Research and Construction of Media Image of Virtual Digital Humans in the Intelligent Media Era: Post-print

Authors: Wang Zhenxiong, Yalong Liu

Date: 2023-10-08T00:00:00+00:00

Abstract

[Purpose] With the deep development of intelligent media convergence and digital technology, virtual digital humans as "digital twins" of real persons are continuously evolving, and the primary objective of this article is to investigate and construct their media image. [Method] This study takes virtual AI, virtual employees, and virtual idols as research subjects, integrates relevant theories from scholars such as Benjamin and Baudrillard, and summarizes the characteristics of virtual digital humans' media image in the intelligent media era to innovatively explore construction pathways for future virtual digital human media images. [Result] As a re-emergent "digital aura" and a consumable "cultural symbol", the future media image of virtual digital humans should be constructed across dimensions including diversified technologies, diversified styles, and diversified scenarios to facilitate "localized" communication for virtual digital humans. [Conclusion] The research and construction of virtual digital humans' media image can, on one hand, provide referential virtual image templates for humans to navigate the metaverse in the future; on the other hand, it can promote human-machine coupling and advance human-machine communication through the power of media image construction.

Full Text

Preamble

Research on and Construction of Media Images of Virtual Digital Humans in the Intelligent Media Era

Wang Zhenxiong[1], Liu Yalong[2]
(1. Meishi Film Academy, Chongqing University, Chongqing 400044; 2. School of Arts, Renmin University of China, Beijing 100872)

Abstract

[**Purpose**] With the deep development of intelligent media convergence and digital technology, virtual digital humans—serving as "digital twins" of real people—continue to evolve and iterate. The primary objective of this article is to investigate and construct their media image. [**Method**] Focusing on virtual artificial intelligence, virtual employees, and virtual idols as research subjects, and drawing upon theories from scholars such as Walter Benjamin and Jean Baudrillard, this study synthesizes and summarizes the media image characteristics of virtual digital humans in the intelligent media era to explore innovative pathways for constructing future virtual digital human media images. [**Results**] As a reemerging "digital aura" and a "cultural symbol" to be consumed, future virtual digital human media images should be constructed across multiple dimensions—diverse technologies, diverse styles, and diverse scenarios—to facilitate "localized" communication. [**Conclusion**] Research on and construction of virtual digital human media images can, on one hand, provide referenceable virtual image models for humans to navigate the metaverse in the future; on the other hand, it can promote human-machine coupling and advance human-machine communication through the power of media image construction.

Keywords: Intelligent Media Era; Virtual Digital Human; Media Image; Digital Aura; Image Construction

---

As a tool capable of simulating, extending, and expanding human thinking, consciousness processes, and intelligent behavior, artificial intelligence is becoming widely applied across various industries. According to data from the *2022 Virtual Human Industry Research Report* [1], the overall market size of domestic virtual digital human applications will reach 270 billion yuan by 2030. As a new attempt at applying AI in the cultural and creative industries, virtual digital humans can be traced back to the "Boeing Man" developed by Boeing in the 1960s to assist aircraft cockpits. Since then, virtual digital humans have gradually emerged in gaming, medical, entertainment, and other fields, such as the virtual singer "Lynn Minmei" from *Super Dimension Fortress Macross* in the early 1980s, the first virtual TV anchor "Ananova" in the early 21st century, and the recent virtual student "Hua Zhibing." Currently, virtual digital human applications in the public eye mainly fall into three categories: intelligent-driven virtual AI figures, service-oriented virtual digital humans, and technology-driven or human-driven virtual idols. However, from the perspective of virtual digital human typology, these three categories each possess distinct media image characteristics and design focal points. Analyzing the media images of each type can effectively reduce the "cyborg traces" of virtual digital humans and achieve technological breakthroughs in the digital era.

1. Virtual Artificial Intelligence: The Re-emergence of Digital "Aura"

Marshall McLuhan once made a prescient prediction: "After a century of development, the human central nervous system has been extended once again; the final stage of human extension is rapidly approaching—the stage of technically simulating consciousness" [2]. Artificial intelligence makes this technical simulation of consciousness possible. This type of virtual digital human refers to virtual AI that not only operates against an algorithm-driven background but also possesses an AI core, enabling it to continuously update its knowledge graph within big data-driven models. Taking China' s original virtual student "Hua Zhibing" as an example, as an important member of China' s largest pretraining model "Wu Dao 2.0," Hua Zhibing can not only create poetry, music, and paintings but also possesses certain reasoning and interactive capabilities that continue to be updated. Similar capabilities are found in "Xia Yubing," driven by the Avatar Framework, and "Amber Xu Yan," an AI virtual life born from the "GAVE" engine. Unlike mechanically reproduced virtual employees and virtual idols, the autonomous learning and creative capabilities of these virtual digital humans allow audiences to perceive their identity as "human" and provide new directions for their development.

Although digital existence is not a physical entity, it possesses many essential characteristics of physical things [3], such as substantiality, extensibility, and holistic integrity (through Martin Heidegger' s theory). Although digital beings lack material foundation, they can provide people with perception or universal passive perception (through Edmund Husserl' s theory) and the "return gaze" of the subject (through Walter Benjamin' s theory). Currently, a veil of "aura" exists between digital humans and real humans, revealing the "unpredictability" and "mysterious omen"in human-machine relationships. This aligns with Walter Benjamin's interpretation of aura in *Paris, Capital of the 19th Century*: "The experience of aura fundamentally transfers the echo found in human relationships to the relationship between inanimate (or natural) objects and humans. Between people and between people and objects, there exists a relationship of mutual gaze." Before the emergence of virtual AI, the industrial production process of virtual digital humans involved 3D modeling and replication of real humans, or redesigning and encoding based on human models, followed by algorithmic or real-time human driving—forming an industrialized production process and becoming a tool for enterprise digital transformation. Consequently, in this continuous "mechanical reproduction," the most important digital "aura" of virtual digital humans as artistic works was broken. However, as a mysterious omen of "fascination," the "digital aura" that brings about distance between humans and machines and generates in viewers key needs such as worshipful intention, object return gaze, imagination, and perceptual capacity should not be broken. Otherwise, future virtual digital humans would merely become a "toy" in Paul Levinson' s terms, still focused on the technology itself without inner substance, and the uniform cyborg images produced on assembly lines would also create

aesthetic fatigue for audiences.

The evolution and human-machine coupling of virtual AI like "Hua Zhibing" enable the re-emergence of mutual gaze experiences between humans and mechanical replicas (digital humans) during interaction. Simultaneously, the return of subjective consciousness and "object return gaze" produces a "shock" effect, allowing the media image of virtual digital humans to evolve through the re-emergence of "aura" during its disappearance, thereby rekindling societal attention and interest in virtual digital humans.

2. Service-Oriented Virtual Digital Humans: Visual Symbols of Brand Image

Service-oriented virtual digital humans aim to help humans complete repetitive and time-consuming work, achieving "content production" by generating knowledge graphs through big data and algorithms. In today' s society, the more common categories are virtual hosts and virtual customer service. Some enterprises and organizations have introduced virtual digital employees to promote digital transformation, but most virtual employees cannot fully replace human positions; their value remains largely in image endorsement and brand promotion. From the global first fully-simulated intelligent AI anchor "Xin Xiaohao" released at the Fifth Internet Conference to the first 3D synthetic anchor "Xin Xiaowei"; from the world's first digital aerospace reporter "Xiao Zheng"to Hunan Satellite TV' s intern digital host "Xiao Yang" ("young"), although people can perceive the technological iteration and progress of virtual hosts—for example, upgraded visual effects, diversified scenario applications, and enriched identity image shaping—their fixed-style cyborg traces and lack of genuine emotion have not yet gained recognition from most audiences. When deployed at scale, digital humans will fundamentally transform the business landscape. Within the next decade, most company managers will likely have a digital assistant or digital employee [4]. Examples include SPD Bank' s virtual employee "Xiao Pu," Baixin Bank's virtual employee AIYA, and Vanke's "Best Newcomer""Cui Xiaopan."Unfortunately, because virtual employees (customer service) cannot comprehend audience emotions and attitudes, most can only complete basic homogeneous tasks. Rather than these basic tasks, virtual employees assume more important responsibilities as corporate image endorsements and promotional vehicles.

Mark Sagar first proposed that "digital humans can never know a user' s emotional state." Foreign scholars Seymour, Mike, Lovallo, Dan, and others have focused on emotional interaction issues between digital humans and humans, noting that "future research on digital humans should focus on emotional engagement, considering whether a human-like face can better solve emotional problems in interaction, such as providing comfort or sympathy"[5]. This demonstrates that both virtual hosts and virtual employees have certain limitations in data technology and lack emotional awareness. Based on this, service-oriented virtual digital humans have gradually become an attempt and breakthrough for major enterprises in digital transformation. Although they cannot yet com-

plete all the work of a formal employee, as a visual cultural symbol, their signified "professional" image value can demonstrate corporate culture from the side, personify and extend the brand, strengthen promotional effects, and complete digital empowerment and brand communication for corporate culture.

3. Virtual Idols: Consumed Cultural Symbols

Virtual idols typically refer to virtual digital humans engaged in idol activities, possessing their own virtual settings and creative output. Professor Yu Guoming summarized them as "a new type of communication medium with built-in relationships, an extension of human strong ties" [6]. Fundamentally, in consumerist society, virtual idols, as body landscapes encoded by media, have become cultural symbols to be consumed. The "gaze" and "transgressive interaction" arising from audience desires themselves have triggered the alienation of virtual idol media images. In this article, the author categorizes virtual idols into two major types: technology-driven and human-driven (see Figure 1 [Figure 1: see original paper]), and conducts case studies on some recently popular virtual idols to analyze their media images.

3.1 Technology-Driven Virtual Idols—The "Other"Image of the Gazed-Upon

Technology-driven virtual digital humans are virtual images generated by creators through design conceptualization, 3D modeling, motion capture, and voice synthesis technologies. Based on the visual presentation effects of digital human appearances, they can be divided into two categories: hyper-realistic and two-dimensional anime-style. Two-dimensional anime-style virtual idols include well-known virtual singers such as "Hatsune Miku" and "Luo Tianyi." Data from the *2022 China Virtual Idol Industry Research Report* shows that over 51% of people like virtual idols because they like two-dimensional anime. Two-dimensional anime enthusiasts have higher identification with virtual worlds and greater acceptance and willingness to watch virtual idols [7]. These virtual idols often possess "neotenic maturity" comic-like image characteristics from two-dimensional anime and produce works through synthesized sweet voices. Hyper-realistic virtual idols are represented by the recently popular "Liu Yexi" launched by Chuangyi Technology. Debuting on Douyin as a "virtual beauty guru who catches demons," the character garnered over 2 million followers and likes on its first video within just three days. Wearing red traditional costume with a topknot and featuring an exquisite appearance with strong Eastern characteristics, "Liu Yexi" has shaped a "heroine" image who upholds justice and vanquishes demons in unit drama plots filled with traditional Chinese cultural elements.

Jean Baudrillard pointed out that "consumption is the consumption of the body; the body is the body of consumption; the body becomes a consuming body, a consuming symbol" [8]. In the intelligent media era, virtual digital humans are encoded by media into body landscapes for people' s appreciation

and consumption, pleasing consumers through voice, performance, and exquisite appearance, and becoming objects of societal scopophilia. Viewers suspend reflection on their own existence and, in the mutual gaze with the digital human' s virtual eyes, obtain a tangible embodied experience [9]. Precisely because increasingly realistic virtual images and maturing interactive technologies bring audiences this "real and tangible"embodied experience, technology-driven virtual idols like "Luo Tianyi"and "Liu Yexi"have become commodities to be consumed. Whether in exquisite cyberpunk style or "neotenic maturity" comic style, the media images of virtual idols have gradually become "otherized."While audiences seemingly do not directly participate in the encoding process, they profoundly influence capital' s image construction of virtual idols. The development of their media images will inevitably be influenced by audiences and tend to satisfy audience aesthetics and tastes.

3.2 The "Naka no Hito" –Interactive Rituals of "Digital Presence"

Currently, the virtual idol market is dominated by human-driven virtual anchors. These virtual idols are typically portrayed by "naka no hito" (the person behind the character). First, the virtual image is manually designed, drawn, and modeled; then motion capture technology captures the "naka no hito' s" body posture, expressions, eye contact, and other details; finally, the "naka no hito" performs and interacts behind the scenes, conducting live streaming activities such as music, dance, gaming, and chatting. Most of these images also feature two-dimensional anime "neotenic maturity" characteristics, but unlike technology-driven virtual idols, the existence of the "naka no hito" narrows the distance between fans and virtual idols, with interactivity and flexibility unmatched by technology-driven virtual idols. Taking the world' s first virtual YouTuber "Kizuna AI"as an example, it gained over one million followers within about a year of its launch. However, compared to Kizuna AI' s cute image, the lively and innocent personality displayed by the "naka no hito" behind the character won greater affection from fans. This contradicted the creative team' s original intention of making the virtual idol "Kizuna AI" an unshakable central subject, and the misrecognition of the center by fans ultimately resulted in Kizuna AI indefinitely ceasing activities and retiring after its final online concert on February 26, 2022.

Undeniably, the portrayal by "naka no hito" somewhat weakens the subjectivity of virtual idols. However, as digitally present "naka no hito" and virtually present fans interact in live streaming rooms, they generate an "interaction ritual chain" of "discussion—watching—rediscussion" [10]. Fans bring their personal beautiful fantasies into the live room and, through "discussion" with the anchor and other fans, obtain affirmation, thereby generating group solidarity and identity recognition, thus establishing deep emotional connections. Human-driven virtual idols become companion symbols that can harbor emotions, serving as spiritual sustenance for Generation Z two-dimensional anime fans. However, some virtual anchors, in order to deepen emotional connections with fans in the

live room, have engaged in endless alienation phenomena through vulgar and borderline content, resulting in "transgressive" interactions with audiences that violate norms. In response, Bilibili issued the *Special Governance Announcement for Virtual Anchors* on July 21, 2022, conducting special governance of inappropriate content and speech, strictly controlling borderline and revealing phenomena, curbing uncivilized behavior by virtual idols, rectifying "transgressive" actions, and maintaining the positive image of virtual idols as ritual myths in their interactions with fans.

4. Construction Pathways for Future Virtual Digital Human Media Images

4.1 Diverse Technologies to Cross the "Uncanny Valley"

Due to cost considerations, most current virtual digital humans employ "Deepfakes" technology in appearance design—broadly speaking, using deep learning technology for precise "face swapping" [11]. Currently, this technology cannot make digital human appearances as vivid as humans domestically, nor can it display non-verbal behaviors that easily generate intimacy, making it difficult to conduct deep-level dialogue and interaction with audiences and even creating a sense of horror. As early as the 1970s, Japanese scholar Masahiro Mori proposed the "Uncanny Valley effect," suggesting that as robot appearance and behavior become more human-like, positive emotions increase. However, after similarity reaches a certain peak, a sense of horror emerges; but when similarity becomes increasingly close to healthy humans, positive emotions re-emerge [12-13]. It is not difficult to see that some current hyper-realistic virtual digital humans have achieved the leap from two-dimensional to three-dimensional, reaching high anthropomorphism in appearance. However, due to immature technology, equipment, and post-production, their lack of image texture makes it difficult to evoke audience empathy. To cross the "Uncanny Valley" in the future, besides carving human facial expression details after "Deepfakes" face swapping, it is also necessary to ensure fluent language and natural movements for digital humans, enhancing their anthropomorphism through technological empowerment across multiple dimensions, thereby crossing the horror "critical point" and achieving healthy image development.

4.2 Diverse Styles to Build IP Brands

Images possess both objectivity and subjectivity. The former refers to any image being formed based on the inherent "form and appearance" of things themselves, while the latter means that images are the sum of people' s perceptions of things or persons [14]. What makes virtual digital humans different is that both the "inherent 'form and appearance' " and the impressions left on the sensory system are determined and controlled by coders at technology companies. Domestic technology companies are gradually finding their direction in style exploration. Due to the rise of "Guochao" (national trend) culture, current virtual digital

human styles are gradually leaning toward national style. Besides possessing exquisite features that conform to Eastern aesthetics in appearance, they also incorporate Chinese history and traditional culture in character background settings. In addition to the popular "Liu Yexi" on Douyin, there are also the "Mei Lanfang Digital Twin" developed by the Beijing Institute of Technology research team, the national-style digital human "Tian Yu" who performed "Qing Ping Yue" on Shaanxi Satellite TV' s Spring Festival Gala, and the national-style virtual KOL Ling Deng from CCTV' s program *Youthful China*. The popularity of "national-style virtual digital humans" has revealed the charm of Chinese-style cyberpunk distinct from Western styles. The vivid utilization of traditional cultural elements combined with new media platform traffic to build unique IP brand images provides a feasible path for constructing future commercial value for virtual digital humans. However, as national-style digital humans emerge endlessly, audience aesthetic fatigue gradually sets in. This demonstrates that domestic virtual digital human image style innovation relying solely on national style is insufficient. As a country with rich cultural resources, China should leverage the resource advantages of traditional Chinese culture, attempt diverse styles based on traditional culture, and shape the "cool medium" image described by McLuhan [15] to build digital human brands with national characteristics. This will continuously stimulate freshness and curiosity about digital humans, fully utilize the plasticity of virtual digital human images, release their innovative vitality, help virtual digital humans keep pace with the times, and achieve further expansion in commercial tracks.

4.3 Diverse Scenarios to Enrich Image Dimensions

The application and flow of virtual digital humans across diverse scenarios can help enrich their image dimensions, making characters more vivid and three-dimensional. Scenarios have become another core element of media, following content, form, and social interaction [16]. In the era of scenarios, with the integrated development of large and small screens, diverse scenarios can provide more display platforms for virtual digital humans, and technological empowerment enables virtual digital humans to achieve cross-scenario leaps.

Due to their perfect replicability, digital beings can exist simultaneously in multiple locations [17]. Taking "AYAYI," the nation's first Metahuman (hyper-realistic digital human) created by Ranmai Technology as an example, it pioneered multi-scenario applications for virtual digital humans, conducting commercial cooperation with famous brands such as Guerlain, Porsche, and Ambrosial, shooting promotional photos at Winter Olympics ski venues, and joining Alibaba as a "digital employee" to become the digital manager for Tmall Super Brand Day. With its hyper-realistic appearance, it shuttles through various scenarios like a real person and takes photos interacting with celebrities and ordinary people. Also shuttling through various scenarios is the digital host "Xiao Yang" ("young") from *Hello, Saturday*. In the program, it interacts with real hosts in real-time and participates in program recording. Outside the program, it also appears

in various exhibition halls and New Year' s Eve live broadcasts. Under the technical empowerment of Hunan Broadcasting System, "Xiao Yang" ("young") has already built its own IP brand. As a new communication medium with built-in relationships, virtual idols' content is more easily noticed and recognized [18]. The application and experimentation of virtual digital humans across diverse scenarios and different fields can, for brands and enterprises, leverage their "built-in relationship" traffic attributes to gain more attention for content, products, or activities. For the construction of virtual digital human media images, multi-scenario application and experimentation can not only reduce their fictitiousness and enhance social attributes but also effectively improve their own interest and "presence" sense, making their image as digital "humans" more full and rich.

As American scholar N. Katherine Hayles once said: "Machines and humans together form a binary system, making them measures of each other." On the one hand, virtual digital humans help humans discover themselves and measure the distance between humans and machines; on the other hand, as one of the gateways for contemporary humans to the metaverse, virtual digital humans serve as connectors to this "mimetic world." Along with technological transformation, the media images of virtual digital humans have found their evolutionary direction through continuous alienation and trial-and-error processes: more realistic appearances, more diverse style characteristics, and more diversified application scenarios. The construction of virtual digital human media images represents humanity' s identity imagination for freely navigating the metaverse in the future, providing referenceable virtual image settings for humans and promoting human-machine coupling and human-machine communication through the power of media image construction.

Notes:

Naka no hito (中之人), derived from Japanese, refers to the person hidden behind the virtual image who manipulates it to perform activities.

Generation Z, derived from the English term "Generation Z," typically refers to the youth group born between 1995 and 2010.

References:

[1][7] *2022 China Virtual Idol Industry Research Report* [R/OL]. https://page.om.qq.com/page/OxbFHgG8n8he 2022-09-30/2023-03-23.

[2][Canada] Marshall McLuhan. *Understanding Media: The Extensions of Man* [M]. Translated by He Daokuan. Nanjing: Yilin Press, 2019: 20.

[3] Yu X, Shi Y, Yu H, et al. Digital human modeling and its applications: Review and future prospects. *Journal of X-ray Science and Technology*, 2015(3): 385-400.

[4] Seymour M, Lovallo D, Riemer K, Dennis AR, Yuan L (Ivy). AI with a Human Face. *Harvard Business Review*. 2023(2): 49-54.

[5] Seymour M, Lovallo D, Riemer K, Dennis AR, Yuan L (Ivy). AI with a Human Face. *Harvard Business Review*. 2023(2): 49-54.

[6] Yu Guoming, Yang Mingyi. Virtual Idol: A New Communication Medium with Built-in Relationship Attributes [J]. *News and Writing*, 2020(10): 68-73.
[8][France] Jean Baudrillard. *The Consumer Society* [M]. Translated by Liu Chengfu, Quan Zhigang. Nanjing: Nanjing University Press, 2014: 123.
[9] Wang Ying, Liang Xueyuan. Virtual Subjectivity and Virtual Materiality: On Digital Human Images [J]. *Journal of Fujian Normal University (Philosophy and Social Sciences Edition)*, 2022(5): 49-58, 170.
[10][USA] Randall Collins. *Interaction Ritual Chains* [M]. Translated by Lin Juren, Wang Peng, Song Lijun, et al. Beijing: The Commercial Press, 2017.
[11] Ji Deqiang. Deepfakes: Visual Politics in the Age of Artificial Intelligence [J]. *Journalism University*, 2020(7): 1-16, 121.
[12] Mori, Masahiro. The Uncanny Valley. Translated by Karl MacDorman and Norri Kageki. *IEEE Robotics & Automation Magazine*, 2012(2): 98-100.
[13] Du Yanyong. Analysis of the Uncanny Valley Effect [J]. *Social Sciences in Yunnan*, 2020(3): 37-44, 187.
[14] Fan Hong. Multi-dimensional Shaping and Communication Strategies of National Image [J]. *Journal of Tsinghua University (Philosophy and Social Sciences Edition)*, 2013(2): 141-152, 161.
[15][Canada] Marshall McLuhan. *Understanding Media: The Extensions of Man* [M]. Translated by He Daokuan. Nanjing: Yilin Press, 2019: 33.
[16] Peng Lan. Scenario: A New Element of Media in the Mobile Era [J]. *Shanghai Journalism Review*, 2015(3): 20-27.
[17] Yu X, Shi Y, Yu H, et al. Digital human modeling and its applications: Review and future prospects. *Journal of X-ray Science and Technology*, 2015(3): 385-400.
[18] Yu Guoming, Yang Mingyi. Virtual Idol: A New Communication Medium with Built-in Relationship Attributes [J]. *News and Writing*, 2020(10): 68-73.

Author Information:
Wang Zhenxiong (2000-), male, from Changsha, Hunan, master' s student at Chongqing University, research direction: journalism and communication; Liu Yalong (2000-), male, from Changde, Hunan, master' s student at Renmin University of China, research direction: digital cultural industry.

(Responsible Editor: Zhang Xiaojing)

Note: Figure translations are in progress. See original paper for figures.