

Review

The Development and Application of Emerging Interaction Technologies in Human-Computer Interaction

Siqi Li ^{1,*}¹ Sichuan International Studies University, No. 33 Zhuangzhi Road, Shapingba, Chongqing, China

* Correspondence: Siqi Li, Sichuan International Studies University, No. 33 Zhuangzhi Road, Shapingba, Chongqing, China

Abstract: With rapid technological advancements, the field of Human-Computer Interaction (HCI) has shifted from traditional interface mechanisms, such as keyboards, mice, and touchscreens, to more natural, multimodal, and intelligent interaction systems. Emerging interaction technologies, such as Brain-Computer Interfaces (BCIs), gesture tracking, voice interaction, eye-tracking, and haptic feedback systems, are transforming how humans interact with digital systems. While these technologies are being applied across diverse fields like neurorehabilitation, immersive education, smart manufacturing, and extended reality, challenges such as accuracy, robustness, and adaptability remain. Furthermore, issues related to data privacy and ethics are barriers to the widespread deployment of these technologies. This paper reviews the development of these technologies, explores their application trends within HCI, discusses their strengths and limitations, and addresses the challenges and future directions, with a particular focus on the convergence of AI, neuroscience, and multimodal interaction systems.

Keywords: Human-Computer Interaction; Brain-Computer Interface (BCI); gesture tracking; voice interaction; eye-tracking

1. Introduction

Over the past few decades, the field of Human-Computer Interaction (HCI) has experienced a rapid shift from conventional interface mechanisms, such as keyboards, mice, and touchscreens, toward more natural, multimodal, and intelligent interaction systems [1]. This transformation is not merely about improving usability; it reflects a fundamental change in how humans expect to communicate with digital systems in increasingly complex and dynamic environments. Today's users seek seamless, intuitive, and context-aware interactions that respond to their physical behavior, emotional states, and even cognitive intentions.

Emerging interaction technologies are at the heart of this transformation. Among the most notable are Brain-Computer Interfaces (BCIs), which interpret neural signals for direct brain-to-machine communication; gesture tracking systems, which use cameras or sensors to detect and classify body movements in real time; voice-based interaction, powered by advanced speech recognition and natural language processing; eye-tracking technologies that follow users' gaze to infer attention and intent; and haptic feedback systems that simulate tactile sensations to enhance realism in virtual experiences [2]. These technologies have moved beyond laboratory settings and are being actively developed and deployed in diverse domains, including neurorehabilitation, immersive education, smart manufacturing, and extended reality (XR) entertainment.

The appeal of these technologies lies in their potential to create more immersive, inclusive, and efficient interactions. For example, BCIs can restore communication abilities for individuals with motor impairments by enabling them to control external devices

Received: 18 April 2025

Revised: 21 April 2025

Accepted: 12 May 2025

Published: 16 May 2025



Copyright: © 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

through mental commands. Gesture tracking offers touchless control interfaces in hygiene-sensitive environments such as operating rooms [3]. Voice interfaces have become increasingly robust in handling real-time multilingual communication, although they still face challenges related to accents and contextual ambiguity, and gaze-based systems are improving assistive technologies for people with limited mobility.

However, these advances also bring significant challenges. Many systems still suffer from limitations in accuracy, robustness, and adaptability across different user populations. For instance, EEG-based BCIs are highly sensitive to noise and require extensive calibration, while gesture tracking systems may fail in poor lighting or with occlusions. Additionally, ethical concerns such as the collection of sensitive neural or behavioral data, as well as the lack of standardized protocols for evaluating interaction quality, remain unresolved. These issues hinder the widespread deployment and long-term reliability of such technologies.

This paper aims to provide a structured overview of the development and application of emerging interaction technologies within the context of HCI. It examines the technical principles behind each major modality, traces their historical evolution, and highlights current application trends. Furthermore, it compares their respective strengths and limitations and discusses foreseeable challenges, including usability, scalability, and privacy risks. Finally, it explores future directions, with a particular emphasis on the convergence of AI, neuroscience, and multimodal interaction systems. By offering a critical synthesis of current progress and open problems, this study contributes to a deeper understanding of how new forms of interaction are shaping the future of human-machine collaboration.

2. Literature Review

2.1. Brain-Computer Interfaces (BCIs)

Brain-Computer Interfaces (BCIs) are systems that enable direct communication between the human brain and external digital devices without the need for traditional muscular input [4]. By capturing and interpreting neural activity, BCIs allow users to control computers, prosthetics, or other interactive systems using only their thoughts. This technology is particularly significant for individuals with severe motor disabilities, and its potential also extends to areas such as gaming, cognitive training, and immersive virtual environments.

Most current BCI systems rely on non-invasive brain signal acquisition methods, particularly electroencephalography (EEG), which measures electrical activity on the scalp. Although EEG provides relatively low spatial resolution compared to invasive techniques like electrocorticography (ECoG) or implanted microelectrode arrays, it is widely used due to its safety, affordability, and portability. The neural signals are typically processed using machine learning algorithms to identify patterns associated with specific mental states or intentions, such as motor imagery (imagining hand movements), attention shifts, or emotional responses [5].

The development of BCIs has evolved from early clinical neuroscience proof-of-concept studies to real-world prototypes backed by both academic and commercial efforts. Companies like Neuralink and Kernel are investing in high-bandwidth invasive BCIs, aiming to create seamless neural interfaces for both medical applications and non-clinical cognitive enhancement. On the other hand, platforms like OpenBCI have made non-invasive systems accessible for research and hobbyist experimentation.

Applications of BCIs are growing rapidly. In healthcare, BCIs are used for neurorehabilitation after stroke, where patients use brain activity to interact with virtual limbs, stimulating cortical plasticity [6]. In assistive technology, BCIs empower patients with conditions such as amyotrophic lateral sclerosis (ALS) to communicate via brain-controlled typing interfaces. Beyond medicine, BCI is gaining interest in adaptive gaming, cognitive

state monitoring (e.g., detecting fatigue or stress), and XR environments where users can navigate or manipulate digital objects using mental commands.

2.2. Gesture Recognition and Tracking

Gesture recognition enables users to interact with digital systems through body movements, most commonly using hand and arm gestures. These systems interpret motion data to trigger commands, manipulate virtual objects, or communicate intent without the need for physical controllers. This type of interaction is particularly valuable in immersive environments such as augmented and virtual reality (AR/VR), and in situations requiring contactless control [7].

Two main categories of gesture tracking systems exist: vision-based and sensor-based. Vision-based systems use cameras, often combined with deep learning models, to detect and classify gestures in real-time [8]. Google's Media Pipe Hands and Meta's hand-tracking in Quest headsets are prominent examples, capable of recognizing complex finger movements with impressive accuracy. Sensor-based approaches rely on devices like inertial measurement units (IMUs), gloves, or wristbands to capture motion data via accelerometers and gyroscopes. These systems generally offer higher precision and are less susceptible to lighting or occlusion issues compared to vision-based systems, although they require wearable hardware and may be influenced by sensor placement accuracy.

Gesture recognition has seen growing adoption in fields such as healthcare (e.g., physical therapy tracking), industrial safety (e.g., hands-free command systems in hazardous environments), and education (e.g., interactive virtual lab environments). However, challenges remain in terms of robustness, scalability, and cross-user generalization. Hand occlusions, motion blur, and ambiguous gestures can reduce system accuracy. Moreover, achieving low-latency performance while running real-time inference on edge devices is technically demanding.

2.3. Voice Interaction and Speech Recognition

Voice-based interaction has become a widely adopted and user-friendly form of communication between humans and machines [9]. With advances in Automatic Speech Recognition (ASR) and Natural Language Processing (NLP), modern systems can not only transcribe spoken words with high accuracy but also understand user intent, respond contextually, and even adapt to emotional tone.

Contemporary voice interfaces leverage large-scale deep learning models trained on multilingual and multi-accent datasets. Tools like OpenAI's Whisper and Google's Speech-to-Text API have significantly improved speech recognition accuracy even in noisy environments. When combined with language models such as ChatGPT, voice interfaces can support dynamic, open-ended conversations that go beyond pre-programmed commands.

Use cases for voice interaction extend across smart homes (e.g., Alexa, Google Assistant), automotive control, healthcare dictation, and accessibility tools for visually impaired users [10]. In high-stakes environments like aviation or emergency response, voice commands can offer speed and efficiency that are often difficult to achieve through physical input.

Nevertheless, voice systems are not without limitations. Recognizing speech in noisy environments or when dealing with diverse accents remains a challenge. Concerns over voice data privacy, especially with cloud-based processing, have led to increased interest in on-device inference. Moreover, cultural and contextual nuances in language interpretation still pose barriers to global deployment.

As edge computing and multimodal AI continue to evolve, future voice interfaces will likely become more emotionally aware, context-sensitive, and capable of adapting in real-time to individual speaking styles and needs.

2.4. Eye Tracking and Gaze-Based Interaction

Eye tracking is a technique that monitors the position and movement of a person's eyes to determine where they are looking. It enables a form of implicit interaction, where systems can respond to visual attention without requiring explicit commands. This technology has gained traction in usability testing, assistive technologies, AR/VR systems, and psychological research.

Most eye tracking systems use infrared light and cameras to measure pupil dilation and gaze direction. Some rely on video-based image processing techniques, while newer systems integrated into AR/VR headsets offer real-time eye tracking with high accuracy and low latency. Tobii and Pupil Labs are among the leading developers of eye-tracking hardware and software platforms.

In practical terms, gaze-based interaction enables hands-free control for users with motor impairments, assists in user attention analysis during web or app usability studies, and allows for foveated rendering in virtual environments, in which the region of the display corresponding to the user's gaze is rendered in high resolution, thereby improving performance and power efficiency [11].

Challenges in eye tracking include calibration drift, user discomfort over long durations, and reduced accuracy due to eye shape variability or eyelash occlusion. Moreover, interpreting gaze data to infer intent requires sophisticated contextual modeling, especially in complex, dynamic environments.

Despite these challenges, eye tracking is regarded as a highly promising complementary input modality. As hardware becomes more miniaturized and affordable, it is likely to be integrated into everyday consumer devices to enable more responsive, personalized, and non-intrusive interactions.

2.5. Haptic Feedback and Tactile Interfaces

Haptic technology provides physical feedback to users through the sense of touch, often via vibrations, forces, or textures [12]. This modality enhances the realism of virtual environments and improves the user's sense of presence and engagement with digital content. In contrast to visual or auditory feedback, haptic interaction engages users in a more physically immersive way, simulating physical properties like pressure, texture, and resistance, simulating physical properties like pressure, texture, and resistance.

There are several types of haptic systems: vibrotactile actuators, which use small motors to create vibrations; force feedback devices, which resist user motion to simulate weight or stiffness; and more advanced approaches such as mid-air haptics, using ultrasonic waves to simulate touch in free space. Examples include wearable haptic gloves (e.g., HaptX), haptic pens for design software, and feedback-enhanced controllers in gaming consoles.

Haptics are particularly impactful in medical simulation and training, where they help replicate the tactile feel of surgical procedures. In remote robotics and telepresence, force feedback allows operators to "feel" distant environments. In consumer applications, smartphones and gaming controllers already use basic haptic cues to convey alerts and immersive feedback.

However, limitations persist in the form of high cost, bulky hardware, and lack of standardization. Replicating the full range of human tactile perception — including softness, temperature, and moisture — remains a technical challenge. Furthermore, user comfort, device fatigue, and power consumption issues must be carefully addressed in design.

3. Domain-Specific Implementation of Emerging Interaction Technologies

The efficacy of next-generation interaction paradigms depends on their alignment with domain-specific operational requirements and human perceptual and physiological constraints. Through four representative case studies, this section delineates how cutting-

edge systems are engineered to satisfy distinct functional imperatives, from neuromotor rehabilitation to precision manufacturing.

3.1. Healthcare: Neural Restoration and Assistive Systems

Contemporary neurorehabilitation systems leverage closed-loop biofeedback architectures to facilitate motor recovery. For instance, Brain-Computer Interface (BCI) systems that decode sensorimotor rhythms (SMRs), such as mu rhythms (8-12 Hz), enable stroke patients to engage in motor imagery tasks, promoting neural plasticity and functional improvement. Studies have demonstrated that such BCI interventions, combined with conventional therapy, can lead to significant improvements in upper limb motor functions [13].

In the realm of assistive communication, eye-tracking technologies, such as those developed by Tobii Dynavox, utilize pupil vector mapping to enable text input for individuals with conditions like amyotrophic lateral sclerosis (ALS). These systems allow users to select characters or words on a screen through eye movements, enabling text-based communication.

Robotic rehabilitation platforms integrate inertial measurement units (IMUs) and actuators to provide adjustable resistance during therapy sessions. These systems assist in reducing muscle spasticity and improving motor control in patients undergoing rehabilitation by providing controlled resistance during movement.

3.2. Education: Cognitive Augmentation through Multimodal Feedback

Immersive learning systems aim to optimize pedagogy through multisensory integration. For example, virtual dissection tables equipped with gesture recognition technology allow medical students to interact with 3D anatomical models, enhancing their understanding of complex structures [14]. Similarly, surgical simulators that incorporate haptic feedback provide trainees with realistic tactile sensations, improving their proficiency in procedures like laparoscopic surgery.

Experimental BCI-enabled attention monitoring systems using functional near-infrared spectroscopy (fNIRS) have been explored to assess cognitive load during learning activities. These systems aim to adapt educational content in real-time based on the learner's mental state, although further research is needed to validate their effectiveness in measuring cognitive load and improving learning outcomes.

3.3. Industrial Automation: Safety-Critical Human-Machine Teaming

Gesture control systems that comply with industrial safety standards, such as ISO 10218, enable contactless operation in hazardous environments. For instance, industrial applications utilize Some industrial applications use capacitive sensing technologies based on 3D electric field tomography to recognize specific gestures, allowing workers to control machinery without physical contact, thereby enhancing safety and efficiency.

Voice interfaces equipped with advanced speech recognition algorithms facilitate hands-free operation in environments with high noise levels. These systems allow workers to input data or control equipment using voice commands, reducing the need for manual interaction and minimizing physical strain. To further enhance situational awareness and task precision, industrial automation increasingly integrates Augmented Reality (AR) systems. Devices such as Microsoft HoloLens 2 incorporate depth sensing and simultaneous localization and mapping (SLAM) technologies to assist in quality inspections. By overlaying digital information onto physical components, these systems help identify defects and streamline maintenance processes, ultimately improving productivity and reducing human error.

3.4. Entertainment: Psychophysical Engagement Optimization

Next-generation entertainment systems enhance user immersion through synchronized visual, auditory, and haptic cues that align with the user's sensorimotor expectations. For example, VR headsets with high refresh rates and precise tracking capabilities provide users with immersive experiences. Haptic suits equipped with electrotactile nodes simulate various tactile sensations, enhancing the realism of virtual environments [15].

Affect-aware interfaces that monitor physiological signals, such as facial expressions or neural activity, enable dynamic adaptation of content based on the user's emotional state. This personalization enhances user engagement and creates more immersive experiences.

4. Advantages and Limitations of Emerging Interaction Technologies

To evaluate the practical viability of emerging interaction technologies, it is essential to analyze their strengths and limitations in relation to core performance criteria. These include signal accuracy, response latency, adaptability across users, hardware cost, and environmental robustness. While each modality introduces unique interaction capabilities, none constitutes a universally optimal solution. Instead, each technology aligns with specific application contexts and imposes distinct constraints on system design and user experience.

The following Table 1 summarizes the key advantages and limitations of five representative interaction modalities:

Table 1. Comparative Analysis of Interaction Modalities.

Technology	Key Advantages	Principal Limitations
Brain-Computer Interface	Enables direct neural input; valuable in motor-impaired contexts	Low signal quality; high training demands; low information throughput
Gesture Recognition	Natural and intuitive; supports contactless control	Sensitive to occlusion and lighting; gesture ambiguity; device wearability issues
Voice Interaction	Hands-free and efficient; widely adopted in consumer systems	Affected by noise and language variability; privacy and latency concerns
Eye Tracking	Low-effort, implicit input; useful for attention inference	Requires calibration; accuracy drift; limited expressiveness
Haptic Feedback	Enhances realism and immersion; supports motor skill learning	High hardware complexity; limited tactile resolution; cost and portability constraints

Brain-Computer Interfaces offer a fundamentally novel input channel by directly interpreting neural signals to infer user intent. This modality is especially valuable in accessibility applications, where conventional input methods are not viable. By decoding neural oscillations such as sensorimotor rhythms, BCIs enable communication and control for individuals with severe motor impairments. However, the signal quality in non-invasive systems is inherently limited by low amplitude, high noise levels, and susceptibility to artifacts caused by muscle movement or electrical interference. The accuracy of classification models often depends on individualized calibration and extended training periods, which may hinder scalability. Furthermore, BCIs tend to exhibit limited information throughput, constraining their usability in complex interaction tasks.

Gesture recognition systems provide an intuitive and natural interaction modality, particularly suited to contexts where contactless operation is essential, such as sterile environments or industrial settings. Vision-based gesture tracking, powered by computer

vision and deep learning algorithms, offers the flexibility of uninstrumented interaction but is highly sensitive to external factors including lighting variation, background complexity, and hand occlusion. Sensor-based alternatives, which rely on devices such as inertial measurement units or wearable gloves, can improve robustness but introduce issues related to user comfort and device maintenance. Additionally, gesture ambiguity and user variability pose ongoing challenges for generalized system performance.

Voice-based interfaces have become increasingly prevalent due to their hands-free, low-effort nature and widespread familiarity among users. These systems are effective in multitasking environments, mobile applications, and industrial workflows. Recent advances in speech recognition algorithms, particularly those incorporating large-scale transformer models, have improved recognition accuracy even under suboptimal conditions. Nonetheless, performance still degrades significantly in high-noise environments or when handling accented, dialectal, or disfluent speech. Concerns related to data privacy and real-time responsiveness further complicate deployment, especially in sensitive or latency-critical applications.

Eye-tracking technologies enable implicit interaction by capturing gaze direction and fixation patterns, offering insights into user attention and intention. This modality has proven effective in applications such as user interface adaptation, accessibility support, and foveated rendering in virtual environments. Since gaze behavior often reflects cognitive focus, it serves as a valuable signal for adaptive systems. However, eye trackers require user-specific calibration and may suffer from accuracy drift due to lighting changes or user fatigue. Moreover, interpreting gaze as meaningful system commands remains a challenging task, potentially increasing cognitive load for users if not designed with care. Haptic feedback systems contribute to interaction by introducing physical sensations, such as vibration, pressure, or simulated texture. This modality is particularly vital in applications requiring tactile realism, such as surgical training, product prototyping, and immersive entertainment. By engaging the somatosensory system, haptics enhance users' sense of presence and embodiment in virtual environments. Despite these advantages, technical limitations remain considerable. Replicating fine-grained tactile properties with high fidelity necessitates complex actuators and advanced control systems, leading to increased costs and energy consumption. Furthermore, current haptic devices are often bulky, which limits their portability and suitability for long-duration use.

Haptic feedback systems contribute to interaction by introducing physical sensations such as vibration, pressure, or simulated texture. This modality is critical in applications requiring tactile realism, such as surgical training, product prototyping, and immersive entertainment. By engaging the somatosensory system, haptics enhances the sense of presence and embodiment in virtual environments. Despite these benefits, technical limitations remain significant. Replicating fine-grained tactile properties with fidelity requires complex actuators and high-resolution control systems, which increase cost and energy consumption. In addition, current haptic devices often remain bulky, limiting their portability and long-duration use.

A comparative analysis of these technologies demonstrates that their strengths and limitations are closely linked to specific use cases and operating conditions. Brain-Computer Interfaces (BCIs) are essential for high-need accessibility applications but are limited by signal quality and processing speed. Gesture recognition excels in enabling hands-free control but requires careful environmental calibration. Voice interfaces are both efficient and scalable, though they are susceptible to acoustic and linguistic variability. Eye-tracking is highly valuable for attention-aware systems but lacks the direct input expressiveness of other modalities. Haptic feedback contributes to realism and interaction depth, yet its development is hindered by hardware complexity and standardization challenges. In conclusion, no single modality can fully meet the demands of modern human-computer interaction. The most promising approach involves multimodal interaction systems that integrate complementary modalities, thereby leveraging their individual strengths and compensating for their respective limitations. When carefully designed, such systems

can provide adaptive, user-centered experiences across a diverse range of application domains.

In conclusion, no single modality offers a comprehensive solution to the demands of modern human-computer interaction. The most promising direction lies in multimodal interaction systems that integrate complementary modalities, thereby combining their respective strengths and compensating for individual weaknesses. Such systems, when carefully designed, can provide robust, adaptive, and user-centered experiences across a wide range of domains.

5. Technological Convergence and Maturity Assessment

The rapid evolution of emerging interaction technologies is occurring in tandem, with these modalities increasingly converging within unified human-computer interaction frameworks. This convergence has led to the development of hybrid systems that integrate neural, visual, auditory, and tactile inputs, facilitating the creation of more adaptive, context-aware, and user-centered applications. However, the degree of technological maturity varies significantly across modalities, which in turn impacts their scalability, standardization, and readiness for real-world deployment.

Currently, the most mature technologies include voice interaction systems and basic eye-tracking mechanisms. These modalities have already been widely integrated into commercial and industrial platforms due to their low hardware requirements, reliability, and well-established developer ecosystems. For instance, voice-based assistants are embedded in smartphones, vehicles, and smart home devices, with robust natural language processing capabilities. Similarly, eye-tracking technology is increasingly used in assistive tools and XR headsets for attention monitoring and visual focus alignment.

In current practice, the most mature technologies include voice interaction systems and basic eye-tracking mechanisms. These modalities have already achieved widespread integration in commercial and industrial platforms due to their low hardware requirements, relatively high reliability, and established developer ecosystems. Voice-based assistants, for example, are embedded in smartphones, vehicles, and smart home devices, supported by robust natural language processing capabilities. Similarly, eye-tracking is increasingly incorporated into assistive tools and XR headsets for attention monitoring and visual focus alignment.

Technologies at the growth stage include surface electromyography (sEMG)-based gesture recognition, affective computing, and haptic feedback systems. These approaches are gaining traction in both academic and commercial sectors. Their growing adoption is driven by advancements in machine learning algorithms, miniaturized sensors, and cross-platform API integration. Nevertheless, these technologies face scalability challenges due to variability in physiological signals, limited interoperability, and the absence of standardized benchmarks.

At the emerging stage, modalities such as bidirectional BCIs and holographic haptic displays remain primarily experimental. These systems promise revolutionary interaction paradigms by enabling closed-loop feedback between users and machines or simulating mid-air tactile sensations. However, they are constrained by high technical complexity, limited clinical or industrial validation, and unresolved regulatory issues concerning data privacy and user safety.

The following Table 2 provides a comparative assessment of current interaction modalities across three dimensions: maturity stage, deployment readiness, and convergence potential.

Table 2. Maturity and Convergence Potential of Interaction Modalities.

Technology	Maturity Stage	Deployment Readiness	Convergence Potential
Voice Interaction	Mature	High (consumer and industrial)	High (e.g., voice + gesture)
Eye Tracking (basic)	Mature	Medium to High	Medium
Gesture Recognition (vision/sEMG)	Growth	Medium	High (e.g., XR systems)
Haptic Feedback	Growth	Low to Medium	High (with VR/AR/robotics)
Affective Computing	Growth	Low	High (e.g., emotion-aware UI)
Brain-Computer Interfaces (BCIs)	Emerging	Low	Very High (AI + neurotech)
Bidirectional BCIs	Emerging	Experimental	Very High
Holographic Haptics	Emerging	Experimental	Medium to High

The convergence potential of these modalities lies in their ability to complement each other across various aspects of human perception and control. For instance, combining BCIs with eye-tracking can enhance intent detection in assistive systems, while integrating gesture recognition with voice input improves command accuracy in smart environments. In immersive experiences, such as extended reality, synchronizing haptic feedback with real-time gaze tracking and spatial audio can significantly enhance user presence and perceptual realism.

6. Conclusion

In conclusion, this study offers a comprehensive analysis of emerging interaction technologies within the field of human-computer interaction. By examining the principles, application domains, comparative advantages, and integration potential of five key modalities — brain-computer interfaces (BCIs), gesture recognition, voice interaction, eye tracking, and haptic feedback — this work provides valuable insights into the evolving capabilities and complexities of interactive systems.

Each modality has unique technical advantages and inherent limitations. BCIs, particularly non-invasive electroencephalography-based systems, provide critical neural input pathways for assistive technologies designed for users with severe motor impairments. However, challenges such as low signal-to-noise ratios, inter-subject variability, and limited communication bandwidth remain. Gesture recognition systems, especially vision-based ones, enable intuitive, device-free interaction but are highly sensitive to environmental factors like lighting, occlusion, and background complexity. Voice interaction has seen widespread commercial deployment due to its natural interface and low cognitive load, yet speech recognition accuracy is still affected by acoustic noise, accent variation, and privacy concerns. Eye-tracking technologies offer low-effort input for attention estimation and adaptive interfaces, but require precise calibration and are limited by user fatigue and expressive control. Haptic feedback systems enhance realism and procedural training in virtual environments, but their deployment is hindered by high costs, actuator complexity, and the lack of standardization.

Application analysis across four sectors — healthcare, education, industrial automation, and extended reality entertainment — demonstrates that the adoption of interaction technologies is context-dependent. In clinical rehabilitation and assistive communication, the focus is on functional restoration. In educational systems and skill acquisition, the priority shifts to cognitive enhancement and embodied learning. Industrial applications prioritize operational safety, environmental adaptability, and reduced manual workload. In

immersive digital experiences, the focus is on perceptual realism and emotional engagement. These variations highlight the importance of aligning interaction design with domain-specific requirements, human factors, and task constraints.

A key takeaway from this analysis is the increasing importance of multimodal convergence. Future systems are expected to integrate neural, visual, auditory, and tactile inputs to create robust, context-aware frameworks. This convergence will require addressing challenges such as signal synchronization, data stream fusion, and developing user-adaptive, regulatory-compliant architectures.

Emerging interaction technologies represent a paradigm shift toward more natural, inclusive, and intelligent human-machine interaction. To realize their full potential, ongoing interdisciplinary collaboration across neuroscience, machine learning, cognitive ergonomics, and systems engineering is crucial. Additionally, establishing international standards and ethical frameworks will be essential to ensure the equitable, secure, and human-centered development of these technologies.

References

1. J. A. Jacko, Ed., *Human Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*. 2012. ISBN: 9780805858709.
2. M. R. Ho, T. N. Smyth, M. Kam, and A. Dearden, "Human-computer interaction for development: The past, present, and future," *Inf. Technol. Int. Dev.*, vol. 5, no. 4, p. 1, 2009. ISBN: 15447537.
3. W. Xu, M. J. Dainoff, L. Ge, and Z. Gao, "Transitioning to human interaction with AI systems: New challenges and opportunities for HCI professionals to enable human-centered AI," *Int. J. Hum.-Comput. Interact.*, vol. 39, no. 3, pp. 494–518, 2023, doi: 10.1080/10447318.2022.2041900.
4. F. Gurcan, N. E. Cagiltay, and K. Cagiltay, "Mapping human-computer interaction research themes and trends from its existence to today: A topic modeling-based review of past 60 years," *Int. J. Hum.-Comput. Interact.*, vol. 37, no. 3, pp. 267–280, 2021, doi: 10.1080/10447318.2020.1819668.
5. D. Wang, E. Churchill, P. Maes, X. Fan, B. Shneiderman, Y. Shi, and Q. Wang, "From human-human collaboration to Human-AI collaboration: Designing AI systems that can work together with people," in *Extended Abstracts CHI Conf. Hum. Factors Comput. Syst.*, Apr. 2020, pp. 1–6, doi: 10.1145/3334480.3381069.
6. J. Preece, Y. Rogers, H. Sharp, D. Benyon, S. Holland, and T. Carey, *Human-Computer Interaction*. Wokingham, U.K.: Addison-Wesley Longman Ltd., 1994. ISBN: 0201627698.
7. Z. Lv, F. Poiesi, Q. Dong, J. Lloret, and H. Song, "Deep learning for intelligent human-computer interaction," *Appl. Sci.*, vol. 12, no. 22, p. 11457, 2022, doi: 10.3390/app122211457.
8. I. S. MacKenzie, *Human-Computer Interaction: An Empirical Research Perspective*. 2024. ISBN: 9780128201943.
9. M. Muthmainnah, P. M. Ibna Seraj, and I. Oteir, "Playing with AI to investigate human-computer interaction technology and improving critical thinking skills to pursue 21st century age," *Educ. Res. Int.*, vol. 2022, no. 1, p. 6468995, 2022, doi: 10.1155/2022/6468995.
10. J. Trost, *The Intersection of Smart Home Technology and the Disabled Population*, Master's thesis, North Carolina Agricultural and Technical State Univ., 2024, doi: 10.18260/1-2--48119.
11. X. L. Chen and W. J. Hou, "Gaze-based interaction intention recognition in virtual reality," *Electronics*, vol. 11, no. 10, p. 1647, 2022, doi: 10.3390/electronics11101647.
12. N. V. Valtakari, I. T. Hooge, C. Viktorsson, P. Nyström, T. Falck-Ytter, and R. S. Hessels, "Eye tracking in human interaction: Possibilities and limitations," *Behav. Res. Methods*, pp. 1–17, 2021, doi: 10.3758/s13428-020-01517-x.
13. Z. Bai, K. N. Fong, J. J. Zhang, J. Chan, and K. H. Ting, "Immediate and long-term effects of BCI-based rehabilitation of the upper extremity after stroke: A systematic review and meta-analysis," *J. Neuroeng. Rehabil.*, vol. 17, pp. 1–20, 2020, doi: 10.1186/s12984-020-00686-2.
14. A. Moinuddin, A. Goel, and Y. Sethi, "The role of augmented feedback on motor learning: A systematic review," *Cureus*, vol. 13, no. 11, 2021, doi: 10.7759/cureus.19695.
15. V. Angelov, E. Petkov, G. Shipkovenski, and T. Kalushkov, "Modern virtual reality headsets," in *Proc. Int. Congr. Hum.-Comput. Interact., Optim. Robot. Appl. (HORA)*, 2020, pp. 1–5, doi: 10.1109/HORA49412.2020.9152604.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of GBP and/or the editor(s). GBP and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.