Organised Sound (2026), 1–13
doi:10.1017/S1355771826101617

CAMBRIDGE UNIVERSITY PRESS

Article

Mapping codependency: Intermedial causality in real-time audiovisual performance

Roberto Alonso Trillo[1], Peter A. C. Nelson[2] and Kratika Jain[3]

[1]Academy of Music, Hong Kong Baptist University, Hong Kong; [2]Academy of Visual Arts, Hong Kong Baptist University, Hong Kong and [3]Indian Institute of Technology Kanpur, India

Abstract

This paper examines how multimodal mapping can function as a collaborative methodology in real-time audiovisual performance. It explores how functional codependency and real-time risk create bonds between artists, becoming fundamental narrative and dramaturgical vectors, and revealing the aesthetic influence of the technological systems shaping contemporary practice. Building on technical research into connectivity and latency, we examine the compositional affordances of those systems and how they shape the experience of performers and audiences. We treat telematic performance as an enabling condition for intermediality and develop our argument through two real-time performances built on a recursive mapping process. The first focused on capturing and transforming gestural data from a violin bow into a robotic system for painting calligraphy within a general constraint of telematic data transfer; the second constructed a longer chain of causality between kinetic sculpture, machine vision, generative music and video mapping. We detail a four-stage multimodal mapping methodology and reflect on how performers and audiences perceive chains of dependency in real-time systems, on the broader apophenic tendency to interpret causation and intention in complex systems, and on what this implies for a reconsideration of post-media aesthetics.

Keywords: apophenia; distributed agency; intermediality; multimodal mapping; telematic performance

(Received 15 January 2026 UTC; revised 26 July 2026 UTC; accepted 21 August 2026 UTC)

1. Introduction

The COVID-19 pandemic intensified the use of telematic systems. In response to the practical constraints of physical distancing, disrupted copresence and restricted mobility, many geographically distributed artists adopted low-latency networks, sensor-based ecosystems and real-time data exchange protocols to sustain collective performance practices (Freeman et al. 2021). This revival prompted the present research; by parsing artistic gestures through telematic systems, we were forced to reconsider the affordances of multimodal mapping and multimodality more generally speaking and to explore the aesthetic possibilities of distributed performance systems. We treated networked transmission as an enabling condition for intermedial compositional practices, framing the difference between distributed and telematic systems as a matter of aesthetics and communication latency. In both cases, mapping processes presuppose and articulate a distance between sources and insist on the distinctness of each medium and linked signal. Mapping thus becomes an act of understanding, mediation and collaboration, one that resists the homogenising pull that would reduce heterogeneous signals to a lowest common denominator. Although these systems enabled new forms of mediated interaction, they also exposed a limitation: technical connectivity delivers function, but what of form? More specifically, how does this type of connectivity shape the content and meaning of the performances it facilitates? And how might multimodal mappings be designed to structure genuine artistic codependency between performers and shape how both artists and audiences perceive chains of causality and meaning?

By analysing the iterative prototyping, rehearsal and execution of two live performances, this paper develops a situated methodology, conceptual vocabulary and dramaturgical strategy for contemporary telematic sound-based performance, articulated through a four-stage curatorial structure – data capture, data transmission, signal transformation and mapping and actuation – that frames multimodal mapping as an organisational logic rather than purely technical procedure. Its contribution is twofold: first, it establishes functional codependency a priori, as a compositional principle built into the system, instead of a relational quality read retrospectively; and second, it reframes apophenia – the audience's inference of pattern and agency within an opaque system – as a positive aesthetic condition that such mappings can deliberately elicit.

2. Related work

2.1. Post-media theory

Telematic and interactive practices in sound art and computer music have long engaged questions of distributed agency, feedback and emergence. Sound in these contexts is frequently understood as a relational process shaped by interactions between performers, technologies and environments (Di Scipio 2003; Waters 2007).

Corresponding author: Roberto Alonso Trillo; Email: robertoalonso@hkbu.edu.hk

Cite this article: Alonso Trillo R, Nelson PAC, Jain K. (2026). Mapping codependency: Intermedial causality in real-time audiovisual performance. *Organised Sound*, 1–13, https://doi.org/10.1017/S1355771826101617

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S1355771826101617 Published online by Cambridge University Press

Multimodal mapping, linking gesture, sound, image and physical actuation, has therefore become a common strategy for structuring exchanges across media (Jordà 2005). Existing research, however, has often focused on the individual instrument or interaction (e.g., Rovan et al. 1997; Doornbusch 2002; Hunt and Wanderley 2002; Hunt, Wanderley, and Paradis 2003; Van Nort, Wanderley, and Depalle 2014) with mapping rarely formalised, a priori, as a curatorial principle capable of shaping relationships across an entire performance ecology and embedding its own techno-narrative logic.

This paper addresses that gap between the functional interdependence produced by transmedia telematic systems and their emergent aesthetic language. We argue that aesthetic significance in such works may arise from their operative mechanisms: how artists are connected, as well as how their exchanges are mapped, is in itself a compositional strategy. Connection alone does not carry that significance, and retrospective curatorial framing can only describe it after the fact (Birringer 2008). We seek to demonstrate that the function and structure of connective systems have significant aesthetic impact; in the performances analysed here, they create functional codependencies between artists through real-time interactive mapping, such that actions in one modality – sound, voice, movement or image – directly condition responses in others.

This emphasis on mechanism and relational structure situates the present study within post-media perspectives, which frame practice around processes once transcoding renders distinct artistic mediums interoperable. Krauss (1999) argued that the breakdown of medium specificity, driven by conceptual art and the broader symbolic homogenisation of market forces, shifted aesthetic value towards recursive structures that emerge from differences within a system. Manovich (2001, 2005) argued that contemporary cultural production is driven by shared computational frameworks that allow distinct inputs – from moving images to sounds, voices, motion data and so on – to be treated as comparable and interoperable signals, so long as they can be normalised and mapped into exchangeable formats. Following Krauss's critique, this interoperability could flatten aesthetic value into pure equivalence, or it could generate differential meaning, extending curatorial structures of orchestration – in the case of sound, beyond pitch, harmony and tempo – into causal relationships across artforms. The performances discussed here pursue the second possibility.

This interoperability, however, does not dissolve the differences between media, and our argument depends on holding two levels apart. Although transcoding renders sound, image, gesture and control data mutually convertible as numerical streams, each medium retains an irreducible material and sensory specificity that shapes its affordances; a robot arm cannot move like smoke, sound is felt in the body where text is read, a projected image collapses depth into surface, etc. The aesthetic significance of the works discussed here lies precisely in the friction between these two levels, between the smooth interoperability of the data and the resistant particularity of each medium. In foregrounding this friction we align with a long line of thought on medium and inter-medium: Higgins's account of intermedia as work falling conceptually between established media (Higgins 2001); the post-medium condition outlined above; Pethő's (2011) reading of intermediality as an aesthetics of the 'in-between'; Smalley's (1997) notion of source bonding, in which listeners bind electroacoustic sound to imagined physical causes; Chion's (1994) synchresis and 'added value', which describe how sound and image transform one another rather than merely coexist; and the ecological concept of affordance (Gibson 1979), which frames each medium as offering and constraining particular actions. Through these lenses, the couplings in our performances become legible as compositional choices rather than arbitrary connections. The violin bow affords continuous, trained gestural nuance, which the UR3e robot arm resists; its joint, force and workspace limits, as well as the calligraphic requirement that translation and rotation be coupled, turn the machine's resistance into compositional structure (as detailed in the first case study below). The kinetic sculpture affords a heavy, periodic tempo and a legible mechanical cause; smoke affords stochastic, non-repeatable turbulence that no performer fully controls; spatialised sound affords an immersive, embodied field and the perception of a causal source; projected video affords simultaneity and the compositing of surfaces; the chroma-key body affords a live, copresent figure that doubles as a portal to generated image; and large-language-model text affords semantic drift and multilingual meaning that stays legible where the other media turn opaque. In each case the mapping is chosen so that a medium's affordances and constraints do compositional work, which is precisely what separates an intermedial composition from an indeterminate flattening of difference.

Within this expanded post-media framework, artistic meaning arises through the orchestration of relations, constraints and feedback across heterogeneous systems. Recent media theory emphasises how such systems operate beyond immediate legibility, producing experiences characterised by opacity, modulation and inference (Hansen 2015; Hayles 2017), aligning artistic experience with broader encounters with computational models that exceed human-scale perception. As agency is redistributed across human and technical actors, the legibility of cause and effect becomes contingent, situational and frequently opaque, and in multimodal and telematic contexts only partially accessible to performers and audiences. Mapping design therefore becomes a critical compositional concern, as mappings determine how signals circulate within a system and how responsibility, intention and narrative coherence are inferred (Baalman 2022).

More recent accounts go further; post-media conditions also participate in the production of subjectivity. Drawing on historical analyses of broadcast, immersive and networked media, Allado-McDowell (2024) argues that media structures operate as interpellative systems, shaping how individuals perceive themselves, others and their capacity for action within a given environment. Interoperability is thus a historically situated condition through which agency is distributed and identities are negotiated, and meaning in post-media contexts emerges from the structural organisation of how signals, bodies and environments are coupled. This reinforces the view of multimodal mapping as a compositional practice that configures relations of dependency and responsibility, rather than a neutral technical affordance enabling cross-media translation.

2.2. Apophenia as aesthetic condition

This shift towards relational orchestration draws both performers and audiences into processes of inference that align with what has been described as apophenia: a mode of sense-making that emerges when agents confront systems whose internal operations outrun human interpretation, whether through deliberate concealment, through distributed opacity or through an unresolvable scalar mismatch between machine operation and human reasoning that persists under full disclosure (Burrell 2016; Steyerl 2016; Steyerl

https://doi.org/10.1017/S1355771826101617 Published online by Cambridge University Press

2018; Apprich et al. 2019). Apophenia here names more than a tendency to misrecognise patterns; it describes how meaning, causality and intention are actively constructed in response to indeterminate or overdetermined technical conditions. Recent theoretical work treats this as a structural feature of engaging with generative and model-driven systems (Bratton and Agüera y Arcas 2022). In such contexts, pattern recognition does not reveal stable underlying structures so much as generate new meaning through probabilistic inference, projection and expectation. In complex orchestrations, part of the aesthetic charge lies in this indistinction; we *feel* the visceral moments of connection mapped into the system alongside the connections and relationships that we *perceive* or *imagine*, without always being able to tell which is which.

Cognitive theories of perception describe the same inferential work. In the lineage running from Helmholtz's *Unbewusster Schluss* (unconscious inference – Helmholtz 1867) to contemporary predictive accounts (Clark 2016; Seth 2019), perceiving is itself constructive; the brain builds models of the world, generates related expectations and revises those against incoming sensory evidence. Perception amounts to a 'best guess' at the root causes of incoming sensory signals, or in Christopher Donald Frith's words, to building 'fantasies that coincide with reality' (Frith 2007: 134). Metahaven (2026: 8–9) sets this alongside Gombrich's beholder's share (Gombrich 1960), the viewer's active supplying of what the image only implies, and argues that art explores such best-guessing 'as a form of culture': perception as an endlessly inconclusive best-guessing process (Metahaven 2026: 24). From this perspective, a system that withholds its causal structure simply lays bare the guesswork perception was already performing. Apophenia, read this way, names an intensification of ordinary perception, one increasing in parallel to the beholder's share.

The prevalence of this condition is evident in large-scale visualisations that claim to render complex climatological data intelligible (Anadol 2023) or the promethean anthropomorphism we use to interpret glitches and inaccuracies in generative AI systems (Alonso Trillo and Poliks 2025). As Steyerl notes, the inability to apprehend intelligible meaning within vast streams of data has become a cultural 'new normal' that is entirely congruent with the perceptual work described above, our propensity to project significance onto partial or random data (e.g., seeing faces in clouds).

Within multimodal performance systems, this foregrounds how relational dependencies are sensed, narrated and negotiated. Multimodal mapping, therefore, operates as a generative constraint – instead of a transparent translation – inviting performers and audiences to infer agency and coherence across distributed processes. In this sense, mapping becomes a strategy that deliberately stages uncertainty, risk and interpretative projection as central aesthetic forces. Although performers operate within tightly interdependent technical networks, audiences may construct multiple and sometimes contradictory interpretations of these systems, including reversals of relational dependency or the perception of connections that are not materially present. For these reasons, this paper treats apophenia as a core aesthetic condition of post-media, data-driven performance; the point at which an audience, unable to fathom the whole system, composes the coherence it perceives.

2.3. Distributed sound practices

Research in telematic performance and distributed sound practice has expanded significantly over the past two decades, in an attempt to enable real-time collaboration across distance through advances in network infrastructure, audio transmission and latency management. Although these developments have made distributed performance increasingly accessible, connectivity has commonly been pursued as a primary aim, with the structural organisation of artistic collaboration (questions of aesthetics, artistic dependency, shared responsibility and perceptual causality) receiving comparatively less attention.

Early examples of explicitly aesthetic engagement with networked codependency can be found in the work of the League of Automatic Music Composers, who from 1978 built performances around interacting microcomputers whose mutual, and often unpredictable, influence was treated as artistic substance (Bischoff et al. 1978), an approach extended by their successor ensemble The Hub (Brown and Bischoff 2002). In a related vein, Di Scipio's ecosystemic practice reframed sound as the emergent product of feedback among performers, systems and environments (Di Scipio 2003, 2008), and Fiebrink and Caramiaux (2018) have argued that mapping and interactive machine learning are themselves creative and aesthetic tools. We build on this lineage and attempt to formalise codependency a priori as a curatorial principle, treating the resulting perceptual opacity as generative.

Early telematic theory foregrounded presence, authorship and collective experience (Ascott 1984, 1990, 2003; Dixon 2007). These prove difficult to sustain across a network, and research in sound-based telematic practice has concentrated on high-quality audio streaming and precise temporal alignment, working to the psycho-perceptual thresholds that set ensemble latency tolerances. Systems such as JackTrip and the LoLa low-latency platform exemplify this approach by treating the network primarily as a transmission substrate whose influence on musical interaction should be minimised through protocol and infrastructure design (Cáceres and Chafe 2010; Drioli et al. 2013; Rottondi et al. 2016; Lakiotakis et al. 2019; Dürre et al. 2022). That minimisation suited the aims of that research, which sought reliable ensemble synchrony. We take these systems as enabling infrastructure and pursue a complementary question of intermedial aesthetics: how the structure of the connection shapes the compositional form it carries?

Parallel strands of practice have instead incorporated delay and instability as compositional material. Works by Oliveros and Chafe reframed latency, echo and temporal displacement as productive musical conditions (Chafe and Gurevich 2004; Oliveros 2005), with more recent examples exploring latency as an aesthetic parameter within networked improvisation (Liloia and Dannenberg 2025). Although these approaches expand the aesthetic vocabulary of telematic music, they remain primarily focused on sonic temporality and do not explicitly address how collaboration across multiple media is structurally organised. As a result, latency may be aestheticised, but agency remains largely individual and musically bounded.

Consistent with the post-media paradigm outlined above, more recent research has extended telematic systems beyond audio to include sensors, physical computing, robotics, immersive audio and environmental data. Within the Internet of Musical Things and related literature, distributed performance is framed as an ecosystem connecting performers, devices, spaces and networks (Turchet et al. 2018; Barthet et al. 2019; Turchet and Tomasetti 2023). However, much of this work remains oriented towards interoperability, scalability and infrastructural robustness. Thus, although aesthetic questions of artistic codependency and shared agency are frequently addressed at the level of individual works, we are concerned with their structural embedding in system design

https://doi.org/10.1017/S1355771826101617 Published online by Cambridge University Press

and how they intersect with broader questions of mapping and intermediality.

2.4. Mapping, interaction and sound as process

In interactive music and sound art, mapping – the correspondence between a performer's control parameters and the sound and media parameters they drive – has long been treated as a central compositional and organisational concern. Jordà (2005) demonstrates that such strategies strongly influence performer agency and the perception of causality, and Hunt et al. (2003) show how mapping choices shape engagement and expressivity. Taken together, these studies treat mapping as the site where relationships between performers, systems and outcomes are actively configured.

Relatedly, sound itself has been theorised as a process. Di Scipio frames it as the product of ecosystemic feedback (Di Scipio 2003, 2008), and Waters discusses performance as an evolving ecology (Waters 2007). Both shift analytical focus from isolated gestures and outputs towards relational processes and system behaviour, with aesthetic form arising from organisational dynamics.

Although frameworks such as EyesWeb (Camurri et al. 2004), Libmapper (Malloch et al. 2013) and Wekinator (Fiebrink 2009) open mapping to exploratory and aesthetic experimentation, we go a step further and treat it, a priori, as a curatorial framework that formalises codependency and creates shared responsibility and aesthetic consequence across a whole collaborative performance. That responsibility is operational; the chain produces work only when every link carries a signal. When a node drops out – a sensor, a stream, a performer's gesture – the failure propagates to every medium downstream, thus keeping the signal alive is a condition of the work existing at all, and a demand made of everyone at once.

In this spirit, the works discussed here treat mapping between data formats as an orchestrating mechanism, determining how agency, causality and communication are spread across participants and systems.

2.5. Distributed agency and audience perception

Scholarship on mediated performance has increasingly described contemporary works as assemblages of human and non-human actors, distributing agency across a spectrum of bodies, software, machines and infrastructures (Birringer 2008; Salter 2010). From this perspective, system architecture actively shapes artistic interaction; our aim is to carry this well-established conceptual shift fully into system design, treating structural decisions (parameter coupling, thresholding, and constraint) as primary aesthetic drivers and compositional affordances in their own right.

Their consequences become particularly evident in how multimodal systems shape performer experience and audience perception. With indirect, layered, or partially opaque mappings, audiences may infer causal relationships across sound, movement and visual change without clear access to how – or even whether – these relationships are being structured. As Steyerl's account of apophenia suggests, meaning in complex systems is often constructed through perceived rather than explicit or legible signals (Steyerl 2016). Although interpretive ambiguity has been framed in sound-based digital art as a productive aesthetic condition (Whitelaw 2008), it is rarely examined in relation to the intentional design of mapping structures themselves.

Through two multimodal performance case studies, this paper examines how mapping structures, understood as compositional devices, shape both artistic collaboration and audience perception.

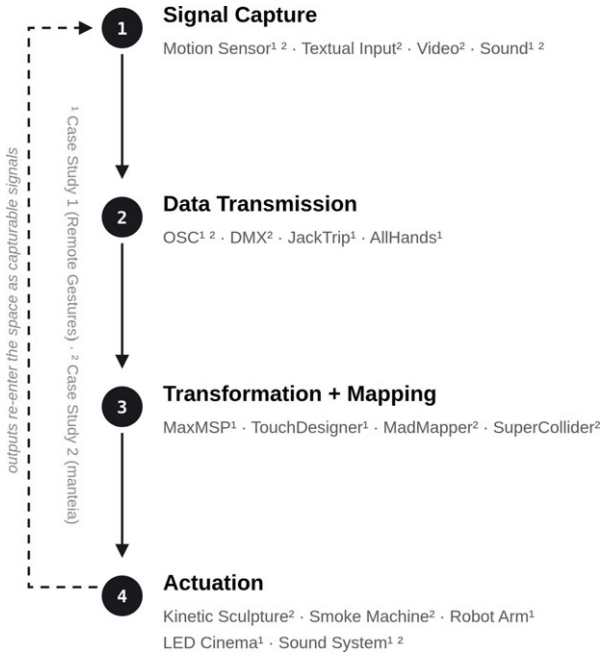


Figure 1. The four-stage multimodal mapping framework. Each actuated output re-enters the space as a capturable signal (dashed return), closing the loop.

It asks how artists negotiate chains of dependency in real time, and whether audiences grasp these mappings directly or arrive at them through inference.

3. Multimodal mapping framework

The multimodal mapping framework presented here formalises the compositional approach used across the performances discussed below (see Figure 1). It specifies how signals are captured, circulated, transformed and enacted to operationalise codependency within real-time performance systems. The framework introduces no new conceptual claims but provides a concrete structure for implementing the mapping strategies discussed earlier, making explicit where agency is coupled, where constraints are imposed, and how causal relations are distributed across media. It is articulated through four interdependent stages – Signal Capture, Data Transmission, Signal Transformation and Mapping, and Actuation – which together define the system conditions under which collaboration, risk and interpretative ambiguity emerge. These conditions link performers together into networks of signal exchange and collaboration, through mappings underpinned by variable degrees of legibility; the framework treats the transparency of the mapping mechanisms as a compositional parameter that can be tuned towards disclosure, opacity or a calibrated alternation.

1. *Signal capture* involves identifying and acquiring aesthetically significant data streams within an artist's existing practice. Our system privileged embodied gestures and material processes – such as violin bow motion, audio dynamics, mechanical rotation or environmental phenomena like smoke turbulence – rather than abstract or arbitrarily generated inputs. By grounding capture in existing artistic vocabularies, expressive intent was preserved while being made available for transmedial coupling.
2. *Data transmission* ensures the reliable and low-latency transmission of captured signals between performers,

https://doi.org/10.1017/S1355771826101617 Published online by Cambridge University Press



Figure 2. A violinist plays when the UR3e robot arm paints in the foreground based on their bow motion. A close-up of the painting process is projected onto the 360° screen.

devices and processing environments. Protocols including Open Sound Control (OSC), Digital Multiplex Signal (DMX) and WebSockets were selected according to the temporal sensitivity and semantic role of each signal. Although technical robustness was essential, transmission was treated as an enabling condition rather than an aesthetic endpoint; the network functioned as a shared substrate through which agency was distributed.

3. *Signal normalisation and mapping* constitutes the central compositional layer of the framework. Incoming signals were normalised to operate as control parameters within another artist's medium, approaching mapping as a curatorial decision that determined affordance correlations across expressive fields. It embedded dependencies through non-linear relations, thresholding and constraint-based mappings designed to balance responsiveness with system stability.
4. *Actuation and output* are where mapped signals are made perceptible as changes in sound, image or physical movement. Outputs were intentionally material and typically legible, allowing performers and audiences to register the effects of distributed action. Actuation closed a feedback loop in which performers adapted their behaviour in real time, reinforcing the relational and ecological character of the performance.

4. Case study 1: multimodal mapping in telematic violin-robot performance

The first case study (Remote Gestures 2024) functioned as an initial testbed for the framework (see Supplementary Video 1). The performance connected two violinists, a composer, a graphics artist and a robotic arm equipped with a calligraphy brush, linking violin performance to generative visual processes and robotic painting, and examining how mapping design established functional codependencies between gesture, visual generation and mechanical actuation (see Figure 2).

4.1. Concept and system architecture

The work was conceived as a distributed duet that displaced conventional instrumental boundaries while preserving embodied musical intent. Two violinists performed synchronously using the MetaBow, an embedded-sensor-augmented violin bow that captures inertial motion data without disrupting traditional playing technique (Alonso Trillo et al. 2023, 2025). Real-time gesture data from the violin bows was transmitted to the sonic and graphics artists as OSC messages using a lightweight Python implementation of the AllHands protocol (Alonso Trillo 2025) whereas high-fidelity audio was exchanged via JackTrip (Cáceres and Chafe 2010) to maintain temporal coherence across locations. The first case study was thus genuinely telematic; the two violinists performed from separate physical locations, with the network operating as a live channel of distributed transmission (see Figure 3).

Rather than mapping gestures locally to sound alone, all incoming data streams were routed into a shared orchestration environment implemented in TouchDesigner, which functioned as the system's compositional core. Motion data, extracted audio features and robotic control parameters were jointly processed and redistributed across three output domains: immersive visuals, robotic calligraphic actuation and live musical response. This configuration positioned the mapping layer as a site of negotiated dependency, where violin gesture simultaneously conditioned sonic articulation, visual behaviour and robotic mechanical inscription, binding performers and systems into a single causal ecology rather than parallel media streams.

4.2. Multimodal mapping

Several mapping strategies were employed to embed functional codependency. The first part of the performance was a duet, with one violinist playing inside the immersive 360-degree cinema and the other off-site. Real-time accelerometer data were mapped to a histogram-like trace, each player's line moving around the circumference to visualise a minute of playing (see Figure 4). This spatialised the individual gestures of the telematic duet while rendering their temporal accumulation collectively visible. In parallel, a root mean square (RMS) power feature extracted from each violin's sound was mapped to the RGBA (Red, Green, Blue, and Alpha) channels of the visual output, so that as a violinist altered their sonic intensity, the brightness and hue of the performance environment shifted in step, coupling sonic articulation to the evolving image.

In the second half of the performance, the now colocated violinists' bow motion was mapped to a Universal Robots UR3e

https://doi.org/10.1017/S1355771826101617 Published online by Cambridge University Press

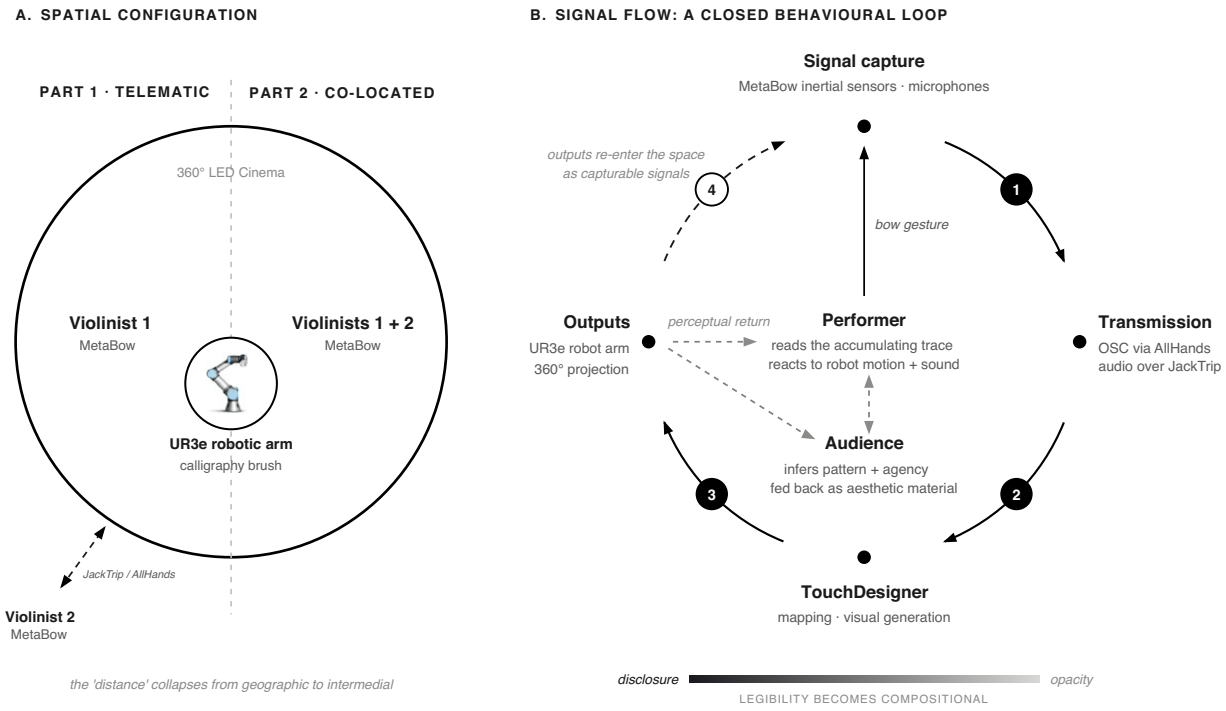


Figure 3. System architecture diagram for *Remote Gestures*, showing the spatial configuration of the performance (left) and the data flow from violin sensors and microphones through AllHands/JackTrip to TouchDesigner, and out to the 360° projection and UR3e robot arm (right).

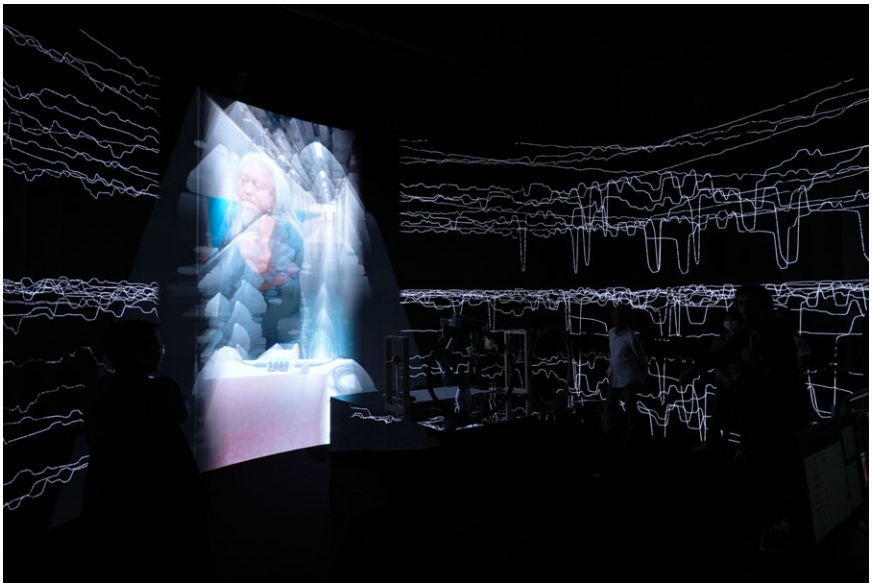


Figure 4. Histogram trace animation driven by motion from the Metabow from performers both inside and outside the 360° cinema from *Remote Gestures*.

robotic arm fitted with a Chinese calligraphy brush. Two accelerometer axes controlled translation and rotation of the robot tool centre point (TCP), deliberately coupled as interdependent parameters. This correlation emulated the embodied constraints of brush calligraphy; rotation without translation produced suspended, abstract motion, when translation without rotation resulted in mechanical plotting. Only when the two were coupled did bow gestures yield strokes matching the violinists' expectations, a finely tuned mapping developed through repeated rehearsal.

The UR3e enforces strict limits on joint configuration and applied force, triggering shutdowns if exceeded. To preserve expressive range while maintaining system stability, rotational freedom was dynamically constrained as a function of the TCP's absolute position. As the brush approached the limits of its reachable workspace, permissible rotation was progressively reduced in the corresponding direction. This position-dependent clamping embedded physical limitations directly into the mapping logic and defined the creative musicians' experience of painting-as-playing. Here the robot's material constraints functioned as a

https://doi.org/10.1017/S1355771826101617 Published online by Cambridge University Press



Figure 5. A violinist performs with the real-time calligraphy system in *Remote Gestures*.

media-specific affordance, doing compositional as well as technical work.

Finally, audio amplitude was mapped to the vertical position of the brush using a threshold-based approach. Contact between the brush and the paper occurred only when the violin's audio level exceeded a predefined value. This allowed performers to position the brush in advance while retaining sonic control over the moment of inscription, reinforcing sound as the primary trigger within the multimodal system. Despite the necessarily mechanical nature of these mappings, the result was a collaborative assemblage that the composer, performers and graphics artist could all, in effect, 'play' together. Subtle shifts in sound normalisation, hue and brush height accompanied the piece as it unfolded, continually reacting and recalibrating to the musicians' evolving gestures (see Figure 5).

4.3. Performance and perceptual observations

The performance unfolded over approximately 15 minutes and was presented several times within the Visualization Research Centre (VRC), a state-of-the-art immersive 360-degree cinema environment at Hong Kong Baptist University. A consistent perceptual asymmetry emerged between performers and audiences; although the former experienced a direct and legible coupling between bow gesture and robotic motion, observers initially perceived the robot's behaviour as indeterminate or partially autonomous.

To work through this perceptual gap and gauge the system's levels of legibility, the violinists instinctively exaggerated their gestures during the opening section, in an unconscious attempt to externalise the mapping logic. As perceptual familiarity with the system developed, gestural articulation became increasingly restrained, a shift that aligns with accounts of telematic performance in which liveness and causality are not self-evident properties of a system but must be actively staged and perceptually reinforced (e.g., Dixon 2007; Auslander 2008).

It is worth distinguishing three levels here: the compositional intention behind a mapping, the performer's lived experience of it and the audience's inference from its effects. These need not coincide (e.g., Gurevich and Fyans 2011). In retrospect, the opening gestures operated in a mode of disclosure and the restraint that followed marked a gradual shift towards opacity, one based on a shared sense of increased awareness; legibility became compositional. Disclosure and its withdrawal emerged through a behavioural loop internal to the event, one more circuit within the assemblage of human and non-human agents that the mapping had bound together. In this process, the audience's inference becomes aesthetic material in its own right, interpretation could expound on the shared awareness of connectivity, but as that connectivity became more complex, the apophenic construction of pattern and agency emerged as a higher-level aesthetic condition created by the underlying logic of our multimodal mapping.

Informal audience feedback ($n = 7$; see Appendix A) consistently identified the robotic calligraphy as the most salient element of the work. From the perspective of the proposed framework, however, the primary outcome lay less in audience comprehension than in the performative consequences of the mapping structure itself. The embedded dependencies constrained the autonomy of individual performers, requiring continuous negotiation and mutual attentiveness. In doing so, the system shaped the improvisational dynamics of the performance and presented codependency as an embodied and perceptually consequential condition rather than a conceptual premise.

In addition, the performative dimension transcended what may at first appear as simple linear or hierarchical causation chains: gesture conditioning sound and image, sound triggering inscription and the robot as a passive reactive constituent. Although the human was typically the first mover, subsequent gestures were already staged within a closed loop; performers played partly by reading the accumulating visual trace on the screen, treated the robot arm rather than the bow as the instrument they were steering and continuously adjusted their improvisation both to the robot's

https://doi.org/10.1017/S1355771826101617 Published online by Cambridge University Press

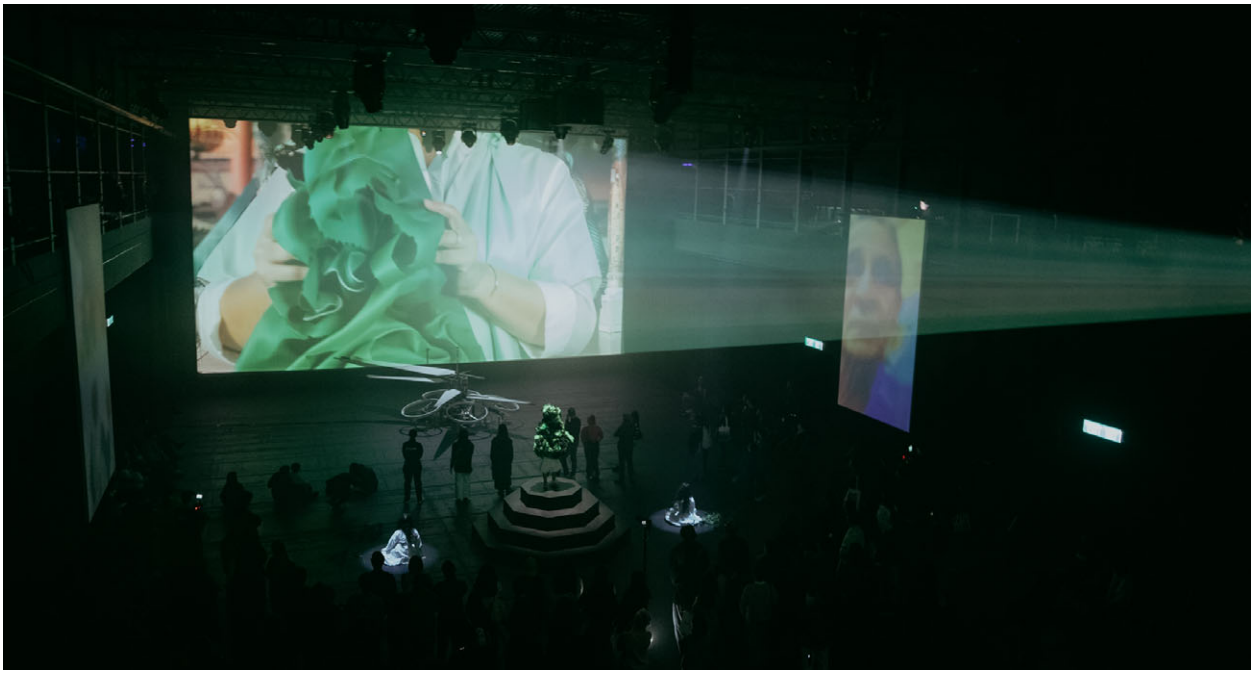


Figure 6. *manteia* saw the creation of a larger ecology of artists and linked works.

movement and to the incidental sounds the arm itself produced. Control and response were thus coconstituted in real time – a form of circular causality in which, as Di Scipio observes of ecosystemic performance, the performer operates inside the loop rather than upon it (Di Scipio 2003, 2008). Comparable arguments recur across cybernetic and human–machine performance, from Pask's Musicolour, whose shifting responses actively recondition the player (Haque 2007), to contemporary robotic musicianship, in which a robot's output measurably reshapes its human partners' timing and phrasing (Bretan and Weinberg 2016). We frame this as a specific form of codependency here: an asymmetric but genuinely reciprocal coupling rather than a symmetrical exchange.

5. Case study 2: multimodal mapping as collective narrative ecology

The second performance (manteia 2024) extended insights from the first by scaling the multimodal mapping framework to a larger ensemble and more heterogeneous media (see Supplementary Video 2). Here, mapping functioned not only to connect performers, but also to produce a collective narrative through layered chains of causality complex enough to exceed immediate legibility. In contrast to the first case study, all performers, instruments and audience shared a single venue (see Figure 6).The distance articulated by the mapping chain was therefore intermedial, produced by couplings between media instead of physical separation.

5.1. From technical exchange to narrative emergence

The project began with a series of workshops in which participants articulated their practices and collaboratively identified sites of potential signal exchange. The performance's conceptual and narrative logic emerged through iterative experimentation, as connections between heterogeneous systems were established, negotiated and reconfigured. This process culminated in a 45-minute performance exploring themes organised around divination, artificial intelligence and apophenia.

5.2. System architecture and mappings

Much like the first performance, signals from a variety of sensors were routed through a central computer running MadMapper, which mapped and redirected the various artistic modalities via OSC and DMX. Like a Rube-Goldberg machine, the performance unfolded as an ecology of causal events, in which the data created by one artist became the input for the next.

As the performance developed, several media were layered into this ecology: a kinetic sculpture, turbulent smoke, generative music, AI-generated sound, video and text, spoken word and a performer whose body doubled as a projection surface. Some of these couplings remained relatively legible to audiences (e.g., the patterns read from the smoke), whereas others (e.g., the sound those patterns generated) resisted stable interpretation, foregrounding the stochastic behaviour of the system (see Figure 7). As these accumulated, the performance thickened into a dense, fast-shifting field of sound, smoke, projection and light. This density made it impossible for viewers to infer linear relationships in real time; yet the anticipation of causality, together with the libretto's themes of oracles and divination, encouraged them to build their own narratives across the media.

Let us consider these couplings in greater detail:

1. Kinetic sculpture to sound: the rotation of the sculpture's radial arm was captured as discrete potentiometer values, alongside a contact microphone on its rotor, transmitted over OSC and mapped to the fundamental, overtone, bass-kick and drone-synthesis parameters of a spatialised composition, with the arm's motion also gating the subwoofer array. The heavy, periodic machine thus became the conductor of the sonic architecture, giving the ecology a

https://doi.org/10.1017/S1355771826101617 Published online by Cambridge University Press

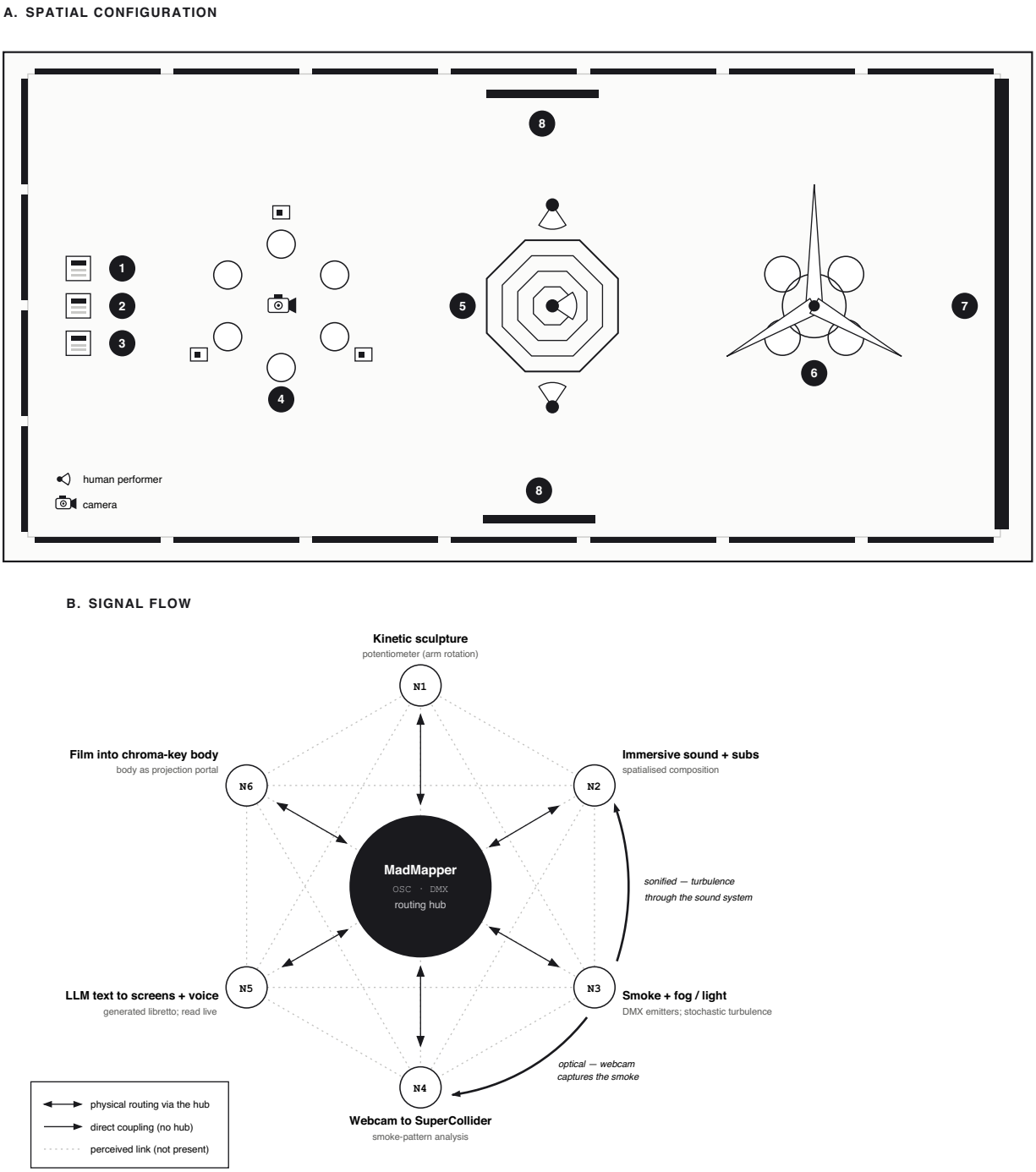


Figure 7. System architecture diagram for *manteia*. Top: the spatial configuration of the performance, keyed as 1 audio computer, 2 central control computer ('brain'), 3 lighting control system, 4 smoke sculpture, 5 central performance platform, 6 kinetic sculpture, 7 main projection screen and 8 side projection screens, with human performers and cameras also marked. Bottom: the data flow from the kinetic artworks with sensors, routed through the central 'brain' via MadMapper and back out to other artworks, screens and speakers.

common ground against which the other, more volatile media could be read.

2. Smoke to sound and light: DMX-controlled smoke machines produced turbulent plumes that were captured by a 360° optical camera and analysed in SuperCollider as particle position, velocity and angle. These features were then used to drive adaptive granular sound synthesis models in SuperCollider and, recursively, the DMX fog, fans and lighting. Air was chosen as an intrinsically unstable, non-repeatable medium precisely because it supplied the system's stochastic, living variability (in contrast to the rotating sculpture) a turbulence that no performer fully controls.
3. Language model to libretto and voice: fine-tuned large language models (LLMs) generated a bilingual (English and Chinese) text projected across two side screens. These were read and sung by a performer in the role of an oracle and filtered through real-time concatenative and granular vocal

https://doi.org/10.1017/S1355771826101617 Published online by Cambridge University Press



Figure 8. Turbulent patterns in smoke are detected via webcam and drawn in real-time via Supercollider in *manteia*.

synthesis approaches in Max. Routing generated language through a synthetic-then-human voice foregrounded the work's theme of machine divination and let semantic drift become audible rather than merely legible. The transformer is here, to use Metahaven's wording, 'opaque as a causal agent', while remaining 'a machine for attention that weighs all parts of its input at once' (Metahaven 2026: 137). Casting it as an oracle made that double character audible: a source whose pronouncements invite interpretation precisely because their causes cannot be traced.

4. Body to image: the oracle performer wore a chroma-key-green costume that was keyed in TouchDesigner into a mobile projection surface onto which an AI live-generated film was composited, turning the body of the performer into a portal between physical presence and the generated moving image.

As we noted earlier, the media were bound in a sustained micro-to-macro dependency; no single medium acted alone, and the removal of any one node would have reconfigured the whole. This reciprocal reliance became a defining constraint on our compositional approach and, in performance, a source of unusual intimacy among the artists, each audibly and visibly dependent on the others. In this scenario, apophenia functioned as an aesthetic approach rather than a perceptual defect. By incorporating complexity and partial opacity into the mapping structure, the framework thematises modern encounters with intelligent and inscrutable technological systems. Multimodal mapping thus operated as connective infrastructure and dramaturgical device, as well as a conceptual metaphor for distributed agency in computational culture. As with the first performance, this dependency was absolute: sound generation would not have been possible without the signals from the kinetic sculpture, the granular sound synthesis could not function without trace signals from the smoke (see Figure 8), the performer needed the LLMs to produce the libretto and the animators needed the performer's body to create the colour-key that revealed their work.

6. Discussion

Taken together, the two case studies demonstrate how the proposed multimodal mapping framework scales from tightly coupled performer–machine interaction to expansive performance ecologies. Across both works, mapping operates as a curatorial and compositional structure that organises relations between heterogeneous media, enforces functional codependency and redistributes agency across human and technical actors. In doing so, the framework directly addresses the gap identified in the related work between technical interdependence and aesthetic consequence, showing how the structure of connectivity itself becomes a primary site of meaning-making. Crucially, codependency becomes a compositional principle; performance then unfolds its consequences, dramaturgical and perceptual, which the discussion below takes up in turn.

The first concerns how technical codependency transforms collaborative practice. By threading extended causal paths between performers, environments and physical actuators, the systems compel participants to develop a situated understanding of one another's gestures, materials and failure modes. Unlike modular or loosely coupled systems, where elements may be exchanged or removed with limited consequence, mapping chains render each component structurally consequential. This produces a condition of shared risk, aligning collaboration with attentiveness, responsiveness and negotiated responsibility. In this sense, 'keeping the signal alive' unifies a diverse set of artists and operationalises post-media accounts of distributed agency by making system architecture itself an active determinant of artistic relation rather than a background condition.

https://doi.org/10.1017/S1355771826101617 Published online by Cambridge University Press

The case studies also demonstrate how narrative coherence can arise as an emergent property of system design rather than as a theme imposed in advance, independently of the form. In both performances, dramaturgical structure evolved through the affordances and constraints of the mapping systems themselves. Decisions regarding parameter coupling, normalisation, thresholding and constraint had direct effects on pacing, tension and rupture, confirming mapping as a compositional strategy through which form is discovered through interaction.

Most significantly, the works reveal how the audience's aesthetic framing shifts as system complexity increases. In both performances, audiences apprehend the work apophenically, inferring pattern, intention and agency from what they perceive; what changes is the register of that inference. In the first case study, the couplings between gesture and outcome could be progressively read, holding inference close to the system's actual structure and grounding a sense of legibility and liveness. In the second, the density and indirection of the mappings pushed the same inference past immediate comprehension, raising it to a different qualitative level in which pattern and agency are attributed to an opaque system. Far from a failure of communication, this apophenic engagement is the work's primary aesthetic condition, shifting audience attention from how the system functioned towards why it behaved as it did. Read together, the case studies mark two chosen points on this deliberate spectrum; the first leans towards disclosure and progressive legibility and the second towards opacity. Neither is privileged; legibility is one compositional strategy among several, and treating it as such aligns the work with the issue's interest in non-narrative frameworks, open works and the phenomenology of perception. At its opaque limit, such a system approaches a non-narratable world, a configuration that 'defies and surpasses the premises of intuitive, everyday experience' and remains cognitively abstract (Metahaven 2026: 171–2). The second performance stopped short of that limit, and it was the audience's apophenic inference that held it back; narrative persisted precisely where causal tracing failed. What the disclosure-opacity spectrum composes, then, is the distance between a system's causal structure and the stories that can still be told about it.

This framing resonates with wider encounters with computational systems that remain experientially consequential even as they exceed comprehension. Multimodal performance thus becomes a site for rehearsing and reflecting upon these conditions, inviting audiences to engage with uncertainty, projection and attribution of agency in embodied and time-based forms. Even when largely invisible, the mapping protocol exerts a formative influence on perception and interpretation, shaping meaning through structure rather than explanation.

7. Conclusion

This paper has introduced multimodal mapping as a compositional methodology for the development of deeply integrated collaborative performance systems operating across sound, image and physical actuation. Through two case studies, it has shown how chains of technical dependency function as structuring mechanisms for artistic exchange across distance, media, and disciplinary boundaries. Mapping is positioned not as an instrumental layer added to an otherwise complete work, but as a generative constraint through which intermedial form, narrative and collaboration are actively produced.

The findings demonstrate that systems organised around codependency create conditions of forced intimacy and shared responsibility between collaborators, eliciting modes of engagement that do not readily arise in modular or hierarchically organised media configurations. By allowing narrative and meaning to emerge through system behaviour and relational interaction, such practices support performances that remain adaptive, contingent and resistant to fixed interpretation. For audiences, the movement from legible causality towards perceptual opacity reframes liveness and agency as experiential phenomena, grounded as much in inference, projection and negotiated attention as in observable control.

By foregrounding mapping protocols that are both structuring and partially inaccessible, this research contributes to ongoing debates in electroacoustic audiovisual composition, intermediality and human–machine collaboration. The aesthetic potential of such systems lies not in transparency, synchronisation or technological display, but in the careful orchestration of dependency, ambiguity and risk within intermedial spaces where sound, image, gesture and computation operate as mutually conditioning forces. Although the implementations discussed here are situated in music, robotics and immersive media, the underlying principles extend to a broad range of intermedial and networked artistic practices operating under post-media conditions. At a more fundamental level, the conditions enabling such systems are grounded in the digital's unique status as a representational environment capable of encoding and operating on other environments of representation, what Handelman and Weatherby (in Fazi et al. 2025) describe as its defining dialectical property 'in that it unifies number as representation (data) and number as operation (syntax, procedure, logic)' (Fazi et al. 2025: 79), a condition that allows heterogeneous symbolic, sensory and material systems to be folded into a single, recursive field of mediation.

8. Future work

The next stage of development for this research will entail extending such mapping techniques into collaborative and ongoing global projects rather than focusing on individual performances. Workshops will be convened to introduce these techniques to artists with varied cross-cultural perspectives and investigate how culturally based assumptions impact system use and understanding.

Further research will also seek to understand how levels of opacity can be composed to offer artists control over the level of legibility versus apophenia. This relates to adapting mapping techniques over the duration of a performance or over multiple performances, and to the learning, forgetting or re-interpretation of causality.

Finally, there is scope to reflect on the ethical and aesthetic dimensions of agency in ever more autonomous systems. As computational processes further enable artistic practice, the multimodal performance framework provides a critical site for exploring how notions of responsibility, authorship and intention are understood. By continuing to address mapping as both a technical and cultural practice, the research aims to further elucidate the nuances of collaborative meaning-making in data-driven artistic environments.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S1355771826101617.

Acknowledgements. The authors acknowledge the use of Grammarly for language editing and grammatical correction during the preparation of this article. The tool was used solely to improve clarity and readability. All scientific content, interpretations and conclusions are authors' own. The authors thank

https://doi.org/10.1017/S1355771826101617 Published online by Cambridge University Press

Karen Yu (performer, percussion), Kung Chi Shing (performer, violin), Davor Vincze (sound artist), Marlen Runz (robotic-system development), Daniel Stempfer (robotic-system development), Taurin Barrera (system design), Samuel Swope (airborne structures and sound design), Alexis Mailles (technical consultant), Vvzela Kook (costume design), current.cam (AI video), Joseph Chan (kinetic sculptures) and Lazarus Chan Long Fung (LLM design) for their contributions to *Remote Gestures* and *manteia*. The authors also thank the audience members who took part in the anonymous feedback survey reported in Appendix A.

References

Allado-McDowell, K. 2024. On Neural Media. *New Models*. https://newmodels.substack.com (accessed 2 July 2026).

Alonso Trillo, R. 2025. Allhands_Python [computer software]. GitHub. https://github.com/MetaBow/allhands_python.

Alonso Trillo, R. and Poliks, M. 2025. Interface After AI. *ECHO: Journal of Music, Thought and Technology* **6**. https://doi.org/10.47041/WXQD9569.

Alonso Trillo, R., Nelson, P. A. C. and Michailidis, T. 2023. Rethinking instrumental interface design: The MetaBow. *Computer Music Journal* **47**(2): 5–20. https://doi.org/10.1162/comj_a_00683.

Alonso Trillo, R., Nelson, P. A. C. and Vincze, D. 2025. MetaBow: Gesture Mapping in Immersive Sonic Environments. *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME 2025)*. Canberra, Australia.

Anadol, R. 2023. Large-Scale Data-Driven Visualisation Projects [Exhibition Works]. Various international venues.

Apprich, C., Chun, W. H. K., Cramer, F. and Steyerl, H. 2019. *Pattern Discrimination*. Minneapolis, MN: University of Minnesota Press.

Ascott, R. 1984. Art and Telematics: Toward a Network Consciousness. *Leonardo* **17**(3): 147–55.

Ascott, R. 1990. Is there Love in the Telematic Embrace? *Art Journal* **49**(3): 241–7.

Ascott, R. 2003. *Telematic Embrace: Visionary Theories of Art, Technology, and Consciousness*. Berkeley, CA: University of California Press.

Auslander, P. 2008. *Liveness: Performance in a Mediatized Culture*, 2nd edn. London and New York: Routledge.

Baalman, M. A. J. 2022. *Composing Interactions: An Artist's Guide to Building Expressive Interactive Systems*. Rotterdam: V2_ Publishing.

Barthet, M., Fazekas, G. and Sandler, M. 2019a. Distributed Music Systems and the Internet of Musical Things. *IEEE MultiMedia* **26**(1): 44–55.

Barthet, M., Turchet, L., Fischione, C. and Fazekas, G. 2019b. The Internet of Musical Things. *Computer Music Journal* **43**(3): 35–51.

Birringer, J. 2008. *Performance, Technology, and Science*. Baltimore, MD: PAJ Publications.

Bischoff, J., Gold, R. and Horton, J. 1978. Music for an Interactive Network of Microcomputers. *Computer Music Journal* **2**(3): 24–9.

Bratton, B. H. 2013. The Black Stack. *e-flux Journal* **53**. https://www.e-flux.com.

Bratton, B. H. and Agüera y Arcas, B. 2022. The Model is the Message. *Noema Magazine*, 12 July.

Bretan, M. and Weinberg, G. 2016. A Survey of Robotic Musicianship. *Communications of the ACM* **59**(5): 100–9.

Brown, C. and Bischoff, J. 2002. Indigenous to the Net: Early Network Music Bands in the San Francisco Bay Area. Crossfade/Walker Art Center. http://crossfade.walkerart.org (accessed 2 July 2026).

Burrell, J. 2016. How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms. *Big Data & Society* **3**(1): 1–12.

Cáceres, J.-P. and Chafe, C. 2010. JackTrip: Under the Hood of an Engine for Network Audio. *Journal of New Music Research* **39**(3): 183–7. https://doi.org/10.1080/09298215.2010.481361.

Camurri, A., Hashimoto, S., Ricchetti, M., Ricci, A., Suzuki, K., Trocca, R. and Volpe, G. 2004. EyesWeb: Toward Gesture and Affect Recognition in Interactive Dance and Music Systems. *Computer Music Journal* **24**(1): 57–69.

Chafe, C. and Gurevich, M. 2004. Network Time Delay and Ensemble Accuracy. *Proceedings of the International Computer Music Conference*.

Chion, M. 1994. *Audio-Vision: Sound on Screen*, trans. C. Gorbman. New York: Columbia University Press.

Clark, A. 2016. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford: Oxford University Press.

Di Scipio, A. 2003. Sound is the Interface: From Interactive to Ecosystemic Signal Processing. *Organised Sound* **8**(3): 269–77.

Di Scipio, A. 2008. Emergent Sound Phenomena in Feedback Systems. *Computer Music Journal* **32**(1): 45–60.

Dixon, S. 2007. *Digital Performance: A History of New Theatre, Dance, Performance Art, and Installation*. Cambridge, MA: MIT Press.

Doornbusch, P. 2002. Composers Views on Mapping in Algorithmic Composition. *Organised Sound* **7**(2): 145–56.

Drioli, C., Allocchio, C. and Buso, N. 2013. Networked Performances and Natural Interaction via LOLA: Low Latency High Quality A/V Streaming System. *Proceedings of the Sound and Music Computing Conference*. Berlin: Springer, 24–30.

Dürre, J., Riedel, T., Rottondi, C. and Sarti, A. 2022. In-depth Latency and Reliability Analysis of Networked Music Performance Over Public 5G Infrastructure. *Proceedings of the 153rd AES Convention*.

Fazi, M. B., Galloway, A. R., Handelman, M. and Weatherby, L. 2025. *Digital Theory*. Minneapolis, MN: University of Minnesota Press.

Fiebrink, R. 2009. Real-Time Human Interaction with Supervised Learning Algorithms for Music Composition and Performance. PhD dissertation, Princeton University.

Fiebrink, R. and Caramiaux, B. 2018. The Machine Learning Algorithm as Creative Musical Tool. In R. T. Dean and A. McLean (eds.) *The Oxford Handbook of Algorithmic Music*. Oxford: Oxford University Press, 181–208.

Freeman, J., Ramakrishnan, C. and Zepeda, J. 2021. Networked Music Performance: Practices, Tools, and Aesthetics. *Journal of New Music Research* **50**(1): 1–15.

Frith, C. D. 2007. *Making Up the Mind: How the Brain Creates Our Mental World*. Malden, MA: Blackwell.

Gibson, J. J. 1979. *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.

Gombrich, E. H. 1960. *Art and Illusion: A Study in the Psychology of Pictorial Representation*. London: Phaidon.

Gurevich, M. and Fyans, A. C. 2011. Digital Musical Interactions: Performer–System Relationships and Their Perception by Spectators. *Organised Sound* **16**(2): 166–75.

Hansen, M. B. N. 2015. *Feed-forward: On the Future of Twenty-first-century Media*. Chicago, IL: University of Chicago Press.

Haque, U. 2007. The Architectural Relevance of Gordon Pask. *Architectural Design* **77**(4): 54–61.

Hayles, N. K. 2017. *Unthought: The Power of the Cognitive Nonconscious*. Chicago, IL: University of Chicago Press.

Helmholtz, H. von 1867. *Handbuch der physiologischen Optik*. Leipzig: Voss.

Higgins, D. 2001 [1966]. Intermedia. *Leonardo* **34**(1): 49–54.

Hunt, A. and Wanderley, M. M. 2002. Mapping Performer Parameters to Synthesis Engines. *Organised Sound* **7**(2): 97–108.

Hunt, A., Wanderley, M. M. and Kirk, R. 2003. Towards a Model for Instrumental Mapping in Expert Musical Interaction. *Proceedings of the International Computer Music Conference*.

Hunt, A., Wanderley, M. M. and Paradis, M. 2003. The Importance of Parameter Mapping in Electronic Instrument Design. *Journal of New Music Research* **32**(4): 429–40.

Jordà, S. 2005. Digital Instruments and Players: Part I—Efficiency and Apprenticeship. *Proceedings of the International Conference on New Interfaces for Musical Expression*.

Krauss, R. 1999. *A Voyage on the North Sea: Art in the Age of the Post-medium Condition*. London: Thames & Hudson.

Lakiotakis, E., Liaskos, C. and Dimitropoulos, X. 2019. Improving Networked Music Performance Systems Using Application-Network Collaboration. *Concurrency and Computation: Practice and Experience* **31**(18): e4730. https://doi.org/10.1002/cpe.4730.

Liloia, A. and Dannenberg R. 2025. Exploiting Latency in the Design of a Networked Music Performance System for Percussive Collective Improvisation. *Proceedings of the International Conference on New Interfaces for Musical Expression* 473–480.

Malloch, J., Sinclair, S., Wanderley, M. M. and Hueber, T. 2013. From Musical Gestures to Musical Structures: The Libmapper Framework. *Proceedings of the International Conference on New Interfaces for Musical Expression*.

https://doi.org/10.1017/S1355771826101617 Published online by Cambridge University Press

Manovich, L. 2001. *The Language of New Media*. Cambridge, MA: MIT Press.
Manovich, L. 2005. Post-Media Aesthetics. Transmedia Research Institute. http://manovich.net (accessed 6 January 2026).
manteia (μαντεία). 2024. [Video]. YouTube. https://www.youtube.com/watch?v=7DtoAFPGB8M (accessed 6 January 2026).
Metahaven. 2026. *The Hard Question of Art: Cognitive Futures*. London and New York: Verso.
Oliveros, P. 2005. *Deep Listening: A Composer's Sound Practice*. New York: iUniverse.
Pethő, Á. 2011. *Cinema and Intermediality: The Passion for the In-Between*. Newcastle upon Tyne: Cambridge Scholars Publishing.
Remote Gestures. 2024. [Video]. Vimeo. https://vimeo.com/963014756 (accessed 6 January 2026).
Rottondi, C., Chafe, C., Allocchio, C. and Sarti, A. 2016. An Overview of Networked Music Performance Technologies. *IEEE Access* **4**: 8823–43. https://doi.org/10.1109/ACCESS.2016.2628440.
Rovan, J. B., Wanderley, M. M., Dubnov, S. and Depalle, P. 1997. Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance. In *Kansei — the Technology of Emotion: Proceedings of the AIMI International Workshop*. Genoa: AIMI, 68–73.
Salter, C. 2010. *Entangled: Technology and the Transformation of Performance*. Cambridge, MA: MIT Press.
Seth, A. K. 2019. From Unconscious Inference to the Beholder's Share: Predictive Perception and Human Experience. *European Review* **27**(3): 378–410.
Smalley, D. 1997. Spectromorphology: Explaining Sound-Shapes. *Organised Sound* **2**(2): 107–26.
Steyerl, H. 2016. A Sea of Data: Apophenia and Pattern (Mis-)recognition. *e-flux Journal* **69**. https://www.e-flux.com.
Steyerl, H. 2018. A Sea of Data: Apophenia and Pattern (Mis-)recognition. In C. Apprich, W. H. K. Chun, F. Cramer and H. Steyerl (eds.) *Pattern Discrimination*. Lüneburg: Meson Press, 1–22.
Turchet, L. and Tomasetti, M. 2023. The Internet of Musical Things: Vision and challenges. *Computer Music Journal* **47**(1): 5–24.
Turchet, L., Fischione, C. and Barthet, M. 2018. Towards the Internet of Musical Things. *Proceedings of the Sound and Music Computing Conference*.
Van Nort, D., Wanderley, M. M. and Depalle, P. 2014. Mapping Control Structures for Sound Synthesis: Functional and Topological Perspectives. *Computer Music Journal* **38**(3): 6–22.
Waters, S. 2007. Performance Ecosystems: Ecological Approaches to Musical Interaction. *Organised Sound* **12**(2): 109–25.
Whitelaw, M. 2008. System Stories and Model Worlds: A Critical Approach to Generative Art. *Proceedings of the International Conference on Generative Art*.

APPENDIX A. AUDIENCE FEEDBACK SURVEY (REMOTE GESTURES)

Following the *Remote Gestures* performance (Visualization Research Centre, Hong Kong Baptist University, June 2024), a short, voluntary, anonymous questionnaire was distributed to audience members; seven responses were received. It collected no personal or identifying data and, as an informal survey, required no formal ethics clearance. Given the small sample, responses are reported as raw counts and treated as informal, corroborative feedback rather than statistically representative data. The survey instrument and a summary of responses are reproduced below.

Survey instrument.

1. On a scale of 1 to 10, how would you rate your overall experience at the performance?
2. Which element of the performance did you find most engaging?
3. How effective was the integration of technology (e.g., low-latency video/audio, robotic arm) in enhancing the performance?
4. Did the performance help you understand the potential of remote collaboration in creative processes?
5. How would you rate the visual elements (e.g., line tracings of the bow, robotic calligraphy) of the performance?
6. How did you feel about the use of the 360-degree cinema environment for this performance?
7. Which part of the performance did you find the most innovative?
8. Did you feel a sense of connection or interaction between the performers despite the physical distance?
9. Would you be interested in attending similar performances that integrate music, live visuals and robotics in the future?
10. Do you have any suggestions for improving this type of performance in the future?

Summary of responses.

Survey item	Responses ($n = 7$)
Overall experience (1–10)	Mean 8.7 (range 5–10)
Most engaging element	Robotic calligraphy 4; 'richness of it all' 1; robot arm's solo 1; the music 1
Technology integration	Very effective 2; effective 4; neutral 1
Understood potential of remote collaboration	Yes, very much 4; yes, somewhat 1; no, not really 2
Visual elements (1–10)	Mean 7.6 (10: 1, 8: 4, 7: 1, 4: 1)
360° cinema environment	Greatly enhanced the experience 6; neutral 1
Most innovative part	Robotic calligraphy 4; robot arm's musical performance 2; low-latency link 1
Sense of connection despite distance	Yes, very much 3; yes, somewhat 2; neutral 1; no, not really 1
Interest in future events	Yes, definitely 6; not sure 1
Suggestions for improvement	No substantive suggestions returned

https://doi.org/10.1017/S1355771826101617 Published online by Cambridge University Press