Artificial Intelligence in Broadcasting and Hosting: Current Application Status and Implications (Postprint)

Authors: Wang Shuya

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

Abstract

The rapid advancement of artificial intelligence technology and its extensive application in the media industry, alongside the upgrading of communication methods, have simultaneously spurred the production of diverse content. These new technologies have already found numerous applications in the field of broadcasting and hosting, even giving rise to many substitutive functions. Consequently, many practitioners have raised such questions and concerns: Will artificial intelligence completely replace broadcasting and hosting? How should broadcasting and hosting respond to this swift tide of artificial intelligence? This paper takes as its entry point the systematic review of the latest application status and future development prospects of artificial intelligence in the field of broadcasting and hosting, and contemplates what challenges and insights the development of artificial intelligence will bring to the future of the broadcasting and hosting industry.

Full Text

The Application of Artificial Intelligence in Broadcasting and Hosting

Abstract: The rapid advancement of artificial intelligence technology and its widespread application in the media industry have transformed communication methods and catalyzed the production of diverse content. These new technologies have already found numerous applications in the field of broadcasting and hosting, even giving rise to many alternative functions. Consequently, many practitioners have raised questions and concerns: Will AI completely replace human hosts? How should broadcasters and hosts respond to this powerful wave of artificial intelligence? This article examines the latest application status and fu-

ture development prospects of AI in broadcasting and hosting, reflecting on the challenges and insights that AI will bring to the industry' s future development.

Keywords: broadcasting and hosting; artificial intelligence; speech synthesis technology; immersive mixed reality; AI synthetic anchor

1. Overview of Artificial Intelligence and Speech Synthesis Technology

Artificial Intelligence (AI) refers to the research and development of theories, methods, technologies, and application systems that simulate, extend, and expand human intelligence. Since its inception, AI theory and technology have matured increasingly, with expanding application domains. Speech synthesis represents a key component of voice interaction, which has become a hotspot with the proliferation of intelligent assistants and smart customer service applications. The three core technologies in voice interaction are speech recognition, speech synthesis, and semantic understanding, with speech synthesis playing a crucial role.

While the public is relatively familiar with Automatic Speech Recognition (ASR), or "speech-to-text" technology, Text-To-Speech (TTS) technology converts written text into spoken words. Today, TTS is widely applied in smartphones and smart products—voices such as Apple' s Siri, Samsung' s Bixby, Xiaomi' s Xiaoai, and Vivo' s Jovi are all generated using TTS technology. Speech synthesis works by extracting appropriate phonetic elements from a speech corpus, modifying and integrating them according to human language patterns using TTS technology, and ultimately outputting speech that conforms to human auditory habits.

At the 2019 CCTV Online Spring Festival Gala, the AI virtual anchor "Xiaoxiaosa" made its debut, performing alongside the human host Sa Beining. Xiaoxiaosa broke the stereotype of robot hosts merely reciting pre-inputted fixed scripts by adapting spontaneously to control the program flow, interacting with Sa Beining and the audience with humor and wit. The technology provider revealed that creating such an AI bionic host requires only 30 minutes.

2. Current Applications of AI and Speech Synthesis in Broadcasting and Hosting

2.1 News Broadcasting: Maximizing Efficiency and Minimizing Costs

AI speech synthesis technology has been applied in news broadcasting for a longer period and across broader domains. To liberate people' s eyes during fragmented time, in 2016, CCTV News and Xinhua News Agency clients pioneered intelligent voice broadcasting, allowing users to listen to complete news reports

smoothly by simply tapping a small button on the interface. This widespread application in new media clients offers audiences greater convenience.

First, users demonstrate considerable tolerance toward AI-generated speech. When acquiring core news information during fragmented time, even if the vocal expression lacks the aesthetic quality of human anchors, few would spend additional time critiquing the delivery after obtaining the information. Second, AI satisfies users' demand for personalized content and voice types. Traditional audio news programs placed users in a passive receiving position, forcing them to choose between freeing their eyes for other tasks and listening to news of interest. For new media clients, low cost, high efficiency, unlimited working hours, and zero error rates represent clear advantages where AI can demonstrate its capabilities now and in the near future.

In video synthesis, at the 5th World Internet Conference in 2018, Sogou and Xinhua News Agency jointly launched the world' s first AI synthetic male anchor "Xin Xiaohao," followed by the first AI female anchor "Xin Xiaomeng" in February 2019. This technology employs cutting-edge AI, representing the latest achievement in advancing from media convergence to in-depth media development. The new AI synthetic anchors utilize Sogou' s "Avatar" technology to create more emotionally expressive voices, more realistic facial expressions and lip movements, and more appropriate paralinguistic elements, making AI anchors warmer and more aligned with human broadcasting habits. The technology also accommodates diverse presentation styles, including seated and standing broadcasts, greatly enriching AI anchors' hosting modalities and expanding application domains. In terms of work efficiency, AI anchors can provide 24-hour service, master multiple languages, and generate high-quality broadcast videos in a short time simply by inputting text, delivering information with posture, expressions, and voices nearly indistinguishable from human anchors. Immediately after launch, these AI synthetic anchors were deployed in news reporting, including coverage of the 5th World Internet Conference and the 2019 Spring Festival travel rush, leveraging AI' s superior information integration capabilities to combine images, videos, and text into highly readable converged media products.

2.2 Voice Restoration: Recreating Deceased Voices

In 2018, CCTV produced the world' s first documentary fully dubbed using AI and intelligent voice technology—*Innovative China*—which brought the voice of the late renowned voice actor Li Yi back to the screen, meticulously recreating his deep, magnetic vocal quality. This breakthrough involved using precise algorithms and quantitative methods to extract and reorganize sound materials from Li Yi' s previous documentary work, with related technical algorithms achieving breakthroughs in better matching the original voice' s pausing and phrasing patterns.

AI voice synthesis technology' s simulation and reshaping of human voices had

already gained widespread application before *Innovative China*, such as Gaode Map' s navigation using voice materials from stars like Lin Chi-ling and Yang Yang, weather broadcasts featuring Angela Baby' s voice, and audiobooks that liberate readers from visual reading. This series of AI dubbing applications fully demonstrates AI' s extensive use in voice-over work. In low-budget advertising voice-over markets, where rigid language and fixed intonation share characteristics with the repetitive patterns generated after establishing voiceprint models, AI speech synthesis can significantly reduce costs and production cycles. Through algorithmic adjustments, AI can perfectly meet advertisers' personalized needs and even create exclusive brand sound signatures by adjusting voiceprint models. Thus, AI may soon impact low-end advertising voice-over professionals.

2.3 Immersive Integration: Reshaping Scenarios for Perfect Presentation

Some AI derivative technologies applied in media communication have created realistic scenarios in the U.S. Weather Channel by combining AR and MR technologies, termed IMR (Immersive Mixed Reality). This approach integrates 360-degree HD video, real-time data-based augmented and virtual reality elements, and experienced live broadcasters to lead audiences through authentic weather conditions, enabling them to imagine how to protect themselves if actually present. Perfect visual effects combined with hosts' acting skills and explanations immerse audiences in severe weather scenarios, creating engaging broadcasts rather than rigid, boring presentations.

The core internet mindset involves transforming from traditional propaganda thinking to guiding, dialogic, and service-oriented thinking within an integrated environment. Indeed, in the AI era, we should prioritize this mindset transformation and upgrading. From the Weather Channel' s immersive mixed reality technology combined with host performance to Jiangsu Satellite TV and other major networks introducing "converged media news studios," and interactive presentations by Ouyang Xiadan in the "Belt and Road" documentary, we can observe a new trend in broadcasting and hosting: "script-reading machines" who simply sit and recite face potential replacement. Appropriately integrating performance elements according to program types and offering performance courses in broadcasting and hosting education are necessary measures to address current and future program format upgrades.

3. Emotional Connection and Human Warmth

Phoenix TV host Wu Xiaoli once said: "Robots may replace my body temperature, but they cannot replace the temperature of my words." Indeed, AI hosts are "intelligent" but not "capable." No matter how intelligent, they remain machines running code, requiring much more time to achieve human emotional depth. The irreplaceable aspect of human hosts lies in their warmth and rich humanistic care.

During the 2017 *First Lesson of the School Year*, host Dong Qing interviewed the renowned translator Xu Yuanchong, who had limited mobility and used a wheelchair. Dong Qing knelt down several times during the interview, even lowering herself further to show respect. This detail not only accommodated the elderly scholar but also demonstrated profound respect, earning unanimous praise from audiences. In the context of AI' s rapid development, "becoming a warmer host" is an essential focus for the broadcasting industry and academic training. Paralinguistic expressions such as smiles, nods, posture, eye contact, and physical touch are indispensable qualities for excellent hosts. How to convey inner humanistic care to create warmer audience experiences represents a key element for enhancing contemporary hosts' core competitiveness.

4. Integrating Performance Elements and Cultivating Acting Skills

Currently, AI anchors' expressions and movements depend entirely on initially recorded motion-capture materials from human hosts, which are then integrated through self-learning. This limitation means that single-source motion materials may suit programs with limited body language but remain highly restrictive. Therefore, in terms of integrating and cultivating performance elements for hosts, we find that performance techniques are now widely used not only in variety shows but also in seemingly serious program types like news broadcasting and weather forecasting. Zhu Guangquan, the "CCTV jokester," delivers news through carefully designed rhyming statements, generating enthusiastic audience responses that prompt reflection on "host performativity."

In live broadcast settings, hosts frequently encounter unexpected situations, and their primary function is to ensure seamless program continuity. When Hunan Satellite TV host Wang Han faced Sun Nan' s unexpected withdrawal during the *I Am a Singer* finals, he delivered an emotionally reasonable impromptu monologue after just a few seconds of preparation, reassuring the station, the singer, and the audience while maintaining high morale—setting a model for the hosting profession. For AI, even with extensive training samples and high self-learning capabilities, perfectly and compassionately handling such unexpected live incidents remains extremely difficult.

5. New Perspectives and In-Depth Content

For the same news topic, different angles produce different effects. Although AI will improve its self-learning abilities in the future, achieving the logical flexibility to adapt strategies with changing factors like humans remains challenging. In an era of extremely convenient communication tools, the importance of compelling angle selection becomes evident. This places higher demands on media professionals' abilities in angle selection, logical structuring, and converged media awareness.

In a People' s Daily interview, Bai Yansong observed that good journalists are

woodpeckers, not magpies—not simply making people happy daily, but stepping beyond the "small self" to focus on the era, anxiously identifying problems and hoping for change. Indeed, hosts can approach the same news from different angles to create varied broadcast effects or excavate facts behind news reports to enhance depth, multidimensionality, and impact, providing audiences with alternative analytical perspectives. This emphasis on "novelty" and "depth" requires hosts to continuously improve their abilities in content discovery, analysis, and excavation.

6. Developing Personal Style and Building Brand Identity

Current AI hosts remain in the stage of imitating human anchors. For social education, entertainment, and other program categories, AI cannot replace hosts' status in audiences' hearts or fully and realistically imitate and innovate their hosting styles and humanistic care. In daily life, when thinking of female variety show hosts, Xie Na may come to mind first; for cultural programs, Dong Qing; for news commentary, Bai Yansong; for documentary narration, Zhao Zhongxiang. This "first response" reflects a host' s unique style and characteristics developed over years.

Famous host Wang Han once summarized his hosting career in three sentences: "Learn to listen; don' t be infatuated with applause; be willing to be a green leaf that sets off the guests." This reveals that an excellent host' s personal charisma, linguistic warmth, and skill often represent a brand or even a city or nation' s image—soft power that advanced AI cannot replace. Meanwhile, hosts with uniform styles who cannot integrate reporting, editing, and broadcasting exhibit mechanical and inflexible characteristics that align with AI' s strengths. Therefore, to enhance core competitiveness, contemporary hosts should develop distinctive personal styles and build their own brands.

Conclusion

Against the backdrop of the AI era, the introduction of new AI technologies has brought both opportunities and challenges to broadcasting and hosting. Facing these challenges, practitioners and talent cultivation should strengthen their confidence, recognize their irreplaceability in certain domains, actively engage in artistic creation, and continue exploring their potential where AI falls short. By seeking unique advantages, they can better respond to challenges. Simultaneously, they must adapt to development trends with an open and inclusive attitude, correctly facing AI technology' s promotion and application within the industry, strengthening collaboration between hosts and AI, and forming complementary advantages. Only by leveraging AI' s strengths in data aggregation and converged media presentation can we achieve "using people to their full potential, using materials to their fullest, and using them at the right time."

References

[1] Xia Haibo. Research on Public Communication and Social Evaluation of China' s AI Technology [D]. Zhejiang: Zhejiang University of Technology, 2017(16).

[2] Wang Desheng. Global AI Development Trends [J]. Competitive Intelligence, 2017(4).

[3] Biaobei Technology & CSDN AI Expert Li Xiulin: The Burgeoning Speech Synthesis Technology and Applications [J/OL] https://blog.csdn.net/biaobei/article/details/83183181.

[4] Sohu News: When Weather Forecasts Use AR Effects, Both Hosts and Audiences Go Crazy [J/OL] http://www.sohu.com/a/291996336_{104421}.

[5] Zhang Lijun. The Irreplaceability of Hosts in the AI Era [J]. Science and Technology Communication, 2017(23): 145-146.

[6] Huang Haitao. Internet Thinking Profit Model [M]. Beijing: People' s Posts and Telecommunications Press, 2015.

(Author' s Affiliation: School of Journalism and Communication, Hunan Normal University)

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