Politecnico di Torino

ScuDo
Scuola di Dottorato ~ Doctoral School
WHAT YOU ARE, TAKES YOU FAR

Doctoral Dissertation
Doctoral Program in Electric Electronic and Communication Engineering (38.th cycle)

Low Latency, Network Resilience, and Accessibility in Networked Music Performance

Leonardo Severi

* * * * *

Supervisors
Prof. Cristina Rottondi, Supervisor
Prof. Andrea Bianco, Co-supervisor

Politecnico di Torino

Summary

Networked Music Performance (NMP) is an interaction between geographically distributed musicians who perform together in real time over telecommunications networks. NMP systems must satisfy stringent technical requirements that distinguish them from conventional videoconferencing applications, including the minimization of latency and the preservation of high audio quality. Beyond audio, musicians interact through visual cues, primarily represented by gestures, which are essential for coordination and musical expression.

This thesis presents several solutions attempting to make NMP practical for every kind of user, exploring technologies from dedicated hardware to accessible software platforms. At the hardware level, a system based on FPGA technology is presented, featuring a custom processor architecture designed for ultra-low-latency audio handling. This specialized approach demonstrates how dedicated hardware can minimize local processing delays. Complementing this, software-based solutions are explored, including web-based implementations that enable users without technical expertise to participate in high-quality NMP sessions through intuitive interfaces. The practical viability of these systems was validated through a real distributed concert between Turin and Wrocław, demonstrating that current technology can support actual artistic performances across international distances.

Network impairments pose significant challenges to audio quality. To support research addressing these issues, the thesis introduces DUST, a dataset containing UDP audio traces that captures fine-grained information about packet arrivals and playback conditions under diverse network configurations. Building on this foundation, packet loss concealment techniques are assessed for both MIDI and sampled audio. For MIDI, different configurations of MIDI streaming over Real-Time Protocol are benchmarked. For sampled audio, a sparse linear prediction algorithm achieves effective reconstruction with minimal computational requirements suitable for embedded hardware.

Beyond audio transmission, the thesis addresses visual communication essential to musical interaction. A Virtual Reality-based approach is presented where conductor gestures are tracked and transmitted to remote musicians as avatar representations, achieving lower bandwidth than video streaming, addressing the limitation of traditional audio-focused NMP systems that neglect visual conducting cues.

For blind and visually impaired musicians unable to perceive visual cues, haptic feedback offers an alternative. Through interviews with musicians, specifications are developed for wearable devices that translate conductor movements into vibrotactile patterns, with body placement varying according to instrument requirements. The accessibility discussion extends to musicians with auditory or mobility impairments, exploring multimodal feedback and personalized rendering as enabling technologies.