Retraction

Retraction: Multimodal Analysis of Visual Language Based on Artificial Intelligence (J. Phys.: Conf. Ser. 1915 022054)

Published 9 September 2022

This article has been retracted by IOP Publishing following an allegation that raises concerns this article may have been created, manipulated, and/or sold by a commercial entity. In addition, IOP Publishing has seen no evidence that reliable peer review was conducted on this article, despite the clear standards expected of and communicated to conference organisers.

The authors of the article have been given opportunity to present evidence that they were the original and genuine creators of the work, however at the time of publication of this notice, IOP Publishing has not received any response. IOP Publishing has analysed the article and agrees there are enough indicators to cause serious doubts over the legitimacy of the work and agree this article should be retracted. The authors are encouraged to contact IOP Publishing Limited if they have any comments on this retraction.

Retraction published: 9 September 2022
Multimodal Analysis of Visual Language Based on Artificial Intelligence

Wenbo Feng1,*
1Xi’an Polytechnic University, Xi’an, Shaanxi, China, 7100481
*Corresponding author e-mail: fengwenbo@xaut.edu.cn

Abstract. This paper studies the human development performance in the visual communication of multimodal channels within the artificial intelligence environment. Starting from the environmental changes of human cognition, the main characteristics of the dissemination of visual information in the context of artificial intelligence are analyzed, and the transformation from passive to active human cognitive activities is discussed. In the process of understanding the diversified visual information, the contradiction in the human memory has also become prominent. It is proposed that in the transmission of multimodal information within the context of artificial intelligence, the human experience of visual perception has changed, the forms to express visual information content and the means to transmit visual information are diversified, and the interactive relationship between people and information has been strengthened. It also makes the aesthetic experience of the aesthetic subject more and more personalized and stylized.

Keywords: Artificial Intelligence, Visual Language, Multimodality, Visual Perception

1. Artificial intelligence affects changes in visual forms
With the continuous development of artificial intelligence and the popularization of informatization and digitization of graphics and images, the visual forms we face in various electronic terminals have also changed. The most important change is brought by artificial intelligence in the transmission and reception of information and the increase of human-computer interaction. The multimodal transformation of visual information in the interactive process has also become the main means of communication. We all know that monomodal information dissemination is mainly completed by a single means or form, such as graphic mode, image mode, sound mode, etc.[1]; and the information dissemination through the multimodality based on artificial intelligence can complete information transmission by two or more channels as dissemination methods and information carriers.

The main reason for this change is that audiences in the artificial intelligence environment have higher and higher requirements for information carriers and forms of communication, and the forms of information dissemination and communication have become more complex and diverse. In the communication of this diversified form of visual information, it is necessary to consider not only the expressive ability, expression methods, aesthetics of the form, and the content contained in the form,
but also the human response to the visual perception in human-computer interaction, the group psychological characteristics, and individual’s life experience and visual experience. At the same time, digital technology factors based on big data and cloud platforms are also important causes of this change.

Therefore, multimodal information dissemination is the form of information dissemination in this artificial intelligence environment. With artificial intelligence and mobile Internet becoming an important part of people’s daily life, this intelligent lifestyle has gradually changed the way of information dissemination. The originally single-modal, linear, isolated, and static information transmission form has changed into a multimodal, non-linear, network interconnected, and dynamically interactive form. Artificial intelligence technology is extending the utilization of space and the expressive power of human information dissemination.

2. Human visual language in artificial intelligence

2.1. Changes in the cognitive environment

Visual perception is the main means for human beings to cognize the world. “Visual perception includes image perception, image recognition, spatial perception, etc. It is the ability to sense three-dimensional space by distinguishing certain patterns from many stimuli in the objective environment, recognizing familiar images[2].” Whether it is graphic perception, image recognition, or spatial perception, all are the human experience of visual perception towards objects. This perception can be the shape and color of the graphic image or the movement of the object. It would be wrong to think that this is the subject's visual perception of the object. In terms of human visual cognition, in addition to cognizing the object itself, the environment in which the object is located is also an important part. Different environments produce different cognitive effects. For example, black circles of the same size, located in a square environment of different sizes, or surrounded by circles of different sizes, are perceived by people that they have different shapes and sizes (Figure 1). In sum, the environment influences people's visual perception.

Figure 1. The illusion of a circle

With the help of artificial intelligence, the impact on human cognition and perception has been unprecedentedly strengthened, and the visual form of information has also been developed (monomodal to multimodal). Similarly, the environment in which humans perceive things has gradually changed. Under traditional circumstances, human visual perception is produced in a real environment of light and shadow, and light and shadow are the preconditions for visual perception. As Arnheim said, “Lightness and darkness are an attribute of the individual and independent object, and light exposure provides a common basis for the existence of all objects. On this common basis, objects and parts of objects can be revealed in the abyss of darkness.” Light forms a certain environmental atmosphere through brightness, color, strength, etc., which affects the subjective visual perception of the audience. In the process of visual communication, some environments are deliberately formed by humans. The purpose is to reinforce the strength of information expression and the intensity of visual perception. In other words, the environment of information has become a "significant form." The visual cognition of the environment is mainly realized by the color, the symbol of light and shadow, and the contrast between the subject and the background. In the actual communication process, this background is mostly static and isolated. In the context of artificial intelligence, the human visual
perception environment has undergone a radical change, mainly characterized by intelligence, interaction, virtualization, digitization, and big data (Figure 2). The background of intelligent information transmission means that in the process of information dissemination, the background can not only change according to the information content, information form, and information means, but also according to the characteristics of the information receiving subject, the subject’s environment, and the subject’s subjective requirements[3]. Interactive information communication refers to the process of information communication by adding subjective participation and involving a certain degree of human-computer interaction. In contrast with the real environment, virtualization is the basis of artificial intelligence. The visual cognitive environment under the digitalized virtual environment relies more on computer, data helmet, circular screen, data glove, holographic projection, and other technical equipment. This kind of virtual environment has more intelligent factors. Although artificial design in the real environment also has intelligent factors, most of them just make the best use of the situation. However, the intelligent design operation in the virtual environment makes the transmission of information more purposeful.

Figure 2. Changes of information environment under artificial intelligence

2.2. From passive acceptance to active participation
Human cognitive activities start from perception. Only when the visual information of perception is noticed can recognition be completed. Human attention behavior is usually divided into active attention and passive attention. The former is caused by the object, which stimulates the audience to produce attention behavior. In the traditional environment, the visual information under passive attention is edited and designed in advance by the information coder, and the audience can only accept it passively. In the cognitive process, they are in a passive acceptance position, and then put forward criticism, such as logo design, posters, slogans, image ambassadors, promotional films, and so on. The audience of this kind of visual information designed in advance have no right to modify and transform, and can only accept it passively[4]. Active attention refers to the conscious behavior of the audience to cognize the object consciously. Under the environment of artificial intelligence, the audience will actively participate in the process of information dissemination. In the process of information coding and decoding, the audience gradually becomes the commander of visual language information, which leads to the change of the audience's behavior mode and makes them actively participate in the process of information dissemination. For example, in the artificial intelligence environment, we can not only understand a place through the official website, microblog, WeChat, and information circle of the place pushed by the computer according to the user's requirements in the artificial intelligence network platform, but also understand the cultural significance of the place through graphics and images, video animation, communication, and interaction, etc. Moreover, we can actively participate in it by
expressing our understandings of the place through digital means. Although the personal expression may be biased, it cannot be denied that it has improved the visual language of the place to a certain extent, and even create another culture based on the place. Therefore, the multimodal presentation of visual language in the context of intelligence accelerates the process to the maximum, which is not only reflected in the change of means of visual information communication but also the expansion of the connotation of visual information.

2.3. Diversified understanding of visual information

As artificial intelligence continues to penetrate the process of information dissemination and solves technical issues such as the input and coding of information dissemination, human thinking is infinitely liberated, and gradually participates in the construction of the visual information ontology, and treats their understandings of visual objects as the parts of these objects. It no longer duplicates objects. What follows is that people put forward higher requirements for the expression of visual information, so the multimodal expression of visual information arises at the crucial moment. It should be said that there is a complementary and mutually reinforcing relationship between the two. Multimodal visual information expression methods have enriched human expression ability, and the human pursuit of the ontology of information has further promoted the development of expression means. As far as cognitive psychology is concerned, all human cognitive behaviors can be understood as the problem-solving process concerning information objects, from the input and knowledge of information to the coding and processing of information and to the output and dissemination of information. This process shows their understanding of information. The transmission of visual information in the artificial intelligence environment has shown a diversified trend. People's cognition of information, especially visual information, has reached the stage of active understanding and participation. Being liberated from the complexity of visual information input, visual information coding, and visual information output, human beings focus only on the personalized expression of visual information, which means that people are not only limited to the understanding of visual information itself but also require further design and processing of this information based on their understandings to create visual expressions and then enter the mass communication (Figure 3).

Cognitive psychology is called information processing theory because it uses information processing as the cornerstone to understand and explain the relationship between human high-level thinking activities and information processing. In this relationship, multimodal visual information is received by human visual organs and completed through perception, attention, understanding, thinking activities, visual language, and artificial intelligence.

2.4. The emergence of memory contradictions

General psychology defines memory as the process of experiential imprinting, maintaining, and reacting. The process of remembering is to encode the perceived information into a model that can be stored in the human brain. If the purpose of memory in the traditional environment is storage, then the
The purpose of memory in the multimodal visual language is no longer just storage, but more involves secondary processing of information, which is the process of information from decoding to encoding. Multimodal visual information dissemination makes information sources be recognized and memorized in a more intelligent environment, and its connotation is constantly expanded. With the increase of human perceptual experience and the support of intelligent technology, the information source is abstracted and simplified infinitely by means of human-computer interaction and in a multimodal way[5].

3. The characteristics of visual perceptual experience in a multimodal environment

3.1. Selectivity of visual perception
With the gradual changes in the human cognitive environment, in the artificial intelligence environment, the expression of objects faced by the public is presented in a variety of forms, (as to the representation of graphics and images, there can be several different art forms to express each kind of aesthetic experience and visual experience. Dadaism, expressionism, ink style, paper-cutting, shadow puppetry, embroidery, and other folk styles). These forms and their combinations bring diversity to the audience. In contrast to language, visual information as a transmissive and persuasive sign is more effective.

Of course, in addition to choosing the visual representation of the object, the way of visual cognition is also selective. The way of visual cognition is the means that human beings use to recognize objects in the virtual network environment. Artificial intelligence and various technological equipment provide us with as variety of choices. Any technology, like computers, mobile phones, data helmets, data gloves, touch screens, circular screens, augmented reality technologies, etc., can enhance human visual perception and cognitive ability.

3.2. Interactivity of visual perception
The information dissemination of visual language, with the support of artificial intelligence, has changed from one-way reception of objects to two-way interaction between man and computer. One-way reception refers to the process from encoding to the decoding of visual information, in which the audience can only passively receive the information and complete the interpretation of the information; two-way interaction refers to the two-way interactive process of information transmission between audience and computer from input, to processing, and to dissemination with the help of artificial intelligence. With the intervention of artificial intelligence technology, visual information involves the subject-object interaction. When computer artificial intelligence transmits information to the subject, it can not only consider the difference in the subjects’ (or audience’s) life experience, but also their difference in knowledge, education, cognition, and other cultural backgrounds, and even the difference in the changes of perspective, etc. Moreover, it also collects the audience's feedback and expression of visual information, enriching the changing connotation of the information object, and this change will act on the information object in time with the support of artificial intelligence technology. What’s more, because of some social media like Weibo, WeChat platforms, Internet, and other chat platforms, etc., the interaction between information and audiences has been upgraded to a multi-party interaction of audience-information-audience. This multimodal information exchange has extremely promoted human perception and this way of interaction can explore the essence of things from more angles.

3.3. Personalization of the aesthetic experience of visual perception
As we all know, the aesthetic style of visual art comes from the hands of professional graphic and image designers and artists. But since the late 1960s when Pop Art became popular worldwide by breaking the boundaries between life and art and between elegance and vulgarity, the movement of art has marched towards the direction of popularization, individualization, and popularization. At the same time, this process is infinitely accelerated and amplified by artificial intelligence. All aspects of
people’s life, study, and work, etc. are affected. With the help of artificial intelligence, everyone can become an artist. Any visual style in art history can be realized effortlessly. People don’t even have to worry about their own artistic talents. Everyone is a designer and communicator of visual symbols. The pictures and videos casually taken by mobile phones quickly form an entity surrounded by information clusters. These information clusters are also diversified in terms of visual expressions because of the difference in the aesthetic background of the designers.

The multimodality of visual language has led to differentiation among the masses, forming individual visual art aesthetic group. Members in these groups gather together because their aesthetic perceptions are similar in one way or another although they come from different occupations, differences in beliefs, knowledge, education, backgrounds, nationalities, and regions. But similar social groups also have personalized characteristics presented in the visual expression of information because of individual differences among their members.

4. Conclusion

Today, with the continuous development of artificial intelligence, people’s observation of things and cognition of the essence of events have increasingly relied on more intuitive visual forms (graphics, images, and videos). More and more multimodal modes of language have realized multi-level, diversified, and intelligent information collection, encoding, and dissemination through various modern digital media with the help of artificial intelligence technology. People actively participate in the construction of information sources; they are not only the receiver but also the designer and disseminator of information. This psychological status of actively receiving information also presents a dynamic composition in the connotation of information. A piece of information is not invariable. From information structure to information connotation, all links of information are in a dynamic change. In the corresponding artificial intelligence environment, the multimodal form of visual language also has the characteristics of intelligence, virtualization, interaction, and diversification. Human beings can complete the aesthetic experience of personal visual information through human-computer interaction by choosing among a variety of visual languages.

References
[1] LIU Q. An Artificial Intelligence Based Data-driven Approach for Design Ideation[J]. Journal of Visual Communication and Image, 2019, (6).
[2] HUANG, Ben-liang. The Semantic Level of Interaction Design[J]. Packaging Engineering, 2013 (2).
[3] PENG, Gang. New Media Language in Network Art [J]. Literature & Art Studies, 2013 (4).
[4] Gombrich, E.H. Translated by LIN Xi. Art and Illusion [M]. Changsha: Hunan Science and Technology Publishing House, 1999.
[5] FENG, Wen-bo. Research on Information Graphics in the Digital Media Art Environment [J]. Packaging Engineering, 2020 (6).