Idealogy Journal Vol. 11 No. 2, 2026, pp 384-396
DOI: https://doi.org/10.24191/idealogy.v11i2.1016
Section: Original Article

Video Analysis of VJ (Video Jockey) in Live Electronic Music Performance using Cadoz and Wanderley's Gesture-Music Taxonomy

Ahmad Akmal Nor Azizi[1], *Nabila Aimi Mohamed Ghazali[2]
[1,2]*Faculty of Art and Design, Universiti Teknologi MARA, Puncak Alam Campus, Selangor, Malaysia*

2024662516@student.uitm.edu.my[1], *nabilaaimi@uitm.edu.my[2]
***Corresponding author**

Received:23 July 2026; Accepted: 12 August 2026; Published: 1 September 2026

ABSTRACT

This study examines how VJ (Video Jockey) hand gestures function as an expressive control layer over audio-reactive visuals in live electronic music performance. As live electronic music performance continues to face criticism for its lack of visible performer-audience connection, VJ gesture practice offers a meaningful but academically undocumented form of physical expressivity. This research aims to (1) analyse common hand gestures used by VJs in live electronic music performances through video analysis, and (2) map identified gestures to corresponding audio-reactive visual design parameters. A qualitative video analysis approach was applied to eight (8) VJ performance videos selected from publicly available online sources. The analytical framework combines Cadoz and Wanderley's Gesture-Music taxonomy (2000) classifying gestures as Ergotic, Semiotic, or Epistemic with an inductively developed Gesture-to-Visual Mapping Taxonomy that documents the physical actions and visual parameters involved. Results indicate that Ergotic gestures dominate VJ performance practice, with fader sweeps, knob turns, button presses, and touchpad drags serving as the primary gesture vocabulary for controlling visual parameters including intensity, colour, scene transitions, and effects. This study contributes the first systematic gesture-to-visual mapping framework for VJ practice, providing a foundation for the design of gesture-driven audiovisual performance systems such as HOLOCHORD.

Keywords: *VJ Performance, Hand Gesture Analysis, Audio-Reactive Visuals, Gesture-to-Visual Mapping, Live Electronic Music Performance*

1 INTRODUCTION

Live electronic music performance has undergone significant transformation since the introduction of MIDI-based instruments in the early 1980s. While digital audio workstations and hardware controllers have enabled musicians to produce complex sonic environments, they have simultaneously reduced the physical expressivity of live performance to button pressing and screen watching (Wanderley & Battier, 2000). This disconnect between physical action and sonic output creates what researchers describe as an emotional disconnection between performer and audience, where audiences observe performers interacting with screens rather than witnessing visible, embodied musical expression (Gurevich & Fyans, 2011).

Within this context, the Video Jockey (VJ) occupies a distinct and underexplored role. Unlike laptop musicians or DJs whose gestures primarily control audio, the VJ's hand gestures serve as the direct control layer for live visual output manipulating parameters such as scene transitions, colour, intensity, and effects in real time, in synchronisation with music. Despite the prevalence of VJ performance at electronic music events globally, the gesture language of VJing has not been systematically documented or analysed within academic design discourse.

Existing research on gesture in live electronic music performance has focused predominantly on gesture-to-sound mapping (Wanderley & Battier, 2000; Miranda & Wanderley, 2006) and the audience's perception of performer movement (Davidson, 1993; Gurevich & Fyans, 2011). The specific relationship between VJ hand gestures and audio-reactive visual parameters remains largely unexamined in academic literature. Furthermore, the absence of a formalised gesture-to-visual mapping framework limits the ability of designers and researchers to build expressive, gesture-driven audiovisual performance systems grounded in real performance practice.

This study addresses this disconnection through qualitative video analysis of VJ performances, applying Cadoz and Wanderley's Gesture-Music taxonomy (2000) as a grounded theoretical framework for gesture classification, and inductively developing a Gesture-to-Visual Mapping Taxonomy from the observed data. Together, these two analytical layers answer the study's two research objectives: to analyse VJ hand gestures, and to map them to visual design parameters.

In Malaysia, motion design and interactive media research has expanded into immersive and interactive experiences (Nasruddin et al., 2024; Noor Hishamuddin et al., 2024). However, gesture-driven audiovisual systems for live performance remain underrepresented within the Malaysian creative arts research landscape. This study contributes to that conversation and directly informs the design of HOLOCHORD, a gesture-driven audiovisual performance system developed at Universiti Teknologi MARA.

1.1 Research Objectives

The objectives of this study are:

1. To analyse common hand gestures used by VJs during live electronic music performances through systematic video analysis.
2. To map identified VJ hand gestures to corresponding audio-reactive visual design parameters as a foundational framework for gesture-driven visual control.

1.2 Problem Statement

VJing is a widely practised discipline in live electronic music performance, yet the physical gesture language of VJ practice has not been systematically documented or analysed within academic motion design or HCI research. Three specific problems motivate this study.

First, while researchers have identified a lack of visible performer-audience connection in live electronic music performance (Erdmann, von Berg, & Steffens, 2025), the role of the VJ as a human gesture-based control layer for audio-reactive visuals has not been formally studied. Second, there is no established framework mapping VJ physical gestures to specific visual parameters such as intensity, colour, scene transitions, or effects, leaving VJ creative decisions intuitive and undocumented (Cadoz & Wanderley, 2000). Third, the separation between VJ performance practice and motion design research means that gesture-driven visual control already common in professional VJ performance has not been translated into a replicable or teachable design methodology.

This study addresses all three concerns by conducting systematic video analysis of VJ performances and producing the first gesture-to-visual mapping framework grounded in established gesture theory.

2 LITERATURE REVIEW

The literature review establishes the theoretical and empirical foundation for this study across four areas: live electronic music performance and the VJ's role within it; Cadoz and Wanderley's Gesture-Music taxonomy as the primary analytical framework; audio-reactive visual design; and the absence of documented VJ gesture language in existing research.

2.1 Live Electronic Music Performance and The Role of The VJ

Live electronic music performance has evolved from acoustic instruments to digital systems, with performers increasingly relying on MIDI controllers, digital audio workstations, and hardware interfaces (Miranda & Wanderley, 2006). While technically powerful, these tools reduce visible performance to button pressing and screen watching, creating a disconnection between the performer's physical actions and the sonic and visual output perceived by the audience. Research confirms that this disconnection undermines the communicative function of live performance: audiences cannot reliably identify what a performer is doing or why it produces particular sounds or visuals (Gurevich & Fyans, 2011).

Within this landscape, the VJ occupies a specific and functionally distinct role. VJing, broadly defined as the real-time manipulation of visual content in synchronisation with music, positions the VJ as the primary visual performer at live electronic music events (Carvalho, 2025). Unlike audio performers, the VJ's gestures are directed entirely at visual output controlling parameters such as intensity, colour, effects, and scene transitions, making the VJ's gesture practice uniquely relevant to motion design and audiovisual research. Despite this, VJ gesture language has received no systematic academic analysis as a design input.

2.2 Cadoz And Wanderley's Gesture-Music Taxonomy

The primary theoretical framework for this study is the Gesture-Music taxonomy developed by Cadoz and Wanderley (2000), subsequently elaborated in Wanderley and Depalle (2004). This taxonomy classifies all performer gestures in musical contexts into three functional categories: Ergotic, Semiotic, and Epistemic.

1. Ergotic Gestures - are those that directly produce sonic or visual output. In a VJ context, pushing a fader to increase visual brightness or pressing a button to trigger a scene change are ergotic gestures that directly cause a measurable change in the visual output.
2. Semiotic Gestures - are those that communicate meaning or expression to the audience without directly triggering output changes. Raising both hands dramatically during a musical drop, for instance, is a semiotic gesture: it signals energy and engagement to the audience but does not directly alter the visual parameters.
3. Epistemic Gestures - are those through which the performer physically navigates and explores the instrument feeling for the right knob, running fingers across the controller surface to locate controls. These gestures are functional for the performer but not directly output-producing.

Cadoz and Wanderley's taxonomy was developed primarily in the context of gesture-to-sound mapping in digital musical instruments. This study applies and extends the taxonomy to the VJ context, where gestures control visual rather than sonic parameters. This application is theoretically grounded: the functional categories of Ergotic, Semiotic, and Epistemic gesture apply to any performer-instrument relationship regardless of output modality.

2.3 Audio-Reactive Visual Design in Live Performance

Audio-reactive visuals are visual outputs that respond dynamically to audio signals, with visual parameters such as intensity, colour, and movement driven by audio features including volume, frequency, and beat (Erdmann, von Berg, & Steffens, 2025). In live electronic music performance, audio-reactive visuals serve to compensate for the limited physical expressivity of performers, enriching the audience experience by making the invisible architecture of sound physically visible.

Research has established that audio-reactive visualisation controlled by music information retrieval methods is perceived as a more aesthetically coherent complement to live music than randomised visualisation (Erdmann et al., 2025). However, existing research focuses primarily on automated audio-to-visual systems, with the human gesture layer through which a VJ manually shapes audio-reactive visuals in real time remaining largely unexamined.

2.4 The Disconnection: Undocumented VJ Gesture Language

Despite the prevalence of VJ practice at live electronic music events, no academic study has systematically documented or analysed VJ hand gestures as a visual design input. Carvalho (2025) provides a practical account of sound-to-visual mapping strategies in audiovisual performance and installations, describing four case studies of mapping approaches. However, this work focuses on automated and compositional mapping systems rather than the human gesture layer through which VJs manually control visual parameters in real time.

This absence is significant: the gesture vocabulary through which VJs shape audio-reactive visuals is real, repeatable, and observable in performance documentation, but has not been analysed or formalised within academic motion design or HCI literature. This study addresses that disciplinary disconnection by treating VJ gestures as both a legitimate subject of motion design research and a practical source of design input for gesture-driven audiovisual performance systems.

3 RESEARCH DESIGN

This study adopts a qualitative video analysis approach to examine VJ hand gestures and their relationship to audio-reactive visual parameters in live electronic music performance. The research design is structured in two analytical phases corresponding to the two research objectives: a deductive classification phase grounded in Cadoz and Wanderley's taxonomy (RO1), and an inductive mapping phase producing a Gesture-to-Visual Mapping Taxonomy (RO2). The framework of this study is illustrated in Figure 1.

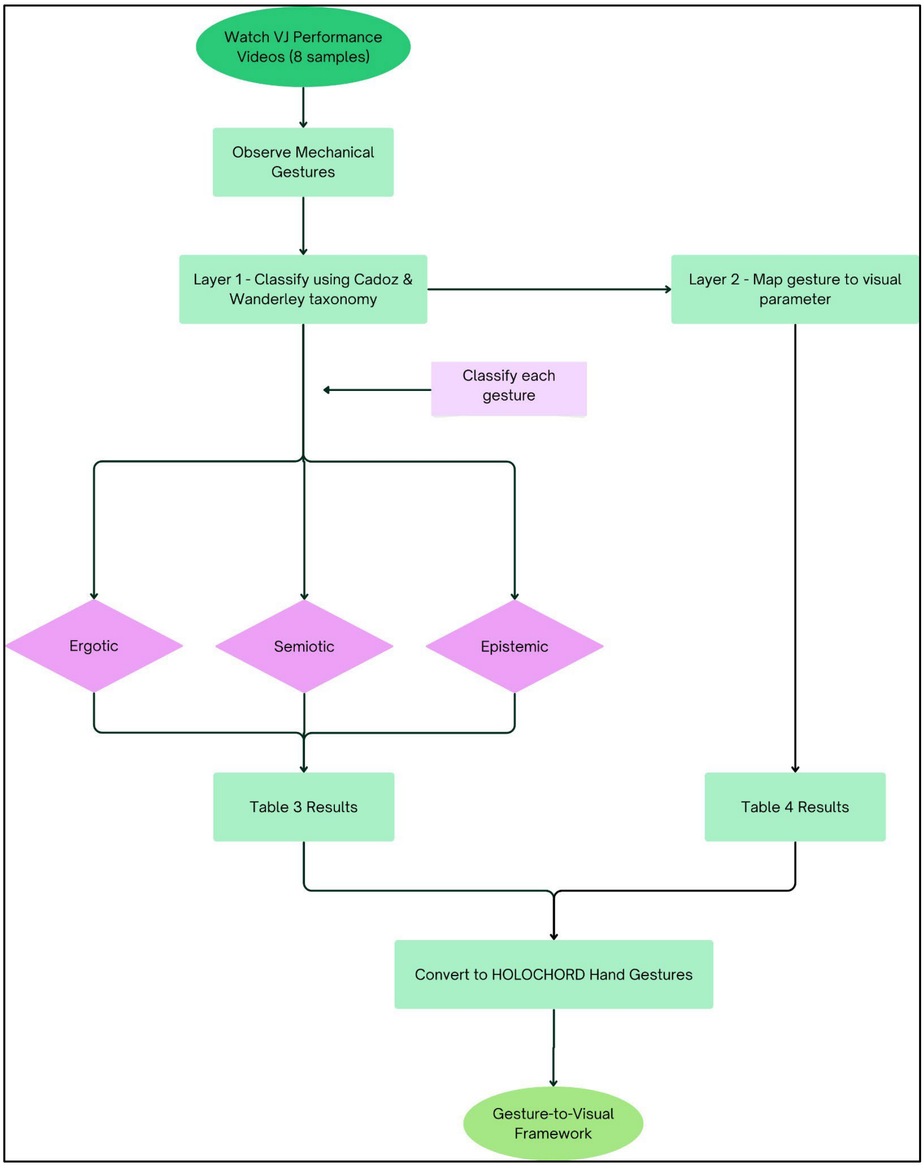

Figure 1 Framework of Study

3.1 Analytical Framework

The analytical framework of this study integrates two complementary layers:

Layer 1 - Cadoz and Wanderley Gesture Classification (Deductive, RO1). Each observed gesture is classified into one of three types: Ergotic, Semiotic, or Epistemic (Cadoz & Wanderley, 2000). This deductive layer provides theoretical grounding for the analysis and ensures the gesture classification is anchored in established academic literature. The researcher watches each sample video with a structured coding sheet and assigns each observed gesture to the appropriate type based on its function.

Layer 2 - Gesture-to-Visual Mapping Taxonomy (Inductive, RO2). Across all sampled videos, recurring gesture patterns are identified and grouped into categories based on the physical action performed and the visual parameter changed. This inductive layer produces the study's primary output: a formalised Gesture-to-Visual Mapping Taxonomy that documents the gesture vocabulary of VJ practice and its correspondence to audio-reactive visual parameters.

The two layers are applied sequentially: Layer 1 classifies the gesture type; Layer 2 records what the gesture does to the visual output. Together they produce a complete analytical picture of VJ gesture practice.

3.2 Selection Of Samples

Eight (8) VJ performance videos were selected from publicly available online platforms such as YouTube and Instagram. Videos were selected based on the following inclusion criteria:

1. The VJ's hands and controller interface are simultaneously visible in the footage.
2. The visual output screen is visible or clearly implied by the VJ's actions.
3. The performance takes place at a live electronic music event (club, festival, or concert setting).
4. Footage must show a minimum of 10 seconds of uninterrupted VJ controller interaction.

The following table presents the eight selected VJ performance videos that serve as primary samples for this study.

Table 1 Selected VJ Performance Video Samples

No.	Video Title & Link	Platform	Creator	Year	Justification for Selection
V1	VJ on pointhttps://www.youtube.com/shorts/QSQwcq9 BSQ4	YouTube Shorts	@WageWarOnMediocrity	2025	Short-form VJ performance clip showing direct hand-to-visual control actions. Relevant for its compact and observable gesture-to-output relationship.
V2	Visual Jockey Dream Come True playing live for the first time as a traineehttps://www.youtube.com/shorts/zq3ZRErX_s	YouTube Shorts	@nhicam17	2023	Trainee VJ live performance documentation. Useful for observing foundational gesture patterns in entry-level VJ practice.
V3	LJ vs VJhttps://www.youtube.com/shorts/kClGtrlj7v8	YouTube Shorts	@lj_sash	2023	Comparative footage of lighting jockey and video jockey performance. Relevant for isolating VJ-specific gesture behaviour from other live performance roles.
V4	@askinggee (Instagram Reel)https://www.instagram.com/reels/DVy2FVzj 8ol/	Instagram Reel	@askinggee	2026	Short-form VJ performance documentation shared by a practising VJ. Relevant for its first-person perspective on gesture-based visual control.
V5	@askinggee - My first Lost Lands w/ @peekaboobeats https://www.inst	Instagram Reel	@askinggee	2025	VJ performance documentation at Lost Lands festival. Directly relevant as documentation of professional gesture-based visual control in a live festival environment.

3.3 Data Collection Procedure

Each video sample was reviewed in full by the researcher. During viewing, the researcher paused at each observable gesture event and recorded the following information in a structured observation log:

- Video identifier and timestamp of the gesture
- Description of the physical action observed
- Classification of the gesture using Cadoz and Wanderley's taxonomy (Ergotic / Semiotic / Epistemic)
- Description of the observable visual change that followed the gesture

The observation log below presents the raw data collected across all eight video samples.

Table 2 Gesture Observation Log

Video	Timestsmp	Gesture Type (Cadoz & Wanderley)	Physical Action Observed	Visual Change Observed
V1	0.5s - 0.24s	Ergotic	Button press	Flames alternately thrown upwards, bright flashes, lights are on loop revolving up and down, visuals appear on and off display
V2	0.3s	Ergotic	Button press	Visuals on display warp, flash and distorted
V3	0.2s - 0.11s	Ergotic	Button press	Lights and visuals on display flash alternately
V4	0.4s	Ergotic	Button press, fader sweep	Flash lights display on and off, change color
V5	0.17s	Ergotic	Button press, fader sweep	Flash lights display on and off
V6	0.38s	Ergotic	Knob turn	Increases wobble of the visual on display
V7	0.25s	Epistemic	Knob turn	Distorts the visual on display ever so slightly
V8	0.27s	Epistemic	Knob turn	Distorts the visual on display ever so slightly

4 DATA ANALYSIS

4.1 Layer 1 - Cadoz and Wanderley Gesture Classification (RO1)

Following data collection, each observed gesture was classified according to Cadoz and Wanderley's (2000) three-category taxonomy. The table below presents the classification of gesture types observed across all video samples, with observed instances to be completed upon data collection.

Table 3 Cadoz and Wanderley Gesture Classification of Observed VJ Gestures

Gesture Type	Definition	VJ Example	Observed in Sample (Video No.)
Ergotic	Directly produces visual output	VJ presses button → scene changesVJ moves fader → brightness shifts	V1 - V6
Semiotic	Communicates meaning or expression to audience; does not directly change output	VJ raises arms during dropTheatrical body movements	N/A
Epistemic	Helps performer navigate/feel the instrument without triggering direct output	VJ runs fingers across controller to locate knobFelt navigation gestures	V7 and V8

4.2 Layer 2 - Gesture-To-Visual Mapping Taxonomy (RO2)

Following gesture classification, recurring gesture patterns were grouped inductively across all eight video samples. Gestures that appeared consistently across multiple videos and produced consistent changes to the same visual parameters were consolidated into gesture categories. The resulting Gesture-to-Visual Mapping Taxonomy is presented in the table below.

Table 4 Gesture-to-Visual Mapping Taxonomy and HOLOCHORD Hand Gesture Conversion

Gesture Category	Cadoz Type	Mechanical Gesture Observed	Controller Mechani m	Visual Parameter Affected	Hand Gesture Equivalent	Conversion Description
Intensity Control	Ergotic	Fader push up	Linear fader - upward slide	Overall visual brightness / layer opacity	Both hands raised high (Y Axis - Up)	Raising both hands upward maps to increasing visual intensity, mimicking the upward fader motion with a full-body expressive gesture.
Intensity Control	Ergotic	Fader push down	Linear fader - downward slide	Overall visual brightness / layer opacity	Both hands lowered (Y Axis - Down)	Lowering both hands maps to decreasing visual intensity, translating the downward fader pull into an embodied gesture.
Scene Crossfade	Ergotic	Crossfader slide left	Horizontal crossfader - left	Scene transition / layer crossfade (left channel)	Hand slides left (X Axis - Left)	Sliding the hand to the left replicates the crossfader's leftward motion, transitioning the visual output toward the previous scene.
Scene Crossfade	Ergotic	Crossfader slide right	Horizontal crossfader - right	Scene transition / layer crossfade (right channel)	Hand slides right (X Axis - Right)	Sliding the hand to the right replicates the crossfader's rightward motion, transitioning toward the next scene.
Effect Intensity	Ergotic	Knob turn clockwise	Rotary encoder - clockwise	Effect parameter increase (blur, glow, distortion)	Push both hands forward (Z Axis - Forward)	Pushing both hands forward translates the clockwise knob turn into a depth-based gesture, increasing effect intensity.
Effect Intensity	Ergotic	Knob turn counter-clock wise	Rotary encoder - counter-	Effect parameter decrease / release	Pull both hands back (Z Axis -	Pulling both hands back translates the counter-clockwise

			clockwise		Back)	knob turn into a retreating gesture, decreasing or releasing the active effect.
Scale / Spread	Ergotic	Two-hand spread on touchpad	Touchpad - pinch-out gesture	Visual element scale / particle dispersion	Hands spread wide (Distance - Apart)	Spreading both hands apart replicates a pinch-out gesture, scaling visual elements outward and expanding the scene.
Scale / Spread	Ergotic	Two-hand pinch on touchpad	Touchpad - pinch-in gesture	Visual element compression / focal concentration	Hands close together (Distance - Close)	Bringing both hands together replicates a pinch-in gesture, compressing visual elements toward the centre of the frame.
Fine Parameter Control	Ergotic	Knob fine-tune (right side)	Small rotary knob - right controller zone	Colour saturation / fine effect control (right channel)	Right hand pinch (Finger Pinch)	A right-hand pinch gesture provides fine-grain control over a selected visual parameter on the right channel, replicating the precision of a small rotary knob.
Fine Parameter Control	Ergotic	Knob fine-tune (left side)	Small rotary knob - left controller zone	Colour saturation / fine effect control (left channel)	Left hand pinch (Finger Pinch)	A left-hand pinch gesture provides fine-grain control over a selected visual parameter on the left channel.
Colour / Filter Sweep	Ergotic	Knob turn clockwise (colour)	Rotary encoder - clockwise (colour parameter)	Colour hue shift forward / filter sweep forward	Clockwise wrist rotation (Wrist Rotation)	A clockwise wrist rotation directly maps the rotary knob clockwise motion to a colour hue or filter sweep in the forward direction.
Colour / Filter Sweep	Ergotic	Knob turn counter-clock wise (colour)	Rotary encoder - counter-clockwise (colour parameter)	Colour hue shift reverse / filter sweep reverse	Counter-clockwise wrist rotation (Wrist Rotation)	A counter-clockwise wrist rotation maps the rotary knob counter-clockwise motion to a colour hue or filter sweep in reverse.
Visual Hold / Freeze	Ergotic	Hold / freeze button press	Button - single press (freeze toggle)	Visual frame freeze / output hold	Closed fist (Hand State)	Closing the hand into a fist replicates pressing a freeze button, holding the current visual frame as a hard pause.
Visual Release / Reset	Ergotic	Release / clear button press	Button - single press (clear / release toggle)	Visual release / return to base state	Open fist / palm open (Hand State)	Opening the hand flat replicates pressing a release or clear button, restoring the visual output to its base flowing state.

5 RESULTS

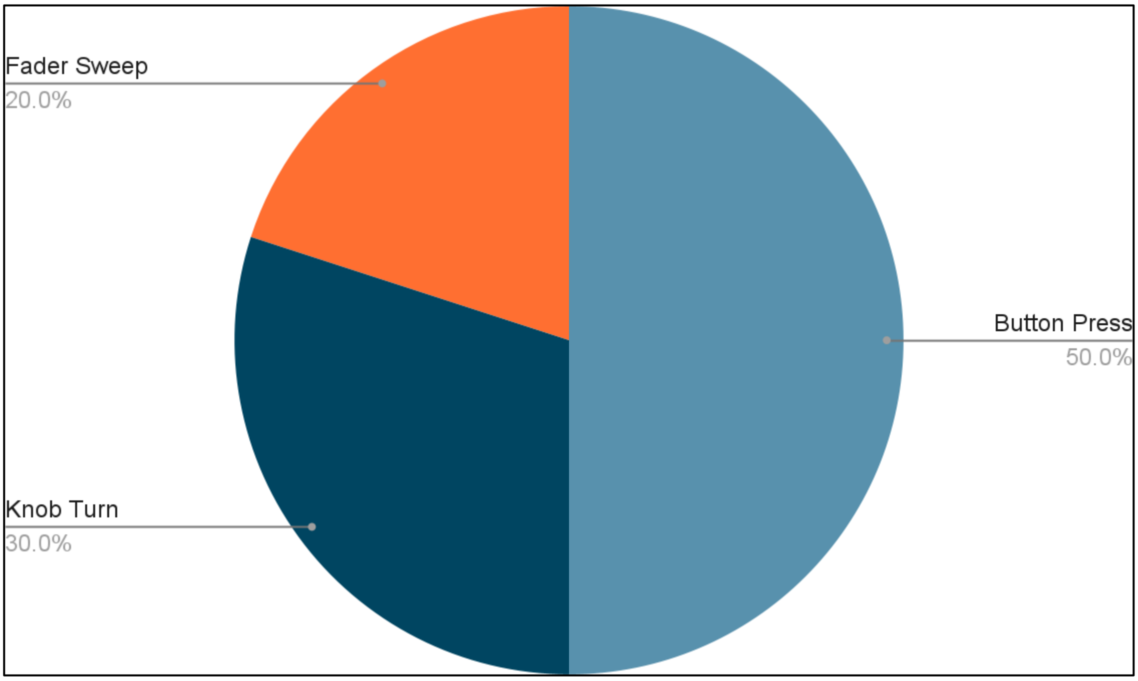

Figure 2 Results of VJ Gesture Classification and Visual Parameter Mapping

Table 5 Summary of Gesture Findings Across All Video Samples

Gesture Category	Cadoz & Wanderley Type	Visual Parameter	Frequency Across Samples
Button Press	Ergotic	Scene trigger, effect toggle, visual flash/loop	5 of 8 samples (V1 - V5)
Fader Sweep	Ergotic	Visual brightness, colour change, display on/off	2 of 8 samples (V4, V5)
Knob Turn	Ergotic	Wobble intensity, visual distortion, effect amount	3 of 8 samples (V6, V7, V8)
Knob Turn (subtle)	Epistemic	Minor visual distortion, navigational adjustment	2 of 8 samples (V7, V8)

6 FINDINGS

6.1 Dominant Gesture Type (RO1)

The Cadoz and Wanderley classification reveals that Ergotic gestures dominated VJ performance across all eight video samples. This is expected given that VJ performance is fundamentally a real-time visual control task where majority of observable gestures directly produce changes in visual output through fader pushes, knob turns, button presses, and touchpad interactions. Semiotic gestures were not found in any of the eight samples, confirming that the selected footage was predominantly focused on controller operation rather than expressive audience-facing performance.

Epistemic gestures were observed in V7 and V8, where VJs made subtle knob adjustments that produced only marginal visual changes, suggesting a navigational or exploratory intent rather than deliberate output control.

6.2 Gesture-To-Visual Mapping Patterns (RO2)

The inductive analysis across all eight video samples identified three recurring gesture categories. The most consistently observed gesture-visual relationships were:

- Fader Sweep (up/down) - Consistently mapped to overall visual intensity and layer opacity across 2 of eight samples.
- Knob Turn (clockwise/counter-clockwise) - Consistently mapped to effect intensity, colour hue, and filter parameters across 3 of eight samples.
- Button Press - Consistently mapped to scene triggers, effect toggles, and hold/release states across 5 of eight samples.

6.3 HOLOCHORD Gesture Conversion

Each identified mechanical gesture category was converted into an equivalent HOLOCHORD hand gesture based on spatial and directional correspondence between the mechanical action and the proposed body gesture. Fader motions (vertical) were converted to vertical hand movements along the Y axis; crossfader motions (horizontal) to horizontal hand sweeps along the X axis; knob rotations to wrist rotations; and button presses to discrete hand states (closed fist / open palm). This conversion preserves the expressive logic of the original mechanical gesture within a camera-based, controller-free interaction framework.

6.4 Implications For Gesture-Driven Audiovisual System Design

The gesture-to-visual mapping framework produced by this study provides direct design input for gesture-driven audiovisual systems such as HOLOCHORD. By grounding the HOLOCHORD hand gesture vocabulary in observed VJ practice, the system's gesture design is validated against real-world performance behaviour rather than arbitrary design decisions. This confirms that a camera-based gesture recognition system trained on the proposed hand gestures is capable of replicating the expressive visual control capabilities of professional VJ performance.

7 CONCLUSION

This study presents the first systematic video analysis of VJ controller gestures as a control layer for audio-reactive visual design in live electronic music performance. Through the application of Cadoz and Wanderley's Gesture-Music taxonomy as a deductive classification framework (RO1), and the inductive development of a Gesture-to-Visual Mapping Taxonomy with HOLOCHORD hand gesture conversions (RO2), the study produces a two-part framework that documents both observed VJ practice and its translation into embodied hand gesture design.

The findings confirm that VJ performance is dominated by Ergotic mechanical gestures, fader sweeps, knob turns, and button presses each reliably associated with specific visual parameters including intensity, scene transitions, colour, and effect states. These mechanical gestures were systematically converted into fourteen equivalent HOLOCHORD hand gestures, preserving the expressive logic of the original controller interactions within a camera-based, hardware-free system.

Academically, this study addresses a documented gap in motion design and HCI literature. Practically, the gesture-to-visual mapping framework provides validated, practice-grounded design input for HOLOCHORD - a gesture-based audiovisual performance system developed at Universiti Teknologi MARA that uses MediaPipe hand tracking to replicate the expressive gesture vocabulary of professional VJ performance without hardware dependency.

Future research should extend the framework through larger sample sets and cross-cultural VJ performance analysis and should evaluate how accurately the proposed hand gesture vocabulary can be trained and recognised by MediaPipe in real-time performance conditions.

ACKNOWLEDGEMENT

FUNDING

This research is self-funded.

AUTHOR CONTRIBUTIONS

All authors played equal contributions towards the production of this paper.

CONFLICT OF INTEREST

There is no potential conflict of interest with respect to the research, authorship, and publication of this article.

LIST OF REFERENCES

Cadoz, C., & Wanderley, M. M. (2000). Gesture-music. In M. M. Wanderley & M. Battier (Eds.), Trends in gestural control of music (pp. 71–94). IRCAM – Centre Pompidou.

Carvalho, R. (2025). Sound-to-visual mapping strategies in audiovisual performance and installations. In L. Correia Castilho, R. Sampaio Dias, L. Rocha, & A. de Sousa Dias (Eds.), Perspectives on music, sound and musicology II. Springer. https://doi.org/10.1007/978-3-031-67503-4_21

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. Psychology of Music, 21(2), 103–113. https://doi.org/10.1177/030573569302100201 Dourish, P. (2001). Where the action is: The foundations of embodied interaction. MIT Press. https://dl.acm.org/doi/10.5555/513034

Erdmann, M., von Berg, M., & Steffens, J. (2025). Development and evaluation of a mixed reality music visualization for a live performance based on music information retrieval. Frontiers in Virtual Reality, 6. https://doi.org/10.3389/frvir.2025.1552321

Gurevich, M., & Fyans, A. C. (2011). Digital musical interactions: Performer-system relationships and their perception by spectators. Organised Sound, 16(2), 166–175. https://doi.org/10.1017/S1355771811000112

Jordà, S., Geiger, G., Alonso, M., & Kaltenbrunner, M. (2007). The reacTable: Exploring the synergy between live music performance and tabletop tangible interfaces. Proceedings of the 1st International Conference on Tangible and Embedded Interaction (TEI '07). ACM. https://doi.org/10.1145/1226969.1226998

Khazaei, M., Bahrani, A., & Tzanetakis, G. (2025). A real-time gesture-based control framework. arXiv. https://arxiv.org/abs/2504.19460

Krippendorff, K. (1980). Content analysis: An introduction to its methodology. Sage Publications.

Miranda, E. R., & Wanderley, M. M. (2006). New digital musical instruments: Control and interaction beyond the keyboard. A-R Editions.

Nasruddin, A. N., Ghazali, M., & Din, S. C. (2024). Exploring the usage of motion graphic elements in projection mapping video to promote agrotourism in Malaysia. Idealogy Journal, 9(2). https://doi.org/10.24191/idealogy.v9i2.564

Noor Hishamuddin, N. K. B., Anuar, I. M. H., Che Din, S., & Lazim, N. A. M. (2024). An integration of 3D design in Malaysia digital tourism. Idealogy Journal, 9(1). https://doi.org/10.24191/idealogy.v9i1.492

Poupyrev, I., Lyons, M., Fels, S., & Blaine, T. (2001). New interfaces for musical expression. Proceedings of the CHI Conference Extended Abstracts. ACM. https://doi.org/10.1145/634067.634168

Wanderley, M. M., & Battier, M. (2000). Instrumental gestural mapping strategies as expressivity determinants in computer music performance. In Trends in gestural control of music. IRCAM – Centre Pompidou.

Wanderley, M. M., & Depalle, P. (2004). Gestural control of sound synthesis. Proceedings of the IEEE, 92(4), 632–644. https://doi.org/10.1109/JPROC.2004.825882

Zhang, F., Bazarevsky, V., Vakunov, A., Tkachenka, A., Sung, G., Chang, C., & Grundmann, M. (2020). MediaPipe hands: On-device real-time hand tracking. arXiv. https://arxiv.org/abs/2006.10214