Imagined Movement as Sonic Gesture: Auditory Expression from a Deep Learning-Based Motion Decoding BCI

Niall McShane
PRACT Research Centre
School of Arts and Humanities
Ulster University
Derry~Londonderry, N.Ireland
n.mcshane@ulster.ac.uk

Dr. Karl McCreadie
Intelligent Systems Research Centre
School of Computing, Engineering and Intelligent Systems
Ulster University
Derry~Londonderry, N.Ireland
k.mccreadie@ulster.ac.uk

Dr. Attila Korik
The Bath Institute for the Augmented Human
University of Bath
Bath, U.K.
ak3825@bath.ac.uk

Professor Damien Coyle
The Bath Institute for the Augmented Human
University of Bath
Bath, U.K.
dhc30@bath.ac.uk

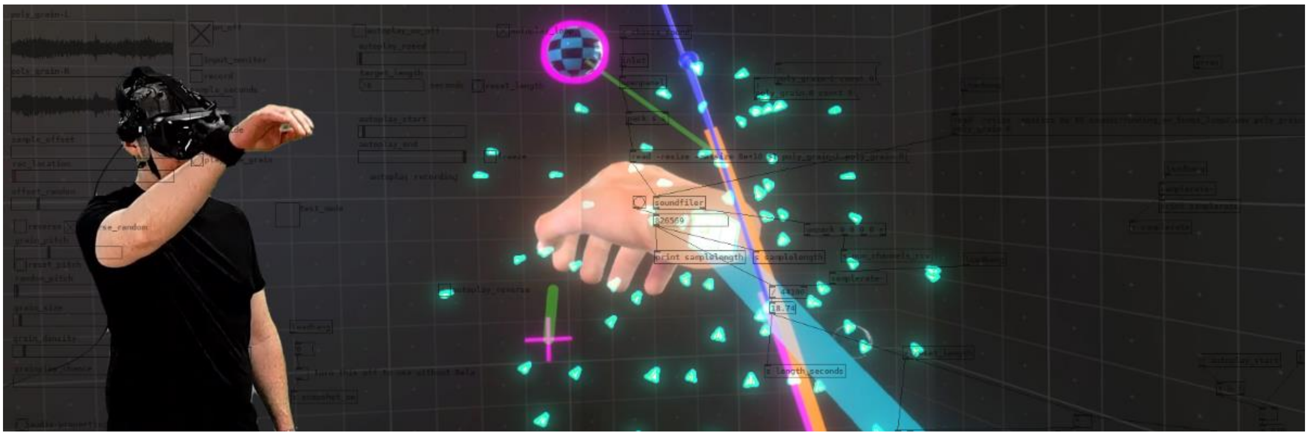

Figure 1: MTD-BCI sonification interface - a participant performs reaching movements decoded from EEG to drive real-time granular synthesis and embodied visual feedback.

Abstract

Motion trajectory decoding (MTD) brain-computer interfaces (BCIs) can translate imagined limb movements into continuous control signals, traditionally presented through visual feedback such as virtual limbs. This paper extends the multimodal expressivity of MTD-BCIs by introducing embodied sonification as a primary interaction modality. Using previously trained CNN-LSTM decoders for three-dimensional imagined arm movement, we mapped decoded motion and velocity signals in real time to a layered granular synthesis system. The framework employs velocity magnitude to modulate textural density and spectral characteristics, temporal accumulation to shape harmonic evolution, and rest-state detection to define acoustic boundaries between imagined gesture phases. Rather than treating sound as supplementary feedback, this approach positions decoded motion as sonic gesture; a continuous expressive articulation of imagined movement grounded in embodied cognition. The multi-layered synthesis architecture demonstrates that complex, many-to-many parameter mappings preserve perceptual coherence between movement dynamics and auditory change, creating a unified motion-audio-visual loop. This proof-of-concept establishes sonification as a viable interaction paradigm for motion-based BCIs, with implications for expressive musical performance, accessible sound-based control, and neurocognitive research into multimodal feedback in embodied human-computer interaction.

Keywords

Sonification, Embodied Cognition, BCI, Motion Trajectory Decoding, Gestural Control, Granular Synthesis

1 Introduction

Brain-computer interfaces (BCIs) enable direct communication between neural activity and external systems, typically using non-invasive electroencephalography (EEG) to infer user intent from patterns of brain activity [20]. Within BCI paradigms, motor imagery is accepted as a widely adopted control strategy, whereby users imagine bodily movements to modulate sensorimotor rhythms and generate control signals [31]. Classical motor imagery-based BCIs rely on discrete classification paradigms and lateralised brain activity, mapping imagined actions to symbolic commands such as left versus right hand movement. While effective for certain applications, these approaches impose artificial constraints on interaction and diverge from the continuous, dynamic nature of human movement. In response, motion trajectory decoding BCIs (MTD-BCIs) have been developed to reconstruct continuous kinematics from neural signals, allowing imagined limb movements to be decoded as time-varying control trajectories rather than categorical states [2,16,21]. By translating neural activity associated with imagined movement into continuous control outputs, MTD-BCIs offer a more intuitive and embodied interaction paradigm than conventional classification-based systems [26]. To date, most MTD-BCI implementations focus on visual feedback, commonly employing paradigms ranging from cursors on a screen to virtual limbs, to reinforce motor imagery and facilitate learning in the context of functional control tasks and accessibility interventions.

Sound has long functioned as a medium for gesture, expression, and agency, offering a temporal and dynamic representation of action that aligns closely with human movement [37,39]. In interactive systems, sonification provides a means of rendering continuous processes perceptible through auditory change rather than visual displacement. Within BCI research, sound has been used for discrete feedback, neurofeedback training, and musical exploration, yet it has predominantly served a supportive or evaluative role rather than acting as the primary medium of interaction [24]. Motion decoding BCIs, generate rich time-varying control signals that are potentially suited to auditory representation. Mapping imagined movement trajectories to sound enables neural activity to be experienced as an unfolding gesture rather than as a symbolic command, reframing motion decoding as not only functional control, but an expressive interaction.

In auditory design, synthesis techniques such as granular synthesis, envelope-driven modulation, and dynamics-based processing are particularly compatible with EEG-derived motion control signals. These approaches can accommodate noise, variability, and non-stationarity while preserving a meaningful structure. Moreover, auditory feedback can complement or partially replace embodied visual representations, offering an alternative pathway for reinforcing motor imagery and supporting sensorimotor engagement [19]. Using MTD-BCI systems in the sonic domain opens a design space in which imagined movement, sound, and perception form a closed expressive loop, bridging neural decoding research with musical interface design.

1.1 Embodiment and Motion Decoding BCI

Embodied feedback, particularly through spatially congruent virtual limbs, has been shown to support learning, performance, and neural adaptation in MI-BCIs [35,36]. This relationship between embodiment and feedback has become especially relevant in continuous BCI paradigms that aim to move beyond discrete command selection to provide direct brain control of an embodied virtual limbs [17,21].

MTD-BCIs, reconstruct upper-limb kinematics from EEG by training decoders on using both neural and movement data. These systems typically decode imagined movement as variable velocity information. Early MTD-BCI implementations primarily employed linear regression techniques to estimate motion trajectories from sensorimotor rhythms [2,3,16,28]. More recent work has demonstrated that deep learning approaches, including convolutional and recurrent architectures such as CNN-LSTM models, can more effectively capture the spatial-temporal structure of imagined movement signals, enabling more robust and expressive decoding [1,15,22,30]. These advances have made real-time decoding of imagined movement more stable, reinforcing the potential of MTD-BCIs within continuous control applications. In practice, decoded motion signals are embedded within feedback loops that shape how users experience and engage with imagined movement. Embodied visual representations, such as first-person virtual limbs, have been used to render decoded trajectories perceptible and interpretable, situating imagined movement within an embodied context.

1.2 Sonification and Sound-Based Interaction in BCIs

Sound has been used in BCI systems a range of functional and artistic contexts, including discrete feedback, neurofeedback training, and exploratory musical applications [13,24,27]. In motor imagery BCI paradigms, auditory signals commonly confirm classification outcomes, indicate task states, or represent neural features such as frequency band power [4,5,19]. Congruent auditory feedback, where sounds are semantically related to the imagined action, has been shown to improve offline classification accuracy in motor imagery tasks [4], while musical biofeedback devices such as the Encephalophone have demonstrated that users can achieve conscious control over pitch through modulation of posterior dominant or mu rhythms [7,8]. In affective brain–computer music interfaces (aBCMIs), continuous sonification has been used to reflect and modulate emotional states, with systems translating EEG-derived measures of arousal and valence into musical parameters such as tempo and loudness [6,9,32].These approaches establish closed-loop interactions in which sound functions as both feedback and stimulus, yet they predominantly render internal affective or oscillatory states rather than decoded volitional motor commands. Within the NIME community, BCMIs have been explored as a distinct creative research tradition since the early 2000s, encompassing EEG-based composition systems [24,25], physiological instrument interfaces [38], and collaborative bio-signal performance [23]. The present work differs in that it treats sound not as a representation of spontaneous neural activity or emotional state, but as an articulation of intentional imagined movement decoded from sensorimotor signals.

Musical interface design foregrounds continuous control, temporal variation, and dynamic change as central to expressive interaction [14,39]. In contrast to systems that treat signal fluctuation as noise to be minimised, digital musical instruments often incorporate variability as an integral part of gesture, phrasing, and timbral nuance [12,40]. Parameter mapping the relationship between input control signals and synthesis variables is a defining characteristic of electronic instrument design, with research demonstrating that complex, many-to-many mappings yield more engaging and expressive instruments than simple one-to-one correspondences [14]. Within this framing, the continuous velocity estimates produced by MTD-BCIs present a suitable control source for sound synthesis. Decoded motion along spatial axes can be mapped to sonic parameters including grain density, spectral distribution, amplitude envelope, and temporal evolution, enabling imagined movement to be articulated as auditory change rather than visual displacement. When such mappings preserve the spatial and dynamic structure of embodied limb movement, sound can convey not only the intensity of imagined action but also its directional and temporal qualities. Embodied approaches to sonification suggest that listeners interpret auditory information through schemas grounded in physical experience [34], and that gestural mappings informed by embodied cognition can provide intuitive access to complex sonic parameters [10,11]. In this context, decoded imagined movement trajectories may function as sonic gestures interpreted as control signals that remain grounded in the kinematics of bodily action. This potentially supports perceptual coupling between imagined movement and sound, affording interaction paradigms that prioritise expressivity and agency over task accuracy alone.

2 BCI Implementation

The sound interface builds on an existing MTD-BCI pipeline that has been trained to decode imagined three-dimensional arm movements from EEG using embodied virtual hand feedback. In the present work, the previously trained CNN–LSTM decoder is reused without retraining and extended with an audio layer that maps decoded motion signals to sound in real-time. This approach allows the sonification framework to be evaluated independently of decoder training, framing sound as an interaction and feedback modality rather than as part of the learning process.

The decoder outputs continuous velocity estimates along x, y, z, spatial axes, which are streamed in real time into a Unity virtual environment. Unity acts as both the visualisation platform and incorporates the custom audio engine, enabling decoded motion trajectories to drive visual feedback and sound synthesis simultaneously.

2.1 System Overview

The system comprises four coupled components: (1) EEG acquisition and arm motion reference recording used during decoder training, (2) a cued motor imagery protocol with embodied virtual hand feedback used to generate paired neural and kinematic data, (3) a CNN–LSTM-based motion trajectory decoder trained on these data, and (4) a real-time Unity-based interaction layer that receives decoded motion signals and maps them to visual and auditory outputs.

During decoder training, EEG signals were paired with arm motion data captured during executed and imagined reaching movements, with a first-person virtual hand providing spatially congruent visual feedback (Figure 2). Training followed a structured, cued motor imagery protocol based on naturalistic three-dimensional reaching actions. Each trial was organised into fixed phases of rest, target indication, movement, and return, ensuring consistent temporal alignment between neural activity, intended movement, and visual feedback. During the imagined-movement trials, participants were instructed to imagine reaching toward a cued and visually highlighted spatial target while controlling the virtual limb via an inverse kinematic (IK) rig to perform the corresponding movement. This establishes a closed-loop training interface in which imagined motor intent is continuously associated with a spatial limb representation. Discrete target cues provide consistent reference points during training.

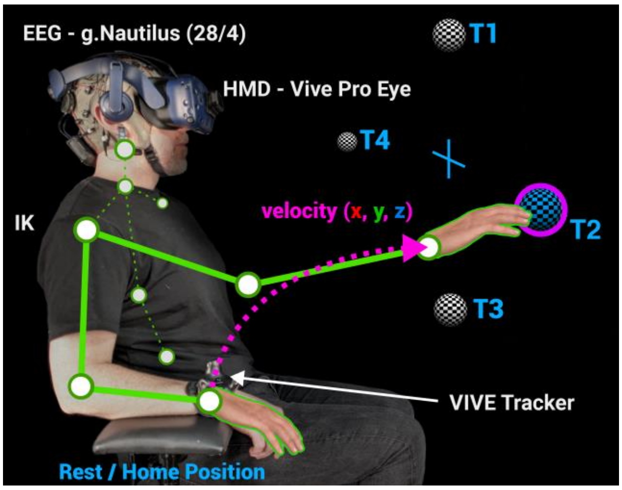

Figure 2: MTD-BCI control interface during decoder training. EEG signals (g.Nautilus, 17 channels) are paired with upper-limb kinematics captured via HTC Vive Tracker. Participants imagine reaching towards cued targets whilst controlling a first-person virtual hand via inverse kinematics, establishing a closed-loop mapping between neural activity and embodied visual feedback.

The trained decoder produces continuous velocity estimates along x, y, and z, streamed into Unity and interpreted simultaneously as visual and auditory control inputs. In this configuration, sound functions as an additional interaction modality layered onto an existing MTD-BCI framework, rather than as a component of the decoder training process.

2.2 Hardware Setup and EEG Acquisition

EEG data was acquired using a g.tec g.Nautilus wireless EEG system, recorded at a sampling rate of 250 Hz and referenced to the left earlobe with a common ground electrode. A reduced 17-channel montage was employed in the real-time setup, positioned according to the international 10-20 system, comprising electrodes F3, FZ, F4, FC5, FC1, FC2, FC6, C3, CZ, C4, CP5, CP1, CP2, CP6, P3, PZ, and P4.

Upper-limb kinematics were recorded during executed movement trials captured directly within Unity at 60 Hz as three-dimensional coordinates that were translated to velocity. Kinematic motion was recorded from a HTC Vive Tracker attached to the participant's right wrist captured by four SteamVR base stations arranged to ensure non-occluded coverage. Kinematic recordings were used as data for the MTD-BCI decoder training but were not utilised during imagined-movement control.

Data was synchronised via UDP, with trial markers embedded directly into the EEG and kinematic data streams ensuring precise temporal alignment between EEG, kinematic recordings, and decoder outputs.

2.3 CNN-LSTM Motion Decoder

Imagined arm movement was decoded using a deep learning architecture that maps time-varying spectral features of the EEG signal to three-dimensional velocity. Rather than operating on frequency band filtered EEG, the decoder is trained on event-related spectral perturbation (ERSP) representations, which capture changes in neural oscillatory activity relative to a resting baseline. This representation provides a compact and interpretable description of motor imagery dynamics that is well suited to continuous control.

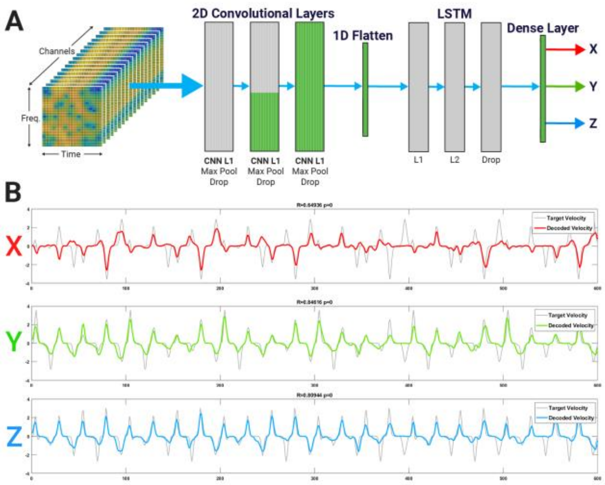

Figure 3: CNN-LSTM motion decoder architecture. Event-related spectral perturbation (ERSP) images computed across 17 EEG channels (0-40 Hz frequency range) are input to convolutional layers for spatial–spectral feature extraction, followed by stacked LSTM layers for temporal modelling. The network outputs continuous velocity estimates (x, y, z) corresponding to imagined upper-limb trajectories.

For each of the 17 EEG channels, ERSP images were computed over a frequency range of 0-40 Hz and segmented into overlapping temporal windows aligned to trial onset and movement execution. These time-frequency representations encode the evolution of sensorimotor rhythms during imagined reaching actions and are stacked channel-wise to form a multichannel input volume. This allows spatially distributed neural activity associated with motor planning and execution to be learned jointly across channels (Figure 3A).

The decoder combines convolutional layers for spatial–spectral feature extraction with recurrent layers for temporal modelling. Convolutional layers learn local structure within ERSP representations, while stacked long short-term memory (LSTM) layers model temporal dependencies across successive windows, maintaining context over the duration of imagined movement. A final dense layer outputs continuous velocity estimates along the x, y, and z axes, corresponding to the imagined upper limb (Figure 3B). In this work, a previously trained decoder is reused without modification, and the resulting velocity signals are treated as continuous control streams rather than performance or system accuracy metrics.

Using ERSP representations and producing smooth, time-varying velocity outputs, the decoder provides control signals that are compatible with continuous interaction and sonification. These decoded movement dynamics formed the basis for the subsequent mapping to sound, allowing imagined motion onsets, direction, and intensity to be articulated as modulated sonic characteristics rather than visual feedback alone.

3 Imagined Movement as Sonic Gesture

In the proposed system, decoded imagined movement trajectories are reframed as sonic gestures. Continuous imagined movement control signals articulate effort, direction, and temporal flow through sound rather than visual feedback displacement alone. The mapping configuration follows established principles from digital musical instrument design, wherein gestural inputs are translated to sonic outputs through structured correspondences [14].

The synthesis engine receives continuously decoded signals from the CNN-LSTM motion decoder at 60 Hz (Table 1). Primary decoded outputs (x, y, z velocities) provide direct spatial information concerning arm movement direction at the wrist endpoint, whilst derived control features (velocity magnitude, acceleration strength, and rest state detection) capture broader gesture dynamics and movement phases.

Several design decisions guide the mapping implementation. Velocity signals, rather than position, serve as the primary control source for two reasons. First, the decoder produces time-varying rate-of-change estimates as its natural output rather than absolute spatial coordinates [21,22]. Second, positional data is fixed to the virtual coordinate space and therefore contingent on the participant's home position and target layout, whereas velocity reflects the dynamic qualities of movement independently of spatial context. From a sonification perspective, this makes velocity the more appropriate control source, as auditory change is itself a temporal and dynamic phenomenon rather than a static one [39]. Each spatial axis maps to a distinct sonic dimension, preserving the directional structure of imagined arm movement: horizontal movement affects temporal progression, vertical movement affects textural density, and depth movement affects spectral spread. Velocity magnitude functions as a global intensity modulator across all synthesis layers, ensuring that overall movement energy is consistently represented in the sonic output. Rest state detection triggers envelope decay and layer attenuation, providing clear auditory distinction between active imagined movement and baseline neural activity.

Table 1: Decoded motion control signals and synthesis parameters. Primary outputs from the CNN-LSTM decoder (x, y, z velocity) provide direct spatial information; derived features (velocity magnitude, acceleration strength, rest state) capture broader gesture dynamics and movement phases

Signal Type	Parameter	Description	Derivation
Primary Decoded Outputs	x-velocity	Horizontal movement (lateral arm displacement)	Direct decoder output
	y-velocity	Vertical movement (arm elevation)	Direct decoder output
	z-velocity	Depth movement (forward/backward reach)	Direct decoder output
Derived Control Features	Velocity magnitude	Overall movement energy	Euclidean norm of velocity vector
	Acceleration strength	Rate of velocity change; gesture dynamics and transients	Temporal derivative of velocity
	Rest state detection	Distinction between active movement and baseline neural activity	Velocity threshold comparison

The design prioritises expressive legibility over instrumental precision. Unlike discrete classification-based BCIs that trigger categorical events, the continuous velocity outputs of the motion decoder provide time-varying control signals suited to gradual sonic modulation. This approach accommodates the inherent noise and variability of EEG-derived signals whilst preserving meaningful gestural structure.

3.1 Synthesis Architecture

The sonification framework employs a layered synthesis architecture implemented using Pure Data (via libPD integration [18]) for granular synthesis, with envelope shaping and effects processing applied through real-time parameter mapping. Decoded motion signals are mapped directly to synthesis parameters across three functional layers: granular texture generation, envelope and shape control, and spatial effects (Figure 4). This layered organisation distributes expressive control across three independent sonic dimensions (timbre, dynamic contour, and spatial character), following the principle that complex, many-to-many parameter mappings produce more expressive and engaging instruments than simple one-to-one correspondences [14].

Layer 1: Granular Texture (Primary Gesture Carrier)

The granular layer functions as the primary gestural surface, rendering continuous imagined movement as evolving sonic texture. Granular synthesis is particularly suited to BCI control signals due to its inherent tolerance of noise and parameter instability. Control signal fluctuations produce subtle timbral variations rather than disruptive artefacts [33]. Each spatial axis of decoded velocity maps to distinct granular parameters: horizontal movement (x-velocity) traverses the source recording via sample offset, vertical movement (y-velocity) simultaneously modulates grain density and grain size, and depth movement (z-velocity) controls pitch dispersion, although tuned for minimal pitch range fluctuations (Table 2). Overall movement energy scales grain play probability, ensuring that gesture intensity directly influences textural density and layer presence. The layer is designed for continuous transformation rather than discrete triggering, articulating imagined movement as unfolding gesture.

Layer 2: Envelope and Shape (Gesture Mass and Sustain)

Envelope and shape control provides dynamic contouring of the granular texture in response to movement qualities and gesture duration (Table 2). Velocity magnitude directly controls amplitude envelope depth, shaping the dynamic profile of grain articulation. A dual-stage low-pass filter cutoff mapping enables both immediate responsiveness and longer-form harmonic development: instantaneous velocity controls a base cutoff range (300-2,500 Hz), whilst sustained movement duration progressively applies an additional brightness boost (0–1,500 Hz) over approximately two seconds of continuous gesture. Rest state detection triggers envelope release, creating gradual amplitude fade during movement cessation.

Layer 3: Effects Processing (Spatial Colouration)

The effects layer shapes the perceived acoustic space through reverb and diffusion processing (Table 2). Velocity magnitude controls both reverb mix level and decay time, situating the granular output within a dynamic spatial environment that responds to movement intensity. Acceleration strength modulates reverb diffusion, increasing spectral complexity during rapid gesture changes. A secondary subtle delay component provides some additional tail but was not dynamically controlled and uses a barely perceptible eight percent wet signal. Rather than introducing additional pitched material, the effects layer adds environmental coherence, positioning the sonification within an acoustic context that evolves dynamically with gesture qualities.

3.2 Implementation Details

The synthesis engine operates entirely within Unity, receiving decoded velocity signals from the CNN-LSTM decoder at 60 Hz. Control signals are smoothed using exponential moving averages to reduce high-frequency noise whilst preserving gesture dynamics. Smoothing ensures that rapid fluctuations in the decoded signal do not produce jarring sonic artefacts whilst still allowing gesture onsets and dynamic changes to be articulated clearly in the audio output.

The granular texture layer is implemented as a Pure Data patch integrated into Unity via the libPD library [18], enabling efficient real-time granular processing with direct parameter control. A custom C# class (MotionToAudio) continuously maps decoded motion signals to granular parameters, sample offset, grain density, grain size, pitch dispersion, and play probability, transmitting control values to enable real-time modulation of granular output based on decoded kinematics. This architecture provides robust, low-latency communication between the decoder output and synthesis layer whilst maintaining the flexibility and sonic sophistication of Pure Data granular processing.

Initially the envelope shaping layer was intended to drive a polyphonic drone type synthesis using OSC and SuperCollider, however this was remapped to the granular synth as the tuning between melodic drone and granular shape limited the auditory cohesion. Focusing purely on a single granular patch allowed the motion to have a more immediate and tangible sonic influence.

Envelope shaping and spatial effects processing are implemented as custom Unity audio filter components applied to the granular synthesis output. The envelope control modulates amplitude depth in response to velocity magnitude and movement cessation. Reverb and delay effects are applied through custom filter components, with reverb mix, time, and diffusion responding to velocity magnitude and acceleration strength as specified in Table 2.

An additional particle-based visual system driven by boid flocking behaviour mirrors the motion-to-synthesis mappings: movement magnitude accelerates boid velocity, acceleration strength introduces variation and cohesion, and rest state detection creates visual stillness that parallels audio envelope decay, offering performers an intuitive visual counterpart to the sonic gesture. (Figure 4).

Table 2: Motion-to-audio mapping schema. Decoded velocity signals are mapped across three synthesis layers: granular texture, envelope shaping, and effects processing with specific parameter ranges and sonic effects. Each spatial axis preserves directional information: horizontal movement (x) traverses sample offset, vertical movement (y) modulates density and grain size, and depth movement (z) controls pitch dispersion.

Layer	Decoded Signal	Synthesis Parameter	Mapping Range	Sonic Effect
Granular	x-velocity	Sample offset	0-1 (normalised)	Horizontal movement traverses source recording, creating timbral progression across playback position
	y-velocity	Grain density	0.01-0.4	Vertical movement modulates textural density; higher velocity increases inter-onset rate
	y-velocity	Grain size	0.05-1.0	Vertical movement expands or contracts individual grain duration, affecting textural cohesion
	z-velocity	Random pitch dispersion	0.01-0.03	Depth movement controls spectral spread; increased z-velocity widens pitch variation across grains
	Velocity magnitude	Grain play chance	0.4-1.0	Movement energy modulates stochastic grain triggering; rest states reduce grain density
Envelope/Shape	Velocity magnitude	Amplitude envelope depth	0-1 (normalised)	Movement energy directly shapes dynamic contour; absence of motion reduces layer amplitude
	Sustained movement (time-based)	Low-pass filter cutoff	300-4,000 Hz	Prolonged gestures progressively increase harmonic brightness via dual mapping: base cutoff from velocity (300–2,500 Hz) + sustained boost (0–1,500 Hz)
	Rest state	Envelope release	0-1 (proportional)	Absence of movement triggers gradual fade via smoothed rest amount scaling
Effects	Velocity magnitude	Reverb mix	0.1-0.5	Movement energy increases spatial diffusion; greater velocity enhances wet signal
	Sustained movement + acceleration	Reverb feedback	0.2-0.8	Prolonged, accelerated gestures extend decay tail; combined influence of sustained movement and acceleration strength

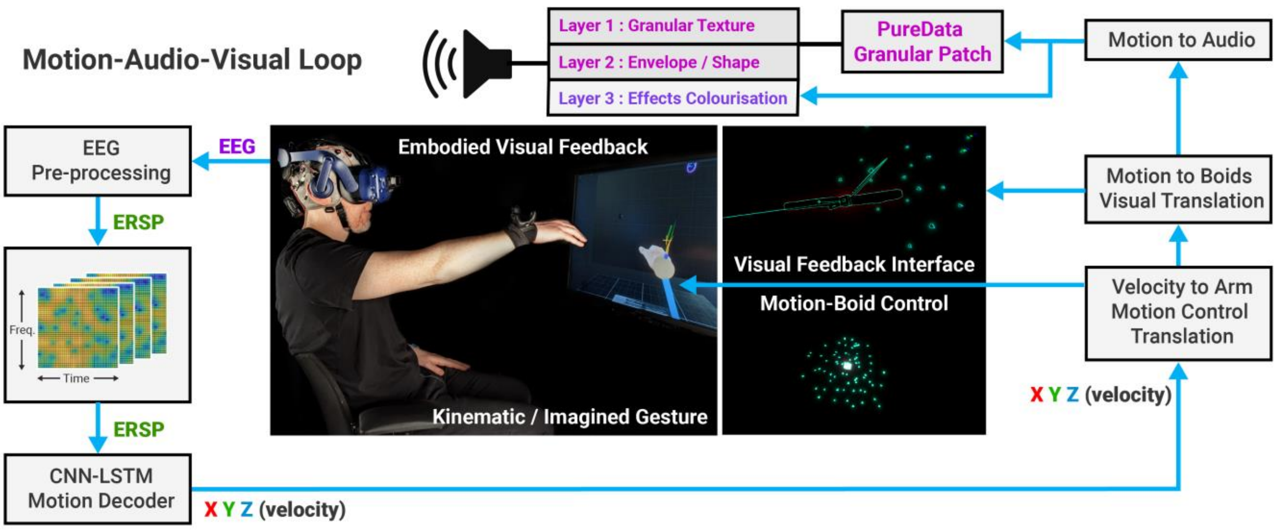

Figure 4: Imagined motion-audio-visual loop architecture. EEG signals are pre-processed and converted to event-related spectral perturbation (ERSP) representations, which are decoded by the CNN-LSTM motion decoder to produce continuous three-dimensional velocity estimates (x, y, z). These velocity signals drive two parallel pathways: (1) the audio synthesis engine, comprising three layers (granular texture, envelope shaping, and effects processing) mapped via the 'MotionToAudio' controller; and (2) a boid flocking system and virtual limb that translates motion parameters into particle-based and embodied visual feedback.

3.3 Evaluation Approach

Evaluation of the sonification framework follows a design-led methodology consistent with iterative DMI development, in which assessment is integrated throughout the design process rather than deferred to a formal end-stage study [29]. At this proof-of-concept stage, formal perceptual evaluation with participants requires a fully closed-loop real-time system in which auditory feedback can actively influence neural decoding and performer behaviour, conditions not yet met by the present implementation. Evaluation therefore focused on three criteria drawn from established DMI assessment principles [29]: (1) *expressive legibility* - whether imagined movement states (rest, onset, and sustained gesture) produce perceptually distinct and appropriately differentiated sonic outputs; (2) *perceptual coherence* - whether the correspondence between movement dynamics and auditory response is sufficiently consistent to support a sense of gestural agency across the motion-audio-visual loop; and (3) *sonic reliability* - whether noise and instability inherent in the decoded velocity stream produce disruptive artefacts in the audio output, or are instead absorbed as subtle timbral variation by the synthesis architecture and signal smoothing. These criteria were assessed through structured author observation across repeated system interactions, with design decisions refined iteratively in response to observed discrepancies between intended and perceived sonic behaviour. The outcomes of this evaluation are discussed in Section 4, with formal participant-based assessment reserved for the live closed-loop implementation described in Section 4.1.

4 Discussion

This proof-of-concept demonstrates that decoded imagined arm movement can effectively drive continuous sonic synthesis, with granular synthesis providing a particularly suitable method for translating kinematics into expressive auditory gesture. The system is evaluated here as a sonification design framework: the primary question is whether imagined movement dynamics can be rendered as perceptually coherent and expressively legible auditory gesture, rather than whether the system functions as a performance instrument in real time. The motion-to-audio mappings exhibit strong correspondence with intended movement dynamics: rest states generate auditory silence and flock cohesion, motion onsets produce rapid textural emergence, and sustained movement creates progressive harmonic brightening. This alignment suggests that sonification can effectively convey the temporal and dynamic qualities of imagined gesture in ways that complement visual feedback.

The choice to focus on granular synthesis rather than combined granular and polyphonic approaches reflects an important methodological finding. Whilst polyphonic synthesis offers pitch-centric control, the coupling between decoded motion and melodic content proved perceptually loose and required additional cognitive mapping between gesture and tonal outcome. Granular synthesis, by contrast, operates independently of pitch hierarchy, instead foregrounding timbre, texture, and spectral character. These dimensions map naturally to the continuous nature of motion trajectories. The reflection aligns with established findings in digital musical instrument design, where complex, many-to-many parameter mappings that preserve gestural structure produce more engaging and coherent interaction than simple one-to-one correspondences [14,39]. The present work extends this tradition by introducing decoded motor imagery as the gestural substrate for sonic interaction, bridging the NIME BCMI lineage [23,25,38] with continuous neural decoding approaches from BCI research. By maintaining direct correspondence between motion dynamics (velocity, acceleration, sustained duration) and granular texture evolution, the system preserves perceptual coherence across the motion-audio-visual loop.

The multi-layered envelope and filtering architecture enhanced the overall sonic impression. Resting states and the acoustic silence between motion onsets provide structural clarity, distinguishing successive gestures and preventing perceptual fatigue. The addition of an accumulating rest state within the motion to audio control provided greater dynamic range, tone and space. The dual-stage filter cutoff mapping, combining immediate velocity response with slower, time-accumulated brightness boost, creates a temporal unfolding that mirrors the intentional structure of reaching gestures: sharp onset, sustained intensity, and gradual release. This temporal organisation echoes principles from embodied cognition and gesture-based sonification, wherein listeners interpret auditory change through schemas grounded in physical experience [34]. Alignment between movement phases and synthesis response creates a potentially intuitive perceptual link between imagined action and sonic consequence, potentially supporting greater connection, expression and agency between human performer and neural interface.

4.1 Limitations and Future Work

This work presents a proof-of-concept based on offline-modelled motion data rather than real-time closed-loop training. Whilst the MTD-BCI demonstrates robust motion decoding for imagined arm movement and the audio synthesis layer exhibits responsive sonic musical output, the system has not been validated as a real-time performance interface. The audio layer has not yet directly influenced BCI learning or decoding performance, and the boid-based visual system's contribution to training and engagement remains unevaluated. The system currently occupies the role of a sonification framework rather than a performance instrument; the transition to the latter is contingent on the closed-loop implementation.

Future development will transition this proof-of-concept into a live, closed-loop BCI-audio interface wherein audio serves as real-time training feedback alongside the virtual limb representation. The real-time configuration opens the potential for performer agency and emergent expression when audio feedback is directly influencing the neural activity and imagined movement decoded signals. This will enable investigation of whether sonic expressivity enhances sensorimotor learning, performer engagement and determine whether performers develop stable movement vocabularies that translate to expressive musicality.

This framework potentially supports new forms of musical performance, accessibility-oriented sound design, and neurocognitive research into multimodal feedback and embodied cognition in creative practice.

Ethical Standards

This study received ethical approval from the Ulster University Research Ethics Filter Committee, Faculty of Computing, Engineering and the Built Environment (reference: CEBE_RE-21-011-A). Ethical approval covered the full duration of the longitudinal data collection period (2021–2024). All experimental procedures involving human participants were conducted in accordance with the Declaration of Helsinki and institutional guidelines for research involving human subjects. Participants provided written informed consent prior to participation and were informed of the study procedures, data recording methods (including EEG, kinematic and eye-tracking data), their right to withdraw at any time without penalty, and the intended research use of anonymised data. All data were anonymised prior to analysis and dissemination. No personally identifiable information is included in this publication.

Acknowledgments

This work was supported by internal funds from the University of Bath and Ulster University; by access to the Tier 2 High Performance Computing resources provided by the Northern Ireland High Performance Computing (NI-HPC) facility, funded by the UK Engineering and Physical Sciences Research Council (EPSRC) under Grant number EP/T022175; the Spatial Computing and Neurotechnology Innovation Hub funded by the Department for Economy Northern Ireland and the UK Research and Innovation (UKRI) Turing AI Fellowship 2021-2026 funded by the EPSRC under Grant number EP/V025724/1.

References

1. Davide Borra, Valeria Mondini, Elisa Magosso, and Gernot R. Müller-Putz. 2023. Decoding movement kinematics from EEG using an interpretable convolutional neural network. *Computers in Biology and Medicine* 165: 107323. https://doi.org/10.1016/j.compbiomed.2023.107323
2. Trent J. Bradberry, Rodolphe J. Gentili, and José L. Contreras-Vidal. 2010. Reconstructing three-dimensional hand movements from noninvasive electroencephalographic signals. *Journal of Neuroscience* 30, 9: 3432–3437. https://doi.org/10.1523/JNEUROSCI.6107-09.2010
3. Trent J. Bradberry, Feng Rong, and José L. Contreras-Vidal. 2009. Decoding center-out hand velocity from MEG signals during visuomotor adaptation. *NeuroImage* 47, 4: 1691–1700. https://doi.org/10.1016/j.neuroimage.2009.06.023
4. Emmanuel Christophe, Jérémy Frey, Richard Kronland-Martinet, Jean-Arthur Micoulaud-Franchi, Jelena Mladenovic, Gaelle Mougin, Jean Vion-Dury, Solvi Ystad, and Mitsuko Aramaki. 2018. Evaluation of a Congruent Auditory Feedback for Motor Imagery BCI. In *Proceedings of the 7th International Brain-Computer Interface Meeting*.
5. D Coyle, Á Carroll, J Stow, K Mccreadie, and J Mcelligott. 2013. Visual and Stereo Audio Sensorimotor Rhythm Feedback in the Minimally Conscious State. *5th International BCI Meeting*: 2–3. https://doi.org/10.3217/978-3-85125-260-6-18
6. Ian Daly, Duncan Williams, Alexis Kirke, James Weaver, Asad Malik, Faustina Hwang, Eduardo Miranda, and Slawomir J Nasuto. 2016. Affective brain–computer music interfacing. *Journal of Neural Engineering* 13, 4: 046022. https://doi.org/10.1088/1741-2560/13/4/046022
7. Thomas A. Deuel, Juan Pampin, Jacob Sundstrom, and Felix Darvas. 2017. The Encephalophone: A Novel Musical Biofeedback Device using Conscious Control of Electroencephalogram (EEG). *Frontiers in Human Neuroscience* 11: 213. https://doi.org/10.3389/fnhum.2017.00213
8. Thomas A. Deuel, James Wenlock, Alana McGovern, James Rosenthal, and Juan Pampin. 2025. Musical auditory feedback BCI: clinical pilot study of the Encephalophone. *Frontiers in Human Neuroscience* 19: 1592640. https://doi.org/10.3389/fnhum.2025.1592640
9. Stefan Ehrlich, Cuntai Guan, and Gordon Cheng. 2017. A closed-loop brain-computer music interface for continuous affective interaction. In *2017 International Conference on Orange Technologies (ICOT)*, 176–179. https://doi.org/10.1109/ICOT.2017.8336116
10. Richard Graham and Brian Bridges. 2015. Managing Musical Complexity with Embodied Metaphors. In *Proceedings of the International Conference on New Interfaces for Musical Expression* (NIME 2015), 103–106.
11. Richard Graham, Christopher Manzione, Brian Bridges, and William Brent. 2013. Exploring Pitch and Timbre through 3D Spaces: Embodied Models in Virtual Reality as a Basis for Performance Systems Design. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
12. Michael Gurevich and Jeffrey Treviño. 2007. Expression and Its Discontents: Toward an Ecology of Musical Creation. In *Proceedings of the 2007 Conference on New Interfaces for Musical Expression (NIME07)*.
13. Elisabeth Hildt. 2021. Affective Brain-Computer Music Interfaces—Drivers and Implications. *Frontiers in Human Neuroscience* 15: 711407. https://doi.org/10.3389/fnhum.2021.711407
14. Andy Hunt, Marcelo M. Wanderley, and Matthew Paradis. 2002. The importance of parameter mapping in electronic instrument design. In *Proceedings of the 2002 Conference on New Interfaces for Musical Expression* (NIME '02), 1–6.
15. Attila Korik, Naomi Du Bois, Jose Sanchez Bornot, Niall McShane, Christoph Guger, Alessandra Del Felice, Olive Lennon, and Damien Coyle. 2025. Decoding the Variable Velocity of Lower-Limb Stepping Movements From EEG. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 33: 3511–3523. https://doi.org/10.1109/TNSRE.2025.3603635
16. Attila Korik, Ronen Sosnik, Nazmul Siddique, and Damien Coyle. 2017. Imagined 3D hand movement trajectory decoding from sensorimotor EEG rhythms. *2016 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2016 - Conference Proceedings*, 76: 4591–4596. https://doi.org/10.1109/SMC.2016.7844955
17. Attila Korik, Ronen Sosnik, Nazmul Siddique, and Damien Coyle. 2019. Decoding imagined 3D arm movement trajectories from EEG to control two virtual arms-a pilot study. *Frontiers in*

Neurorobotics 13, November: 1–22. https://doi.org/10.3389/fnbot.2019.00094
18. LibPdIntegration. 2026. LibPdIntegration/LibPdIntegration. Retrieved April 24, 2026 from https://github.com/LibPdIntegration/LibPdIntegration
19. Karl A. McCreadie, Damien H. Coyle, and Girijesh Prasad. 2013. Sensorimotor learning with stereo auditory feedback for a brain-computer interface. *Medical and Biological Engineering and Computing* 51, 3: 285–293. https://doi.org/10.1007/s11517-012-0992-7
20. Dennis J. McFarland and Jonathan R. Wolpaw. 2002. Brain-computer interfaces for communication and control. *Clinical Neurophysiology* 113, 6: 767–791. https://doi.org/10.1145/1941487.1941506
21. Niall McShane, Dr. Karl McCreadie, Dr. Darryl Charles, Dr. Attila Korik, and Professor Damien Coyle. 2022. Online 3D Motion Decoder BCI for Embodied Virtual Reality Upper Limb Control: A Pilot Study. In *2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)*, 697–702. https://doi.org/10.1109/MetroXRAINE54828.2022.9967577
22. Niall McShane, Karl McCreadie, Darryl Charles, Attila Korik, and Damien Coyle. 2023. Decoding Motion Trajectories in an Upper Limb BCI: Linear Regression vs Deep Learning. In *2023 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)*, 1039–1044. https://doi.org/10.1109/MetroXRAINE58569.2023.10405752
23. Sergio Mealla, Aleksander Väljamäe, Marina Bosi, and Sergi Jordà. 2011. Listening to Your Brain: Implicit Interaction in Collaborative Music Performances. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
24. Eduardo Miranda and Andrew Brouse. 2005. Toward Direct Brain-Computer Musical Interfaces. In *Proceedings of the 2005 Conference on New Interfaces for Musical Expression*.
25. Eduardo Reck Miranda. 2006. Brain-computer music interface for composition and performance. *International Journal on Disability and Human Development* 5, 2. https://doi.org/10.1515/IJDHD.2006.5.2.119
26. Gernot R. Müller-Putz, Reinmar J. Kobler, Joana Pereira, Catarina Lopes-Dias, Lea Hehenberger, Valeria Mondini, Víctor Martínez-Cagigal, Nitikorn Srisrisawang, Hannah Pulferer, Luka Batistić, and Andreea I. Sburlea. 2022. Feel Your Reach: An EEG-Based Framework to Continuously Detect Goal-Directed Movements and Error Processing to Gate Kinesthetic Feedback Informed Artificial Arm Control. *Frontiers in Human Neuroscience* 16, March: 1–21. https://doi.org/10.3389/fnhum.2022.841312
27. Ryu Nakagawa and Shotaro Hirata. 2017. AEVE: An Audiovisual Experience Using VR HMD and EEG. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*. Retrieved from https://www.nime.org/proceedings/2017/nime2017_paper0097.pdf
28. Patrick Ofner and Gernot Müller-Putz. 2012. *Decoding of velocities and positions of 3D arm movement from EEG*. https://doi.org/10.1109/EMBC.2012.6347460
29. Sile O'Modhrain. 2011. A Framework for the Evaluation of Digital Musical Instruments. *Computer Music Journal* 35, 1: 28–42. https://doi.org/10.1162/COMJ_a_00038
30. Sidharth Pancholi, Amita Giri, Anant Jain, Lalan Kumar, and Sitikantha Roy. 2022. Source Aware Deep Learning Framework for Hand Kinematic Reconstruction Using EEG Signal. *IEEE Transactions on Cybernetics*: 1–13. https://doi.org/10.1109/TCYB.2022.3166604
31. G. Pfurtscheller, C. Brunner, A. Schlögl, and F. H. Lopes da Silva. 2006. Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. *NeuroImage* 31, 1: 153–159. https://doi.org/10.1016/j.neuroimage.2005.12.003
32. Rafael Ramirez, Manel Palencia-Lefler, Sergio Giraldo, and Zacharias Vamvakousis. 2015. Musical neurofeedback for treating depression in elderly people. *Frontiers in Neuroscience* 9. https://doi.org/10.3389/fnins.2015.00354
33. Curtis Roads. 2004. *Microsound*. MIT Press, Cambridge, Mass.
34. Stephen Roddy and Brian Bridges. 2020. Mapping for meaning: the embodied sonification listening model and its implications for the mapping problem in sonic information design. *Journal on Multimodal User Interfaces* 14, 2: 143–151. https://doi.org/10.1007/s12193-020-00318-y
35. Filip Škola and Fotis Liarokapis. 2018. Embodied VR environment facilitates motor imagery brain–computer interface training. *Computers and Graphics (Pergamon)* 75, June: 59–71. https://doi.org/10.1016/j.cag.2018.05.024
36. Filip Škola and Fotis Liarokapis. 2022. Reducing Brain-computer Interaction Training Time with Embodied Virtual Avatar: In *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 7–17. https://doi.org/10.5220/0011049100003124
37. Denis Smalley. 1997. Spectromorphology: explaining sound-shapes. *Organised Sound* 2, 2: 107–126. https://doi.org/10.1017/S1355771897009059
38. Atau Tanaka. 2000. Musical Performance Practice on Sensor-based Instruments. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, 1–6.
39. M.M. Wanderley and P. Depalle. 2004. Gestural Control of Sound Synthesis. *Proceedings of the IEEE* 92, 4: 632–644. https://doi.org/10.1109/JPROC.2004.825882
40. David Worrall. 2014. Can Micro-Gestural Inflections Be Used to Improve the Soniculatory Effectiveness of Parameter Mapping Sonifications? *Organised Sound* 19, 1: 52–59. https://doi.org/10.1017/S135577181300040X