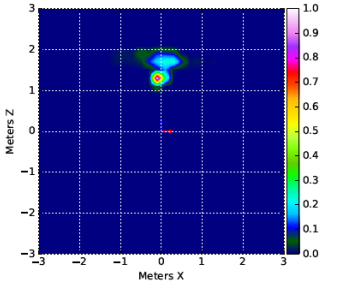


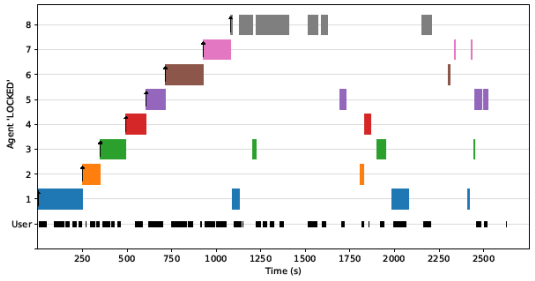


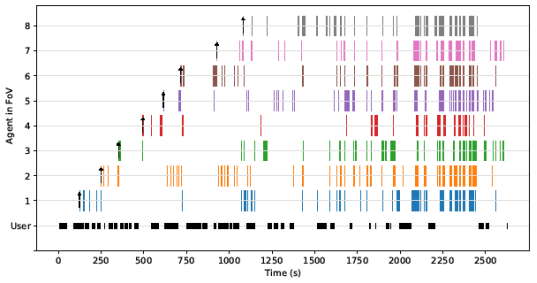


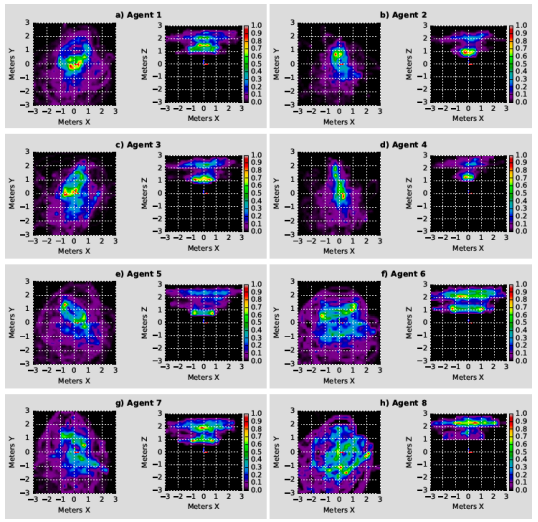






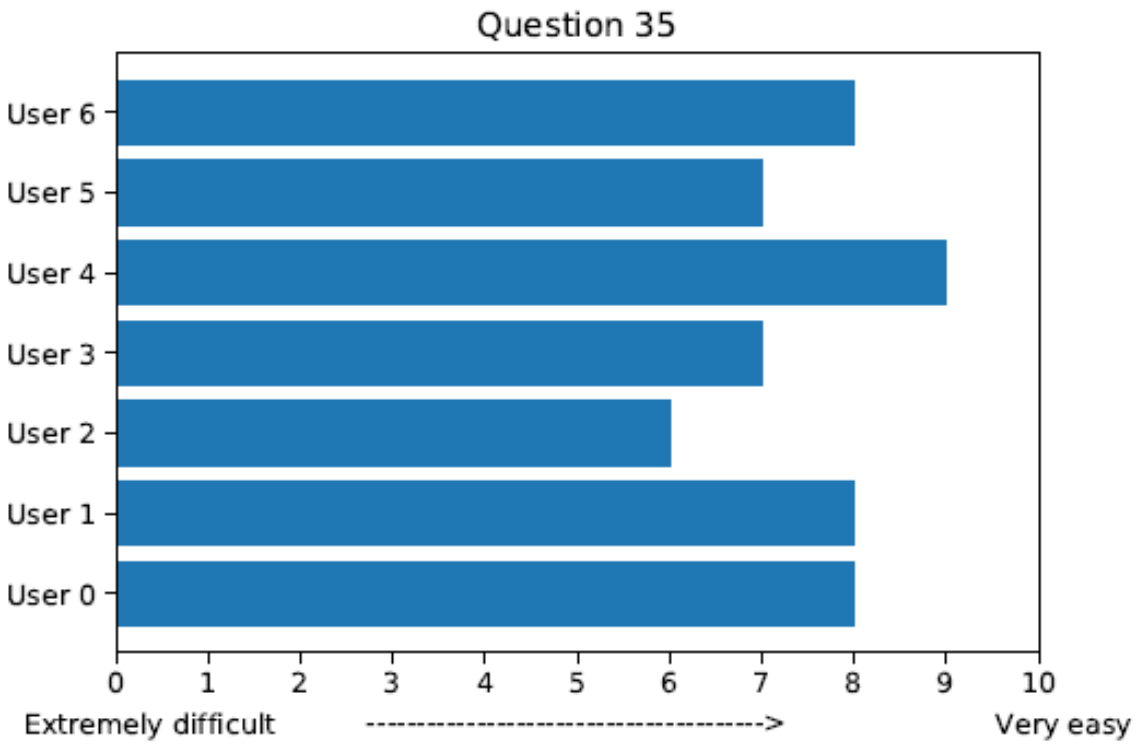




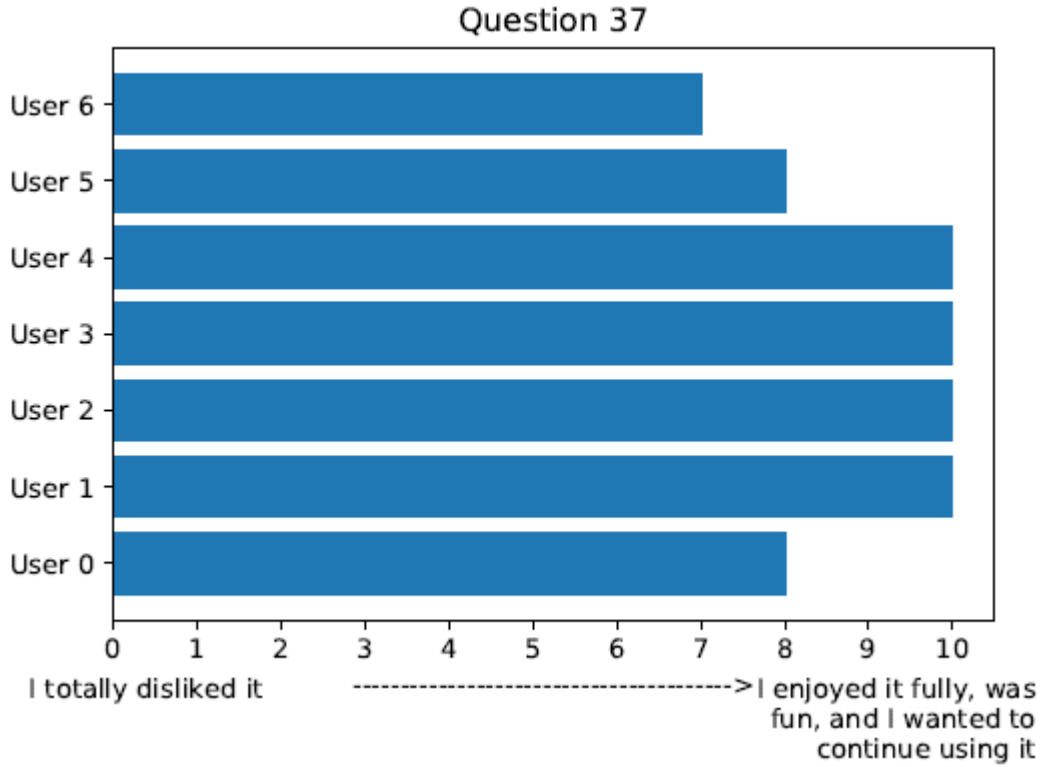


Results

User Evaluation (7 musicians): Survey



In general, how easy was to use the whole system?



How much did you enjoy the performance?

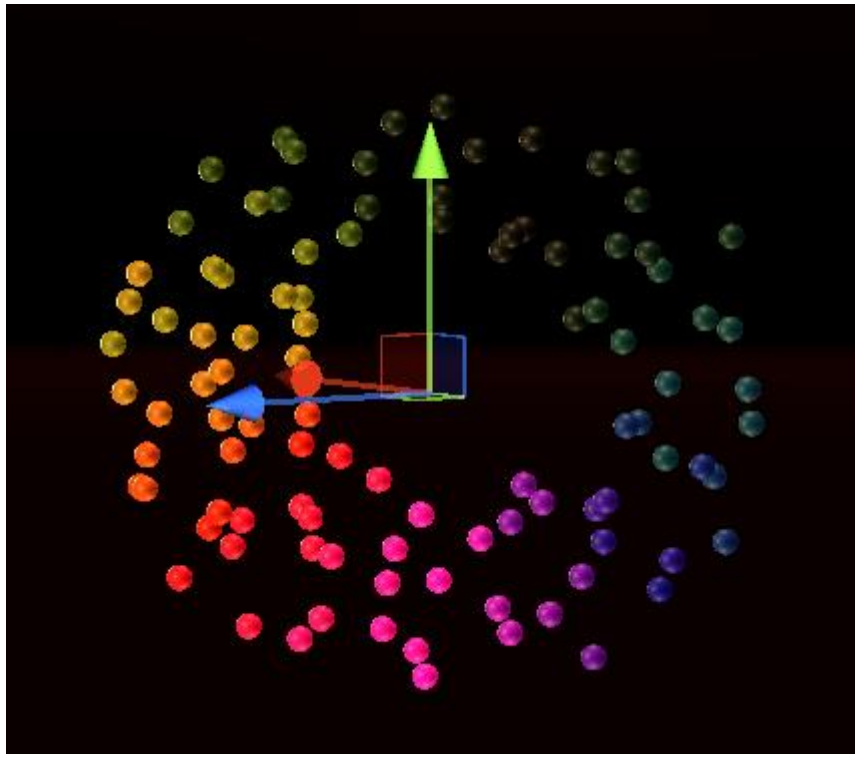
Agenda

- Introduction
- System Design
- Implementation
- Autonomous Behavior
- Results
- **Conclusions**

Conclusions

- The results validate the *way-of-making* for the system.
- There were limitations related with: Equipment, measurement tools, and time restrictions.
- Recommendation: Address limitations in terms of latency, packet loss, and more efficient implementations.
- Future work: Exploration of other algorithms, more participants, and deeper evaluations for human aspects.

Sync-Swarm Exploration: Swarmalators (3D)



$$\dot{\mathbf{x}}_i = \frac{1}{N}\sum_{j\neq i}^{N}\left[\frac{\mathbf{x}_j-\mathbf{x}_i}{\left|\mathbf{x}_j-\mathbf{x}_i\right|}\left(1+J\cos\left(\theta_j-\theta_i\right)\right)-\frac{\mathbf{x}_j-\mathbf{x}_i}{\left|\mathbf{x}_j-\mathbf{x}_i\right|^3}\right],$$

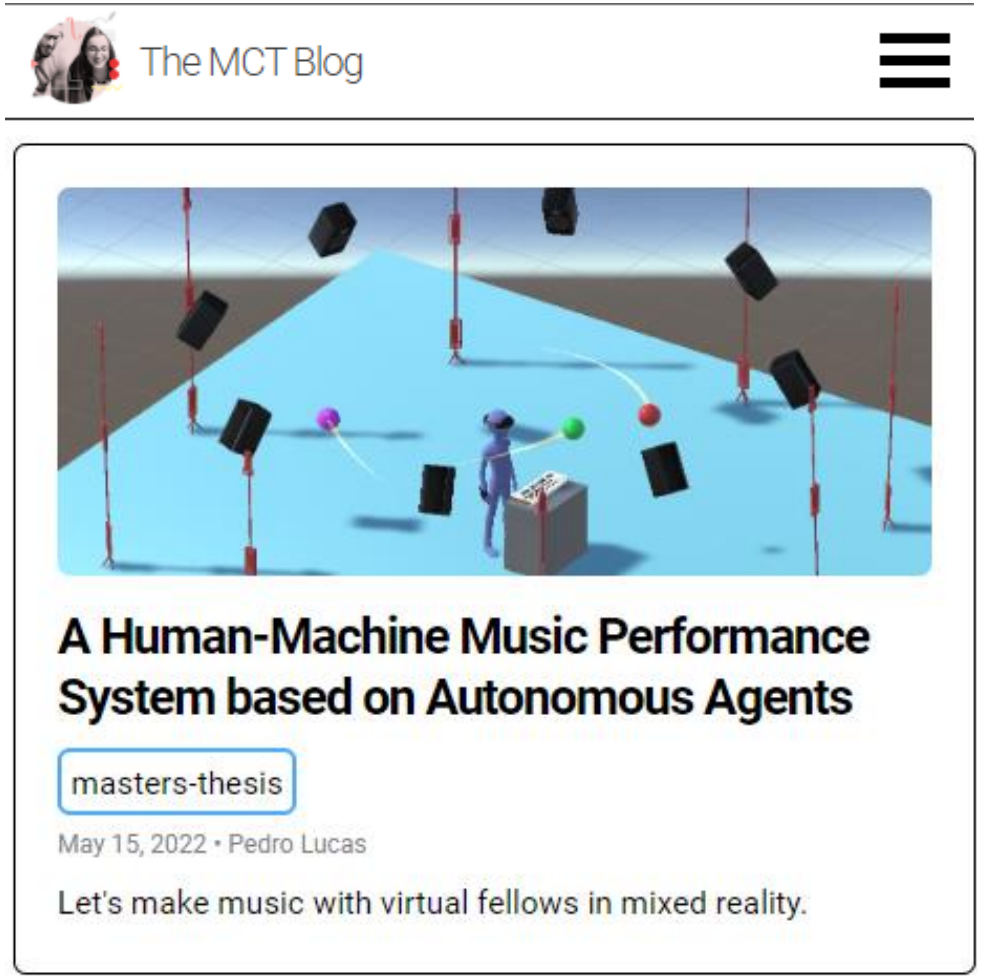
$$\dot{\theta}_i = \frac{K}{N}\sum_{j\neq i}^{N}\frac{\sin\left(\theta_j-\theta_i\right)}{\left|\mathbf{x}_j-\mathbf{x}_i\right|},$$

O'Keeffe, K.P., Hong, H. & Strogatz, S.H. Oscillators that sync and swarm. *Nat Commun* **8**, 1504 (2017). https://doi.org/10.1038/s41467-017-01190-3

For more details check the blog post on the MCT Blog:

- Summary.
- Full manuscript.
- Demo video.
- Audio excerpts.
- Source code.

https://mct-master.github.io/



https://mct-master.github.io/masters- thesis/2022/05/15/pedropl-human-machine-impro.html