



The Automation of Style: Seeing Photographically in Generative AI

Daniel Palmer, Katrina Sluis

► To cite this version:

Daniel Palmer, Katrina Sluis. The Automation of Style: Seeing Photographically in Generative AI. Media Theory, 2024, 8 (1), pp.159-184. <hal-04702682>

HAL Id: hal-04702682

<https://hal.science/hal-04702682v1>

Submitted on 19 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License

The Automation of Style: Seeing Photographically in Generative AI

DANIEL PALMER

RMIT University, AUSTRALIA

KATRINA SLUIS

The Australian National University, AUSTRALIA

Media Theory
Vol. 8 | No. 1 | 159-184
© The Author(s) 2024
CC-BY
<http://mediatheoryjournal.org/>

Abstract

In histories of photography, the notion of ‘seeing photographically’ is associated with so-called ‘straight’ photographers such as Edward Weston. It means to train the human-camera eye – to ‘previsualise’ – by conforming to the conditions of what Vilém Flusser calls the “photographic program”. At the same time, seeing photographically is a constant dynamic between a protean technical practice and aesthetic conventions or styles. This paper revisits the notion of seeing photographically to contextualise how notions of photographic style operate in generative imaging practices and AI discourse. We argue that since not only photorealism but the look of photographs from any period or style can be generated with the right combination of prompts and parameters, it has become possible to ‘previsualise’ and ‘see photographically’ by treating the history of photography as a style market. The concept of ‘style’, outdated in art history, becomes newly important and photography becomes a memory to be evoked in ahistorical stylistic evocations.

Keywords

photography, style, generative AI, text-to-image models, aesthetics

There's a new hot trend in AI: text-to-image generators. Feed these programs any text you like and they'll generate remarkably accurate pictures that match that description. They can match a range of styles, from oil paintings to CGI renders and even photographs (Vincent, 2022).

With the rise of AI image synthesis, the history of art and images more broadly are now in the service of statistical, prediction-based image generation. Given the importance of descriptive language for both computer scientists in image processing and for users who seek to conjure images with prompts, one outcome is that a certain kind of art history is being popularised and operationalised in generative models through the application of an old-fashioned idea. That idea is *style* – the “constant form” or “qualities shared” in the art of an individual or group, in art historian Meyer Schapiro's classic definition (1953: 287). But as technology reporter James Vincent writes in the breathless promotional quote above, text-to-image generators have the capacity to evoke a range of styles: “even photographs”. The medium of photography has become a style, and the term style has become an elastic one – a synonym for ‘the look of something’.

Style has in fact been important to the development of image processing research for some time. Consider the technique of *style transfer*, which takes two images – a content image and a reference image – and blends them together with the goal of preserving the *content* of the original image while applying the visual *style* of the other. In contrast, in text-to-image models, style is strictly speaking a kind of fiction: a desired ‘look’. This is made explicit by the person prompting these models, conjuring statistical predictions based on an interpolation of training data. In popular models, such as Midjourney and Stable Diffusion, users can currently generate images with variations of the phrase ‘in the style of’ to approximate historical or otherwise familiar imagery by well-known artists whose work forms part of that training data (one of the reasons for the controversy over the lack of permission, recognition and payment). In this essay, we are specifically interested in how style operates in the discourse of generative AI in relation to the history of photography and notions of photographic visuality. We ask: What does it mean that the ‘medium’ has become a style and the popular history of photography is now available as a set of imitable aesthetics in generative AI? How do

we make sense of photography's residue in the prompt economy, and what does this mean for 'seeing photographically'?



Figure 1: Screenshot of the website Artvy, an “art generation tool” which positions photographers and artists as “AI art styles”, <https://www.artvy.ai/ai-art-style/edward-weston>, January 2024.

Seeing photographically

Although the term ‘style’ occupies an ambiguous place in professional photographic discourse, discussions of style proliferate in amateur blogs and how-to guides that encourage an individual approach to subject matter. In an article titled ‘Finding Your Own Photographic Style, and Why It’s So Important’, the popular Fstoppers blog urges its readers to experiment to “find your unique style” on the basis that “First and foremost is your photographic eye” (Rackham, 2022). Under this paradigm, style and

eye are intimately connected. Perhaps this is not surprising: throughout twentieth-century modernist photography the development of a photographer's eye and way of seeing stands in for discussions of style. This idea reproduced itself in how photography was typically taught. It was common practice in photography education to ask first year students to go out and take photographs without any film loaded in their cameras. The rationale was that a student would get used to framing the world as photograph, tuning their eye to the camera lens. The camera sees the world differently to the human eye: for example, cameras can keep everything in focus with the right depth of field, both things in the distance and up close at the same time, unlike the unaided human eye. Similarly, while human eyes inevitably focus on something in the field of vision, the camera lens sees indiscriminately. This results in a contingency of unintentional detail initially seen as a disadvantage but later embraced by modernists and proponents of photography's 'reality effect'. Furthermore, the photographic image is two-dimensional, lacking the three-dimensional depth perceived by a human with binocular vision. Finally, a photograph is a slice of the world at a particular moment, as opposed to the continuous dynamism of vision offered by ordinary sight. To 'see photographically' therefore is to naturalise what the world looks like when viewed through a camera lens, or – as the photographer Garry Winogrand famously quipped – the practice of taking photographs is precisely “to see what the world looks like in photographs” (MoMA).

To see photographically in this sense is premised on standardisation. Camera technology has been constantly evolving since photography was conceived in the 1830s, but much of photographic technology is enduring – the design of lenses (their focal lengths and apertures) and the light-sensitivity rating of film (ISO, or its digital substrate counterparts). As Daniel Chávez Heras and Tobias Blanke put it, “the intermediate steps of interaction involved in producing photographic images are in fact heavily mediated by numerical parameters and standardised metrics” (2021: 1165). These pre-programmed “calculations” are “needed to render space visible” (Heras and Blanke, 2021: 1165) and this standardisation helps to account for why camera vision is at the heart of modernist 'straight' or 'pure' photography associated with early twentieth-century artists like Paul Strand and Edward Weston. While late nineteenth- and early twentieth-century Pictorialists had sought to distinguish their work from

both mechanical records and the emerging masses of amateur snap shooters by concentrating on soft focus and elaborate printing techniques, straight photographers embraced the camera's technical capacity for sharp focus and detail. Thus, for Weston:

the photographer's most important and likewise most difficult task [...] is learning to see photographically—that is, learning to see his subject matter in terms of the capacities of his tools and processes, so that he can instantaneously translate the elements and values in a scene before him into the photograph he wants to make (Weston, 1943).

This is the influential North American modernist ethos in photography, further embodied by the idea of 'previsualisation' enthusiastically promoted by Weston's West Coast colleague Ansel Adams. Adams published influential manuals for photographers that remained in print for five decades and established the famous Zone System for standardising the translation of light into greyscale values (Cubitt et al., 2013).

So-called 'straight' photographers did not usually understand their approach as a style.¹ What mattered was seeing clearly and 'honestly'; in other words, a kind of rugged individualism of photographic vision that disavowed anything as superficial or frivolous as style in favour of formal clarity. But to the extent that iteration is possible, style can be identified and copied. Weston's aesthetic is now an "AI Art Style" (Fig. 1). In another influential articulation of seeing photographically, French photojournalist Henri Cartier-Bresson focused on the temporality of the Leica camera shutter, proposing the notion of the decisive moment (1952) to describe the creative fraction of a second when a picture is taken. Cartier-Bresson's privileging of the intuitive composition of events in the rangefinder – he was famously against cropping the negative – once again serves to naturalise the fusion between the human eye and camera lens and once again became a style. Even those who promoted more radical approaches to photography were unable to avoid becoming a style. For instance, European avant-garde photographers associated with New Vision foregrounded camera angles in an attempt to refresh ways of seeing the world tied to a political vision, but this was easily imitated and "disarmed" as a "style" in postwar North American "subjective" photography (Solomon-Godeau, 1989). The philosopher Vilém Flusser described the camera as a "seeing machine" (Flusser, 2000: 23). For

Flusser, photographers internalise a “photographic program”: “The camera is programmed to produce photographs, and every photograph is a realization of one of the possibilities contained within the program of the camera” (2000: 26). Flusser did not write of style, but from this perspective, ‘seeing photographically’ is a human-machine assemblage producing a style that appears unstylised, in which the photographer’s eye nonetheless becomes disciplined to the apparatus in a manner that can be imitated.

Style in art and photography history

Whilst the concept of style may have been disavowed by canonical photographers, it was nonetheless used by art historians and curators to legitimate photography as an artform. By the nineteenth century, style had become one of the foundations of the new discipline of art history. Historians studied epochal style (notably the Renaissance and the Baroque) and art historical movements (Impressionism et al.) as well as the style of individual artists. Amanda Wasielewski writes of the “systematic method of analyzing artworks developed by Giovanni Morelli in the mid-nineteenth century, where the style in which ears or hands were depicted by particular artists is taken as telltale signs of authorship and authenticity” (Wasielewski, 2023: 194). Later, the Swiss-German art historian Heinrich Wölfflin (1864–1945) turned a generation away from the analysis of the content of the works of art towards its form in his influential *Principles of Art History* (1915). Paradoxically, as Michelle Henning reminds us, style became such a dominant category for art history in large part due to photographic reproduction (Henning, 2015), which helped to enable the widespread comparison and analysis of artworks. Here, as for every other evidentiary use of the medium, photography’s apparent absence of style – its seemingly objective record-like nature – proved useful. But this is also, of course, one of the reasons photography has been considered a minor art.

In order for photography to enter the art museum, early photography historians had to work overtime to persuade sceptical art historians that photographers could develop a style that transcended the mechanical nature of the medium. In his 1942 book *New Photo Vision*, one of the first great historians of the medium, Helmut Gernsheim, tried to spell it out, but reduced the idea of photographic style to composition:

“considerations of style, of composition, play an important role in ‘objective photography’ in addition to technical considerations” (1942: 9). Beaumont Newhall, appointed the first curator of photography at the influential Museum of Modern Art (MoMA) in New York in 1940, was insistent that although photographic style *depends* on technology, it is driven by the photographer’s unique vision. For Newhall, it was simply art historians’ ignorance about photography that made them unable to distinguish differences between photographic styles (Henning, 2015: 590). Similarly, for Gernsheim:

To the critic acquainted with the work of great photographers there is evidently as much difference in style, treatment of and preference for certain subjects, as a connoisseur of painting finds in the works of painters (1962: 19).

Note that Gernsheim here updates his understanding of style from mere composition to the “treatment of and preference for certain subjects”. Indeed, photographic style is usually associated with a photographer’s preferred choice of subject matter *and* their aesthetic approach to it. Take, for instance, Richard Avedon, the portrait and fashion photographer whose work is easily recognised by his use of white backgrounds and even lighting. His white backgrounds are both subject matter and approach, content and form. Likewise, style in photography is often conflated with *genres* such as photojournalism or fashion. In her 1978 book *Photographers at Work: A Sociology of Photographic Styles*, Barbara Rosenblum demonstrated that, unlike conventional art historical mediums, style in photography “is not the outcome of the history of the rules of a form” but is shaped by a photographer’s social and working context (1978: 111).

The concept of style fell out of fashion in art history and theory in the late twentieth century, coinciding with the new centrality of photography to contemporary art. With the rise of postmodernism and its embrace of eclecticism, style was seen as a conservative, formalist and superficial way of understanding art, too individualistic and too linear in its conception of influences.² Since then, attention to an artist’s style has been largely reserved for art dealers or connoisseurs, serving to authenticate (and inflate) an artist’s originality in the art market. But it remains possible to look back at the history of photography as a series of different formal styles, including movements

(Pictorialism, Surrealism, f/64) and genres (photojournalism, fashion, etc.) as well as the preoccupations and predilections of individual photographers. The definitive three-volume *Encyclopedia of Twentieth Century Photography* includes an entry called “Stylistic Pioneers”, in which it simply says: “The most renowned photographers of the century introduced new styles” (Warren, 2005: 723). By this, the writer means that a certain canon of photographers helped to introduce a new aesthetic approach – such as Pictorialism or ‘straight photography’. Indeed, since the middle of the twentieth century, an art historical approach has been applied to photography through the canonisation of ‘master’ photographers in exhibitions and books. Photographic style, here, is what is *recognisable*, and from a commercial point of view, *intentional* and *regular*, no matter how idiosyncratic. Every camera has the potential to make interesting photographs; the question is who is responsible. The tension between the camera and photographer as the stylistic originator was obvious in exhibitions like *The Photographer’s Eye* (1964), in which John Szarkowski at MoMA theorised photography’s distinctive aesthetic character based on both camera vision *and* an individual photographer’s eye. Szarkowski’s aesthetic approach to the qualities of the medium made him an early adopter of the now widespread connoisseurship of photographs by anonymous artists. In her book *On Photography*, Susan Sontag rather cynically surmised that the “presence of a coherent photographic style” in any one photographer was largely a superficial product of photography’s career in the art museum and aspirations to being a legitimate art (1977: 135).³

Style in image processing

In the field of computing, a parallel discourse of photographic style has emerged in relation to image classification and optimisation. The question of how machines might evaluate photographs and determine their aesthetic value has animated a strand of research over the past two decades, initially to improve image search and support the surfacing of “beautiful” images at scale on social media platforms (Joshi, 2020). In undertaking this task, photographers have become central. Early work by Datta et al. (2006) sought to address the “highly subjective task” of judging image aesthetics, suggesting researchers might move beyond the task of conjuring quality images from sensor data to the problem of photographic connoisseurship. A series of benchmark

datasets were developed for aesthetic evaluation, including the Aesthetic Visual Analysis dataset (Murray et al., 2012), which harvested annotations and ratings from online photography communities including photo.net and dpchallenge.net as training data – positioning amateur photographers rather than art historians as ‘domain experts’. This in turn spurred further work in the field on sentiment analysis, “interestingness” (Gygli et al., 2013) and memorability (Isola et al., 2014). Unsurprisingly given the reliance on amateur photography communities as the source of “ground truth” data, the overall aesthetic value of an image in computing became closely related to compositional attributes that could be mathematically modelled (for example, the rule of thirds, the compositional guideline that breaks an image down into thirds both horizontally and vertically).

While some researchers were seduced by the Sisyphean task of the computational definition of “beauty”, the equally elusive problem of how machines might “recognise image style” was taken up by researchers at Adobe and the University of California. In their canonical paper, Karayev et al. (2014) argued that although “visual style” is central to how human-authored images communicate meaning, it had largely been overlooked by the field. To help clarify the nebulous concept of style and develop datasets for style classification, researchers once again turned to online photographic communities. Flickr Groups, as “community-curated collections of visual concepts”, became a standing reserve of imagery for stylistic concepts, such as the “Film Noir Mood group” and the “Geometry Beauty group” (Karayev et al., 2014: 4). According to the schema developed, photographic style could thus be understood in terms of camera modes (Macro, HDR), composition styles (minimal, geometric), genres (Vintage, Horror) and types of scenes (hazy, sunny). Through this process, ‘style’ became an elastic category capable of accommodating a range of genres, moods and techniques. For the researchers, it proved empirically that style was a product of “low-level statistics, colour choices, composition and content” (Karayev et al., 2014: 5). Thus, as the concept of style migrated from the museum to the computer science lab, it became associated with predictable abstractions and distinctive clusters of pixels, filtered through the photographic cultures of image sharing.

Having established a set of parameters for stylistic classification, the concept of “neural style transfer” was later proposed in a seminal paper demonstrating how deep learning

might generate a new image by combining the *content* of one image with the *style* of another image, typically from the history of painting (Gatys et al., 2016).⁴ While turning New York skylines into a contemporary rendering of van Gogh's starry night seemed like a leap forward, the question of "deep photo transfer" remained more elusive: without brush marks and other recognisable textures and patterns, style transfer in photography was limited to "time of day, weather, season" (Luan et al., 2017: 4490). Rather than "stylize", the key challenge of style transfer when applied to the photographic domain was to preserve "photorealism" (Gatys et al., 2016: 2421). By 2018, researchers at technology company NVIDIA had developed a style-based generator architecture for Generative Adversarial Networks (StyleGAN) that could generate convincing photorealistic images of humans in a process described as 'style mixing' (Karras et al., 2018). In doing so, it showed how the paradigm of style could be utilised beyond the production of derivative artworks and operationalised in the production of synthetic photorealism, which would ultimately usher in the era of the deepfake and experiments like *This Person Does Not Exist*, a website able to generate random human faces that do not actually exist (Wang, 2019).

The mainstreaming of text-to-image generators since OpenAI released DALL•E in 2021 has since shifted attention towards autoregressive and diffusion models as the primary technique for generating photorealistic images. Here, text prompts enable users to synthesize images referencing any artistic style that the model has been exposed to during training. As many researchers have shown, the vast training data for such models tends to be indiscriminately scraped, incorporating huge swathes of web imagery representing different forms and aesthetics, from Flickr photos to design portfolios, eBay images to medical imaging. As a result, prompting the model with style modifiers from the technical and cultural history of photography has emerged as a primary method to nudge the model towards familiar forms of photorealism from different historical periods. As Sontag already suggested nearly half a century ago, photography's *metastyle* is *realism*: "Photography's commitment to realism can accommodate any style, any approach to subject matter" (Sontag, 1977: 93). But it is crucial to understand however that text-to-image generators produce forms of stochastic photorealism extracted from the regularity of the training data the model has been exposed to. As many have observed, this represents the final nail in the coffin

of indexicality: Hito Steyerl (2023: 82) rightly refers to generated images as “statistical renderings” that “shift the focus from photographic indexicality to stochastic discrimination”. For these models, as Roland Meyer argues, “the ‘photographic’ seems to be just another ‘style’, an aesthetic, a certain ‘look’, not a privileged mode of indexical access to the world” (2023: 108). Furthermore, as Meyer observes, these images depart from other forms of synthetic photorealism like CGI or architectural visualisations as they are not based on the simulated optics of a virtual camera or the physics of light: its modelling of the world is “flat”, based on “visual patterns” (2023: 108). Despite the ability to prompt using camera settings, the modelling of “photorealistic style” in image generators privileges the “*visual* rather than *optical* aspects of the photographic” (Meyer, 2023: 108). These are images that are neither indexical, nor simulations: they are *predictive*.

Photographic styles and generative AI

If we are to comprehend the residue of photographic style(s) in the visual predictions of generative image models, it is salient to remember that software has facilitated a nostalgia for dated or obsolete photographic aesthetics for some time. For instance, nostalgic interest in Polaroid and analogue film cameras among hobbyists is matched by so-called ‘filters’ in digital software that also regurgitate photographic histories. Taking their name from the filters that professional photographers attached to their lens to adapt to lighting conditions, digital ‘filters’ are today applied as a kind of executable aesthetic wash. Initiated as basic practicalities (sharpening and so on) in professional imaging software like Adobe Photoshop, professional photographer software Adobe Lightroom made it possible to easily imitate the tonal range and colour balance of classic analogue film stock like Kodachrome or Tri-X. Filters entered the popular imaginary when they found their way into nostalgia-driven iPhone photo apps like Hipstamatic (2009) and – at a time when the image quality of camera phones remained low – became a selling point to improve and distinguish snapshots on social media platforms such as Instagram (2010), whose square format itself imitates the style of Polaroid photographs (Palmer, 2012). Soon after, smartphone cameras incorporated a range of popular filters – ‘dramatic’, ‘noir’ and so on – to evoke particular moods. With the addition of multiple lenses to the iPhone, new modes like ‘portrait’ could

digitally simulate the soft-focus bokeh of traditional camera lenses. Today, professional cameras like Fujifilm include on-board film simulations to mimic analogue-era aesthetics, while Apple iPhones enable the user to preset a “style” based on colour and contrast preferences to “personalise the look directly in Camera” (Fig. 2). Built-in software applies adjustments to different parts of an image based on conventions about what makes “good” photography.

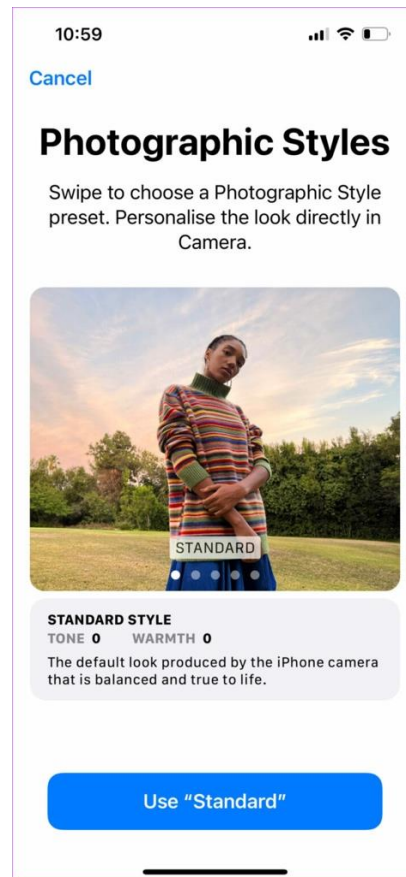


Figure 2: Screenshot of Apple iPhone ‘Photographic Styles’ feature in the Camera app, 2023.

In generative imaging, the application of photographic style moves beyond pre-set aesthetic settings to make available the entire history of photography and its techniques for new images. The superficial *look* of photographs from any period or style can be conjured with the precise combination of prompts and parameters. At the simplest level, one can simply prompt the model with the phrase ‘in the style of <photographer>’ or ‘by <photographer>’. However, prompting an image using a generative model like Midjourney (v5.2) with only a photographer’s name inevitably

produces a pale imitation. Requesting “a street in Paris” by a canonical photographer like Cartier-Bresson, discussed above, leads to images that suggest silhouetted figures on damp streets (both of which feature in some of his most famous images), but no decisive moment (Fig. 3). The image looks staged, more like a film set or advertisement. The same problem of staginess and commercial gloss is true of most of the photographer style prompts we tested, especially where people are involved. For instance, when asked to produce an image by Diane Arbus in a New York park, Midjourney generated images of attractive young women who look like fashion models rather than the “freaks” Arbus famously loved (1972: 3) (Fig. 4). The portrait format and blurred trees in the background are reminiscent of her work, but it is well known that Arbus preferred outsider figures – her best-known park photograph features a boy with a hand grenade, and others include teenage lovers, lesbians, black kids and women wearing too much makeup.



Figure 3: Midjourney prompt: “A street in Paris in the style of Cartier-Bresson”



Figure 4: Midjourney prompt: “A park in New York in the style of Diane Arbus”

Prompting in the style of a particular artist or photographer has raised obvious issues around copyright and the moral rights of artists whose work has been used as training data without their permission. Critically, as Meyer and many others have argued, style is above all “a source of value” (2023: 109). In generative image tools, the value of an artist’s style is captured and reprocessed into a generic look, as Ben Davis has argued (2023). As a result, efforts are being made to protect or prevent the tools from training

on artists' work. For instance, Glaze is a tool that enables artists to apply “style cloaks” to their art before sharing online (Shan, 2023). And given the legal dangers involved, developers of generative image tools are also increasingly seeking to limit training data to copyright-free or otherwise approved images – notably Adobe Firefly and Getty’s so-called “commercially safe” generative tool, released in 2023. The Getty tool specifically prohibits any prompt with the name of an actual person, because “it doesn’t want to manipulate or recreate real-life events” but also to avoid copyright or authenticity issues (David, 2023). This rules out both ‘a photo of Donald Trump being arrested’ and ‘A street in Paris in the style of Cartier-Bresson’. At the time of writing, a more sophisticated prompt to produce a photorealistic image will include not only the subject matter, but specific camera models (Canon 5D, etc.) and formats (‘DSLR’), focal lengths, aperture, film stock, aspect ratio (‘AR’), lighting cues (‘natural light’), poses and so on. Detail is also prompted (‘UHD’ – ultra high definition). Given the provenance of model training data, photographic prompts are also able to conjure the aesthetics of photosharing platforms (‘2014 era Tumblr’, ‘Junglecore’, ‘trending on Flickr’ and so on).⁵ Advisory guides have proliferated as the prompting community documents their discoveries of styles and photographic techniques latent in the model and how best to deploy them (Fig. 5).

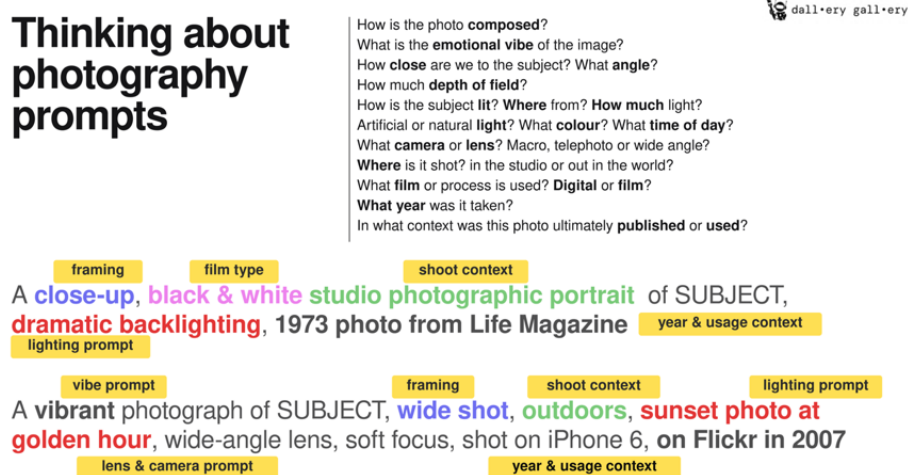


Figure 5: Screenshot from *The DALL-E 2 Prompt Book v 1.02* by Guy Parsons (2022).

One resource, developed by professional photographer Andrei Kovalev, is the ‘Midjourney AI Styles Library’, a “non-profit educational initiative” run by volunteers who have amassed “a carefully curated collection of hand-picked Midjourney styles”

containing over 3900 examples, with benchmarks to compare how the styles “behave” when applied to different environments and subjects (Kovalev, 2023). At the time of writing, the library contains 660 photographers-as-styles, from Anna Atkins, a canonical nineteenth-century English botanist known for her cyanotypes, to Deutsche Borse Photography Prize winner Richard Mosse. Biographical details provided on each photographer are simply offered as “style observations”. The results of sample prompts – from ‘cyberpunk character by Guy Bourdin’ to ‘tech genius teenage girl by Berenice Abbott’ – are carefully documented in a contemporary form of previsualisation that seeks to educate the aspiring prompter about how each style modifier operates, privileging photographers’ names as a specialist language for navigating the model. Rather than the eye, the lens and the world, the iterative nature of the prompt – as a form of searching latent space – attunes prompt and outcome.

In his ‘Midjourney Guide for Photographers’, Kovalev (2023) offers a disclaimer for his readers that “Midjourney cannot replace your vision”. In a curious evocation of modernist values, he emphasises Midjourney’s value⁶ as a tool of previsualisation in the professional photographer’s pipeline: “With just one prompt, you can previsualize a shooting style, a location, a makeup concept, a scene’s atmosphere, and so on – the possibilities are literally limitless!” Following Flusser, if the task of the photographer is to exhaust the program of the camera, the task of the “promptographer” is to exhaust the latent possibilities of the model. But while creating images is simple, predicting the output of a prompt and producing a convincing imitation of photography is more difficult. The tendency of models to default to a painterly aesthetic might explain why prompts aimed at generating photorealism also routinely contain “adjectives that often emphasize hyperbolic image aesthetics – hyper-maximalist, hyper-realist, hyper-detailed, and more” (Munster and Rossiter, 2024: 63). While this rhetoric echoes the hyperreal aesthetics of contemporary capitalism, it also speaks to a desire to push the models to produce images that are distinctive simply by virtue of their detail and resolution. In this sense it is reminiscent of the fetishization of lens quality in the history of amateur photography. Kovalev concludes his guide by urging photographers not to shy away from the AI image revolution, stating: “it’s not the style that makes an artist. It’s the stories that the artist tells with that style.” Nevertheless, in an age of

hypercirculation, the quest to find one's unique "vision" from generic software remains pressing.

Styles are clearly commodifiable. It is possible, in short, to 'see photographically' with generative imaging tools by treating the history of photography as a style market. For example, the website Aftershoot⁷ invites wedding photographers to pay for "hand-crafted signature AI Styles" and "find the ones that suit you the best." Photography thereby becomes a memory to be evoked in self-conscious stylistic evocations and photographic clichés. As Andrew Dewdney reminds us: "the afterlife of photography, residual as it might technically be, also maintains a powerful representational hold on culture and upon reality" (2021: 4). Paradoxically, in AI image generation, photography is both challenged and glorified – in an example of what Dewdney has polemically described as photography's "zombie condition" in which its annexation by computation ensures it remains an "animated corpse [...] caught paradoxically between life and death" (2021: 10). Paul Frosh puts a more positive spin on a similar observation. Writing about the screenshot as a neglected but pervasive element of photography's contemporary expanded field, he suggests that "photography survives radical change by being systematically remembered and reproduced, discursively and materially, in accordance with contemporary conditions" (2023: 187). To be sure, photography remains a reference medium in generative imaging models. The medium's persistence reflects a shift in its valorisation from the singular enframed modernist gesture to photography as a set of relations to be mined. What is being 'remembered' in the invocation of photography in the use of image generators is not the archive itself, but what can be statistically extracted and generated.

Aesthetic conventions remain powerful. This is brought into sharp focus with Meta's AI image generator, whose marketing infuses the model with human connoisseurship to bring authenticity back into a dead statistical output and ensure the afterlife of photographic connoisseurship in the model. The generator is trained on 1.1 billion Instagram and Facebook photos (Edwards, 2023). Its underlying model, Emu, performs what Meta's computer scientists call "quality tuning" on the basis that "following certain photography principles in curating the quality-tuning dataset leads to improved aesthetics for a broad set of styles" (Dai et al., 2023: 7). The "human evaluation" is done by "*specialist* annotators" who "have a good understanding" of

“certain principles of professional photography composition, including the ‘Rule of Thirds’, ‘Depth and Layering’, and more” (Dai et al., 2023: 5-6). Likewise, aesthetic conventions inform Google’s model Imagen 2:

We trained a specialized image aesthetics model based on human preferences for qualities like good lighting, framing, exposure, sharpness, and more. Each image was given an aesthetics score which helped condition Imagen 2 to give more weight to images in its training dataset that align with qualities humans prefer. This technique improves Imagen 2’s ability to generate higher-quality images (Google DeepMind).

Here the human-in-the-loop provides domain expertise, nudging the model toward stylistic photographic norms, in a move which seeks to emphasise its quality engineering (not unlike how companies compete on the engineering of their cameras).

Google also allows users to prompt with images to “condition Imagen 2 to generate new imagery that follows the same style” in a process they call “Fluid style conditioning” (Google DeepMind). In a parallel move, in November 2023, Midjourney added a feature called “style tuner” – which the company describes as a “tool that controls our model’s personality” (Midjourney, 2023). This feature encourages users to develop their own distinctive visual style and apply it to future images they generate, bringing aesthetic consistency to ‘unstable’ diffusion models which would otherwise require workarounds. Midjourney marketing heralds a new future of “promptography” commodification: “Explore aesthetics like never before and share resulting style codes and tuning URLs with friends” (Midjourney, 2023). Reference to the “model’s personality” is also telling: each text-to-image model and its various versions have their own conventions and limitations, as enthusiastic users like to point out on their social media feeds.

To the extent that style is promoted as offering a means to customise the generic aesthetics of text-to-image models, engineers follow a trend in recent years towards the “personalisation” of “taste” in image production (see for example Shaji and Yildirim, 2017). They seek to cater not just to social media users (under pressure to aesthetically optimise their Instagram feeds) but also to brands, for whom “style forms

part of strategic communication” by helping them “tell stories within recognisable genres” (Schroeder, 2013). Or, as one marketer on LinkedIn AI suggests:

The emergence of AI in defining photographic styles represents a paradigm shift in content creation. Moving away from traditional stock photography, AI now allows for the generation of unique, style-specific images. This capability is invaluable for brands and creators seeking distinctive visual content that aligns with specific themes or audience preferences, offering unparalleled customization in image creation (Shmoylov, 2024).

The task of ‘tuning’ then is to produce images which have a repeatable aesthetic signature that can communicate brand value. From a business perspective, marketing for Getty’s new generative tool claims that “customers” can “add their own data to train the model and generate images with their brand style” (David, 2023). For the individual user, just as AI writing assistant tools like the Chrome extension Compose AI are designed to imitate a person’s writing style, image generator tools will increasingly ‘learn’ a photographer’s ‘way of seeing’ and suggest how an image might be generated accordingly.

Conclusion:

Seeing photographically in a post-photographic era

Seeing photographically using text-to-image generators involves the reproduction of photography’s aesthetic conventions. The prompt can conjure photographic processes (cyanotypes, autochromes, collodion prints, daguerreotypes), genres (stock photography, fashion photography, candid street portrait, corporate headshot, award winning photo), camera types and perspectives (macro shot, drone shot, GoPro camera, tilt shift photography, fisheye, pinhole, iPhone), lighting (golden hour, studio lighting, lens flare) and film type (Polaroid, Kodak infrared film, Kodak Ektar 100, Ilford Delta 3200). In this sense, style goes well beyond individual artists or groups, and becomes a pre-set modifier for conditioning the model’s output. Everything, in short, “becomes a recognizable and marketable ‘style’, a repeatable visual pattern extracted from the digitally mobilized images of the past” (Meyer, 2023: 100).

Furthermore, although each of the generative AI models have their own aesthetic qualities, the default generative AI look has now itself become a style: a hyper-real ‘democratic’ fantasy art. Careful and even counter-intuitive prompting is required to avoid the overly glamorous gloss we saw in our test examples and period style clichés. Like stock photography, with which generative imaging is intimately linked – both in terms of its training data and in terms of its end use – the dominant ‘style’ of the generators, for now at least, is necessarily nostalgic and kitsch.⁸ As the art critic Clement Greenberg wrote in 1939, against the backdrop of fascism:

The precondition for kitsch, a condition without which kitsch would be impossible, is the availability close at hand of a fully matured cultural tradition, whose discoveries, acquisitions, and perfected self-consciousness kitsch can take advantage of for its own ends. It borrows from its devices, tricks, stratagems, rules of thumb, themes, converts them into a system, and discards the rest. It draws its life blood, so to speak, from this reservoir of accumulated experience (Greenberg, 1939: 10).

As the concept of style migrates from art history to the computer science lab, the history of photography is likewise disconnected from its social, political and cultural contexts. Instead, the history of photographic techniques and the ghosts of its ‘stylistic pioneers’ are harnessed as a programming syntax – establishing a new grammar of the image.

What does all this mean for photography and photographers working today, for whom establishing a defined style remains a critical part of establishing a ‘brand’? If the camera once conditioned the eye of the photographer, language conditions the ‘promptographer’ in image synthesis. There is no internalisation of a physical camera, whose technical residue is simply a linguistic device for securing the truth claims of photorealism in a predicted image. Instead, the syntax of the prompt mutates the cultural and technical history of photography into a creative programming language, apprehended and re-constituted in how-to guides as explanatory detail rationalising the relations between input and output. In the process, as Meyer points out, “‘style’ ceases to be a historical category and becomes a pattern of visual information to be extracted and monetized” (2021: 107). Photographic style becomes, in short, an ahistorical

programming language in competition with the photographers that the models nonetheless depend on.

References

- Arbus, D. (1972) *Diane Arbus*. Millerton: Aperture.
- Blom, I. (2007) *On the Style Site: Art, Sociality, and Media Culture*. Berlin: Sternberg Press.
- Cartier-Bresson, H. (1952) *The Decisive Moment*. New York: Simon and Schuster.
- Cubitt, S., D. Palmer and L. Walkling (2013) 'Enumerating Photography from Spot Meter to CCD', *Theory, Culture & Society* 32(7-8): 245-265. DOI: 10.1177/0263276412472377.
- Dai, X. et al. (2023) 'Emu: Enhancing Image Generation Models Using Photogenic Needles in a Haystack'. arXiv. Available at: <http://arxiv.org/abs/2309.15807> (Accessed: 4 February 2024).
- Datta, R., D. Joshi, J. Li and J.Z. Wang (2006) 'Studying Aesthetics in Photographic Images Using a Computational Approach', in A. Leonardis, H. Bischof and A. Pinz (eds.) *Computer Vision – ECCV 2006*. Berlin: Springer, pp.288-301.
- David, E. (2023) 'Getty made an AI generator that only trained on its licensed images', *The Verge*, 25 September. Available at: <https://www.theverge.com/2023/9/25/23884679/getty-ai-generative-image-platform-launch> (Accessed 4 October 2023).
- Davis, B. (2023) 'Is Crafting "Super Prompts" for A.I. Generators the Art of the Future? Probably Not', *Artnet*, 27 April. Available at: <https://news.artnet.com/opinion/ai-prompt-engineer-2288620> (Accessed 4 May 2023).
- Dewdney, A. (2021) *Forget Photography*. London: Goldsmiths Press.
- Edwards, B. (2023) 'Meta's new AI image generator was trained on 1.1 billion Instagram and Facebook photos', *Ars Technica*, 12 July. Available at: <https://arstechnica.com/information-technology/2023/12/metas-new-ai-image-generator-was-trained-on-1-1-billion-instagram-and-facebook-photos> (Accessed 8 December 2023).
- Flusser, V. (2000) *Towards a Philosophy of Photography*, trans. A. Matthews. London: Reaktion Books.

- Frosh, P. (2023) 'Screenshots and the Memory of Photography', in W. Gerling, S. Möring and M. De Mutiis (eds.) *Screen Images In-Game Photography, Screenshot, Screencast*. Berlin: Kulturverlag Kadmos, pp.173-192.
- Gatys, L.A., A.S. Ecker and M. Bethge (2016) 'Image Style Transfer Using Convolutional Neural Networks', in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, USA: IEEE, pp.2414-2423. <https://doi.org/10.1109/CVPR.2016.265>.
- Gernsheim, H. (1942) *New Photo Vision*. London: Fountain Press.
- Gernsheim, H. (1962) *Creative Photography: Aesthetic Trends 1839–1960*. New York: Bonanza Books.
- Google DeepMind. 'Imagen 2: Our most advanced text-to-image technology', *Google DeepMind*. Available at: <https://deepmind.google/technologies/imagen-2/> (Accessed 12 December 2023).
- Greenberg, C. (1961) *Art and Culture: Critical Essays*. Boston: Beacon Press.
- Gygli, M., H. Grabner, H. Riemenschneider, F. Nater and L. Van Gool (2013) 'The Interestingness of Images', in *2013 IEEE International Conference on Computer Vision*, Sydney: IEEE, pp.1633-40. <https://doi.org/10.1109/ICCV.2013.205>.
- Henning, M. (2015) 'With and Without Walls Photographic Reproduction and the art Museum' in M. Henning (ed.) *The International Handbooks of Museum Studies: Museum Media*. Oxford: Wiley-Blackwell, pp.577-602.
- Heras, D.C and T. Blanke (2021) 'On Machine Vision and Photographic Imagination', *AI & Society* 36: 1153-1165. <https://doi.org/10.1007/s00146-020-01091-y>.
- Isola, P., J. Xiao, D. Parikh, A. Torralba and A. Oliva. (2014) 'What Makes a Photograph Memorable?' *IEEE Transactions on Pattern Analysis and Machine Intelligence* 36(7): 1469-1482. <https://doi.org/10.1109/TPAMI.2013.200>.
- Joshi, B. (2020) 'Image Aesthetics at Scale: Web 2.0, Flickr and its legacy'. Presentation at The Photographers' Gallery, London, 9 December 2020.
- Karayev, S., M. Trentacoste, H. Han, A. Agarwala, T. Darrell, A. Hertzmann and H. Winnemoeller (2014) 'Recognizing Image Style', in *Proceedings of the British Machine Vision Conference 2014*. Nottingham: BMVA Press, pp.122.1-122.11. <https://doi.org/10.48550/arXiv.1311.3715>.

- Karras, T., S. Laine and T. Aila (2018) 'A Style-Based Generator Architecture for Generative Adversarial Networks', *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. California: IEEE, pp.4396-4405.
- Kovalev, A. (2023) 'Midjourney for Photographers', *Midlibrary*, 5 January. Available at: <https://midlibrary.io/midguide/midjourney-ai-for-photographers> (Accessed: 4 February 2024).
- Luan, F. et al. (2017) 'Deep Photo Style Transfer', in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu: IEEE, pp.6997-7005. <https://doi.org/10.1109/CVPR.2017.740>.
- MoMA (Museum of Modern Art). Collection text for 'Garry Winogrand, *New York, 1968*'. Available at: <https://www.moma.org/collection/works/111128> (Accessed 15 August 2023).
- Meyer, R. (2023) 'The New Value of the Archive: AI Image Generation and the Visual Economy of 'Style'', *IMAGE: The Interdisciplinary Journal of Image Sciences* 37(1): 100-111.
- Midjourney [@midjourney] (2023) 'We're now testing V1 of our Midjourney "Style Tuner". Type /tune and render a custom web tool that controls our model's personality. Everything from colors to character detail. Explore aesthetics like never before and share resulting style codes and tuning URLs with friends', *Twitter*. Available at: <https://twitter.com/midjourney/status/1719897351735967981> (Accessed: 4 February 2024).
- Munster A. and N. Rossiter (2024) 'Performing the Automated Image' in R.A. Trillo and M. Poliks (eds.) *Choreomata: Performance and Performativity after AI*. Florida: CRC Press, pp.47-73.
- Murray, N., L. Marchesotti and F. Perronnin (2012) 'AVA: A large-scale database for aesthetic visual analysis', in *2012 IEEE Conference on Computer Vision and Pattern Recognition*. Rhode Island: IEEE, pp.2408-2415. <https://doi.org/10.1109/CVPR.2012.6247954>.
- Palmer, D. (2012) 'iPhone Photography: Mediating Visions of Social Space', in L. Hjorth, J. Burgess and I. Richardson (eds.) *Studying Mobile Media: Cultural Technologies, Mobile Communication, and the iPhone*. New York: Routledge, pp.85-97.
- Rackham, I. (2022) 'Finding Your Own Photographic Style, and Why It's So Important', *Fstoppers*, 14 April. Available at:

- <https://fstoppers.com/education/finding-your-own-photographic-style-and-why-its-so-important-591442> (Accessed: 20 June 2023).
- Rosenblum, B. (1978) *Photographers at Work: A Sociology of Photographic Styles*. New York and London: Holmes and Meier Publishers.
- Schapiro, M. (1953) 'Style', in A. L. Kroeber (ed.) *Anthropology Today: An Encyclopedic Inventory*. Chicago: University of Chicago Press, pp.287-312.
- Schroeder, J.E. (2013) 'Snapshot Aesthetics and the Strategic Imagination', *Invisible Culture* (18). Available at: <https://papers.ssrn.com/abstract=2377848>.
- Shmoylov, V. (2024) 'AI & Visual Culture: 6. AI & Photography', *LinkedIn*, 15 January. Available at: <https://www.linkedin.com/pulse/ai-visual-culture-6-photography-vladimir-shmoylov-kauke/> (Accessed: 4 February 2024).
- Shan, S., J. Cryan, E. Wenger, H. Zheng, R. Hanocka and B.Y. Zhao (2023) 'Glaze: Protecting Artists from Style Mimicry by Text-to-Image Models', arXiv. <https://doi.org/10.48550/ARXIV.2302.04222>.
- Shaji, A. and G. Yildirim (2017) 'Personalized Aesthetics: Recording the Visual Mind using Machine Learning', *NVIDIA Developer Blog*, 29 March. Available at: <https://developer.nvidia.com/blog/personalized-aesthetics-machine-learning/> (Accessed 27 February 2021).
- Solomon-Godeau, A. (1989) 'The Armed Vision Disarmed: Radical Formalism from Weapon to Style', in R. Bolton (ed.) *The Contest of Meaning: Critical Histories of Photography*. Cambridge, MA: MIT Press, pp.82-107.
- Sontag, S. (1977) *On Photography*. New York: Farrar, Straus and Giroux.
- Sontag, S. (1982) 'On Style', in *A Susan Sontag Reader*. New York: Penguin Books, pp.137-158.
- Steyerl, H. (2023) 'Mean Images', *New Left Review* 140/141: 82-97.
- Vincent, J. (2022) 'All these images were generated by Google's latest text-to-image AI', *The Verge*, 24 May. Available at: <https://www.theverge.com/2022/5/24/23139297/google-imagen-text-to-image-ai-system-examples-paper> (Accessed 20 June 2022).
- Wang, P. (2019) *This Person Does Not Exist*. Available at: <https://thispersondoesnotexist.com/> (Accessed: 4 February 2024).
- Warren, L. (ed.) (2006) *Encyclopedia of Twentieth-Century Photography*. New York: Routledge.

- Wasielewski, A. (2023) 'Authenticity and the Poor Image in the Age of Deep Learning', *photographies* 16(2): 191-210. DOI: 10.1080/17540763.2023.2189158
- Weston, E. (1943) 'Seeing Photographically', *The Complete Photographer* 9(49): 3200-3206.

Notes

- ¹ Rather unusually, North American photographer Walker Evans used the term "documentary style" to describe his work, partly to distance his art from the dominant strand of social documentary associated with the Farm Securities Administration in the 1930s.
- ² What matters more to critics and theorists is an artwork's intertextuality and its social and political context of production (including the identity of the artist, for instance as a racialized subject). Ina Blom, in her book *On the Style Site: Art, Sociality, and Media Culture*, also notes that while "the term *style* has all but disappeared from art critical or art historical terminology" (2007: 11), it has become prevalent in everyday life and political culture – and links it to contemporary social identity and subjectivity.
- ³ In Sontag's 1965 essay 'On Style' she had already sought to make a distinction between "true" artistic style and "superficial" stylization (see Sontag, 1982).
- ⁴ This technique captured public imagination in 2016 with the Prisma app, which enabled users to transform their photos into "art masterpieces". Here style was a shorthand for stylization, in which an image was stylized according to artistic input, typically with a painterly outcome (not unlike Photoshop's familiar "stylize" filter).
- ⁵ A key resource for the prompt engineering community is the 'Aesthetics Wiki' fandom (https://aesthetics.fandom.com/wiki/Aesthetics_Wiki) – an ever-growing community-led database dedicated to "the identification, observation and documentation of visual schemata" as a resource for the proliferating image aesthetics of the web.
- ⁶ <https://midlibrary.io/midguide/midjourney-ai-for-photographers>
- ⁷ <https://marketplace.aftershoot.com>
- ⁸ As Meyer notes, "Both stock photography databases and text-image generators rely on text descriptions of visual content, but while stock photography searches can only find what has already been produced and described, prompts are used to find what exists only as a latent possibility" (2023: 100).

Daniel Palmer is a Professor in the School of Art at RMIT University. His books include *Installation View: Photography Exhibitions in Australia 1848–2020* (Perimeter Editions, 2021) with Martyn Jolly; *Photography and Collaboration: From Conceptual Art to Crowdsourcing* (Bloomsbury, 2017); *Digital Light* (Open Humanities Press, 2015), edited with Séan Cubitt and Nathaniel Tkacz; *The Culture of Photography in Public Space* (Intellect, 2015), edited with Anne Marsh and Melissa Miles; *Twelve Australian Photo Artists* (Piper Press, 2009), co-authored with Blair French; and *Photogenic* (Centre for Contemporary Photography, 2005).

Email: daniel.palmer@rmit.edu.au

Katrina Sluis is Associate Professor and Head of Photography and Media Arts at The Australian National University (ANU), where she leads the Computational Culture Lab

in the School of Art and Design. Previously, she was Senior Digital Curator at The Photographers' Gallery London and founding co-director of the Centre for the Study of the Networked Image at London South Bank University. Her research addresses the politics and aesthetics of the image in computational culture, its automation, social circulation and cultural value. She is the co-editor of *The Networked Image in Post-Digital Culture* (Routledge, 2022).

Email: Katrina.Sluis@anu.edu.au

