Doctoral Dissertation
Doctoral Program in Control and Computer Engineering (38^{th} cycle)

Experiencing the Metaverse: Designing Immersive and Interactive Virtual Worlds

Alessandro Visconti

* * * * *

Supervisors
Prof. Fabrizio Lamberti (Supervisor)
Dr. Alberto Cannavò (Co-supervisor)

Doctoral Examination Committee:
Prof. Lucio Tommaso De Paolis, University of Salento, Italy (Referee)
Prof. Sofía Seinfeld, Universitat Oberta de Catalunya, Spain (Referee)
Prof. Minjuan Wang, The Education University of Hong Kong, Hong Kong, China
Prof. Abdulmotaleb El Saddik, University of Ottawa, Canada
Prof. Valentina Gatteschi, Politecnico di Torino, Italy

Politecnico di Torino
September 18, 2026

I hereby declare that, the contents and organization of this dissertation constitute my own original work and do not compromise in any way the rights of third parties, including those relating to the security of personal data.

. . . . . .

Alessandro Visconti

Turin, September 18, 2026

Summary

The Metaverse, combining "meta" (beyond) and "universe", refers to a decentralized, persistent, and immersive three-dimensional online environment where users, represented by avatars, can interact, create, and collaborate within virtual spaces that are not bound by the physical world. Rather than denoting a single platform or application, the Metaverse is increasingly conceptualized as a persistent network of interconnected virtual worlds supporting multiple forms of human activity, including socialization, cultural participation, education, professional collaboration, and entertainment. Within these worlds, often leveraging immersive technologies such as Head-Mounted Displays for immersive Virtual Reality (VR) experiences, users are not merely passive observers of digital content, but active participants who inhabit spaces, communicate with others, and engage in collective activities. As these enabling technologies become increasingly accessible, virtual worlds are increasingly used as platforms capable of hosting social and collaborative experiences. Despite this potential, current implementations of immersive social virtual worlds still lack consolidated design principles capable of ensuring usability, presence, and socially meaningful interaction, and the success of these virtual worlds depends on whether they can support interactions that users perceive as engaging, coherent, and meaningful. This doctoral thesis addresses this gap by investigating two tightly interconnected challenges: how the structural design configuration of social virtual worlds influences user perception and engagement, and how social interactions in avatar-mediated communication can be effectively reproduced and intentionally augmented.

The first challenge concerns the structural configuration of virtual worlds, specifically how spatial layout, architectural organization, visual composition, and navigational affordances shape user perception and social presence. Unlike physical architecture, virtual environments are not limited by real-world constraints. Virtual environments characterized by high visual richness but poor spatial legibility can impair users' sense of orientation and consequently constrain exploratory behavior. In contrast, virtual environments that are deliberately structured, through coherent layouts, clear navigational cues, and recognizable spatial hierarchies, can guide navigation and promote engagement. These strategies not only support interaction with virtual objects, but also encourage social behaviors, including user

co-presence, communication, and participation in shared activities. The second challenge concerns avatar-mediated communication. Avatars act as the primary interface through which users express identity, emotions, and intentions, yet current consumer VR systems typically track only a limited set of body movements, creating a gap between how avatars look and how they behave. This can make interactions feel artificial or uncomfortable, undermining social presence and behavioral plausibility, and making the improvement of avatar expressiveness essential for fostering trust and engagement in immersive social environments.

Building on the discussion above, Chapter 1 introduces the research context and motivations. Chapter 2 focuses on the design of immersive virtual worlds and its influence on user experience. Within this context, MetaLibrary is introduced as a case study: a social virtual world designed to support cultural experiences centered on reading, knowledge exploration, and community interaction, deployed at the central public library of Turin, Italy. From this system, a more systematic study combines a literature review, an analysis of existing social VR platforms, and expert interviews, leading to a set of design guidelines for immersive social virtual worlds. A user study then evaluates alternative configurations of key environmental parameters, including lighting conditions, spatial density, and scene richness. Chapter 3 addresses avatar-mediated communication and non-verbal behavior through four studies. A first study examines emotional expression by comparing a hardware-based facial tracking solution with a software-based approach that approximates facial motion from audio input. A second study investigates handshake interaction techniques using handheld controllers, evaluating solutions ranging from abstract interface-driven approaches to gesture-based simulations. A third study introduces a machine learning-based framework to improve inverse kinematics-driven avatar animations using hugging as a representative use case. Finally, a study on Artificial Facial Mimicry (AFM) examines how the controlled amplification of users' emotional facial cues influences dyadic communication in immersive VR.

In conclusion, the research presented in this doctoral thesis demonstrates that the quality of a virtual social experience depends on the interaction of multiple design factors whose combined effect is stronger than any single factor alone. From the perspective of virtual environment design, the studies confirm that increased visual richness does not linearly translate into better experience; socially effective environments emerge from balanced trade-offs between navigability, perceptual load, and interaction-supporting spatial affordances, and that virtual environment design is a matter of coherent compositional choices rather than the optimization of isolated variables. Regarding avatar-mediated communication, facial tracking yields significant perceptual advantages over approximation-based approaches for emotions such as sadness and disgust, while differences are less pronounced for states where speech and body movement provide sufficient contextual cues, suggesting that trade-offs between complexity and expressiveness depend on the context. Gesture-based interactions that resemble real-world behavior are consistently

perceived as more natural than abstract alternatives, reinforcing the importance of supporting users' natural motor behaviors in social VR. The machine learning-based animation framework demonstrated that data-driven approaches can effectively mitigate the limitations of affordable tracking systems, improving the realism of close-contact avatar interactions while highlighting motion smoothness as an area requiring further development. Finally, the AFM study showed that the subtle augmentation of users' facial expressions can significantly enhance interpersonal closeness, trust, and social presence without negatively affecting perceived naturalness, positioning social augmentation as a promising design strategy for immersive environments. All these findings advance a human-centered framework for creating socially meaningful Metaverse experiences, demonstrating that addressing both environmental and behavioral dimensions in an integrated manner is essential for the design of effective immersive social systems.

This Ph.D. thesis has been typeset by means of the TEX-system facilities. The typesetting engine was pdfLATEX. The document class was `toptesi`, by Claudio Beccari, with option `tipotesi=scudo`. This class is available in every up-to-date and complete TEX-system installation.