Research on the Effect of Al Synthetic Anchor Video Communication from the Perspective of Body Communication

Mengdi Zhang*
Communication University of China,Nanjing,Nanjing 211172,China
***Corresponding author E-mail: estherrr@foxmail.com**

Abstract: With the application of artificial intelligence in the field of media, information dissemination has entered the era of intelligent media. In practical application, Al composite anchors show basically the same news broadcast effect as live anchors. This study studied the influence of physical factors on Al composite anchors' willingness to watch videos through questionnaire survey. It is found that, although Al composite anchors are constantly updated and iterated, the domestic audience's contact and recognition of this new thing is not high, and the physical factors of Al composite anchors, such as micro-expressions, movements, timbres, etc., are not as rich as those of live anchors, especially micro-expressions and timbres, which have a significant impact on the video communication effect. Based on this discussion, Al synthetic anchors can't just stay at the initial stage, and live anchors should complement each other's strengths and weaknesses, and share each other's glory in the upsurge of Al synthetic anchors.

Keywords: body communication; Al synthetic anchor; Sound and color; Willingness to watch

1 Introduction

With the development of Internet technology and the change of times, Al synthetic anchors are constantly appearing in innovative applications of the press with a new look. In recent years, Xinhua News Agency, CCTV and other mainstream media have successively launched auxiliary news reports of Al synthetic anchors. On November 7, 2018, the world's first Chinese Al synthetic news anchor-"Xin Xiaohao" came out. It is trained by combining modeling with techniques such as voice, lip-shape, expression synthesis and deep learning. It can automatically generate videos with corresponding contents according to the input text, and ensure that the audio in the videos is naturally consistent with the expressions and lip movements, showing a broadcast effect almost similar to that of live anchors. We have observed that Al composite anchors are no different from live anchors in broadcasting professional ability. However, compared with traditional live anchors, Al composite anchors have defects in expression communication, action communication and emotional expression. These factors also lead to differences in communication effects between Al composite anchors and live anchors. Under this background, this paper attempts to analyze the differences between Al anchors and live

anchors in micro-expressions, movements, sounds and other aspects, as well as the influence of these differences on the news video dissemination effect, and discuss the irreplaceability of live anchors after the upsurge of Al synthesis anchors recedes.

2 Literature Review

2.1 Review of research on body transmission

Throughout the history of communication studies, physical problems have been absent in communication studies for a long time. Classical school, classical theory, research on radio, television, newspapers and even today's Internet and intelligent terminals, focus on people only in the aspects of communication effect, social control, communication role, etc., and pay little

attention to the human body itself. In the article "Body Problems in Communication and the Future of Communication Research" written by Liu Hailong, the intellectual history of body problems in communication research is combed, and the body theories related to communication research in theoretical resources such as McLuhan, media archaeology, cybernetics and post-humanism are preliminarily discussed. It is believed that body provides a rare logical breakthrough point for reflecting the existing two communication concepts and taking care of the practice of new media technology, thus opening up a domestic research field of reflecting the classical communication theories from the perspective of body [1].

Studying the body from the perspective of semiotics, the expressions, movements and psychological activities of the body can be spread and expressed in the form of symbols, and the content carried by the media is the research object. However, the body research of Al anchors is more limited to the discussion of body presence, while the research on nonverbal symbols such as expressions and movements is less. Shao Peng and Yang Yu's research shows that the theoretical perspective of embodied communication makes journalism and communication focus on the spread and practice of the body. Taking Al virtual anchors and live presenters as the research objects, they think it is very important to coordinate the relationship between the presenters' physical presence and their absence in the era of artificial intelligence, expound the irreplaceability of live presenters, and propose that presenters should give full play to the role of the body as a "communication platform" and improve their comprehensive quality and ability to adapt to "[2].Zhao Xin analyzed the deficiencies of Al news anchors in eyes, expressions and movements from the creativity and diversity of news broadcasting, and then provided the direction of sustainable development for live anchors[3].

Through combing the literature, it is found that the theoretical research foundation of body itself is relatively weak in China. Although the extensive application of artificial intelligence makes people pay more attention to the body, there is still little discussion about body communication in the field of Al composite anchors, and the influence of body communication on video communication is absent. Therefore, it is of certain academic value to study the

influence of Al composite anchors on video news dissemination effect from the perspective of body communication.

2.2 Research on Al synthetic anchor

In 1956, John McCarthy formally put forward the concept of "artificial intelligence". After many times of stagnation and progress, artificial intelligence gradually appeared in the field of host communication in the evolution of technology. At present, the application of artificial intelligence in the field of broadcasting, hosting and communication in academic circles is still in its infancy. Currently, most of the researches and discussions on Al synthetic anchors are still focused on whether it can replace announcers and hosts.

Yin Lei makes a comparison from technical innovation, abstract reasoning, artistic aesthetics, emotional analysis and other dimensions. She thinks that in abstract reasoning, information programs can give full play to the strengths of artificial intelligence, but traditional anchors are better than analytical programs. In terms of artistic aesthetics, the current artificial intelligence anchors still can't feel ,explain and express beauty; In emotional analysis, artificial intelligence anchors often present dull and boring images in programs, lacking humanized communication, which gives viewers a sense of distance[4].

Wang Yiting and Du Peng believe that "artificial intelligence" needs to go through three spaces of language construction from its appearance to its large-scale application: survival-standardization-aesthetics, that is, the process from information sharing to pleasant resonance. From the technical level of speech emotion recognition, they find that the performance of artificial intelligence in the realm of broadcasting and hosting is still in the primary stage, and the related technologies are still immature. In the process of language flow, some issues still need to be improved, such as the synthesis of facial expressions, the expression of lips and the expression of attitudes and emotions in broadcasting[5].

From the above research summary, it can be seen that the general view of the development of Al synthetic anchor in academic circles is that it is still in the primary stage of development, and the related technologies are not mature enough. Although new Al synthetic anchors continue to appear, they have not yet reached the level of strong artificial intelligence engaged in human creative activities. In addition, at present, most of the researches on Al synthetic anchors focus on theoretical discussion and application analysis of artificial intelligence technology in the media industry, while the research on communication effect lacks objective evaluation and empirical research as a whole. Based on this, from the audience's perspective, this paper analyzes the video communication effect affected by the lack of functions such as micro-expressions, movements, sensuality, etc. of Al synthetic anchors by questionnaire survey and content analysis, so as to explore the irreplaceability of live anchors after the upsurge of Al synthetic anchors recedes.

3 Research methods

Based on the review of the existing literature, this paper puts forward the following three questions: Can Al composite anchors convey the emotions of news reports like live anchors?

What is the difference between Al synthetic anchor and live anchor in terms of micro-expression, movement and sound and color? How do these differences affect the communication effect of news videos?

This paper attempts to answer the above questions by questionnaire survey.

3.1 Questionnaire survey method

3.1.1 Hypothesis establishment

Based on the above questions, the corresponding assumptions are put forward here:

H1: The richer the micro-expressions of Al composite anchors, the stronger the audience's willingness to watch the video news of Al composite anchors.

H2: The richer the action of Al composite anchor, the stronger the audience's willingness to watch the video news of Al composite anchor.

H1: The richer the tone and color of the Al composite anchor, the stronger the audience's willingness to watch the video news of the Al composite anchor.

In order to prove the hypothesis, this study adopted the questionnaire survey method, sampled the whole population and made a questionnaire.

3.1.2 Sampling method

It is assumed that the corresponding research object is the audience who contact the Al synthesis anchor. This survey completed the process of data collection, collation and analysis on October 17th, 2022. The main contents of the survey are as follows: (1) Confirmation of basic information, including gender and age. (2) Differences in micro-expressions, movements, and sounds between Al anchors and live anchors; (3) Audiences' cognition, attitude, and behavior of Al synthesized anchor video news.

This questionnaire is distributed online, mainly through WeChat group, friends circle, Weibo and Douban.

4 Research findings and discussion

A total of 216 valid questionnaires were collected in this survey, of which 53.2% of the respondents indicated that they had never been exposed to the video news broadcast by All composite anchors, 95.4% of the respondents indicated that they preferred live anchors, and

more than 90% of the respondents indicated that they were more likely to resonate with live anchors. From the audience's perception of the micro-expressions, movements and sounds of live anchors and Al anchors, live anchors have richer micro-expressions, movements and sounds than Al composite anchors. Especially in terms of sensuality, respondents scored 4.5 points on the sensuality of live anchors and 2.7 points on the sensuality of Al composite anchors. From the data, it can be seen that there is still a wide gap between Al composite anchors and live anchors in emotional transmission.

4.1 The richness of Al's composite anchor micro-expression positively affects the communication effect of video

Pearson correlation analysis shows that there is a significant negative correlation between the unsmooth micro-expression of Al composite anchor and the audience's viewing intention (r=-0.175,P=0.010), which indicates that the smoother and richer the micro-expression of Al composite anchor is, the stronger the audience's viewing will be. The research hypothesis H1 is supported. Now, lip shape, eye changes and smiles of Al synthesized news anchors are all visual effects of generating corresponding sequences based on voice-driven expression sequences by using multimodal synthesis model, which is difficult to achieve the richness of micro-expressions of live anchors when broadcasting. In william james's Principles of Psychology, two concepts, knowledge of acquisition and Knowledge-about, were compared. The former depends on the perception and intuitive response in the contact process; The latter depends on concept and rational judgment. Live news anchors usually form "knowledge of acquaintance" through perception and intuitive reaction, which is why live news anchors naturally accompany the corresponding eyes and expressions in the broadcast process. However, Al synthetic anchors have no physical perception and intuitive reaction between their own subjects and the outside world, and their Verbal communication is different from the "psychology-physiology-psychology" of live anchors. They "extract the voice, lip shape, expression, movement and image of the live prototype in news broadcast, and use intelligent synthetic technology and deep learning and other joint modeling databases to nonlinearly and automatically generate intelligent audio and video of the corresponding content from the input Chinese and English texts"[6]. So themicro-expression of the Al synthetic anchor is rather Stiff.

Micro-expression, as a non-verbal symbol, plays an important role in the process of information transmission. Albert Merabian, an American communication scientist, quantified the effect of people's emotional expression into a formula: 100% of information transmission = 7% of language +38% of voice +55% of posture. Therefore, in interpersonal communication, the expression of voiced language only plays a small part of the role in transmitted information, while a large part of it is accomplished through nonverbal symbols. Non-verbal symbols are synergistic and accompanying in the process of communication, which complement each other and are important factors in the process of information transmission. Micro-expressions of Al anchors, that is, nonverbal symbols are absent in the process of information transmission, and

the corresponding transmission process may be incomplete for the audience, which leads to the audience's weak willingness to watch the video news broadcast by Al anchors.

4.2 The richness of Al composite anchor action has no effect on the communication effect of video

Through Pearson correlation analysis, it is found that there is no significant correlation between the action richness of Al composite anchors and the audience's viewing intention.(r=-0.109,P=0.111), the research hypothesis H2 is not supported. The reason is that news programs are more serious, and there is not much room for body language to play. The movements that can be examined are only the small changes of the head, shoulders and eyebrows with the broadcast content. Body posture, body movements and etiquette movements all belong to body language. Body language refers to the nonverbal symbols that use body changes in communication, such as expression, movements, body posture, body space distance, etc. as auxiliary tools to convey information and exchange thoughts and feelings[7].Compared with the voiced language, the movements as body language are more vivid and intuitive, which is more conducive to the transmission of information. However, in News Network and the selected Al anchors broadcast videos, the anchors sit in front of the anchors to express their voices, but the use of body language is not obvious. Therefore, the hypothesis that actions affect the audience's willingness to watch videos has not been confirmed in the study.

4.3 The richness of Al synthetic anchor's timbre positively affects the communication effect of video

Pearson correlation analysis shows that there is a significant positive correlation between the richness of Al composite anchor's timbre and the audience's viewing intention (r=-0.303,P=0.000), which indicates that the richer the Al composite anchor's timbre, the stronger the audience's viewing intention. The research hypothesis H3 is supported. The broadcast of Al anchor relies on digital technology processing. Although this technology has been constantly updated and iterated in recent years, it is still in the "toy" stage of technological development. According to Paul levinson, technological development usually has to go through three stages, and its dialectical logic can be expressed as "toy-mirror-art" or "pre-reality-reality-post-reality". The technology in infancy has obvious toy attributes, and the public's pleasure lies not in the result of the technology, but in the technology itself. At present, the Al synthetic anchor has not become a tool for news production process reengineering, but is more like a novel "media technology toy"[8]. This kind of technology that stays in the "toy" stage is always opposite to emotion. We found that although Al anchors show the same information transmission effect as live anchors, the richness of their sound and color affects the audience's willingness to watch. In the questionnaire, 95.4% of the respondents indicated that they prefer live anchors, and more

than 90% of the respondents indicated that they are more likely to resonate with live anchors. This reliance on technology is quite different from the creative method of broadcasting in broadcasting science: the so-called creative method of broadcasting is a process in which the creative subject knows, feels, understands, induces, organizes and structures the object. At this latitude, the growth of emotion is not a stylized mechanical treatment, but a surging and ever-changing emotional ups and downs[9].In the "toy" stage, AI synthetic anchor's voice, lip-shape, expression synthesis and deep learning technology can only make it have an image similar to a real person, and can't further convey emotions. The lack of emotional expression makes it difficult for Al synthetic anchor to reach the "mirror" stage of technological development, even more difficult to reach the "art" stage, and the audience's willingness to watch only stays at the attention of technology. After the "toy" craze receded, the attention of Al synthetic anchors plummeted, and the irreplaceability of live anchors was once again highlighted.

5 Summary and reflection

This study proposes that there is a big difference between Al composite anchors and live anchors in conveying emotions, which will further affect the video transmission effect. The rapid popularization of mobile internet and smart devices enables the public to receive, forward and comment on all kinds of information anywhere, and the professional value of Al synthetic anchor lies here. In the process of effective information dissemination, the accuracy of Al synthesized anchor language, the timeliness of information and the highest working efficiency at the lowest cost are beyond the reach of live anchors. On the other hand, the continuous updating and iteration of technology also provides a development direction for the grow of Al synthetic anchors in the artificial intelligence context of modern China to better adapt to the media convergence environment.

However, it is worth noting that this study also found that the audience's contact with the AI composite anchor is low and their recognition is not high. The technical support of the AI composite anchor makes it difficult for the audience to reach the deep requirement of "expressing feelings", which makes the audience's willingness to watch the video of the AI composite anchor not strong. While empowering news, it is worth thinking deeply about the emotional transmission problems caused by the lack of micro-expressions, movements and sounds. Even though the video transmission effect of Al synthetic anchor is not ideal, the continuous development of artificial intelligence technology and its application in the broadcasting and hosting industry still bring unprecedented pressure and challenges to live anchors. The "toy" stage of technology is a prerequisite for the subsequent stage. To jump out of the initial stage, Al synthetic anchors need to add more innovations in the media industry during the transition period, and constantly innovate themselves to reach the subsequent stage.

Correspondingly, live anchors also need to constantly innovate themselves in human-machine coupling, and complement each other's strengths and weaknesses, communicate with

each other and enjoy each other's prosperity in the upsurge of Al synthetic anchors.

However, due to the limitation of time and research conditions, this study still has the shortage of sample sampling and sample size. In the future research, researchers in the field of Al synthetic anchors may need to pay more attention to the physical phenomenon of Al synthetic anchors and expand the research and practice of Al synthetic anchors, so as to transform and upgrade live anchors, complement each other with Al synthetic anchors, and constantly build their own irreplaceable power.

References

[1] Liu Hailong. (2018). Physical problems in communication and the future of communication research. International Press, (02),37-46.(in Chinese)

[2] Shao Peng & Yang Yu. (2020).Al Virtual Anchor and Host's Body Communication. Journal of China Radio and Television, (06),71-74.(in Chinese)

[3] Zhao Xin. (2022). Affection: My humble opinion on the body communication of Al composite news anchors. China Media Science and Technology, (04),65-67. (in Chinese)

[4] Yin Lei. (2019). Research on the application of Al technology in the field of media. Master's dissertation. Harbin:Harbin Normal University.(in Chinese)

[5] Wang Yiting & Du Peng. (2019). Artificial Intelligence Recognition of Host Emotion. China Journal of Radio and Television (01),24-27.(in Chinese)

[6] Yu Chun. (2020). Off-body and on-body in communication: cognitive interaction of artificial intelligence news anchors. International Press (05),35-50.(in Chinese)

[7] Bi Jiwan. (1999). Intercultural nonverbal communication. Beijing: Foreign Language Teaching and Research Press.(in Chinese)

[8] Yi Yangang. (2019). As the anchor of Al synthesis of "toys". Young reporter (03), 112.(in Chinese)

[9] Zhang Song. China Broadcasting Science. Beijing: Beijing Broadcasting Institute Press. (in Chinese)

[10] Tan Jie. (2020). Research on the influencing factors of young audience's willingness to use Al composite anchor video. Master's dissertation. Jinan:Jinan University.(in Chinese)

[11] Arjun Chaudhuri. (1998). Product class effects on perceived risk: The role of emotion. International Journal of Research in Marketing, 15(2):157-168.

[12] Venkatesh V., Thong J., Xu X.(2012).Consumer Acceptance and Use of Information Technology:Extending the Unified Theory of Acceptance and Use of Technology. MIS Quarterly, 36(1) :157-178.

[13] Venkatesh V,Morris M G,Davis G B,et al.(2003).User Acceptance of Information Technology:Toward a Unified View . MIS Quarterly, 27(3):425-478.

[14] Davis F D.(1989).Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly, 13(3):319-340.

[15] Al-Jabri I M, Sohail M S. (2012). Mobile banking adoption: Application of diffusion of innovation theory. Journal of Electronic Commerce Research, 13(4): 379-391.

[16] Clerwall C. (2014). Enter the Robot Journalist. Journalism Practice, (05):519-531.

[17] Rodriguez L., Dimitrova DV.(2011).The levels of visual framing.Journal of Visual Literacy, 30 (1): 48-65.