

**Nikoleta Kerinska**

## **The Destiny of the Technical Image: A Classification Proposal**

The apparent objectivity of technical images is illusory, because they are as symbolic as all other images. They must be deciphered by anyone who wishes to grasp their meaning. Indeed, they are extremely abstract symbols: they encode texts into images; they are metacodes of texts. Imagination, from which they originate, is the capacity to encode texts into images. To decipher them is to reconstruct the texts that such images signify. When technical images are correctly deciphered, a conceptual world emerges as their universe of meaning. What we see when contemplating technical images is not “the world,” but certain concepts about the world, despite the apparent automaticity with which the world impresses itself onto the surface of the image.<sup>1</sup>

Vilém Flusser

### **1. Introduction**

The 1970s and 1980s were marked by an unprecedented expansion of audiovisual possibilities in Western societies, driven primarily by cable television, which allowed for far more refined audience segmentation and opened space for specialised channels. One of the most emblematic cases was the emergence of MTV in 1981, whose continuous programming of music videos transformed the relationship between music, image, and cultural consumption, making it a central vector of advertising and lifestyle direction for young audiences. In parallel, technologies such as teletext, the Walkman, fax machines, and answering machines became part of everyday life, broadening the circulation of content and the portability of the audiovisual experience, within a context marked by profound social and economic transformations – including the advance of neoliberalism.

The circulation of electronic images accelerated thanks to the consolidation of televisual infrastructures, accompanied by increased capacity for content archiving and distribution. Television established itself as the primary medium of audiovisual diffusion, while the growth of catalogues and databases began to transform the way audiovisual content was organised and accessed, anticipating practices of algorithmic selection that would only become dominant decades later. The 1980s also witnessed an intensification of the consumption of images as a form of cultural identity, with music videos, targeted advertising, and media products circulating on an international scale.

---

<sup>1</sup> Flusser, V. (2002). *A filosofia da caixa preta: ensaios para uma futura filosofia da fotografia*. Rio de Janeiro: Ed. Relume Dumará, Coleção Conexões, p. 14-15. In the original: “A aparente objetividade das imagens técnicas é ilusória, pois na realidade são tão simbólicas quanto são todas as imagens. Devem ser decifradas por quem deseja captar-lhes o significado. Com efeito, são elas símbolos extremamente abstratos: codificam textos em imagens, são metacódigos de textos. A imaginação, à qual devem sua origem, é capacidade de codificar textos em imagens. Decifrá-las é reconstituiu os textos que tais imagens significam. Quando as imagens técnicas são corretamente decifradas, surge um mundo conceitual como sendo o seu universo de significado. O que vemos ao contemplar as imagens técnicas não é “o mundo”, mas determinados conceitos relativos ao mundo, a despeito da automaticidade da impressão do mundo sobre a superfície da imagem.”

A new regime of visual communication thus emerged, in which the technical image came to structure behaviours, desires, and modes of cultural participation — a direct harbinger of the hyperconnected audiovisual culture that would follow.

During this period, Vilém Flusser was already questioning the functioning and circulation of images in our society. For this philosopher, defining and understanding the technical image — and the emergence of a “techno-imagination” — was an essential step towards understanding our own civilisation: “If we take into account the incredible variety of multicoloured surfaces that have surrounded us since the Second World War, any attempt to find a definition capable of encompassing such diversity seems doomed to failure. At first glance, there appears to be nothing in common between the innumerable coloured symbols that demand our attention on streets, shop windows, the walls of houses and public squares, in books and on television, on T-shirts and bottle labels, on tools and gadgets.”<sup>2</sup>

For Flusser, this ‘visual jungle’ that characterises the present is not an accidental phenomenon, but a profound expression of the *Zeitgeist*<sup>3</sup> — that is, of the way in which Western culture has come to symbolically organise the world through codified surfaces. The philosopher identified in these manifestations an epistemological pattern which, despite its apparent visual heterogeneity, constitutes what he describes as “the first true style of the West since the Gothic period”<sup>4</sup>. This is a ‘style’ defined not by formal coherence, but by the common logic determined by apparatuses, which generate this coded world. The analytical force of this observation lies precisely in the perception that the unity of our time is located not in form, but in the mode of codification — in the programmatic structure that underlies technical images, rendering them simultaneously multiple in appearance and homogeneous in operation.

Yet the fundamental reason for the importance of these images, for Flusser, was their communicative power: “(...) we discover, know, and evaluate the world in which we live through messages received from surfaces, and this is why we are now submerged in a sea of colours and gadgets. Because we are no longer primarily programmed by black-and-white printed texts, but by technical images in techno-colours.”<sup>5</sup>

---

<sup>2</sup> Flusser Vilém, (2022) *Mutations dans les Relations Humaines : De la Communicologie*, Ed. HD/Le Bon Voisin / AKA, p. 253. In the original: « Si nous tenons compte, toutefois, de l’incroyable variété de surfaces multicolores qui n’ont cessées de nous environner depuis la seconde guerre mondiale, toute tentative de trouver une définition susceptible d’englober tant de diversité paraît vouée à l’échec. A première vue, il semble qu’il n’y ait rien de commun entre les innombrables symboles colorés qui s’imposent à notre attention dans les rues, les vitrines, sur les murs des maisons et des places publiques, dans les livres et sur le petit écran, sur les tee-shirts et les étiquettes des bouteilles, sur les outils et les gadgets. »

<sup>3</sup> *Ibid.*, p. 253. In the original: « l’esprit du temps ».

<sup>4</sup> *Ibid.*, p. 254. In the original: « On peut dire que ce premier véritable style occidental depuis le gothique est ce qu’il y a de commun à tous les éléments hétérogènes dans notre monde codé ».

<sup>5</sup> *Ibid.*, p. 256. In the original: « (...) nous découvrons, connaissons et évaluons le monde dans lequel nous vivons à travers des messages reçus des surfaces, et que c’est là la raison pour laquelle nous sommes à présent submergés par

One has the impression that today the philosopher's voice resonates with even greater clarity, for we are more than ever surrounded by 'multicoloured surfaces' emerging from all manner of screens and gadgets. Consequently, we are programmed by them – by these technical images that shape perceptions, affects, and modes of being in the world, and which since the mid-twentieth century have never ceased to evolve from the standpoint of their generative apparatuses. Transposed to the contemporary context, Flusser's words reveal a disquieting pertinence, and this observation is the starting point of my reflection, which revisits the concept of the technical image to examine its most recent manifestations.

The following text adopts the perspective of an artist, researcher and lecturer in computer art, rather than that of a philosophical reflection in the strict sense. My interest focuses, in particular, on images produced by generative artificial intelligence, which are increasingly present on digital platforms, social media, search engines, creative work environments, advertising, cultural consumption interfaces and, increasingly, in all spaces where visibility has become a dominant vector of communication. This is, therefore, an attempt to understand the ways of life of these images, to try to penetrate their nature and origin, which remain opaque, almost magical, to the naked eye.

This reflection proposes to examine images produced by artificial intelligences through the lens of their specific method of generation, seeking in Flusserian thought anticipations and prophetic echoes that allow for a critical engagement with these emerging technologies. The aim is to understand how these images constitute *a new generation of technical images*, whose most distinctive structural characteristic is the *logic of sampling*, specifically the *statistical recombination of vast sets of visual data*, which replaces direct visual references to the world with a process of probabilistic synthesis.

Through comparative analysis, I will attempt to delineate the principal characteristics of AI-generated images and, by proposing a method of classification within the broad family of technical images, to understand how they situate themselves within this universe. The term 'classification' seems the most appropriate for this study, as it refers to a set of images sharing well-defined functional characteristics, generally based on technical, perceptual, or operational criteria. From the outset, it is important to mark a conscious choice, a positioning: photographic images, computer-generated images, and generative images may be considered distinct classes of the technical image, since each of these groups corresponds to a specific mode of production and a particular technical regime.

---

une mer de couleurs et de gadgets. Parce que nous ne sommes plus prioritairement programmés par des textes imprimés en noir et blanc, mais par des techno-images en technicolor. »

Following this line of reasoning, I will first mobilise the Flusserian definition of the technical image; I will then address the computer-generated image<sup>6</sup>; and subsequently compare its specificities with those of generative images, to better elucidate the differences between these two classes. In conclusion, I will open some avenues for future reflection.

## 2. The Starting Point: The Flusserian Definition of the ‘Technical Image’

Vilém Flusser defines the technical image as “an image produced by apparatus”<sup>7</sup>. Such apparatus, in turn, operate as a “black box”<sup>8</sup>, that is, a device that receives input to generate the image, and whose main characteristic lies precisely in the fact that the process of image generation takes place inside this black box, which is opaque to the user. The consequence of this fact is significant: “Apparatus are products of technology, which, in turn, is applied scientific text. Technical images are, therefore, indirect products of texts – which gives them a historical and ontological position distinct from that of traditional images.”<sup>9</sup> In this sense, technical images do not extend the human hand, but rather the very scientific rationality that structures them. They arise from devices programmed by theories, and it is this mediation that redefines the act of seeing and of producing visibility. The recognition of the radically distinct ontological and genealogical status of technical images in relation to traditional images has occupied intellectuals, artists, researchers and thinkers who have questioned their expressive, ideological and artistic power since the advent of photography.

The invention of photography inaugurates the era of technical images, marking a decisive rupture in the universe of the image: for the first time, reality seems to inscribe itself, dispensing with the artist’s skill. This technical realism is also a moment of liberation for the visual arts, which are freed from the burden of faithful representation and open themselves to other forms of sensibility – more subjective, symbolic, or experimental. At the same time, the image is democratised, circulates, multiplies, infiltrates everyday life and scientific, journalistic, and police practices, acquiring the ambiguous status of proof, testimony, and condensed narrative of an instant.

Photography also profoundly alters our relationship with time, with the gaze, and with the very act of knowing. It captures the ephemeral, fragments the continuous, prepares the ground for

---

<sup>6</sup> A computer-generated image (CGI) is a picture or visual element created entirely with digital software rather than captured by a camera.

<sup>7</sup> Flusser, Vilém (2002), *Op. Cit.* p. 13.

<sup>8</sup> *Ibid.*, p. 15.

<sup>9</sup> *Ibid.*, p. 13.

modern sensibility and for the proliferation of the great variety of technical images. It finds a genealogy of the automatic image – subsequently cinematographic, televisual, and videographic.

The transition from the line to the pixel and then to the bit, as analysed by Flusser in *Philosophy of the Black Box*, describes an escalation of abstraction in the history of technical images. The line represents the original human gesture, the manual inscription that anchors the image in bodily experience and subjective intention. With the advent of devices, this gesture is replaced by the pixel, the smallest unit of the electronic image, which no longer derives from the hand, but from calculations that fragment the continuum into discrete points. The bit, in turn, takes this abstraction to the limit: it dissolves the pixel into pure information, reducing the image to binary operations that can be recombined indefinitely by programmes<sup>10</sup>. Thus, the technical image ceases to be an inscribed surface and becomes an informational event, whose ‘materiality’ depends entirely on computational processes that remain invisible to the observer.

It is precisely at this juncture that the Flusserian definition of technical images gains importance: it allows us to understand photography not merely as a new medium, but as the beginning of a programmed imagetic regime, in which the apparatus comes to mediate, condition, and orient the gesture of creating. Flusser shifts the discussion from representation to operability, showing that these images do not result from a direct expressive intention, but from a game between the human and the programme inscribed in the device.

Flusser’s definition allows us, on the one hand, to consider all images generated by devices or machines as belonging to the same family and, on the other, encourages us, with each new device, to revisit the question: ‘What kind of image is this, and what is its relationship with the device that produced it?’. For example, while analogue photography records light as a continuous inscription on a physical medium – a chemical trace sensitive to the world – digital photography converts that same light into pixels and, ultimately, into bits, transforming the image into calculable and programmable information. Following Flusser’s ideas and intuition, it is possible to answer this question only by deciphering the modes of operation of the devices and their methods of automation.

### **3. From the Computer-generated Image to the Statistical Synthesis of Information**

The term ‘computer-generated image’, or CGI, refers to images produced by code and mathematical models, which are considered in this text as the second class of technical images. CGI have been extensively analysed by authors such as Edmond Couchot, Philippe Quéau and Paul Virilio,

---

<sup>10</sup> Flusser, Vilém (2002), *Op. Cit.* p. 63-65.

among others, who highlight the shift from the photographic record to computer simulation. A brief analysis of CGI is necessary to understand the images generated by artificial intelligence, as well as to situate them as the third class of technical images. CGI chronologically precedes images generated by artificial intelligence, the areas of which can be traced back to the first experiments with generative neural networks in the 2010s, and which form the central focus of this article.

In the context of CGI, it is important to recognise that the development of programming languages and the first three-dimensional modelling software, established what was then a new regime of visual production based on mathematical formalisation and the algorithmic manipulation of form<sup>11</sup>. Languages such as C, Pascal and FORTRAN, initially designed for scientific and engineering purposes, became the structural framework for the creation of computer-generated images, whilst the first 3D modelling systems – initially restricted to laboratories, universities and pioneering studios – opened up the possibility of constructing entirely virtual environments. These devices not only expand the field of visual creation, but also establish the technical image as a calculated, navigable and transformable entity.

The universe of CGI underwent exponential expansion in the 1990s, accompanied by a palpable sense of euphoria among artists, thinkers and researchers who saw in this new visual paradigm the promise of a profound transformation in the field of imagery, as well as in the field of art. The CGI and its most complex manifestation – virtual worlds – were hailed as the art of the 21st century<sup>12</sup>.

Pierre Lévy envisaged virtual worlds as the concrete expression of cyberculture: “The canonical genre of cyberculture is the virtual world. We should not understand this term in the strict sense of a computer simulation of a three-dimensional universe explored using a stereoscopic headset and data gloves. Let us instead grasp the broader concept of a digital reservoir of sensory and informational virtualities that only come to life through interaction with human beings. Depending on the devices, this actualisation is more or less inventive and unpredictable, leaving a variable element to the initiatives of those who immerse themselves in it. Virtual worlds can eventually be enriched and explored collectively. In this case, they become a meeting place and a means of communication between their participants. The world engineer thus emerges as the great artist of the

---

<sup>11</sup> Computer graphics first appeared in the early 1950s, with experimental electronic images like Laposky’s Oscillons, and became a formal technological field in the 1960s with Sutherland’s Sketchpad.

<sup>12</sup> Some examples of such publications: *Cyberarts: Exploring Art & Technology* de Linda Jacobson (1992); *Esthétique des arts médiatiques* de Louise Poissant (1995); *The Art of Human-Computer Interface Design* de Brenda Laurel e S. Joy Mountford (1990); *Arte no Século XXI: A Humanização das Tecnologias* de Diana Domingues (1997).

21st century. He provides the virtualities, designs the spaces of communication, organises the collective apparatus of cognition and memory, and structures sensorimotor interaction with the universe of data.”<sup>13</sup>

For this philosopher, the art of the coming millennium is conceived collectively by multiple creators in shared digital environments, where the work is no longer an object and becomes a process, a space and an experience<sup>14</sup> – in other words, an immersive, connected, shared spatio-temporal image.

In this context, the artwork as a process is characterised by its evolutionary and open-ended nature, continually updated by the interactions of participants and by the internal transformations of the system, such that creation takes place in a state of constant flux, with no point of conclusion. As spaces, virtual worlds constitute navigable and habitable environments, symbolic territories that organise perception and action, establishing forms of presence which, though immaterial, structure modes of circulation, encounter and collective construction.

The experience of virtual worlds is realised through the user’s sensory and cognitive immersion, whereby the user not only observes but participates, intervenes and co-produces the meaning of the environment, making each experience unique and unrepeatable. Thus, for Lévy, virtual worlds constitute connected and shared artistic propositions, in which creation, interaction and experience intertwine, profoundly redefining the status of the work, the author and the public in digital culture.

With the artistic practices of the 1990s, there arose a need for a specific vocabulary to describe computer-generated images. Terms such as 3D images, synthetic images, infographics, computer graphics, virtual images and, finally, computer-generated images began to emerge in the search for appropriate terminology that better captured their nature. The term ‘computer-generated image’ is particularly revealing; it refers to images produced through the synthesis of abstract information, the construction of mathematical models and the algorithmic manipulation of form. This affirms the identity of a computationally conceived image, which derives from computational procedures and mathematical operations.

---

<sup>13</sup> Lévy P. (1999). *Cibercultura*, (Trad. de Carlos Irineu da Costa) São Paulo: Ed. 34, 1999, p. 144. In the original : O gênero canônico da cibercultura é o mundo virtual. Não devemos entender esse termo no sentido estrito da simulação computacional de um universo tridimensional explorado com um capacete estereoscópico e datagloves. Vamos antes apreender o conceito mais geral de uma reserva digital de virtualidades sensoriais e informacionais que só se atualizam na interação com os seres humanos. De acordo com os dispositivos, essa atualização é mais ou menos inventiva, imprevisível, e deixa uma parte variável para as iniciativas daqueles que nela mergulham. Os mundos virtuais podem eventualmente ser enriquecidos e percorridos coletivamente. Tornam-se, nesse caso, um lugar de encontro e um meio de comunicação entre seus participantes. O engenheiro de mundos surge, então, como o grande artista do século XXI. Ele provê as virtualidades, arquiteta os espaços de comunicação, organiza os equipamentos coletivos da cognição e da memória, estrutura a interação sensório-motora com o universo dos dados.

<sup>14</sup> *Ibid.* p. 135 - 145.

According to Edmond Couchot, the computer-generated image constitutes a unique category within the realm of technical images, precisely because it breaks away in an unprecedented way from visible reality and from pre-existing representational processes: “Generally speaking, all representational processes based on optics generate images that ‘adhere’ to reality, images in which every point is linked to reality through the projective logic of representation.”<sup>15</sup> In the analogue photographic image, each point corresponds to a point on the real object, defining a physical continuity, a direct connection between the world and its visual representation. In the computational realm, however, this correspondence disappears: “no point of any pre-existing real object corresponds to the pixel. The pixel is the visual expression, materialised on the screen, of a calculation performed by the computer, in accordance with the programme’s instructions.”<sup>16</sup> If anything pre-exists the computer-generated image, it is not the object, but the programming language and the calculation that engender it. It is, therefore, a distinct figurative logic, in which the image does not represent reality, but *simulates* it.

The concept of simulation is equally central to the work of Philippe Quéau<sup>17</sup>, who emphasises the decisive role of mathematical models in the generation of computer-generated images: “The computer-generated image necessarily refers back to its model. It can only be understood through its relationship with it.”<sup>18</sup> Here, the model is not an external reference point, but a formal structure that precedes and organises the image. The CGI is the visualisation of its own system of rules.

A profound epistemological rupture with the first technical images thus becomes apparent: the image becomes a product of abstract operations without any visual reference. It no longer bears witness, but projects; no longer records but reconstitutes. The consequence is a distinction in the mode of functioning between the first class of technical images and the second generation, in which the field of representation is replaced by modelling, inaugurating a regime of visibility in which the image is always the result of a computation.

The CGI also distinguishes itself by its capacity to incorporate any type of pre-existing image transcribed into binary code. The description in binary code of the CGI determines one of its most important characteristics: flexibility and openness to all forms of hybridism. Photographs, videos, drawings, three-dimensional models, or signals captured by sensors can be translated into

---

<sup>15</sup> Couchot, Ed. (1999). « Da representação à simulação » in *Imagem Máquina*, André Parente (org.), Rio de Janeiro: Éd. 34, p. 41. In the original : “(...) de maneira geral, todas as operações de figuração fundadas na ótica geram imagens que ‘colam’ ao real, imagens das quais cada ponto está ligado ao real, pela lógica projetiva da representação.”

<sup>16</sup> *Ibid.*, p. 42. Obs.: In this text, Edmond Couchot refers to images generated directly on a computer using specific software and programming languages. It appears that the author does not take digital photography into account.

<sup>17</sup> See Quéau, Ph. (1989) *Metaxu : théorie de l’art intermédiaire*, Ed. Camp Vallon et Éloge de la simulation : de la vie des langages à la synthèse des image (1986), Ed. Camp Vallon.

<sup>18</sup> Quéau, Ph. (1999) «O tempo do virtual» in *Imagem Máquina* André Parente (ord.), Éd. 34, Rio de Janeiro, p.92. In the origin: “A imagem de síntese remete necessariamente ao seu modelo. Ela não pode compreender se senão através da sua relação com ele.”

manipulable numerical structures, opening the field of the image to almost infinite transformations and multiplications. In this context, the image ceases to be a stable object or a surface and becomes a set of possible operations: it can be decomposed, recombined, simulated, animated, deformed, or expanded according to logics that no longer belong to the optical domain, but to the algorithmic. The CGI does not merely produce new images; it absorbs and transforms all others, establishing a regime of total *virtuality*<sup>19</sup>.

It is precisely in this terrain of computational simulation that the passage to generative images produced by artificial intelligences is inscribed. If, as Couchot and Quéau show, the CGI no longer represents the real but simulates it from mathematical models, AI-generated images displace this paradigm even further: they do not depend on an explicit geometric models, but on neural networks trained and modelled on large-scale data corpora, capable of inferring patterns and producing new variations. Simulation stops being the execution of a formal model and becomes a *statistical synthesis of information*, whose results are visual objects termed generative images. The expression ‘statistical synthesis of information’ accurately describes the functioning of generative models: they do not create images from nothing, nor do they retrieve pre-existing images, but synthesise unprecedented combinations from statistical patterns trained on large datasets.

#### 4. Images Generated by Artificial Intelligences

Images generated by artificial intelligence represent a relatively new field of production<sup>20</sup>, and research into them and their impact on the creative industries and society is still in its infancy. For Lev Manovich, these images mark a turning point in our history: “The new ‘generative media’ revolution brought about by AI appears to be as significant as any of the previous revolutions. In fact, it may well be as significant as the invention of photography in the 19th century, or the use of linear perspective in Western art in the 16th century.”<sup>21</sup> Before discussing the recurring nature of AI-generated images, I suggest a brief analysis of the operating principles of the programmes responsible for their existence.

It is important to note that AI-driven image generation methods constitute a diverse field, in which different forms of input – text, sketches, images, layouts or semantic structures – guide

---

<sup>19</sup> The term ‘virtuality’ is used to refer to future possibilities, denoting everything that has not yet come to pass but exists as a potential ready to emerge in different directions.

<sup>20</sup> We highlight significant software releases and the underlying algorithms that have shaped the landscape of generative AI from 2021 to 2024 by companies such as Google, Meta, Microsoft and OpenAI.

<sup>21</sup>

Manovich, L. (2023). “AI Image Media through the Lens of Art and Media History” In: IMAGE. Zeitschrift für interdisziplinäre Bildwissenschaft. Generative Imagery: Towards a ‘New Paradigm’ of Machine Learning-Based Image Production, Jg. 19, Nr. 1, S. 34–41. DOI: <https://doi.org/10.25969/mediarep/22323>

distinct modes of production. From *text-to-image*, which translates verbal descriptions into visual representations, to the broad range of *image-to-image* operations (such as *sketch-to-image*, *style transfer*, *inpainting* or *super-resolution*), each procedure reveals a specific relationship between code, model and input data. Added to these modalities are structural approaches, such as *layout-to-image*<sup>22</sup> and *scene-graph-to-image*<sup>23</sup>, which organise visual synthesis based on spatial maps or semantic relationships, as well as hybrid modalities that combine multiple vectors of orientation. However, although the methods mentioned operate through distinct pathways, they all align with Flusser's conception according to which "to function is to permute programmed symbols"<sup>24</sup>. This convergence highlights a continuous technical genealogy, in which the logic of programmed images described by Flusser is extended and transformed, now enhanced by the algorithmic plasticity of generative models, which expand and reconfigure the very regime of programming.

All the approaches mentioned deserve careful examination, but in this text my focus is on the *text-to-image* method. Analysing the *text-to-image* method is not an arbitrary choice, but a strategic one: it exemplifies the tensions inherent in the generation of images by AIs. It is in this procedure that the relationship between language and visibility becomes most explicit, revealing how models translate verbal descriptions into visual representations and, consequently, how their internal grammars, biases and semantic capabilities operate. Other modalities – *sketch-to-image*, *layout-to-image*, *style transfer* – are equally relevant but distribute operational functions between human and computer system in a segmented and contextually delimited manner. Consequently, *text-to-image* condenses the generative gesture in its most radical form, as it directly transforms a statement into an image, simulating a kind of *algorithmic imagination*.

Furthermore, *text-to-image* has become the dominant paradigm in recent research, both due to its cultural prevalence and the impact it has on discussions of authorship, creativity, veracity and aesthetics. It serves as a prime laboratory for observing how models articulate references, synthesise styles and produce visual coherence from abstract instructions. In other words, by studying *text-to-image*, we analyse the point at which generative imagery is most explicit in its mechanisms and most decisive in its implications.

To get to the heart of the issue of AI-generated images, I propose we examine how Midjourney works. This platform is regarded as one of the best for generating images and visual content. Its generative mechanisms are based on diffusion models: a technology that gradually

---

<sup>22</sup> It transforms structural sketches (boxes, positions, sizes and object labels) into a complete, realistic image. The model is given a layout – for example, "a person here, a tree there, a car in the background" – and generates an image that adheres to this spatial arrangement.

<sup>23</sup> It converts a semantic scene graph (objects and the relationships between them, such as 'cat on the table' or 'window behind the chair') into a coherent image. The focus is on the relationships and the semantic structure, not just on the positions. If you wish, I can produce more formal versions, more educational ones for students, or more concise ones for slides.

<sup>24</sup> Flusser, Vilém (2002), *Op. Cit.* p. 25.

transforms random noise into a coherent image guided by a descriptive text. Once the *prompt* has been written and submitted, Midjourney begins by converting each word into a mathematical representation that captures its meaning, style and possible visual associations. Starting from a completely chaotic ‘cloud’ of pixels, the AI applies a series of successive steps to remove the noise and bring forth shapes, textures and colours corresponding to the description. This process is iterative: the image gradually becomes clearer, as if it were materialising point by point, like a jigsaw puzzle in which each pixel finds its place after repeated attempts.

Another key aspect of how Midjourney works lies in its training on vast datasets of images and text. Thanks to this learning process, the model develops a statistical understanding of artistic styles, objects, lighting conditions, visual compositions and even aesthetic trends. It does not copy existing images but combines the patterns it has learnt to produce *something new*. It is this ability to fuse concepts, interpret styles and extrapolate plausible details that gives Midjourney its *creative* and often surprising character.

In the case of Midjourney, as with other platforms that use similar creation methods, the image-generation algorithms are primarily those based on deep neural networks. They operate on the principle of statistical selection drawn from vast visual datasets. Rather than representing a pre-existing object or simulating an explicit geometric model, these systems learn statistical regularities from image elements and then generate new visual configurations by recombining the initial distributions. Each image produced is therefore the result of a probabilistic calculation, in which every pixel or formal fragment is the result of an estimate, mean or variation within a learned latent space. From this perspective, it becomes pertinent to describe these productions as ‘statistical sampling images’: they do not derive from a real referent, nor from a stable formal model, but from a generation process based on the density of past occurrences. This definition highlights their probabilistic nature, their roots in data dynamics and their unique ontological status – at once synthetic and lacking a concrete attributable origin (save for an immense quantity of data).

At the beginning of his book *The Universe of Technical Images: In Praise of Superficiality*, Flusser puts forward a thesis: “New images do not occupy the same ontological level as traditional images, because they are phenomena without parallel in the past. Traditional images are surfaces abstracted from volumes, whilst technical images are surfaces constructed from points”<sup>25</sup>. By distinguishing technical images from traditional images in ontological terms, Flusser provides a decisive key to understanding generative images. If traditional images are ‘surfaces abstracted from volumes’, they still maintain a relationship of continuity with the sensible world, based on the projection of a

---

<sup>25</sup> Flusser, V. (2012) *O universo das imagens técnicas : elogio da superficialidade*, Coimbra : Ed. AnnaBlume, p.16. In the original: “as novas imagens não ocupam o mesmo nível ontológico das imagens tradicionais, porque são fenômenos sem paralelos no passado. As imagens tradicionais são superfícies abstraídas de volumes, enquanto as imagens técnicas são superfícies construídas com pontos.”

gesture, a viewpoint or a perceptual experience. Technical images, on the other hand, being ‘surfaces constructed from points’, usher in an entirely different regime: they do not result from the reduction of a pre-existing reality, but from the composition of discrete, calculable units, whose organisation stems from abstract programmes and operations.

In generative images, this condition is taken to the extreme, as the points no longer correspond to physical or optical features, but to probabilistic values compiled within a space of virtualities. The image is not extracted from the world, but synthesised from learned statistical distributions, emerging as a provisional actualisation of a multidimensional latent field. Thus, its ontology ceases to be founded on representation or resemblance and comes to operate according to a logic of sampling, variation and optimisation, in which each image is at once singular and statistically probable. In this sense, generative images confirm and radicalise the Flusserian thesis: not only do they belong to another ontological level, but they establish a visuality in which the image is not a vestige of the visible, but rather a momentary manifestation of an abstract logic – a programmed aesthetics of probability.

Another term that can be used to describe these images is suggested by Lev Manovich. According to the researcher, ‘prediction’ is in fact the most frequently used technical term in studies describing the methods of image-generation programmes. Thus, even though the word may sound metaphorical or evocative, it precisely describes the scientific process that takes place when we use these tools: “‘Prediction’ is the actual term often used by researchers in their publications describing visual generative media methods. So, while this term can be used figuratively and evocatively, this is also what actually happens scientifically when you use image generative tools. When working with a text-to-image model, the neural network attempts to predict the images that correspond best to your text input. I am certainly not suggesting that using all other already accepted terms such as ‘generative media’ is inappropriate. But if we want to better understand the difference between visual media synthesis methods and other representational methods developed in human history, employing the concept of ‘prediction’ and thus referring to these systems as ‘predictive media’ captures this difference well.”<sup>26</sup>

In this logic, the image is not a singular object in the classical sense, but the provisional realisation of a statistical possibility. Each image is simultaneously unprecedented and probable, singular and redundant, giving rise to an aesthetic of variation and proliferation, in which – according to Flusser’s definition – information is not what is repeated, but that which emerges as a probable situation within a programme. Far removed from an aesthetic of representation or simulation, such images belong to a generative aesthetic, whose paradoxical nature lies precisely in the tension between novelty and repetition.

---

<sup>26</sup> Manovich, 2023, *Op.Cit.*

Following Flusser's logic, if photography translates the world into an image via a machine, the *text-to-image* method translates text into an image via a machine as well, producing a radically technical image, as it does not derive from the concrete world, but from statistical models trained on other technical images. It is a visuality that emerges from learned probabilistic distributions, a level of abstraction even further removed from human agency. In this sense, the operation confirms Flusser's diagnosis that "the purpose of all abstraction is to distance oneself from the concrete in order to grasp it better"<sup>27</sup>, but also that such a movement is regressive – a "reculer pour mieux sauter"<sup>28</sup>. Culture, he says, dances around the concrete, but each new stage paradoxically makes a return to it more difficult. *Text-to-image* embodies this 'final stage' of which Flusser speaks: a concretisation of the absurdity of abstraction, in which the image no longer represents the world, but rather the calculation that replaces it<sup>29</sup>, and where the apparatus produces increasingly autonomous visual surfaces, whilst the human agent is limited to running the programme.

In other words, images resulting from the statistical synthesis of information—which emerge as probabilistic phenomena without a stable or formalised frame of reference—can be defined as images of statistical sampling. They do not convey a gaze or a gesture but are configured as 'projections of concepts' within Flusserian analysis. They emerge from a distributed learning process, in which form is obtained through operations of interpolation or optimisation within a multidimensional latent space<sup>30</sup>. One might consider this latent space as the extreme form of the black box, thus constituting the contemporary and expanded version of the Flusserian 'programme': no longer an explicit set of rules, but an opaque statistical field, inferred through training accessible only via inputs and outputs. The result of this operational logic is an aesthetic of variation, proliferation and ambiguity: each image is at once singular and statistically probable, unprecedented and saturated with reminiscences. Far removed from an aesthetic of representation or simulation, these images belong to a generative aesthetic, the nature of which is paradoxical.

The paradox stems from the fact that images generated by AIs are composed of statistically predominant elements and are therefore the result of mathematical inferences, deduced from logical conclusions based on premises, data and formal models. As such, for the system that generates them, they are entirely predictable, and any originality is inconceivable. For some reason that escapes us, they can be surprising to us. This paradox is also defined by the unprecedented mode of representation they establish, which is perhaps the main reason for the curiosity they arouse.

---

<sup>27</sup> Flusser (2012), *Op. Cit.* p. 20

<sup>28</sup> *Ibid.* p. 20.

<sup>29</sup> *Ibid.* p. 20-21.

<sup>30</sup> Isto significa que a imagem, ou o resultado gerado, surge porque o modelo percorre e combina coordenadas dentro de um espaço matemático muito amplo - o chamado espaço latente - onde cada dimensão representa alguma variação possível dos dados.

In this sense, generative AI takes the logic pioneered by computer graphics to the extreme: it does not simulate a model but statistically reconstructs possible visual models. The result is this unprecedented and recent (at present) regime of figuration in which hybridity intensifies, the relationship with the real is more abstract and indeterminate, perhaps even non-existent – a regime in which the image derives from a generative process that continually reconfigures its own visual field.

Images generated by AI are perfectly captured in Flusser's thinking. When he states that we are programmed by technical images, he precisely anticipates the shift we observe today: the image ceases to be an explicit construction and becomes the result of automatic processes that elude direct human intentionality. In this sense, images produced by generative models definitively bring to life Flusser's thesis that we live submerged in coded surfaces: whilst the computer-generated images still allowed one to trace the model, the algorithm, the geometry, the image generated by AI operates at a much higher level of opacity: it emerges from learned correlations, statistical sampling, and invisible, evolving patterns.

For Flusser, technical images are surfaces that conceal their internal processes, and this formulation is highly relevant today. The generative image is inaccessible to the person who creates it, as well as to the person who contemplates it – its message (or truth) dissolves amidst sampling, calculation and inference. Nor does it refer to a time-space, nor does it possess a time-space of its own. It brings about a shift in the field of figurative logics invented to date. This shift directly invokes Flusser's notion of techno-imagination, which he describes as a rupture between us and the world, mediated by devices that come to imagine on our behalf. The image created by I.A.s intensifies this rupture: it is not merely a matter of recompiling images, but of calculating and combining an unimaginable quantity of data, with the aim of making visible the mechanisms of 'imagination' within the black box.

The 'black box' argument, as formulated by Vilém Flusser, becomes newly intensified in the case of AI-generated images because the opacity at stake is not merely technical but epistemic and ontological. Early photography was opaque to most users - the chemical reactions in the darkroom were invisible, and the apparatus encoded physical and optical processes that few could fully explain - yet its opacity was *bounded*: although hidden, the causal chain between world, light, lens, and image was materially traceable. Similarly, synthetic 3D images produced through software rely on mathematical models that may be difficult for non-experts to read, but the parameters, algorithms, and rendering pipelines remain, in principle, inspectable and reversible. By contrast, generative AI systems introduce a qualitatively different form of opacity because their internal operations are not simply hidden but fundamentally uninterpretable. The weights of a deep neural network do not correspond to human – legible categories, and the training data – often proprietary,

but sometimes unknown and in their raw form – cannot be reconstructed from the output. This means that the causal chain between input prompt and generated image is not only concealed but *non-recoverable*, producing an image whose conditions of possibility cannot be traced back to a stable referent, a physical process, or an abstract model.

The correspondence with early photography therefore breaks down at the level of epistemic access: while both involve invisible processes, the invisibility of chemistry or optics was anchored in deterministic physical laws, whereas the invisibility of AI emerges from statistical abstraction, scale, and non-linearity, to which we can add a series of evolutionary calculations of extreme complexity and unpredictable outcomes.

What we cannot grasp in AI-generated images is not merely ‘how the machine works’ but *what the machine has seen, how it has encoded it, and why a particular visual configuration emerges* from billions of parameters. This produces a new kind of black box – one in which the opacity is not a temporary limitation of user knowledge, but a structural feature of the technology itself. Consequently, the opacity of generative AI is not just deeper but different in kind: it dislocates authorship, referentiality, and accountability in ways that neither early photography nor CGI ever did, because the generative apparatus cannot be fully audited, reconstructed, or even conceptually mapped onto human categories of intention, representation, or causality.

The issue ceases to be merely aesthetic or technical and becomes political, as Flusser had already suggested when he stated that the struggle of the future would be between programmers and the programmed. This political dimension acquires new urgency in the context of generative systems. For Flusser, apparatuses are not neutral tools but techno-cultural machines that shape the field of possible gestures, and AI platforms radicalize this asymmetry by concentrating power in the hands of those who design, train, and control the models. When users generate images without access to the training data, the internal logic, or the decision-making processes of the system, they are not co-authors but operators of a pre-structured apparatus whose rules remain inaccessible.

Today, to ask “What does it mean to generate images?” is to ask who controls the programmes, who sets the standards, and who guides the collective imagination in a world where visibility is increasingly automated. This produces a political economy of vision in which authorship becomes distributed upward - toward corporations, engineers, and opaque infrastructures – rather than outward toward users. The user’s prompt becomes a superficial interface, a pseudo-freedom masking the fact that the generative space is already delimited by decisions made elsewhere: what data was scraped, what biases were encoded, what aesthetic norms were statistically reinforced, what exclusions were silently reproduced.

In this sense, the ‘master’ of the creation is not the user but the apparatus itself, backed by the institutions that configure it. The political stakes are therefore not only about representation

but about governance: who defines the parameters of visual possibility, who controls the epistemic conditions under which images emerge, and who benefits from the opacity that shields these processes of statistical sampling methods. Unlike earlier photographic or synthetic technologies, where users could in principle understand or manipulate the underlying mechanisms, generative AI relocates agency into a domain that is almost structurally inaccessible<sup>31</sup>, transforming the user into what Flusser would call a “functionary” of the apparatus. The result is a new regime of image production in which power is exercised not through visible coercion but through invisible architectures of computation, making the political question of AI opacity central to contemporary debates about authorship, accountability, and the redistribution of creative agency.

Flusser’s words have a particularly strong impact in the context of images generated by AIs. Originally conceived as guidance tools, capable of helping to explore ideas, visualise concepts or support creation, these images often become screens that replace reality. In certain cases, their apparent visual fidelity can give the illusion of a coherent world, when in fact they are based on statistics rather than on experiences of the world. They do not represent reality: they recompose it, smooth it out, normalise it, until it becomes unrecognisable. Thus, what should be a tool for understanding becomes a mechanism that shapes our imagination and our relationship with the world, just as Flusser had written: “Images are mediators between man and the world. (...) Images are intended to represent the world. But in doing so, they come to stand between the world and man. Their purpose is to serve as maps of the world, but they end up becoming screens. Instead of using images to understand the world, man comes to live according to the images.”<sup>32</sup>

With generative images, the shift from map images to screens reaches an unprecedented level of intensity; more than ever before, they function as filters that reorganise our very access to reality. The Flusserian ‘screen’ ceases to be merely a metaphor and becomes a structural principle: the user does not see the world through the image but sees the world through an apparatus that has already filtered, selected and synthesised the visible even before any sensory experience occurs. AI images do not merely divert the gaze from the world; they reconstitute a world never seen or inhabited, thereby establishing a regime of visibility continually reconstructed according to statistical patterns, economic interests and invisible algorithmic decisions. In this sense, Flusser’s theory not only applies to generative AIs, but finds in them its most extreme realisation: the triumph of the screen over the map, of programming over experience, of the image over the world.

---

<sup>31</sup> We can admit that the artist or the user can develop his own algorithm and work with a personal database. In this case, he would have control over the creative process and would effectively become the programmer of his own images.

<sup>32</sup> Flusser (2002), *Op. Cit.*, p. 7. In the original : “As imagens são mediações entre o homem e o mundo. (...) Imagens tem proposito de representar o mundo. Mas ao fazê-lo, interpõem-se entre mundo e homem. Seu proposito é serem mapas do mundo, mas passam a ser biombos. O homem invés de se servir das imagens em função do mundo, passa a viver em função das imagens.”

## 5. Provisional Conclusion

To conclude, I return to Flusser's idea that 'all technical images are metacodes of texts'. In this way, each generation of technical images redefines its own grammar of mediation: photography encoded scientific texts on optics and chemistry; computer-generated images translated mathematical models and formal descriptions of space; however, AI-generated images introduce a decisive break, as they are the first whose generating text is not a scientific formula, but natural language itself. In this sense, they represent not merely a continuity, but the point of saturation of the Flusserian thesis: the moment when the image ceases to be derived from specialised knowledge and begins to emerge directly from human language. In this sense, they represent not merely a continuation, but the culmination of Flusser's thesis: the moment when the image ceases to be derived from specialised knowledge and comes to be produced by human language mediated by technology. It is likely that this relationship between the text - which sets the prompt without translating the imaginary into words - and its role in generating visual proposals is the most important point. In this scenario, the external world, which once served as a reference for the image, is replaced by a textual regime that organises and produces the visible. The image thus becomes the full realisation of the Flusserian programme: a metacode that no longer refers to the world, but to the linguistic universe itself that precedes it.

This transformation raises a range of questions that go beyond the technical and extend into the aesthetic, ethical and epistemological realms. Firstly, it is necessary to investigate how reciprocity and the exchange between language and the imagination are operationalised by prompts; how the relationship between text and image is reconfigured in the creative process; and how the disparity between the two symbolic systems influences visual creation. Secondly, it is crucial to distinguish human interpretation of the prompt – laden with intention, context and subjectivity – from computational reading, based on statistical correlations and formal rules, and to understand how this asymmetry affects processes of signification. Thirdly, we must examine the consequences of a scenario in which the artist comes to guide the machine through instructions, whilst the result tends to exceed or displace their own imagination, establishing a new economy of authorship and creative agency.

Ultimately, with generative imagery, Flusser's 'black box' becomes a veritable Pandora's box, unleashing forces beyond human control and giving rise to an irreversible irruption of the unexpected. We are facing a technological era whose effects have yet to be fully understood. The feeling that we have opened a new Pandora's box is also reinforced by the sense that we are confronted with technologies whose possibilities for image creation are inexhaustible, a fact that also provokes deep anxieties regarding authorship, veracity, power and control. As in the myth, it is not

merely a matter of unexpected forces escaping human dominion, but of recognising that every imagistic innovation brings with it both expansions of the visible and new zones of opacity. Technical images had already displaced the human into a programmed game; generative images radicalise this displacement, multiplying uncertainties and powers that escape direct supervision. And yet, there remains – like at the bottom of Pandora’s box – a glimmer of hope: the possibility of reinventing our relationship with generative models, of personalising them, of subjecting them to human purposes, and of re-inscribing imagination within a horizon where calculation and creation are not opposing forces, but productive tensions that are continually negotiated and generate critical experimentation.

## References

- Couchot, Ed. (1999) «Da representação à simulação» in *Imagem Maquina*, André Parente (org.), Rio de Janeiro : Éd. 34.
- Flusser, V. (2002) *A filosofia da caixa preta: ensaios para uma futura filosofia da fotografia*, Rio de Janeiro: Ed. Relume Dumará, Coleção Conexões.
- Flusser, V. (2012) *O Universo das imagens técnicas: elogio a superficialidade*, Coimbra: Anablume.
- Flusser, V. (2022). *Mutations dans les Relations Humaines : De la Communicologie*, Ed. HD/Le Bon Voisin/AKA.
- Lévy, P. (1999), *Cibercultura*, (Trad. de Carlos Irineu da Costa) São Paulo: Ed. 34, 1999.
- Manovich, L. (2023) : AI Image Media through the Lens of Art and Media History. In: *IMAGE. Zeitschrift für interdisziplinäre Bildwissenschaft. Generative Imagery: Towards a ‘New Paradigm’ of Machine Learning-Based Image Production*, Jg. 19 (2023), Nr. 1, S. 34–41. DOI: <https://doi.org/10.25969/mediarep/22323>.
- Quéau, Ph. (1999). « O tempo do virtual » in *Imagem Maquina*, André Parente (org.), Rio de Janeiro : Éd. 34.