Mutual Reinforcement of Predictions and Artistic Illusions in Generative Language Modeling

Qiaohe Yang

yangqiaohe0408@126.com

Academy of Arts & Design, Tsinghua University, Beijing, China

Abstract. Artificial intelligence systems, represented by ChatGPT, are revolutionizing society and art with a speed and density never seen before. Therefore, when generative language models intervene in art, constructing contexts and contexts of human consciousness and behavior, and creating simulations and generating "predictive" information that reinforces the artist's creative intent, the perception of art's virtual image will be enhanced. The anchoring of data and information selected under the guidance of the artist's subjectivity and the feedback constructed by the Internet crawler system carries a projection of the artist's own. The commonality between consciousnesses is in fact the reading of "illusion". In this paper, we take the relationship between the "prediction" of generative language modeling and artistic illusion as a research content, and explore the unconscious artistic illusion produced when facing the predictive "information". It tries to explore the possibility of creative art practice that grows organically between artificial intelligence and art.

Keywords: Generative Language Modeling; Illusion; Image Psychology; Corpus

1 Introduction

The wave of artificial intelligence technology led by ChatGPT in 2018 has swept over the world, making language model (LM), a key technology of artificial intelligence, highly concerned by academia and the computer information industry, and becoming the focus of the topic pushed into the public's view. However, the development of language model technology does not come out of nowhere, and has gone through a continuous evolution from grammatical rule language model, statistical language model to neural network language model, of which the neural network language model provides a broader technical basis for deep learning. At present, based on Generative Pre-Train and Natural language Processing (NLP) is the cutting-edge development trend in the field. Through the massive text data and corpus on the Internet to learn language rules and patterns, semantic understanding based on context, the use of deep learning architecture to improve the learning ability, to build a "big data + big model" automatic learning to adapt to the downstream tasks of the model paradigm.

It has to be recognized that the speed of the iterative development of artificial intelligence has exceeded human expectations. When Turing used the "imitation game"[1]in 1950 to describe the problem of "whether machines can think", it was an attempt to validate the possibility of machine-intelligence through the preparation of simulated processes and states of mind. Nowadays, generative language models, such as ChatGPT, complete conversations with users

ADDT 2023, September 15-17, Xian, People's Republic of China

DOI 10.4108/eai.15-9-2023.2340896

and have already simulated the basic functions of human socialization. The breadth and effectiveness of the application of large language models have made people realize that it is no longer nonsense for computational programs to have a conscious mind. However, there is no equivalence between simulation and ontology. From the ontological point of view, the intelligent consciousness of generative language model is the technical logic of the model architecture constructed on the basis of strings, computational formulas, programming codes and so on. And from the epistemological point of view, different scholars from the field of psychology, philosophy or art and other areas of its value and significance to explore, is to expand the disciplinary dimensions of the theoretical study of artificial intelligence. Not only that, when the generative language model is regarded as a tool by artists, the artistic creativity inspired by its techniques of language understanding, predictive optimization, and generative fine-tuning extends the application field of the technology with the practicability of art.

2 Consciousness and the "Prediction" of Language Models

Artificial intelligence simulations were not intended to be "new humans". As early as 1977, R. Shanker suggested that the purpose of program simulation is human comprehension. In the study of effective information communication, the act of reading and interpreting text involves not only comprehension and interpretation, but also prediction. Prediction has a significant impact on comprehension. From human experience, it is not difficult to realize that when placed in a specific contextual scene, even if some information is not received in the strict sense, it does not affect people's understanding of the content of the information and make a reply. This is because human beings are able to predict various possibilities based on their personal experience and knowledge, and then select the corresponding information words to be projected into the context to complete the rationality of self-cognition. Hermann von Helmholtz took the visual senses as his research perspective and proposed that human prediction is an unconscious reasoning mechanism with a hierarchical structure (**Figure 1**). Conscious thought is preceded by sensory perception as the preferred processing mechanism, with sensory sensors transmitting information for the brain about the model constructed for the outside world. When discrepancies occur between sensory information and modeled expectations, "perception depends on a priori expectations, i.e., regularities extracted from early sensory events. Regularities perceived in the external world are then translated into predictions of the current perceptual signal through the interaction of E (error units) and R (representation units)." [2] In other words, when the acquired information appears in an incomplete or mismatched state, the neural network starts to autonomously predict the unknowable information on the basis of a priori experience and interprets the visual scene with understanding the information. In this way, it can be seen that information has a selective function in the set of possible states. Of course, the components of the neural network are processed in parallel and run concurrently when the brain carries out the prediction mechanism. Conscious activity is generated instantaneously, so it is difficult to clearly distinguish prediction from perception in self-awareness.

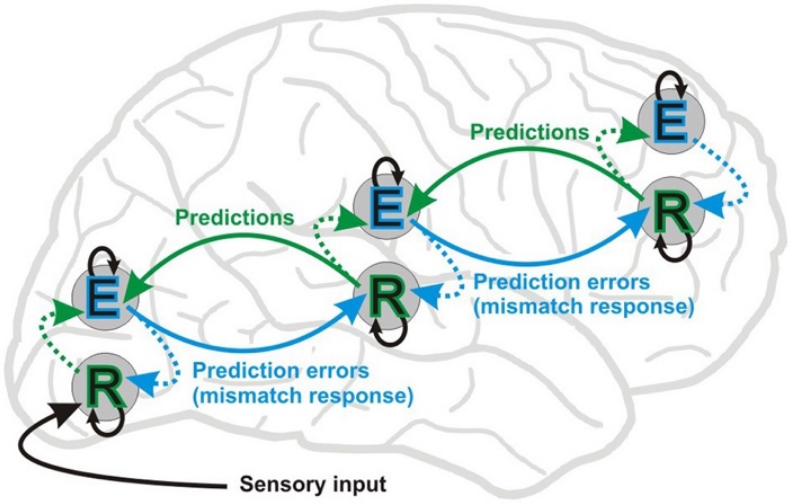

Fig. 1. Simplified scheme of the hierarchical predictive coding framework (Friston, 2005, 2008, 2010).

When modeling the parallel operation of neural networks in the brain to try to solve the data sparsity problem encountered in estimating probabilities in n-grams, in 2003 Yoshua Bengio proposed the classical neurolinguistic model Nerual Network Language Model (NNLM) to parallelize the processing of data, and parameters in order to better preserve word order information. However, the emergence of the NNLM did not start with a higher quality of pre-training. It needs to satisfy the need to "simultaneously model (1) the complexity of word applications (e.g., syntax and semantics) and (2) how to apply multiple variations in different contexts (i.e., modeling polysemy)".[3] Although developed through several models such as NNLM, RNNLM, word2vec, ELMo, etc., the language models developed before ELMo encountered challenges in semantic discrimination of polysemous words.ELMo released by AllenAI in 2018 started from the bi-directional language modeling to be able to efficiently capture contextual information based on contextual contexts to solve the problem of polysemous semantics, and the important language modeling prediction, one of the tasks, began to enter an iterative development stage. In the same year, GPT (Generative Pre-Training) released by OpenAI provides a pervasive semi-supervised learning approach to natural language for language understanding with a one-way language model. In this modeling approach focusing on generative tasks, language comprehension tasks become an evaluation category to study the effectiveness of the language model, which is categorized into four types: natural language reasoning, question answering, semantic similarity, and text categorization. Prediction plays an important role in the semi-supervised and unsupervised training of large-scale language models as part of the "pre-training + fine-tuning" paradigm.

2.1 Pre-training and Fine-tuning

The semi-supervised learning approach is a combination of two stages, Unsupervised pre-training and Supervised fine-tuning. First, pre-training refers to "the first stage of training using generalized tasks and large-scale unlabeled data to allow machine learning models to learn parameters with strong generalization." [4] Secondly, in the supervised fine-tuning phase, the pre-trained model is adapted using a task-specific pre-labeled corpus to accomplish adaptation

to the task. In essence, the "fine-tuning" phase is filling in the gaps in information to accomplish the target task through prediction. In contrast to the unconsciousness of the brain's prediction mechanism, the predictions of the language model are guided by human "supervision", thus reflecting the characteristics of prefabrication. Of course, the simulation of a language model is expected to result in higher quality feedback. Therefore, when a language model predicts the next word from human input, the prediction development will go through the stages of satisfying self-expectations, learning human preferences, and aligning human preferences.

"Satisfying Self-Expectations" is a supervised fine-tuning of pre-labeled data to learn human intentions about the desired answer to a question. Tasks such as common-sense reasoning and sentiment analysis are more often used to deal with processing for contextual understanding, however, human labeling of data is used for the purpose of performing the task. The predictions at this stage are as if humans are creating virtual mirrors of their alter egos by means of artificial language models, and the viability of linguistic dialog mechanisms is utilized to achieve the self-satisfaction of the Creator.

The model of "learning human preferences" moves from the early simulation environment to natural language. The Reward model is trained to optimize the learning of emotions and descriptions by fine-tuning the model according to the labels of human-annotated preference data. At this point, reinforcement learning of predictions begins to attempt to escape the dominance of human self-consciousness by correcting seemingly reasonable errors as well as falsehoods.

"Aligning Human Preferences" aims to align AI to human values and follow human intentions or preferences. The accuracy of learning is reinforced by human feedback training that "scales to larger models, collects more feedback data, closely monitors researcher-labeler protocols, and provides frequent feedback to the labeler".[5] This not only reconstructs the deep learning model, but also begins to investigate the application of multimodal models with input textual information that extends from characters to visual images, speech recognition, and so on. This phase of the research is trying to build the model program as if it were a human being, with a multi-sensory perceptual system for information acquisition mechanisms, and then moving to conscious understanding and pre-determination.

2.2 Corpus Selection

The selection and capacity of the text corpus will have a direct impact on the perplexity and validity of the predictions of the language model. If we understand the correlation between corpus and comprehension accuracy from the perspective of cognitive psychology, we will find that "context" and "memory processing" correspond to it. The "context" is the knowledge or experience that human beings must use in their learning and comprehension processes. Once the quantity and quality of the variable elements change, it will have a direct bias effect on information extraction and cognitive processing. It can be seen that the capacity of the text corpus dataset and the bias of the selected corpus sources are the main factors that bias the completeness and accuracy of the language model's understanding of the world. For example, in the early days of GPT, when natural language inference (NLI) tasks, such as word connotation, co-reference, and grammatical ambiguity, were performed to recognize textual semantics, the selected corpus included datasets such as QNLI, SciTail, or RTE. Subsequently, the GPT-3 model corpus selection was fine-tuned to use text-based WEB browsers to answer the questions,

i.e., collecting data from web pages and predicting human preferences to select the corresponding web source quotes for answering, which is optimized in terms of capacity expansion. However, system programs dominated by the English language family need to construct a brand new corpus model if they want to understand Chinese semantically. The ERNIE language model released by Baidu in 2019 follows the "word granularity" pre-training of the Chinese corpus of the BERT model, and selects the Baidu Encyclopedia, Baidu News and other corpora to expand the a priori semantics of the Chinese text, and to enhance the semantic comprehension and learning ability. When the model parameters and data size are increasing, the learning performance can not only be enhanced, but also learn the implicit knowledge. However, it is obvious that the language model can correlate explicit knowledge with implicit knowledge on the premise of constructing a more mature model, so as to realize more efficient and accurate prediction.

Frederic Bartlett has pointed out that "many of the transformations and distortions that occur in human consciousness are intended to make the story more coherent and plausible"[6], so it is clear that prediction exists and plays the same role both within the consciousness of the human brain and in the generative pre-training of language models. In the learning and comprehension process, on the one hand, it depends on the background of existing experience (corpus), and on the other hand, it is necessary to choose the appropriate mechanism for extracting information depending on the type of task. As the technical capabilities of generative language modeling are widely used in different domains, the challenges and potentials that come with it are the focus of debate among scholars in various fields. Artists who are adept at modeling the world and keenly aware of the characteristics of the times will not miss the intervention of AI technology in the creative practice of art. McLuhan recognizes in this regard that "art is accurate, forward-looking knowledge of the psychological and social consequences of the next technology." [7] In this regard, does the predictive correlation of generative language modeling with perception and consciousness trigger a certain effect on artistic perception? To the extent that language models validate the feasibility of learning and understanding, can artists use them to construct new artistic paradigms? These questions are worth discussing in depth.

3 Predictions and Illusions in AI Art

It is known that humans not only fill in the vacant information and complete comprehension and cognitive judgments through a sequence of predictions, but also construct relatively familiar contextual scenarios that self-maintain the plausibility of the predictions. In this cognitive process, non-retinal stimulus experiences such as hallucinations, mental imagery, associations, perceptual fillers, and many illusions are automatically prioritized for execution, resulting in a Phantom perception.In terms of the memory system, the memory process in which an individual unconsciously extracts that information and executes it, as guided by contextual contextual bias, is known in cognitive psychology as It is easy to realize that in both perceptual and psychological research, the variable factors of prediction inevitably involve the contextual context, the prediction operation mechanism, and the projection feedback of the illusion. Prediction gives rise to illusions.

However the study of illusion is explored by Ernst H. Gombrich by placing it in the field of art. Art constantly creates predictions and illusions in different art forms to stimulate curiosity and

the desire to create and practice art. Compared with cognition and psychology, which are studies of individuals, artistic illusion involves two identities, the artist and the viewer, and the relationship between the two turns from time to time in the process of prediction and illusion in different artistic practices, so artistic illusion will show a duality.

First of all, art itself is a contextual context when the artist consciously creates the world of imaginary images. It is only in the inner context of art that the illusion of conceptual and formal construction can plausibly work wonders. The viewer is subjected to a strong prediction of narrative sequence by the bias of the art, which actually induces an illusion. On the one hand, there is the illusion of perception from the unconscious. From Roman illusionist paintings and Pompeii's panoramic frescoes to virtual art, immersive experiences, and virtual reality art, both classical art and contemporary multimedia art aim to create virtual spatial contexts that put the viewer in a closed state and thus create perceptual illusions. Although the art form focuses more on the study of visual illusion, the human brain itself has multimodal neurons, and when responding to an abstract theme, it does not only respond to a specific vision, while media art has become a new art paradigm with multimodality, and it begins to eliminate the barrier between art and the viewer with a comprehensive perceptual illusion experience. The art phenomenon reflects the fact that in a given contextual context, projected predictions will somehow replace perception, i.e., "the greater the probability of a symbol appearing in any given situation, the less informative it is, and the ability to predict the outcome eliminates the need to listen." [8] Even if the context in which the art is created is not derived from real-world experience, and it is not possible to extract the viewer's pre-existing experience in order to predict vacancies, or even if the viewer is able to consciously recognize the fictional nature of the art, this does not prevent the artistic illusion from still having an effect. This is because the mechanism of projective prediction will operate. Thus, on the other hand, the illusion arises from the viewer's mental anchoring in the self-interpretation of the art. Having stimulated the viewer's imagination, it can trigger the projection of unpresented information. The illusion of art needs to rely on the mutual reinforcement of projections and illusions. In other words, filling in the artist's cues with one's own experience.

Secondly, although the prediction and illusion of art comes from the artist's self-preparation, it still creates an illusion that can easily blur or confuse the original intention of art. In the process of art practice, "contextualization" contains a double semantic meaning. One refers to the "corpus" of artistic materials, that is, the source of artistic expression or the reference of artistic intent; the other points to the artistic action (practical mechanism or formal language), the practical path of creation with what artistic style, media application, or form of expression. The effect of artistic illusion can only be strengthened when the prediction caused by the context of "action" enhances the intention of the artist's "corpus" context. From this, it is easy to realize that in creating the mechanism of prediction, the artist not only needs to pay attention to the match between the concept and practice of art, but also needs to pay attention to whether the content projected by the viewer is in line with the mechanism of prediction. In order to satisfy the effect of artistic illusion, the artist needs to leave space for prediction on the one hand, so that the viewer can have the possibility of projection on the other. Perhaps the artistic vocabulary of "white space" in classical Chinese painting is sufficient to provide the viewer with enough space for perceptual prediction. On the other hand, it is necessary to ensure the feasibility of the viewer's prediction that the individual has the ability to fill in the gaps in the "context" of the known experience. Because "all information exchange consists in a concession to the knowledge

of the receiver, the context that governs it and the awareness of those possible alternate interpretations must be excluded."[9] In the process of creating, predicting, and projecting feedback on artistic illusions, both the artist and the viewer need to engage in role substitution in order to gain a fuller dimension of understanding. This substitution experience is itself an important way of constructing artistic illusions.

In the above discussion, the phenomenon of prediction-induced illusions in the fields of perception, perception and art has been described separately. However, in the age of intelligence and digitization, the study of art forms has long been unable to escape the drastic changes brought about by media and technology. The evolution of the medium of prediction and illusion has entered a new era initiated by Artificial Intelligence. As the iterative development of the technology matures, the exploration of the possibilities of contemporary art and the subversive construction of art theories are imminent. aI interventions in art are still limited to a few experimental projects by a few artists, such as Artist Xu Bing's *Artificial Intelligence Infinite Films* (*AI-IF*) (2017-2020), to be released in 2021, and Cai Guoqiang's *cAI*™ (2023), both of which are based on linguistic language, with the aim of creating a new generation of art. Both of which explore generative art creation based on deep learning of language models. Using these two art projects as research samples, this paper explores the correlation and alienation between the prediction of language models and artistic illusion.

As an art project that also takes AI generative as its starting point, AI technology not only has instrumentality as an art medium, but also, based on its intelligent and autonomous generative nature of deep learning and prediction, will assume the identity of the art creation subject together with the artist and the viewer during the project practice. In other words, the AI art based on generative language model has three "artists". The artist selects the appropriate AI program medium for the expression of artistic concepts, and creates predicted biases and vacant areas under close guidance to construct an interactive mechanism for artistic practice. Subsequently, the guided viewer participates in the project and inputs the corresponding textual information to the AI program, at which time the deep learning of the AI program begins to extract the contextual corpus and outputs the corresponding feedback results through the prediction mechanism within the language model. The "input-output" interaction mechanism between the viewer and the AI program will continue to operate cumulatively during the practice of the project. It is not difficult to find out that generative AI art consists of "rules" constructed by the artist, "art" generated by the mutual collaboration between the viewer and the AI, and the three of them work together to complete the art practice. At the same time, its art form is not presented as a fixed art carrier, including but not limited to the generative algorithmic prediction path, feedback multimodal information and even the narrative logic of AI art, all of which are used as the art form of generative AI to explore the future extension direction and possibility of art practice boundary.

In terms of the selection of the AI program system, the *AI-IF* project is in the era of ChatGPT-2 technology due to its creation between 2017 and 2020. Therefore the initial creation of the *AI-IF* project was based on the open source GPT-2 language model. And the program of the cAI™ project "developed by Cai Studio and the E.I. Art technical team, is an integrated system that learns and generates text, images, audio and video." [10] Being a self-created program for deep learning rather than choosing an existing generative language model, the program construction will show a more artist-conscious dominance. Therefore AI art under the dominance of techno-

logic will also produce different predictions and artistic illusions due to the variability in the structure of the selected language models.

The intention of the artist to create the concept can be seen from the extraction of the contextual corpus dataset. Since the corpus base of the GPT-2 language model comes from the massive and diverse Internet dataset. Therefore the narrative context of the AI-IF project is not consciously constructed by the artist, but unconsciously unfolded based on the external corpus dataset. This is in line with what Xu Bing pointed out at the Pingyao International Film Festival 2021 (Dialogue-Interaction) event, that the project tries to explore, with the help of AI, "something that cannot be created and provided by natural human filmmakers, but is needed by human beings", and pursues an "internal formula" for human image expression. "internal formula" or "functioning structure" of human pictorial expression.[11] The cAI™ project, however, attempts to build a connection between the artist and the unseen outside world, so that the artist's individual consciousness and artistic methodology become the core of the exploration. The main sources of the program corpus developed for Cai Guoqiang are the artist's artistic creations, audio recordings, writings, images and archival materials over the years. The multimodal possibilities of information are formed in the fusion of the constantly iterating AIGC technology. The secondary sources are based on the artist's keyword tags, based on the Internet information with a crawler system to automatically crawl the corresponding corpus, from external data sets to supplement and enrich the corpus of the internal main line. This constitutes a multidimensional construction of the corpus that is internally set and externally supplemented. Moreover, it is the continuous self-iterating corpus data within the generative language model that can extend the infinite growth cycle for the art.

In terms of the prediction mechanism and the construction of artistic illusion, the prediction mechanism of AI-IF involves four model architectures, which are the script model, the video alphabet model, the matching of the generated script with the video subtitles, and the generating dialog audio and background music model. The GPT-2 language model is labeled with six parameters: character, place, number of people, what happens in the video, actions and objects. Based on the existing context of experience between the viewer and the computer operation, the project provides an interactive interface for the viewer, where the movie genre option logo and the text input box are intended to guide the viewer to participate in inputting textual information, and the GPT-2 generates visual feedback of text, image and sound in the prediction model based on the labeled parameters and the viewer's input text. As a movie, the project disrupts the traditional way of creation and the viewing experience of the viewer. The viewer is guided in the context of the action, and the feedback from the language model is in most cases not comprehensible to the human language system, e.g. the narrative between the captured video clips and the generated subtitles is difficult to understand. At this point, the viewer then needs to translate the information seen into the context of the action, realizing that it is a record of a variety of different intentions, predicting accordingly based on the individual's pre-existing experience, and at this point presenting a certain state of readiness ready to start projecting, a state also known as mental orientation. AI-IF, as a "private cinema" project in which the viewer actively participates and generates in real time, creates for the viewer the illusion of art, first and foremost, of creative initiative. Whether due to the performance of the language model algorithm or the limitations of the project's own generative mechanism, the narratives and logics of the viewer-driven films are not as good as they could have been. The viewer is passively guided by the predictive mechanism, but is actually trapped in the illusion of self-direction.

Another dimension of the illusion is reflected in the viewer's projection of the feedback. The viewer's projections of vacant, illogical narrative episodes are filled in with appropriate interpretations, and even this inevitably distorts the message of the feedback image. What's more, for this project, whether or not the generated image narrative or artistic demand meets the viewer's expectations is not the artistic purpose; rather, it is the practicality of the generative process itself that is the focus. The viewer is no longer the recipient of the art, but rather one of the creative subjects who complete the process.

In contrast to the collaboration between AI-IF and the viewer, cAI™ constructs another intelligent virtual image of the artist, even embarking on an experimental "collaboration" with Cai Guoqiang. This can be seen as a collaboration between human and AI, but also as a collaboration between the artist's multiform self. Because the model's corpus is selected to serve the individual artist, the feedback predicted by the cAI™ program is more closely aligned with the artist's consciousness. And with the rapid iterative development of the corpus dataset, it is indeed able to transcend the cognitive dimensions of being an individual human being in terms of perception and vision, and provide more diversified revelations for human beings. In the experimental exploration of the project, the artist provides input to the program through text or real-time image viewing, and the feedback visual information provides imaginative inspiration for the artist's creation (**Figure 2**). Cai Guo-Qiang describes in his transcript, "cAI™ i.e. returned to me some images. I was inspired by it: the so-called canvas on the moon, waiting for aliens to give information to human beings, is not as good as the human beings' search for themselves, the shadow of the earth society, the canvas that is the mirror of the human society, and the cAI™ is also Cai's mirror ....."[12] Whether the cAI™ is capable of predicting and conforming to human consciousness, or can this validate that an AI is capable of autonomous thought? This is clearly an illusion when practiced by artists. Interpretation of visual images of feedback is practically inseparable from an attempted projection effect, and polysemy is clearly the key to that problem. In life as in art, it is the ability to project rather than the conceptual context that shapes our understanding and judgment. Although the cAI™ project is currently releasing publicly only some of the results of its initial explorations, it will continue to present Cai's artistic philosophy and cognitive thinking in the future. Admittedly, it remains to be seen whether the highly personalized prediction model of the simulated artist will tend to become fatigued under the limitations of the corpus and prediction mechanism, generating feedback in a cyclical manner, and how far the boundaries of AI art exploration will extend.

Fig. 2. *cAI*™, Selected images based on Cai's narration, 2023

4 Conclusion

In human perception and consciousness, the logical process from the triggering of a prediction to the creation of an illusion is not yet absolutely conclusive. Still, we recognize that human prediction is a synthesis of perception and reasoned knowledge. Considering a generative language model either as a systematic program or, divorced from the technological trappings of artificial intelligence, simply as a linguistic system, the elements are hooked up to each other in an unnatural, formally logical relationship. The predictions in the model itself operate according to a framework and logic. Can the reasoned predictions of a language model generate perceptions, if their generativity is seen only as an inevitable product of the act of operating within the framework of the model? When the generative language model intervenes in the collaborative creation of art, the first thing we feel is the sense of urgency that there is no place to hide in the age of artificial intelligence, followed by the exploratory and forward-looking nature of AI art creation. The artistic perception embodied in AI art comes more from the artist's prediction and projection than from AI. Because the prediction of generative language model is not a set of a priori logic. As the iterative capacity of a language model gets larger, its rational-knowledge properties get stronger, while the tangibility of perception gets weaker. Language models are only the mainstream models that can be widely used in the current stage of AI development, and the future development of the field of AI will not be stagnant, and the practice of integration with art will not end. Therefore, the direction of AI art practice needs to focus on the fact that although the AI medium at the technical level has a predictive logic structure of its own, it still needs to study and explore the conversion mechanism (decoding) when it is applied in different fields. In art, only by relying on a wider range of art practice experience can we realize the possibility of AI technology carrying the perception of reasoning, and only then can AI as an art medium work in a real sense.

References

[1] Turing A M.:Computing Machinery and Intelligence.Mind. Vol. 59 ,pp.433-460. (1950)

[2] Stefanics G, Kremláček J, Czigler I, et al. :Visual mismatch negativity: a predictive coding view. Frontiers in Human Neuroscience. Vol. 8,p.666. (2014)

[3] Peters M E, Neumann M , Iyyer M, et al.: Deep contextualized word representations. Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Association for Computational Linguistics.p.2227. (2018)

[4] Chaoyang Z,Guibo Z, Jinqiao W.:ChatGPT's inspiration to linguistic macromodeling and new development ideas of multimodal macromodeling. Data Analysis and Knowledge Discovery.p.27. (2023)

[5] Stiennon N, Ouyang L, Wu J, et al. :Learning to summarize from human feedback. NIPS'20: Proceedings of the 34th International Conference on Neural Information Processing Systems. pp. 3008-3021.(2020)

[6] Best J.: Cognitive Psychology. Translated by Huang Xiting et al. China Light Industry Press, Beijing. p.110.(2000)

[7] McLuhan,M. :Understanding the Medium: On the Extension of Man. Translated by He Daokuan. Yilin Publishing House, Nanjing. p.97.(2019)

[8] Gombrich,E H.: Art and Illusion. Translated by Chengkai,Y., Benzheng,L.,Jingzhong,F. Guangxi Fine Arts Press, Nanning.p.172.(2012)
[9] Gombrich,E H.: Art and Illusion. Translated by Chengkai,Y., Benzheng,L.,Jingzhong,F. Guangxi Fine Arts Press, Nanning.p.205.(2012)
[10] Cai Guo-Qiang.:Launches Artificial Intelligence Art Project: cAI™.Cai Guo-Qiang.com. [2023-05-11].https://mailchi.mp/caiguoqiang/aicaipressrelease?e=fa3d4dd3f6
[11] In the live video of the event, Xu Bing, in response to the first audience question, mentions that there has been a tendency towards " formulaic" knowledge production within image making. For some of the interviews, see 5th Pingyao International Film Festival | Artificial Intelligence Unlimited Cinema. Tencent News. [2023-05-11] https://vku.youku.com/live/ilproom?id=8133992.
[12] Cai Guo-Qiang.: Launches Artificial Intelligence Art Project: cAI™.Cai Guo-Qiang.com. [2023-05-11].https://mailchi.mp/caiguoqiang/aicaipressrelease?e=fa3d4dd3f6