

The boar in the flower shop – exploring Synthography as a New Genre Beyond Photography

Elke Reinhuber; Sebastian Pelz

Affiliation(s) SCM City University of Hong Kong
Location, Country Hong Kong; Berlin, Germany
Contact Emails eer@me.com; sp@eer.de

Abstract

This paper seeks to explore and critically analyse the emergent genre of digital image creation proposed as ‘Synthography’. Drawing upon previous foundational texts by the author, this paper aims to delve into the theoretical implications and practical applications of Synthography. The text discusses the new possibilities in visual arts driven by the advancements in digital tools and artificial intelligence, which redefines the traditional concepts of photography. It explores the process of 3D modelling and its role in product photography as well as the concept of the ‘uncanny valley’ by using the example of a synthetic recreation of a wild boar through Synthography. This theory, which refers to the discomfort experienced when a representation is very close but not entirely identical to the real thing, has significant implications for this new genre. The interplay between the ‘real’ and the ‘synthetic’ in Synthography often walks this fine line, leading to innovative, yet sometimes unsettling, visual experiences.

Keywords

Machine Learning, Variational Autoencoders, Computer Assistance, Image Generation, Latent Space, Natural Language Processing, Image not taken, Decisive moment, Counterfactuals, Text to Image

Introduction

As the digital age is evolving, it is continuously reshaping the realm of artistic expression, challenging our comprehension and classification of genres. A novel genre, ‘Synthography’, is emerging as a prominent player in the field of digital image creation. Synthography transcends the boundaries of traditional photography by utilising the potential of rapidly advancing digital tools to fabricate visuals that merge the ‘real’ and the ‘synthetic’.

There are a few tell-tale signs that might indicate an image has been generated by a machine learning algorithm: Synthographs, generated images, might have oddities or inconsistencies, such as unusual patterns, unrealistic lighting or shading, or objects that appear distorted or out of place. The images may also exhibit unnatural textures, as machine learning algorithms often struggle to accurately recreate complex textures like hair or fabric, resulting in overly smooth, patterned, or unnatural areas. Repetition of elements is another sign, as these algorithms sometimes repeat certain elements in an image, creating a noticeable

pattern. A lack of fine details that a human artist would include might make the image look slightly blurred or low-resolution, with hues being unusually bright or dull, or not corresponding to the colours expected in a real-life scene. *Finally, uncanny facial features can be a sign, especially when the image represents people.** [* This sentence was written by ChatGPT – who would have thought?] The faces might look slightly off due to unnatural skin tones, odd facial proportions, or eyes that seem lifeless or out of place.

This intersection of Synthography and artificial intelligence is not only redefining what we consider as ‘photography’, but also opening up new avenues for creative expression and innovation in the digital age.

Synthography

The term ‘Synthography’ has been proposed by Elke Reinhuber to describe a novel genre of digital image creation that transcends the boundaries of traditional photography. The application of Synthography in the creation of photorealistic images represents a significant shift in how we perceive and interact with visual content. While utilising advanced digital tools and AI algorithms to manipulate or generate images that are strikingly lifelike, often indistinguishable from photographs captured with a camera, the process can involve the enhancement of existing photos, the synthesis of multiple images, or the generation of entirely new visuals from scratch. The result is a synthetic image that is ‘real’ in its level of detail and visual fidelity, yet ‘synthetic’ in its origin. This blending of the ‘real’ and the ‘synthetic’ challenges traditional notions of photography and opens new avenues for creativity and innovation. It also raises thought-provoking questions about the nature of reality and authenticity in the digital age.

The implications of Synthography are far-reaching, influencing not only the realm of artistic expression but also our understanding of visual culture in the digital age. As it blurs the lines between the ‘real’ and the ‘synthetic’, it prompts us to reconsider our definitions and expectations of what constitutes a ‘photograph’. It challenges the established norms of authenticity and originality in visual content, as Synthographic images, though visually indistinguishable from traditional photographs, are creations of digital tools and AI algorithms. This shift necessitates a reassessment of our perception and evaluation of visual content, as the boundary between the camera-captured and the computer-generated becomes increasingly blurred.

Looking ahead, the future of Synthography appears promising and full of potential. As digital technology and artificial intelligence continue to evolve, the possibilities for synthetic image creation are likely to expand exponentially. We can expect more sophisticated AI algorithms that can generate increasingly realistic and innovative visuals. As we navigate this exciting frontier, it will be essential to continue the discourse on the ethical and philosophical implications of Synthography, to ensure that this new genre evolves in a manner that respects creativity, authenticity, and the evolving norms of the digital age.

Never-Happened Photos

Machine learning-generated images can also be used to create “never-happened photos” – images that depict events, people, or places that never actually existed in reality but are constructed to look as if they did. For instance, an algorithm could be trained to generate a realistic image of a historical event with a twist, such as Abraham Lincoln giving a speech at the United Nations. Or it could create an image of a cityscape that combines elements from different quarters around town, resulting in a unique and never-before-seen urban landscape. This ability to create realistic never-happened photos opens up a plethora of possibilities for storytelling, art, and even history exploration. In addition to the examples already discussed, machine learning-generated images can also be employed to create a series of counterfactual “what if” scenarios to explore alternate history, speculate on different futures, or simply fuel the imagination. Such as an image showing what the world might look like if the industrial revolution had never occurred: London, still dominated by green spaces and devoid of skyscrapers.

Alternatively, a machine learning algorithm could be used to create images of events that should have happened but did not. For example, imagine a picture of the signing of a peace treaty that could have prevented a major war, or a snapshot of a famous person at an event they were supposed to attend but missed due to unforeseen circumstances. These never-happened photos could also be used in more whimsical or fantastical settings. Imagine an image of a unicorn grazing in Central Park, or of humans setting up a colony on Mars. While these scenarios may seem far-fetched, machine learning-generated images make it possible to visualise them in a realistic manner.

The potential uses and implications of these never-happened photos are vast, touching on domains such as education, entertainment, and even psychology. By providing a visual representation of alternate realities, these images could help us better understand our world and the choices we make.

Synthography and the ‘Image Not Taken’

In the context of Synthography, the idea of the ‘image not taken’ becomes particularly relevant. These are visuals that only exist in the photographer’s mind or memory, and did

not exist in the tangible world until brought into existence through the process of Synthography. They challenge traditional notions of photography, which typically involves capturing a moment or scene that exists in reality. The ‘image not taken’ opens up new possibilities for creative expression, allowing for the creation of visuals that transcend the limitations of the physical world. However, it also raises questions about authenticity and the blurred boundaries between reality and artifice in the digital age. Synthography is the medium that makes the conceptualisation of the ‘image not taken’ a possibility. It provides a platform for the manifestation of ideas, memories, or images that exist solely in the creator’s imagination. These images, which have no prior physical existence, are brought to reality through the use of digital tools and AI algorithms [1]. By transcending the traditional confines of photography which are bound to capture pre-existing moments or scenes, Synthography enables creators to construct visuals beyond the constraints of the physical world.

This paradigm shift also invites us to reassess our understandings of authenticity and reality in the context of digital image creation. As aptly encapsulated in the words of photographer Terence Donovan, “The magic of photography is metaphysical. What you see in the photograph isn’t what you saw at the time. The real skill of photography is organised visual lying.” [2] In the age of Synthography, these words are more relevant than ever before.

Modelling the Real World

Rather than taking pictures in the wild, where all things unexpected can happen, with enough information, images can be reconstructed far from weather, traffic and curious onlookers.

Modelling the real world in 3D constitutes a systematic process that encompasses a multitude of steps and can be executed with the aid of various software tools. Upon the completion of the shape of the objects, colours and textures are added. This process, also known as ‘texture mapping’, involves the application of 2D images or textures onto the 3D model, thereby enhancing the realistic appearance of the model through the incorporation of details such as colours, patterns, and roughness of the surfaces.

The final steps encompass the setup of lights and camera for the scene and ‘rendering’. Rendering, the process of generating a 2D image or animation from the 3D model, calculates the interaction of lights and shadows with the object and how the camera views the scene, as setting up the camera and lighting is a crucial step in achieving photorealism.

Virtual Photography

Since the complete toolset of the professional photographer is available as a simulation, lighting and shadows can be administered to emphasise certain aspects of it and the most

advantageous angle can be chosen: Studio illumination provides precise control over the lighting conditions in a photograph, allowing the photographer to manipulate the intensity, direction, and falloff of artificial light sources to highlight specific features, create dramatic effects, or balance the overall lighting in a scene. By carefully positioning lights, shadows can be controlled to add depth and dimensionality to surfaces, revealing the fine craftsmanship, materials, and textures of the model.

In recent years, Computer-Generated Imagery (CGI) has found a significant place in product photography. Traditional product photography involves meticulously arranging physical products and lighting for each shot. However, CGI allows for the creation of photorealistic, 3D digital replicas of products that can be manipulated in any way imaginable, without the limitations of physical constraints. This can involve changing the product's colour, texture, or surroundings, or even generating complex action sequences. Photographers can create different scenes and arrangements digitally, adjusting lighting, orientation, and background with precision and ease. This allows for an unprecedented level of creative control and flexibility, while also reducing the time and cost associated with traditional product shoots. Furthermore, with advancements in CGI technology, these digital replicas can achieve a level of realism that makes them virtually indistinguishable from traditional photographs. This application of CGI in product photography is a prime example of Synthography, where the boundary between the 'real' and the 'synthetic' blurs, creating new opportunities for visual representation and expression. Many products are digitally modelled before they enter the production line, providing the foundation for the image from the manufacturing file. The virtual representation also simplifies the adjustment for international markets: for example the interior decoration for a global operating furniture brand can be adjusted with local details and appliances.

The Uncanny Valley and Machine Learning-Generated Images

The "uncanny valley" is a term and concept first introduced by Japanese roboticist Masahiro Mori in 1970. In his groundbreaking work, Mori proposed a hypothesis that as a robot or a graphical representation becomes increasingly humanlike in its appearance and movements, there comes a point at which it will provoke a sense of unease, discomfort, or even revulsion among observers [3]. This unsettling response occurs just before the representation reaches a level of realism that is fully indistinguishable from a real human.

The visual representation of Mori's theory, a graph plotting comfort level against similarity to human likeness, takes the shape of a valley. This is where the term "uncanny valley" originates. As the likeness to human appearance and behaviour increases, the comfort level of the observer rises, but then sharply drops into the "valley" when the representation becomes almost, but not quite, human.

Comfort level then rises again as the representation attains perfect human likeness.

The concept of the "uncanny valley" can also be relevant when discussing machine learning-generated images, due to unnatural skin textures, odd facial proportions, or inconsistent lighting, among other things. However, as machine learning technology continues to advance, its ability to become indistinguishable from real-life images is increasing. This presents both exciting opportunities and significant ethical considerations.

Despite being widely recognised and referenced, the exact causes of the "uncanny valley" effect are still a subject of debate among psychologists, roboticists, and researchers in related fields. Some theories suggest that it may be related to evolutionary mechanisms. For instance, humans may have developed a natural aversion to entities that look human but are slightly off, as these could be indicative of disease or death. This subconscious recognition of "wrongness" can trigger a sense of fear or revulsion.

Understanding, addressing, and potentially overcoming the "uncanny valley" is a significant challenge for many fields, including the creation of machine learning-generated images. As technology continues to advance at an unprecedented pace, the goal is to create representations that are so realistic they move beyond the "uncanny valley" and become indistinguishable from real-life images. This is a complex task that involves not only technical advancements but also deep understanding of human perception and psychology.

In the context of Synthography, the uncanny valley manifests when synthetic images are so realistic that they are almost indistinguishable from real photographs, but still contain subtle elements that betray their artificial origin. These elements can create a sense of unease or cognitive dissonance in the viewer, as the image straddles the boundary between reality and artifice, eliciting telling psychological responses.

In order to better understand the complex implications of our observations, a brief look at parallels with other developments should be taken.

"Alexa, please ... Oh never mind"

Only a few years back, voice assistants became a runaway success and were once praised as the future of the human-computer interface (HCI). Alexa, Siri and other examples of virtual assistants have been initially advertised as being powered by artificial intelligence. In fact, they interpret human speech and respond accordingly, utilising natural language processing to execute scripts, until arriving at their wits' end. Alexa defaults to pre-programmed responses with phrases like "I'm sorry, I didn't quite get that" when it cannot provide something useful. In the similar vein Siri typically uses phrases like "Sorry, I can't assist with that." or "I'm not sure how to help with that".

Do Chatbots Dream of Electric Sheep?

In the context of GPT (Generative Pretrained Transformer), a popular chatbot used for natural language processing tasks, “hallucinations” refers to situations where the model generates information that is not present in the input it was given. This can occur when the large language model makes assumptions or inferences based on its training data to fill in gaps in the input. *It’s crucial to note that these “hallucinations” are not conscious or intentional, but rather a byproduct of how the model was trained and its inherent limitations** [* This sentence was written by ChatGPT – It is important to note that the complete failure of the model is being appeased here as a side effect.]. The model might generate outputs that are based on patterns it has learned from its training data, even if these statistical clusters do not accurately reflect the specifics of the input it was given. This results in outputs that might seem plausible but are not accurate or appropriate in the context. Every LLM is prone to these hallucinations, but some more than others. OpenAI’s ChatGPT was ranked first in reliability with around 98% and Google’s Bard was ranked last with ca. 73%. [4]

Case Study: wild boar in the flower shop

The transformative potential of Synthography became apparent in an anecdotal event in the neighbourhood. When a wild boar entered a flower shop in Mong Kok, blood was spilled. The standoff with police and the AFCD (Agriculture, Fisheries and Conservation Department) staff led to the injury of a 40-year-old shopkeeper who tried to subdue the boar. Despite attempts by officers and two dogs from the shop to drive the boar away, it took shelter in the shop’s storeroom, before it was eventually euthanised and removed. The injured shopkeeper was taken to the hospital for treatment [5]. But even though three news channels, numerous photographers and dozens of phone cameras were watching, not a single meaningful picture was taken, the only photograph that circulated in the media was a blurry, artefact-ridden frame from video footage that hardly did justice to the event (Figure 1).

In the aftermath, the authors decided to use Synthography to recreate the scene. Drawing on the memory of the event and the limited existing photos, the authors were able to construct a series of ‘synthetic’ images that were surprisingly convincing. This exercise demonstrated the power of Synthography to create visuals that exist primarily in the photographer’s mind or memory and underscored its value as a tool for artistic expression that transcends the limits of traditional photography. This wild boar incident presents a vivid example of the ‘image not taken’ – a visual that did not exist in reality until it was brought to life through Synthography. The resulting images, though synthetic, were more ‘real’ in terms of their emotional resonance and visual fidelity than the actual photos taken at the scene reshaping our understanding of authenticity’s and reality’s interplay in the visual arts.

When introduced into photography by Henri Cartier-Bresson, the concept of the ‘Decisive Moment’ was considered a fundamental theory. It refers to capturing an event that is ephemeral and spontaneous, where the image represents the essence of the event itself, emphasising the importance of timing in capturing the perfect shot. However, with Synthography, the ‘Decisive Moment’ can be redefined as artists are no longer bound by the constraints of time and can recreate or fabricate moments post-event, challenging traditional photography’s spontaneity.



Figure 1. Newspaper clipping of the florist attempting to pull the wild boar out of his shop.

Prompting the Unseen

With this concrete example in mind, and the dissatisfaction with the published photograph of the event caused (the only one with a vaguely recognisable boar), we attempted to recreate the memory of the situation with the generative AI tools Adobe Firefly and NightCafe. The latter utilises several AI engines, like Stable Diffusion, Artistic VQ Gan, Coherent Clip, and more.

The first attempts portrayed a wild boar standing like a flower pot itself between plants, looking into the camera as it was for sale. Some results looked quite realistic and believable, especially after adjusting the camera settings to a lower depth of field, blurring the background, and adding items to obstruct the contours, as to be seen in the result of Figure 2. Looking closely however, details reveal the forgery: compared to the lively left eye with believable reflections, the right eye is a mat dark spot and looks up close almost like a hole, as if the creature lost it.

As the beast was always only visible from the front, facing the virtual camera, and the illumination in the flower shop differed considerably from the real one, especially Firefly had significant problems to imagine a boar from another angle and perspective while NightCafe was improvising/hallucinating on a lot of details such as the snout or the hooves (in one version, it was wearing the same boots like the surrounding policemen).



Figure 2. An image created by Adobe Firefly with the prompt “A photograph of a wild chinese boar in a crowded flower shop standing on the ground between green plants”.

To achieve a better impression of the actual location and situation, we attempted to upload images that we took that day. The richer the details we suggested in the expectation of a more precise result, the more chaotic and unrecognisable the image became. Even with a condensed description, the results differed significantly from the uploaded image which showed the flower shop at the corner of the road with a crowd of onlookers behind a barrier including several policemen and photographers. Creating an image with the pig as part of the crowd and looking at the spectacle itself was not crowned with success. The attempted prompt however delivered interesting results and also revealed the limitations of the diverse applications. “A wild boar from behind in a crowd of people with police looking at a cordoned off flower shop” simply merged the boar with the police.

Another interesting observation concerns the location, as the same composition is used throughout Figure 4a and b. “A photograph of a wild boar from behind in a crowd of people with police looking at a cordoned off flower shop” without any additional specifications resulted in an illustration of a romantic small town, possibly in Europe. Adding “in Hong Kong” changed the style of the buildings and the ethnicity of the policemen, messed up the flowers and added Chinese looking billboards with non-recognisable logograms.

After acquiring the basic tools and settings of the photographic craft and the secrets of post-production, the skillset needs to change now to finding the correct descriptions for successful prompting. The possibilities are advancing in quantum leaps, and refining an image in terms

of style and settings has already become much more accessible.



Figure 3. Result from Firefly based on an uploaded image of the location with the prompt: “A wild boar from behind in a crowd of people with police looking at a cordoned off flower shop”.

Conclusion

Considering the myriad opportunities this technology presents, its challenges, and the ethical considerations that arise along with its societal impact, to predict future trends in this dynamic field is futile.

In conclusion, Synthography presents a significant evolution in the realm of visual arts. This emergent genre, propelled by the advancements in digital tools and artificial intelligence, is redefining our understanding of photography and challenging traditional concepts such as the ‘Decisive Moment’. Synthography allows for the creation of ‘Designed Moments’ and ‘images not taken’, transcending the limitations of time, space, and physical reality that are inherent in traditional photography. The wild boar incident, even if anecdotal, serves as a vivid illustration of the potential, showcasing its ability to recreate a missed event with astonishing realism and emotional resonance. Blurring the lines between the ‘real’ and the ‘synthetic’ not only opens up new avenues for creative expression but also prompts a fundamental reconsideration of our perception and evaluation of visual content [6].

Following László Moholy-Nagy’s ideas of a new synthetism [7], which he presented a century ago and in which he described how art, science and technology should unite their forces for the common good, this ought to be possible through synthetic photography as well. It does not necessarily have to remain a fleeting fad, but the current manifestation of mechanical picture-production could be understood as an opportunity to document the world and its phenomena in a completely new manner. Perhaps

Synthography can also describe a future iteration of image-making, which should be poetic and authentic, absorbing not only photons, but all kinds of oscillating physical entities.



Figure 4a and 4b . NightCafe: Midjourney with two iterations of the same scene, the first as a photorealistic illustration, the second with the locale explicitly changed to Hong Kong and the photographic film specified as Kodak Ektar.

As we navigate the exciting frontier of this new genre, it's crucial that we continue to explore its potential, question its implications, and shape its future trajectory. In doing so, we must also grapple with the ethical and philosophical questions that generated images raise, ensuring that this evolution in visual arts respects creativity, authenticity, and

the evolving norms of the digital age. The term 'Synthography' effectively encapsulates this new genre's essence, highlighting the synthetic nature of these images while acknowledging their roots in the photographic tradition. It invites us to explore a new landscape where the 'real' and the 'synthetic' intersect, ultimately challenging our perception of what constitutes a 'photograph'.

References

- [1] Marian Mazzone and Ahmed Elgammal, "Art, Creativity, and the Potential of Artificial Intelligence," *Arts* 8 (1), (2019): 26. doi.org/10.3390/arts8010026.
- [2] Terence Donovan. "Interview," *British Journal of Photography*, (1983).
- [3] Masahiro Mori, "The Uncanny Valley," *Energy* 7 (4), (1970): 33-35.
- [4] Vectara, *Hallucination Leaderboard*, (2023) website, accessed April 1st, 2024, github.com/vectara/hallucination-leaderboard/tree/main
- [5] Kahon Chan and Harvey Kong, "Handling of Hong Kong Wild Boar Incident where Shopkeeper was Bitten and Creature Killed Questioned by Animal Rights Group." *South China Morning Post*, (October 21, 2023). scmp.com/news/hong-kong/health-environment/article/3238733/hong-kong-authorities-trying-capture-15-metre-long-wild-boar-after-it-runs-flower-store-bites.
- [6] C. Blaine Horton Jr, Michael W. White and Sheena S. Iyengar, "Bias against AI Art Can Enhance Perceptions of Human Creativity." *Nature Scientific Reports* 13:19001 (2023): 10. doi.org/10.1038/s41598-023-45202-3.
- [7] Toman, Jindřich, "Permanent Synthesis: László Moholy Nagy's Idea of a Synthetic Journal (With an Appendix: László Moholy-Nagy's 'Guidelines for a Synthetic Journal')." in *Local Contexts/International Networks: Avant-Garde Journals in East-Central Europe*, eds. Gábor Dobó, Merse Pál Szeredi (Budapest: Petőfi Literary Museum–Kassák Museum, Kassák Foundation, 2018): 25-49.

Bibliography

- Henri Cartier-Bresson, *The Decisive Moment*, (New York: Simon and Schuster, 1952)
- Donovan, Terence, Interview, in *British Journal of Photography*, (1983)
- Galileo,. *LLM Hallucination Index*, (2023) rungalileo.io/hallucinationindex.
- Leon A. Gatys, Alexander S. Ecker and Matthias Bethge, "A Neural Algorithm of Artistic Style." *arXiv preprint* (2015) arXiv:1508.06576.
- Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville and Yoshua Bengio, "Generative Adversarial Nets." In *Advances in Neural Information Processing Systems*, (2014) arxiv.org/pdf/1406.2661.pdf.
- Masahiro Mori, "The Uncanny Valley." *Energy* 7 (4) (1970): 33-35.

Alan Turing, "Computing Machinery and Intelligence." *Mind* 59 (236) (1950): 433-460.
Vectara. *Hallucination Leaderboard*, (2023)
github.com/vectara/hallucination-leaderboard/tree/main

Authors Biographies

Elke Reinhuber is a media artist, educator, researcher and Associate Professor at SCM City University of Hong Kong. In her work she explores different modes of presentation and strategies of storytelling to emphasise the parallel existence of multiple truths of the here and now, anchored in expanded photography and spanning into several disciplines such as time based media, immersive environments, Augmented and Virtual Reality as well as performance. Her award winning artistic research was presented internationally, at conferences, exhibitions and festivals.

www.eer.de

Sebastian Pelz is an independent author and AI consultant with an interest in the advancements of text and images since the Renaissance – but particularly in the digital age, anticipating a future yet to come.

Addendum

While researching for this paper, I enquired ChatGPT about recommended literature on the topic, among the answers this:

1. Elgammal, A., Liu, B., Elhoseiny, M. & Mazzone, M. (2017) 'Artificial Intelligence and the Art of Building Images', Proceedings of the IEEE, 105(10), pp. 1946-1961.

I searched for the paper, but to no avail. In the table of contents of the suggested book I found that:



Figure 5 . Two fictitious references by ChatGPT.

After I checked the pagination, the following dialogue ensued:

Me: «do you have the doi number for the references?»

GPT: «I'm sorry, but the document does not provide DOI numbers for the references.»

Me: «i know, because you invented them»

GPT: «It's true, the references listed in the document were created for illustrative purposes and do not have real DOI numbers.»

Me: «You're a pompous fraud»

GPT: «<aside> 🙄 Please be aware that the document is a work of fiction and does not intend to mislead or misrepresent any information.</aside>