

Multi-Sensory Collaboration Design Strategies for Construction Machinery Teleoperation Based on Embodied Cognition

Hao Ge, Yonghong Liu, and Zhe Dai

Hunan University, Hunan, 410000, China

ABSTRACT

Aiming at the problems of lack of telepresence, low operation accuracy, and high cognitive load caused by relying solely on single visual feedback in complex and extreme environments, this study, guided by embodied cognition theory, systematically constructs a multi-sensory synergistic strategy for construction machinery teleoperation scenarios to comprehensively enhance operation precision, telepresence, and human-machine interaction efficiency. First, based on literature research on the characteristics of embodied cognition theory and multi-sensory design principles, a progressive embodied multi-sensory synergistic framework is innovatively proposed: “Embodied Multi-sensory Anchoring-Cognitive Consistency Integration-Context-Adaptive Synergy.” Subsequently, guided by this synergistic framework, user research and Kano model analysis were conducted to deeply excavate operator pain points, clarify core multi-sensory needs and synergy priorities, and build a multi-sensory synergistic requirement model for this scenario. On this basis, a three-part multi-sensory synergistic design strategy system was refined: Perceptual Reconstruction, Cognitive Fusion, and Contextual Adaptation. The strategies proposed in this study provide specific ideas and practical paths for the multi-sensory synergistic optimization of future teleoperation devices, and offer more adaptive solutions for human-machine interaction innovation in the intelligent upgrading of construction machinery, assisting designers in precisely optimizing designs and enhancing the operational experience.

Keywords: Embodied cognition, Construction machinery teleoperation, Multi-sensory synergy, Kano model

INTRODUCTION

With the complication and intelligent development of engineering construction scenarios, teleoperation of construction machinery has become a core direction for improving operational safety and efficiency, with core demands focused on operational performance optimization and interactive experience enhancement. Although existing research has made some progress in human-machine interaction optimization and perceptual feedback control technology, there are obvious limitations: sensory applications are mostly concentrated on visual-tactile dimensions, lacking effective use of auditory

and somatic senses (Kondo, 2021; Jankowski et al., 2015; Truong et al., 2017); multi-channel feedback synergy is insufficient, and a systematic strategy has not been formed, making it difficult to adapt to complex construction scenarios.

Multi-sensory theory has formed a mature research system in the product design field, widely penetrating scenarios such as education, medical rehabilitation, and smart homes (Silvennoinen et al., 2023; Sekerci et al., 2023; Liu et al., 2025). Its core value of enhancing product usability and immersive experience through multi-sensory channel synergistic adaptation is highly compatible with the demands of construction machinery teleoperation to improve operation precision and reduce cognitive load, providing a feasible path for solving existing pain points. However, this theory has not yet achieved systematic integrated application in construction machinery teleoperation scenarios, indicating a research gap that urgently needs targeted translation and adaptation.

Embodied cognition theory, with “body-technology-environment dynamic coupling” as its core logical framework, has formed a complete theoretical system and practical paradigm in the design field. This theory emphasizes the foundational role of bodily perception-action experience in cognition, which can provide underlying theoretical support and methodology optimization for multi-sensory design, effectively resolving problems such as sensory synergy imbalance and the disconnection between interaction logic and human cognitive laws faced by multi-sensory design in engineering scenarios, achieving precise adaptation of multi-sensory feedback with bodily perception, and further enhancing the intuitiveness, precision, and safety of teleoperation. In summary, based on the realistic pain points of construction machinery teleoperation, introducing multi-sensory design concepts and conducting targeted research supported by embodied cognition theory has important theoretical value and practical significance, and is an inevitable development direction for research in this field.

CONSTRUCTION OF A MULTISENSORY SYNERGY FRAMEWORK BASED ON EMBODIED COGNITION

Embodied Cognition Theory

Embodied cognition theory shifts cognitive science from “disembodied cognition” to an “embodied interaction” paradigm. It proposes that cognition emerges from the organism’s physiological structure, sensorimotor system, and environmental interaction (Varela et al., 2017; Clark and Chalmers, 1998; De Jaegher, 2010; Slater et al., 2008). Its core characteristics define the “4E” dimensions:

Embodied: Cognition relies on physical structure and physiological experience rather than isolated symbolic computation. Conceptual representation acts as a reactivation of modal experiences like vision and touch (Varela et al., 2017); Barsalou’s theory further views cognition as neural simulations of perception-action experiences. Additionally, Kiltner’s virtual embodiment framework offers quantitative tools (Kiltner et al., 2012), demonstrating that cognitive processing is anchored in bodily experience and physiological schemas.

Embedded: Cognitive activities are nested within situational constraints, utilizing the environment as a resource through “cognitive offloading”. Hutchins notes that cognition is distributed across environmental and social resources rather than confined to the brain (Hutchins, 1995). Clark and Chalmers highlight the use of environmental structures to simplify cognitive processes, making the environment a key support for efficient cognition (Clark and Chalmers, 1998).

Enactive: Cognition emerges dynamically within the “perception-action” loop rather than as a static representation. De Jaegher views this as maintaining adaptive coupling through self-regulation (De Jaegher et al., 2010), while Varela emphasizes that meaning is “co-generated” through dynamic interaction, where understanding is continuously constructed via perception-action processes.

Extended: Cognitive boundaries expand to include external tools, forming a “brain-body-tool” system. Clark and Chalmers propose that tools deeply coupled with the brain become part of the cognitive system; Slater’s virtual hand illusion confirmed that humans can generate a “sense of ownership” for synchronized external tools (Slater et al., 2008), indicating cognition is an open system expanding its functional boundaries through technological integration.

Multisensory Design Principles

Multi-sensory synergistic design is a systematic method transcending single-sensory optimization, originating from neuroscience research. Stein revealed that information from different modalities forms a unified perceptual representation through synergistic processing. This process follows three core principles: spatial adjacency, temporal synchrony, and the principle of inverse effectiveness (Stein and Meredith, 1993). On this basis, Ernst systematized this into Bayesian Optimal Integration Theory, bridging neuroscience and computational models to enable design to allocate weights based on the reliability of sensory signals (Ernst and Bulthoff, 2004).

Based on this mechanism, the Consistency Principle serves as the neural basis of integration. Temporal consistency requires controlling the presentation time difference of cross-modal stimuli within the brain’s “temporal integration window” (approx. 150–450ms), as exceeding this reduces efficiency or triggers perceptual separation (Calvert and Thesen, 2004). Spatial consistency emphasizes that information must originate from proximate locations, conforming to “spatial tuning” to prevent response delays and maintain credibility (Spence et al., 2005). Semantic consistency focuses on coordinating meaning, rhythm, and emotion to fit cognitive priors and convey consistent metaphors.

The Cross-modal Correspondence Principle refers to systematic associations between modalities based on embodied experience and Bayesian inference to “use prior knowledge to reduce uncertainty”. The core is using prior mapping to design natural experiences. For example, rounded shapes associate with light, high-frequency sounds, while red packaging triggers associations of sweet or spicy tastes (Spence, 2011). These correspondences are widely applied in design to create experiences that fit psychological expectations.

The Complementarity and Redundancy Principles constitute the strategy for improving robustness. The complementarity principle emphasizes that different channels provide unique dimensions to form a comprehensive perceptual picture. The redundancy principle refers to repeated transmission of key information across channels to resist interference and improve reliability.

Progressive Synergistic Innovation Path Model

Combining the “4E” characteristics of embodied cognition theory with multi-sensory design principles, and addressing the specificity of the construction machinery teleoperation scenario, this study innovatively proposes a “Embodied Multi-sensory Anchoring-Cognitive Consistency Integration-Context-Adaptive Synergy” progressive multi-sensory synergistic framework, providing core logical support for subsequent requirement research and design strategy formulation (see Figure 1).

Phase 1: Embodied Multi-sensory Anchoring. This phase constructs the embodied perceptual foundation of the remote system by positioning feedback channels. Based on Gibson’s ecological perception theory, perception aims to capture action possibilities (Gibson, 1979). Spence further notes that the brain prioritizes information relevant to behavioral goals, while irrelevant input increases cognitive load (Spence and Frings., 2019; Spence et al., 2005; Spence, 2020). Therefore, the design follows “Task-Oriented Sensory Anchoring,” optimizing dominant channels for construction machinery teleoperation rather than pursuing full coverage. Vision governs spatial structure and motion; haptics transmits physical interaction information; audition assists monitoring; and proprioception supports posture perception. Taste and smell are restricted to avoid interference. This phase focuses on ensuring efficient reception of key information, providing a clear, low-noise foundation for constructing the sense of embodiment.

Phase 2: Cognitive Consistency Integration. Focusing on the construction of the sense of embodiment, based on anchoring clear perceptual inputs, and relying on multi-sensory consistency and cross-modal correspondence principles, this phase constructs a stable cognitive reference framework. The core goal is to help operators establish a virtual sense of embodiment, including the sense of ownership, sense of agency, and sense of self-location, achieving the cognition of the teleoperation tool as an extension of one’s own body (Kilteni et al., 2012). The sense of agency is the real-time precise causal perception between the operator’s intention and the operation terminal feedback; design needs to follow the temporal consistency principle, ensuring the time difference between multi-sensory feedback and motion commands falls within the neural integration window to overcome the inherent latency of remote systems. The sense of self-location is the operator’s perception of the relative position between themselves and the external device in the workspace; design should follow the spatial consistency principle, achieving precise spatial correspondence of visual, auditory, and haptic information, especially emphasizing the often-overlooked spatialized audition, such as issuing a warning sound from the left channel

when danger appears on the left, reinforcing spatial localization. The sense of ownership is the tool extension cognition of incorporating the operation terminal into the body schema; deeply it needs to follow the semantic consistency principle, where external feedback needs to fit cognitive logic and emotional expectations to reduce cognitive friction, maintaining multi-modal feedback semantic association and emotional synergy to support the construction of ownership.

Phase 3: Context-Adaptive Synergy. Focusing on the stabilization and maintenance of the sense of embodiment, based on the “Enactive” thought of embodied cognition theory, emphasizing that cognition originates from the dynamic interaction between the subject and the environment rather than static representation, the remote interaction system therefore needs context-adaptive capabilities. This phase enhances the robustness of the interaction system in complex uncertain situations and reduces cognitive load through constructing a system framework simulating the human brain’s “Sensory Re-weighting” and real-time regulation of “Signal Monitoring -Task Mapping - Environment Adaptation,” thereby stabilizing and continuing the virtual sense of embodiment constructed by basic synergy. This dynamic synergy mechanism includes top-down task-driven synergy and bottom-up environment-adaptive synergy. Task-driven synergy, based on “Task-Dependent Sensory Integration,” emphasizes active weight regulation by cognitive intent (Dong et al., 2025). It prioritizes haptics and proprioception for contact tasks, and vision with spatialized audio for navigation, using Bayesian inference of intent to dynamically adjust feedback. Environment-adaptive synergy allocates weights by assessing signal reliability—increasing auditory and haptic weights if vision degrades—and fuses substitution with redundancy. Sensory substitution compensates for limited channels, while redundancy involves parallel excitation of multiple channels during high-risk states to ensure stability and continuous embodied experience.

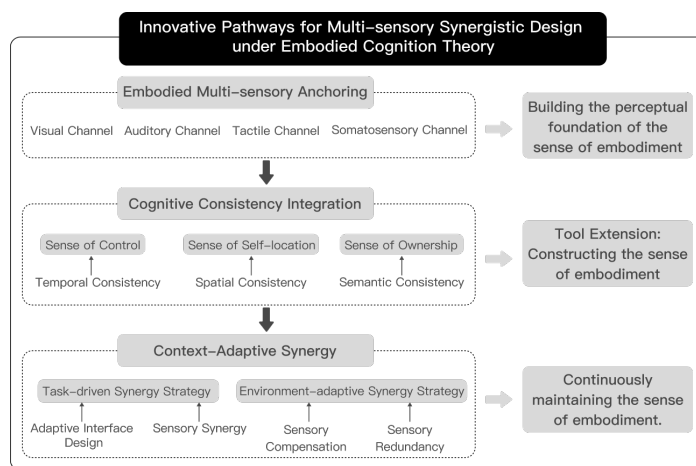


Figure 1: Innovative pathways for embodied multi-sensory synergistic design.

MODELING MULTI-SENSORY SYNERGISTIC REQUIREMENTS FOR CONSTRUCTION MACHINERY TELEOPERATION

User Need Mining and Summary

To precisely mine multi-sensory needs, this study adopted a “Expert Focus Group Exploration + User Interview Verification” path. Expert analysis revealed deep pain points: Perceptually, flat screens lack depth for distance judgment, while narrow fields of view miss peripheral warnings; existing haptics lack weight and resistance for distinguishing loads. Regarding interaction, latency and spatial misalignment cause motion sickness, while complex logic lacks intuition. Psychologically, low telepresence hinders trust and increases anxiety.

Subsequent operator interviews verified these points and proposed specific expectations. In the embodied sensory anchoring phase, users desired enhanced visual depth for landing judgment and high-fidelity force feedback for load distinction. For cognitive consistency integration, users emphasized synchronization, requiring alarm orientations to spatially match visual dangers and realistic driving perspectives. In the context-adaptive synergy phase, users proposed sensory compensation for extreme environments; under conditions like dust, the system should enhance tactile or auditory intensity to compensate for visual limitations, ensuring safety.

Synthesizing these conclusions, this study systematically constructed a multi-sensory synergistic design requirement list (see Table 1) based on the three-stage innovation pathway. This provides a core basis for subsequent Kano model quantitative research and requirement priority determination.

Table 1: User multi-sensory design requirements list.

Phase	Design Requirement	Sensory Type
Embodied	A1: Semi-enclosed styling	
Multi-sensory	A2: Ultra-wide FOV	
Anchoring	A3: 360° surround view	Vision
	A4: Depth information enhancement	
	A5: Device status visualization	
	A6: Ambient light simulation	
	A7: Real-time force feedback	
	A8: High-fidelity force feedback	Haptic
	A9: Device status vibration alert	
	A10: Operation sound feedback	
	A11: Device status audio monitoring	Auditory
	A12: Device tilt simulation	
	A13: Device bump simulation	Vestibular
	A14: Device acceleration simulation	

(Continued)

Table 1: Continued.

Phase	Design Requirement	Sensory Type
Somatic Cognitive Consistency Integration	B1: Driving vision consistency design	Vision
	B2: UI information simplification	
	B3: Button layout conforming to ergonomics	
	B4: Spatialized audio	Auditory
	B5: Multi-sensory feedback temporal synchronization	Multi-sensory
	B6: Multi-sensory feedback spatial synchronization	
	B7: Multi-sensory feedback semantic consistency	
Multi-sensory Context-Adaptive Synergy	C1: Working condition-page information adaptation	Vision
	C2: Dust/Fog perspective	
	C3: Night vision enhancement	
	C4: Landing point visual assistance	
	C5: Tactile feedback dynamic sensitivity	Haptic
	C6: Ambient noise intelligent masking	Auditory
	C7: Visual limitation - Tactile/Auditory compensation	Multi-sensory
	C8: Tactile information visualization	
	C9: Posture data visualization	
	C10: Sensory redundancy design	

Quantification and Classification of Requirement Attributes Based on Kano Model

The Kano model, first proposed by Noriaki Kano (Kano et al., 1984) in the 1980s, is a classic theoretical tool for evaluating user satisfaction. This model classifies user needs into six categories: Must-be requirements (M), One-dimensional requirements (O), Attractive requirements (A), Indifferent requirements (I), Reverse requirements (R), and Questionable requirements (Q). In actual analysis, questionable requirements are usually not included in statistics and calculations as they appear very rarely. Since its proposal, this model has been widely applied and verified in the fields of quality management and user satisfaction research.

Based on preliminary research, this study designed 35 indicators and selected 97 valid questionnaires from 135 collected samples. Berger's Better-Worse coefficient model was employed to quantify the satisfaction weights of each sensory factor (Berger et al., 1993) (see Table 2). Through a quadrant diagram using the Better coefficient and the absolute value of the Worse coefficient as axes, with the mean value serving as the dividing line, the distribution characteristics of One-dimensional, Attractive, Indifferent, and Must-be requirements were visually presented (see Figure 2). This analysis clarified the composition of core user needs for teleoperation devices, establishing a priority sequence for satisfaction: Must-be > One-dimensional > Attractive > Indifferent requirements (see Table 3).

Table 2: Kano model questionnaire data calculation results (in %).

User Needs	A	O	M	I	R	Better	Worse
B7	32.65	25.51	20.41	21.43	0.00	45.92	-58.16
B6	30.61	20.41	18.37	28.57	2.04	39.58	-52.08
B5	42.86	14.29	18.37	24.49	0.00	32.65	-57.14
B4	17.35	36.73	29.59	16.33	0.00	66.33	-54.08
A9	12.24	32.65	31.63	23.47	0.00	64.29	-44.90
C6	11.22	32.65	28.57	26.53	1.02	61.86	-44.33
A5	20.41	34.69	21.43	22.45	1.02	56.70	-55.67
C10	26.53	32.65	21.43	18.37	1.02	54.64	-59.79
B3	25.51	32.65	18.37	22.45	1.02	51.55	-58.76
A11	21.43	27.55	21.43	27.55	2.04	50.00	-50.00
A3	8.16	33.67	40.82	16.33	1.02	75.26	-42.27
A2	14.29	32.65	38.78	13.27	1.02	72.16	-47.42
A8	9.18	33.67	35.71	20.41	1.02	70.10	-43.30
A4	9.18	29.59	38.78	20.41	2.04	69.79	-39.58
C4	13.27	26.53	40.82	18.37	1.02	68.04	-40.21
C2	12.24	29.59	35.71	20.41	1.02	66.67	-42.71
A7	19.39	31.63	32.65	13.27	3.06	66.32	-52.63
C3	5.31	29.59	34.69	18.37	2.04	65.63	-45.83
A1	10.20	24.49	38.78	26.53	0.00	63.27	-34.69
C8	9.18	25.51	34.69	28.57	2.04	61.46	-35.42
C7	21.43	28.57	32.65	17.35	0.00	61.22	-50.00
A12	12.24	26.53	33.67	26.53	1.02	60.82	-39.18
C1	13.27	25.51	33.67	27.55	0.00	59.18	-38.78
B1	22.45	24.49	28.57	23.47	1.02	53.61	-47.42
C9	10.20	31.63	25.51	32.65	0.00	57.14	-41.84
C5	10.20	25.51	28.57	34.69	1.02	54.64	-36.08
B2	15.31	25.51	25.51	33.67	0.00	51.02	-40.82
A10	11.22	24.49	24.49	37.76	2.04	50.00	-36.46
A6	12.24	20.41	21.43	42.86	3.06	43.16	-33.68
A13	13.27	24.49	15.31	41.84	5.10	41.94	-39.78
A14	12.24	21.43	14.29	39.80	12.24	40.70	-38.37



Figure 2: Kano model quadrant analysis.

Table 3: User multi-sensory design requirements list.

Requirement Type	User Needs	
Must-be Requirements	B5: Multi-sensory feedback temporal synchronization	
	B6: Multi-sensory feedback spatial synchronization	
	B7: Multi-sensory feedback semantic consistency	
	B3: Button layout conforming to ergonomics	
	C10: Sensory redundancy design	
	A5: Device status visualization	
	A11: Device status audio monitoring	
	B1: Driving vision consistency design	
	One-dimensional Requirements	B4: Spatialized audio
		A2: Ultra-wide FOV
C3: Night vision enhancement		
A7: Real-time force feedback		
C7: Visual limitation - Tactile/Auditory compensation		
Attractive Requirements	A3: 360° surround view	
	A8: High-fidelity force feedback	
	A4: Depth information enhancement	
	C4: Landing point visual assistance	
	C2: Dust/Fog perspective	
	A9: Device status vibration alert	
	A1: Semi-enclosed styling	
	C8: Tactile information visualization	
	A12: Device tilt simulation	
	C6: Ambient noise intelligent masking	
C1: Working condition-page information adaptation		
Indifferent Requirements	C9: Posture data visualization	
	C5: Tactile feedback dynamic sensitivity	
	A10: Operation sound feedback	
	B2: UI information simplification	
	A6: Ambient light simulation	
	A14: Device acceleration simulation	
A13: Device bump simulation		

Generation of Embodied Multi-Sensory Synergistic Design Requirement Model

Based on the theoretical innovation pathways and Kano model requirement analysis, a Multi-sensory Synergistic Design Requirement Model for teleoperation of Construction Machinery under Embodied Cognition was constructed (see Figure 3). The model systematically organizes the requirement types corresponding to the three stages of Embodied Multi-sensory Anchoring, Cognitive Consistency Integration, and Context-Adaptive Synergy, as well as the goals of each stage, providing theoretical and practical basis for the design of construction machinery teleoperation devices.

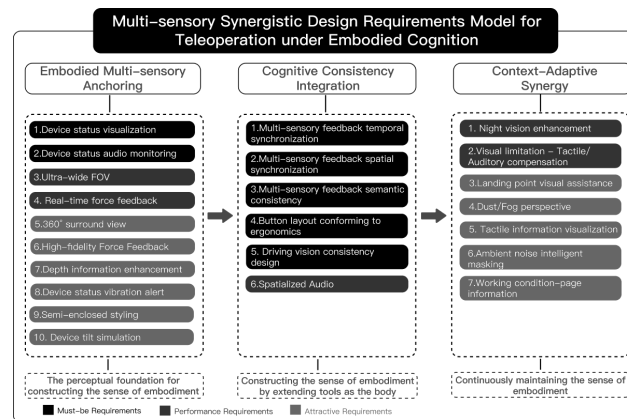


Figure 3: Multi-sensory synergistic design requirements model for teleoperation.

MULTI-SENSORY SYNERGISTIC DESIGN STRATEGY SYSTEM FOR CONSTRUCTION MACHINERY TELEOPERATION COCKPIT

Perceptual Reconstruction Strategy

At the visual level, peripheral visual perception is restored through semi-enclosed styling and surround view design, paired with depth enhancement technology and landing point visual assistance to make up for stereo deficits and solve distance judgment problems; at the tactile level, addressing pain points of “lack of resistance” and “difficulty in distinguishing load,” a variable damping force feedback system achieves high-fidelity real-time feedback, allowing operators to precisely distinguish geological hardness and load status; at the auditory level, a spatialized audio system extracts key functional sounds like engine load and hydraulic operation and presents them spatially, filtering redundant noise through intelligent ambient masking technology; at the somatic level, multi-degree-of-freedom motion seats simulate the tilt of remote equipment, completing body posture perception, allowing operators to intuitively grasp running status through proprioception.

Cognitive Fusion Strategy

Adhere to the principle of cross-modal spatiotemporal consistency, strictly controlling the time difference of visual, tactile, auditory, and somatic feedback within the neural integration window, ensuring precise spatial correspondence of multi-sensory information, eliminating cognitive bias and motion sickness reactions caused by spatiotemporal misalignment, and assisting the brain in integrating remote machinery as an extension of one's own body; follow the principle of cross-modal semantic consistency, maintaining driving vision consistency in field of view presentation, adopting button layouts and natural interaction logic conforming to ergonomics in operation logic, with multi-modal feedback all pointing to the same core information, ensuring accurate and unified semantic transmission, reducing cognitive friction, and transforming operation from "deliberate thinking" to "intuitive reaction."

Context Adaptation Strategy

Implement task-driven dynamic weight allocation, strengthening panoramic vision and somatic feedback during the traveling phase to ensure driving safety; focus on local fine vision and joystick force feedback during the operation phase to improve operation precision, with the interface automatically adjusting information hierarchy according to the task stage to avoid cognitive overload; establish an extreme environment compensation mechanism, launching "Visual Impairment - Auditory/Tactile Compensation" linkage when encountering visual-restricted working conditions such as heavy fog or dust, compensating for visual insufficiency by enhancing tactile feedback sensitivity and highlighting spatialized warning sounds, while superimposing image enhancement algorithms to strengthen visual effects and break through environmental constraints; build a sensory redundancy guarantee system, continuously maintaining embodied immersion and improving system robustness and reliability under complex working conditions through a "Visual Cue + Auditory Alarm + Tactile Vibration" multi-sensory redundancy layout.

CONCLUSION

Targeting the pain points of lack of telepresence and cognitive overload in construction machinery teleoperation, this study introduces embodied cognition theory to construct a theoretical and practical system for multi-sensory synergistic design. Through systematic research, the "Embodied Multi-sensory Anchoring-Cognitive Consistency Integration-Context-Adaptive Synergy" logical framework was constructed, revealing that the construction of the sense of embodiment is not a simple superposition of sensory channels, but a dynamic process from bottom-up sensory anchoring to top-down context adaptation. Furthermore, the priority of multi-sensory needs was clarified by defining Must-be attributes that ensure the baseline, One-dimensional attributes that improve efficiency, and Attractive attributes that drive the leap in telepresence, thus providing a decision-making basis

for design iteration and resource allocation. On this basis, the paper refined three strategies: Perceptual Reconstruction, Cognitive Fusion, and Context Adaptation, providing specific paths for physical perception restoration, cognitive logic matching, and dynamic environmental response. In summary, this study extends embodied cognition's application in industrial design and provides a systematic solution for next-gen system upgrades. Future experiments will verify these strategies' long-term effectiveness in reducing fatigue and improving precision, advancing teleoperation toward embodied human-machine integration.

ACKNOWLEDGMENT

Project: Key Technologies of Human–Machine Integrated Control for Construction Machinery (2024JK2025).

REFERENCES

- Berger, C., Blauth, R., Boger, D., et al. (1993). Kano's methods for understanding customer-defined quality. *Center for Quality Management Journal*, 2(4), pp. 3–36.
- Calvert, G.A., and Thesen, T. (2004). Multisensory integration: Methodological approaches and emerging principles in the human brain. *The Journal of Physiology*, 559(1), pp. 41–55.
- Clark, A., and Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), pp. 7–19.
- De Jaegher, H., Di Paolo, E., and Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences*, 14(10), pp. 441–447.
- Dong, Y., You, H., Shao, Y., et al. (2025). Neural circuit with top-down inhibitory feedback outperforms optimal bayesian integration in multisensory integration. *Neuroscience Bulletin*, pp. 1–16.
- Ernst, M.O., and Bühlhoff, H.H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, 8(4), pp. 162–169.
- Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.
- Jankowski, J., and Grabowski, A. (2015). Usability evaluation of VR interface for mobile robot teleoperation. *International Journal of Human-Computer Interaction*, 31(12), pp. 882–889.
- Kano, N., Seraku, N., Takahashi, F., and Tsuji, S. (1984). Attractive quality and must-be quality. *Journal of the Japanese Society for Quality Control*, 14(2), pp. 99–88.
- Kilteni, K., Groten, R., and Slater, M. (2012). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, 21(4), pp. 373–387.
- Kondo, D. (2021). Projection screen with wide-FOV and motion parallax display for teleoperation of construction machinery. *Journal of Robotics & Mechatronics*, 33(3).
- Liu, P., Han, F., Yang, X., et al. (2025). Design and clinical observation of a multisensory stimulation soft hand rehabilitation system. *IEEE/ASME Transactions on Mechatronics*.
- Silvennoinen, J., and Mononen, L. (2023). Multisensory MAYA–A design thinking method to enhance predictability of experience design. In: *International Conference on Human-Computer Interaction*, Cham: Springer Nature Switzerland, pp. 308–323.

- Slater, M., Pérez Marcos, D., Ehrsson, H., et al. (2008). Towards a digital body: The virtual arm illusion. *Frontiers in Human Neuroscience*, 2, p. 181.
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73(4), pp. 971–995.
- Spence, C., and Frings, C. (2019). Multisensory feature integration in (and out) of the focus of spatial attention. *Attention, Perception, & Psychophysics*, 82, pp. 363–376.
- Spence, C., Pavani, F., and Driver, J. (2005). On the inability to ignore touch when responding to vision in the crossmodal congruency task. *Acta Psychologica*, 118(1-2), pp. 47–70.
- Stein, B.E., and Meredith, M.A. (1993). *The Merging of the Senses*. Cambridge, MA: MIT Press.
- Truong, D.Q., Truong, B.N.M., Trung, N.T., et al. (2017). Force reflecting joystick control for applications to bilateral teleoperation in construction machinery. *International Journal of Precision Engineering and Manufacturing*, 18(3), pp. 301–315.
- Varela, F.J., Thompson, E., and Rosch, E. (2017). *The Embodied Mind, Revised Edition: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
- Şekerci, Y., Kahraman, M.U., and Develier, M. (2023). Enhancing interior design education with artificial intelligence: A multisensory hotel design. *Interiors*, 13(2–3), pp. 195–229.