

The Power Relations of Photography in the Algorithm Age: From Digitization to AI Generation

Rui Zeng *

Photography, San Francisco Art Institute, San Francisco, CA94106, US.

* Corresponding author Email: rustyzen@gmail.com

Abstract: This article explores the fundamental changes in the paradigms of image shooting and viewing after entering the algorithm age, and studies the power structures of images. The history of the algorithm age began with the digitization of images. With the popularity of the Internet and social media, the language turn of images has gradually completed after entering the AI era. The new shooting and viewing methods have given people more power and broken the existing aesthetic discourse structure. However, at the same time, through algorithms, commodification and addiction mechanisms, another scene of power concentration has also formed. This article emphasizes the changes in the power structure of images, and proposes that in the algorithm age, we need to deal with the power changes of images cautiously to ensure that AIGC technologies have a positive and sustainable impact on social development.

Keywords: AIGC; Photography; AI Image Generation; Algorithm Images; Power Relations; Midjourney.

1. Introduction

In early 2023, AIGC officially faced the public in an amazing way, bringing a pioneering way of content creation by generating content through artificial intelligence, greatly changing the mode of work of creation and production. Especially in the field of AI image generation, it introduces a whole new dimension to the visual information network which was mainly based on photography and video before, bringing new changes to power relations.

While basically continuing the basic logic of algorithmic image power relations, it has also made improvements in new contexts. On the one hand, it shocks the original power relations and provides the public with more participation and discourse power. On the other hand, through alienation, homogenization, addiction mechanisms, stereotypes and other means, AI image generation is also building its own system of power relations. Although the industrial structure of AI image generation is not yet perfect, the construction of its power relationship system is still in progress. However, the trends that have already emerged deserve our close attention, in-depth thinking and active discussion.

2. The Digital Transformation of Imagery: Examining and Changing Power Relations

2.1. Mass Replication and Information Flows: Disrupting Existing Power Structures

Since the advent of mass image reproduction, discussions about power dynamics have persisted. Susan Sontag's work *On Photography* highlights how photography empowers vision, serving various entities like the state, industry, and science [1]. This perspective regards photography as a practical knowledge aligned with bureaucratic interests, framing photos as technical mediums for conveying information. Mass image reproduction further bolsters power control. Authorities leverage images for propaganda, advancing image technology and content to maximize their

influence. Concurrently, public awareness of power's manipulation of content and events transforms images into visual evidence, solidifying their role as information carriers.

William Flusser approaches this from a different angle, viewing photography as a precursor to the post-industrial era. It shifts from object-focused to information-focused, rendering ownership obsolete. Channels of circulation and media encode its deeper meaning [2]. Mass image replication profoundly disrupts human perception, erasing photo objectivity and redefining them as entertainment, social rituals, and tools of power. This dissolution of objectivity reshapes how we perceive photos, turning them into relational information situated between capture and observation, sender and receiver, individual and mechanism. This intermediary nature makes photos unstable, susceptible to manipulation, alteration, and exploitation.

2.2. Algorithmization of Images: Opening New Dimensions of Power Relations

The algorithmization of images opens new dimensions for discussing power relations. This is not just because of faster dissemination and more images, but more importantly qualitative changes in shooting, production, dissemination, viewing, and interpretation. Digitalization started algorithmic images. In the late 1980s and early 1990s, digital imaging gained popularity, enabling algorithmic applications for imaging and storage. Digitalization, networking and social media changed the order of images and impacted power relations. Lower technical barriers enabled more participation in image production, communication, viewing and consumption. Dissemination evolution meant authoritative centers were no longer needed to collect, review, edit and publish information. Power decentralized to individuals and media entered the "self-media" era. Algorithms control image meaning by making photographic devices "black boxes". The black box hides internal mechanisms, rendering them invisible and unnecessary to understand. For most, the black box interior is unknown; only results are seen. Algorithms encode image dissemination, making the process a black box too. Audiences are unaware of reasons for the seemingly

random emergence of images. Based on business, political and institutional needs, algorithms push specific images to audiences, disguising it as random. Rules and paradigms in the black box can be manipulated and controlled. Through guiding rules and paradigms, image production, viewing, dissemination and interpretation are all affected, while things outside are ignored and excluded, reinforcing their role. In algorithms, mastering rules and paradigms means mastering power, also becoming the main way for external powers to infiltrate and manipulate.

2.3. The Collective Turn and Linguistic Turn in Photography

In the realm of photography, theorists are delving into the transformative effects brought about by its integration into the algorithmic age. A prominent example is Daniel Palmer's analysis of the platform ThisMoment, showcasing the concept of "super photography." [3] This involves a shift from traditional photography, where images encapsulated singular moments, to a more intricate process shaped by personal and collective memory dynamics. Within this context, users can blend their photographic recollections with snippets from mass media, reflecting shifts in identity and cultural concerns. The act of photography is no longer confined to individual perception or the camera's "black box"; instead, it emerges from engagement with a communal world of creation and experience. Photography in the algorithmic age is fundamentally a collaborative endeavor, wherein production, dissemination, and consumption are interdependent [4]. Simultaneously, the author's identity takes a backseat, transforming photography into an all-encompassing symbol of meaning.

Another pair of theorists, Daniel Rubinstein and Katrina Sluis, introduce the concept of the "undecidability of photography," releasing it from its traditional indexical ties to reality. This notion challenges the assumed cause-and-effect relationship between a photograph and the objective world, granting images multiple ambiguous meanings within diverse contexts and non-linear timelines [5]. This metamorphoses the photograph into an elusive and open entity. Although algorithmic images may not fit into conventional symbolic and iconic frameworks, Rubinstein and Sluis assert that they still function as a form of language. Hence, photography is not moving away from language but rather undergoing a profound linguistic transformation. This transition suggests that we are experiencing an "image shift," moving from language usage to a more image-centric communication form. This shift is grounded in the algorithms themselves, turning images into a new kind of language [6].

This evolving relationship holds substantial implications for power structures. The proliferation of algorithmic images, coupled with their indeterminacy, contributes to a superficial and momentary character, particularly evident on popular social media platforms. These images tend to evoke instant emotions and offer fictional styles, often preventing a deeper analysis. Conversely, the uncertainty embedded in algorithmic images results in a lack of fixed meaning and a dearth of shared interpretations. This renders images agents of emotion, impulse, symbols, and even language. This superficiality, immediacy, and emotional nature render images susceptible to confusion and addictive consumption. These traits also conceal the underlying linguistic transformation. Notably, the algorithmic mechanisms play a pivotal role in power dynamics. Through mechanisms like

labeling, descriptive text, and content analysis, algorithms amass information for commercial, political, and public opinion objectives. They then adjust image creation and distribution strategies based on this data. The combination of surface-level imagery, concealed linguistic shifts, and algorithmic control leads viewers to be influenced by these intentional images, often without being aware of it. This unequal relationship between algorithms and viewers impacts agency, awareness, and resistance.

3. The Emergence of AIGC: A New Turning Point in the Visual Realm

In early 2023, along with the public opening of ChatGPT, the discussion about AIGC became increasingly intense. ChatGPT entered the public vision in the form of question-and-answer text, allowing people to encounter content in this field for the first time in a natural language way. With the wave of AIGC led by ChatGPT, there have also been a series of AIGC tools in the field of visual images. Among them, Midjourney, Dall-E, Stable Diffusion and Firefly are the most widely known and used by people at present, and new tools continue to emerge. The core functionality of these AIGC tools is text-to-image and image-to-image. Text-to-image generates images from text, where users enter a series of keywords (prompts) in the prompt box, and the algorithm generates images related to the keywords based on large language models. Image-to-image allows users to upload an image, and the algorithm first analyzes the image into keywords, and then generates corresponding images. The wave of AIGC marks the entry of the algorithm age into a new stage.

At the same time, AIGC also marks the completion of the algorithm's linguistic and collective shift of images. Although Midjourney is an image generation tool, its basic architecture is still language-centric. In the previous algorithm age, people gave images labels on the Internet. Therefore, images are not connected visually to form an image network, but are connected through language texts such as tags, titles, and descriptions, of which tags are the most important. Putting tags together with keywords in text generation, one can see the similarities between the two, such as brevity, discreteness, generalization, naming, etc. At the same time, one can also see the correspondence between the two. This correspondence constitutes indexing of images and data. Therefore, the images generated from this still have a linguistic framework at the bottom level. In Midjourney, the use of keywords is highly valued. The way images are modified is also focused on modifying keywords, which is more like a literary writing process than an image production process. In this logic, the visual is no longer a process, but a result (or multiple results), and the only action left is to choose.

4. AIGC - Continuity and Change of Algorithmic Image Power

4.1. AIGC Perpetuates the Attributes of Algorithmic Images

On April 19, 2023, artist Boris Eldagsen used AIGC-generated images to win the Sony World Photography Award, but he refused the award, sparking huge controversy. This event involves an important question: should the winning images be regarded as the artist's individual works, or AI's works? At the same time, how should the copyright

ownership of AIGC-generated content be defined? These issues still do not have fair answers to this day. However, these issues reflect a deeper problem, namely the gradual shift towards collectivism in image creation. Prior to the algorithm age, image creation has long been closely related to individual creators. However, the development of the Internet and social media has gradually weakened this personal binding and moved towards collectivism. AIGC demonstrates this more directly. People retrieve images through keywords, and these labels contain countless related images, which together constitute the database for generating content and provide materials for model training. But the authors of each image are obscured, the images are separated from the identities of individuals and authors, and become purely informational, evolving in algorithms.

At the same time, AIGC users are no longer facing a specific image creator, but a complex and extensive subject composed of big data and large language models. The identity of this subject is no longer a person, but the algorithm itself. For users, there must be personal intentions, interests, aesthetics and decisions when using AIGC to generate images, but at the same time, there must also be a collective image from the system. In the dialogue between the two, it can even be said that the collective image occupies an absolutely dominant position, and the individual user can only guide its originally existing possibilities without really transcending its boundaries.

Since AIGC has not been widely available to the public for a long time and has not been widely popularized, it is still impossible to fully determine what changes will take place in the image under the AIGC environment. But it can be seen that AI image generation represented by Midjourney basically continues the changes that occurred after images entered the algorithm age. This is also the reason why incorporating AI image generation into the discussion of the image system. On the one hand, in terms of content, AIGC continues the same context as the digitization of images. In addition to a large number of photos and videos, content such as paintings and film photography that rely on physical media also need to be digitized through recapturing, scanning, etc. in order to enter the algorithmic domain. Even digital drawing and other purely digital workflows that emerged relatively late also embed the logic and methods of collection and digital imaging. That is to say, the visual model framework of AIGC is based on algorithmic images as materials and the linguistic information of algorithmic images as indexes.

On the other hand, from the perspective of viewing and interpretation, AIGC inherits the black box attribute of image tools, and further amplifies this feature. When facing AI-generated images, the image behind is an even larger, unknown system. It just discards the physical reliance of image tools and turns to pure algorithms. At the same time, it no longer limits to optimizing algorithms for a single photo, an object, or a referential relationship between a photo and an object, but arranges combinations of massive materials, and further advances the language of algorithmic images.

In addition, compared with algorithmic images, AIGC has more pronounced characteristics of superficiality and indeterminacy. For viewers, the image generation process is to input keywords, wait for the algorithm to run, and get 1-4 images. This is exactly the same as the psychological process of using a camera. The creator can only choose image results and image objects, except that the objects have changed from objective entities in the real world to strings of characters.

4.2. AIGC Realizes the Dispersion of Image Power

Although AIGC has not been available to the public for a long time, it has caused tremendous reaction and discussion. Compared with previous technological revolutions, this technological innovation is directly relevant to most people and will change the structure of image discourse rights. In Midjourney, the user's basic operation is to enter keywords and wait for the algorithm to generate result images. Other usage functions are also mostly based on this. In addition to the quality of the generated images, the input, waiting, and selection have extremely low learning costs for most people. As long as you can use language, you can turn it into images. This process continues the development trend of imaging technology, which is to lower technical barriers, expand audience reach, and expand business benefits. Since the invention of film, every major change in imaging technology has followed this logic. Low learning costs give users the opportunity to participate in visual expression. From the invention of images to the emergence of AIGC, the public has obtained the right to visual expression and the right to receive expression. Now, the masses can take pictures with their phones, turn photos into videos, and upload and share them on social platforms. The emergence of AIGC adds time and fiction dimensions, which can turn non-real-time and non-physical content (such as emotions and imagination) into images, further expanding the way and content of visual expression.

On the other hand, with the reduced learning costs of visual expression and the increased degree and scope of participation, more professional fields can be integrated. For example, in popular science, literature and business, text information can be transformed into visual content more conveniently and effectively. Midjourney generated images can also be used as materials for further application in design, film and television and other fields. The integration of multimedia will provide new soil for innovation in content and forms of expression, and more content will be expressed with the help of AIGC. The collision of different media will also produce new forms of expression adapted to this era.

In addition, AIGC basically continues the decentralization trend. Usually, the centralization of discourse power requires the establishment of a series of complex value judgment standards to achieve, and to monopolize discourse power and interpretation power through this. However, in Midjourney, the image production process is extremely simple, and contains uncertainty and unpredictability. The space left here is less to establish value judgment standards, and it is difficult to judge high and low levels between two normal AI generated images. This means that discourse power and interpretive power are handed over to general users.

The easy-to-use techniques of AIGC give more people the right to visual expression, but at the same time also limit the experimental nature and creativity of AI-generated images, leaving users no chance to develop personal visual styles. On the other hand, due to the lack of value judgment standards, production and dissemination may become chaotic, and it is difficult to distinguish between quality and inferior content, and quality content is also difficult to obtain encouragement and promotion. This may lead to the formation of standards within different cultural groups, deepening the barriers between circles.

With the explosive growth in the number of AIGC users, the image content they create has also exploded, greatly expanding the original visual information in a new direction. However, at the same time, regulatory and censorship systems and laws have not kept up in a timely manner, leaving AIGC content creation and output in a relatively free environment for the time being. The lack of regulatory and censorship systems although to some extent realizes the right of free speech, but excessive freedom will also lead to many problems, amplifying the impulses in human nature. Currently, content related to sex and violence accounts for a relatively high proportion in AIGC content and is highly concerned. Although there is no evidence so far that these contents have had a significant impact on the real society, referring to similar issues that have arisen with previous emerging media, this point deserves continuous attention and importance. At the same time, AIGC also lacks clear provisions on copyright ownership, involving related issues such as plagiarism judgment and moral responsibility that also lack a clear consensus.

4.3. AIGC Leads to New Concentration of Power

While decentralizing power, AIGC is also in the process of centralizing power, continuing the basic logic of algorithmic image processing in the algorithm era. Although the technology provides convenience and rights on the surface, the algorithmic mechanism is more covert and deceptive in gaining power.

Aesthetic centralization is the most direct manifestation. When using AIGC to generate images, users cannot fully control the whole process. Although some AIGCs, such as Stable Diffusion, allow training of custom models, the underlying algorithms are largely the same. Especially on platforms like Midjourney, although users can choose a certain amount of visual styles, the choices are limited and cannot really achieve personalized creative styles. These style choices are like filters in algorithmic images, providing convenient stylistic needs for most users, while also limiting the diversity and individuality of creative styles.

In order to achieve stylistic effects, AIGC needs to rely on stereotypes. In the original language corpus or image library of AI, there are already a large number of stereotypes. In order to allow viewers to quickly recognize and understand images, AI takes the initiative to extract and integrate these stereotypes, including aesthetic stereotypes and stereotypes of the real world, such as associating crime with blacks and associating Asian women with oriental mysticism. Although some problems have been improved to a certain extent through human intervention, the underlying logic of AI's use of stereotypes remains unchanged. The aesthetic styles and content styles of AI-generated images woven from stereotypes, on the one hand, cater to existing widespread cognition, and on the other hand, reinforce viewers' recognition of their own stereotypes in the images, thus further strengthening the original stereotypes. These stereotypes have been used many times in history as important tools for power manipulation of public opinion and shaping public consensus and attitudes, such as vilification of women, Jews, and blacks. Therefore, AI's use and enhancement of stereotypes deserves high attention and intervention to prevent inappropriate stereotypes from being enhanced and widely disseminated.

In addition, it can be said that AIGC advocates an averaging doctrine, averaging in two directions. First, AIGC integrates the image materials in the database into a unified, self-consistent visual language that conforms to algorithmic logic, and then provides users with creative services through this language. However, excellent human artworks often contain accidents, errors and illogical parts, and these are precisely the places that are more thought-provoking and worth exploring, and are also important ways to expand people's boundaries of visual cognition and aesthetic standards. In contrast, AI's averaging doctrine unifies all accidents, errors and illogicality to the same standard line. Even if errors occur, they are errors of the algorithm rather than intentional ones. This dooms AI-generated images to be aesthetically average and mediocre from the outset. This in turn requires users to accept such homogenized aesthetics, and even cater to such aesthetics to a certain extent.

Secondly, the degree of freedom of user operations and participation is restricted, so that AIGC cannot be regarded as a flexible tool. Users may get images that do not match expectations, and it is difficult to get images that match expectations, let alone realize the images in their minds in an AI way. In this case, AI already has a certain degree of subjectivity. The relationship between human and AI is no longer a simple relationship between use and tool, but a relationship between two subjects. This involves communication, division of labor and mutual understanding. Although the cooperative relationship saves users' mental labor in creation, it also deprives users of the right to freely express themselves to some extent, which seems to be an inevitable direction of development.

On the other hand, Midjourney also has noteworthy features in terms of centralization. Its interface design and workflow design have a certain addictiveness. As mentioned earlier, AI-generated images often defy user expectations, and each mismatch will encourage users to continue investing time and energy to modify keywords or look for reference images in order to obtain results that better match expectations. This mismatch of expectations and the continuous creation process provide users with a sense of stimulation, but this contingency is a response to the user's intentions rather than an accident of the algorithm. Therefore, under the dual pressure of chasing expectations and obtaining stimulation, users will continue the behavior of AI image generation, similar to unpacking blind boxes, which makes users engage in a communication process with AI. On the one hand, AI may get closer to user expectations with correct guidance, and more importantly, after experiencing multiple mismatches of expectations and stimulations provided by AI, users will conduct a kind of psychological compensation, gradually accepting the influence of AI stimulations, and thus modifying their own expectations and moving closer to the direction of AI stimulations.

Another Midjourney feature also enhances this addictiveness. After generating images with keywords, users can get 4 images, each of which can be used as the basis for the next variation creation, allowing users to continuously make such selections and differentiations. However, such operations will cause users to fall into following the logic given by Midjourney, and as the number of repetitions increases, the degree of reinforcement of AI styles will also increase. Users may gradually lose their original intentions, and can only move forward along the path set by AI, gradually forgetting or compromising their original ideas. The

mismatch of expectations and frequent modifications, as well as the logic path constraints, provide conditions for the addictiveness of AI image generation. And addictiveness in turn lays the foundation for the emergence of problems such as attention, harvesting, and commercialization. In the algorithm era, attention is crucial to commercialization, and the addictiveness of AI means it can attract user attention, creating opportunities for commercialization.

As the industrialization of AI continues to upgrade, the alienation of workers will also intensify. At present, the use of artificial intelligence image generation technology is still at the exploratory stage, and users mostly use it out of their own free will. However, as the in-depth development of industrialization, some visual content production companies may start layoffs and use AI to replace manpower, while also recruiting AI operation positions. This will lead to AI operators no longer being able to truly represent personal wishes, but only become "parts" in the content production pipeline, becoming part of the AI mega machine. On the other hand, the industrialization of AIGC leads to an increase in artificial intelligence content products, causing consumers to consume AI content products knowingly or unknowingly, thus pulling consumers into the industrial chain loop. In this way, the creation end and the consumption end of the content both become affiliated to the industrialization of AI. It can be said that the industrialization of AI will expand the alienation of workers to the field of mental labor and even creative fields, further exploiting the rights of workers.

5. Conclusion

AIGC technologies such as Midjourney's image content generation technology inherit and expand the power strategies of images in the algorithm era, affecting the ways of information dissemination, public opinion shaping, and visual expression. Such changes in power structures have caused widespread concern. However, as AIGC technologies continue to evolve, we face both great opportunities and important challenges. In terms of practical applications, clear ethical guidelines and industry standards need to be

established to ensure that the generated images are not abused, mislead others or infringe on the rights and interests of others. At the information dissemination level, algorithmically generated images should be identified and labeled so that users can understand their sources and authenticity, thus enhancing the transparency of information. At the legal level, we need to carefully consider how to adapt to the emergence of AIGC technologies, update copyright laws, privacy protection laws and other laws to adapt to emerging rights disputes. At the same time, regulatory agencies also need to closely monitor the development of AIGC technologies to ensure compliance and protect the public interest. At the moral level, we need to encourage researchers and developers to follow ethical principles, pay attention to the social impact of technology, and ensure positive use of technology. The public also needs to strengthen cognition of AIGC technology, understand its potential risks and benefits, and thus participate in and use these technologies more rationally. In summary, we must work together to intervene and adjust the application of AIGC technologies in society in a timely manner, to ensure that these technologies have a positive and sustainable impact on social development.

References

- [1] Sontag, S. (2019) *On Photography*. London: Penguin books.
- [2] Flusser, V. (2018) *Towards a philosophy of photography*. London: Reaktion Books.
- [3] Palmer, D. (2010). Emotional archives: Online photo sharing and the cultivation of the self. *Photographies*, 3(2), 155-171.
- [4] Palmer, D. (2013). A collaborative turn in contemporary photography?. *Photographies*, 6(1), 117-125.
- [5] Rubinstein, D., & Sluis, K. (2013). Concerning the undecidability of the digital image. *Photographies*, 6(1), 151-158.
- [6] Rubinstein, D., & Sluis, K. (2013). The digital image in photographic culture. *The photographic image in digital culture*, 22-40.