

Lev Manovich

A Medium That Thinks: Generative AI and Media Cognition

| | |
|--|----|
| Introduction | 2 |
| 1. The Trade-off Between Variability and Control | 5 |
| 2. Unprecedented Degrees of Freedom | 6 |
| 3. The Anti-Avant-Garde | 8 |
| 4. A Medium That Thinks | 12 |
| 5. A Medium That Understands All Other Media | 14 |
| 6. The Entanglement of Style and Content | 15 |
| Conclusion: The Challenges of AI Medium | 18 |

Introduction

The conversations and writing about Generative AI often treat it as a social phenomenon to be critiqued, an ethical problem to be debated, or a disruption whose consequences need to be assessed. These are valid concerns — but they don't tell us much about the specific properties of AI as a creative medium. There is also no shortage of discussions about AI in relation to "art" and "creativity" — but these typically take both concepts for granted, as though they were timeless categories rather than relatively recent inventions of the Romantic period.

What we lack is a different kind of inquiry: a careful analysis of AI *as a medium* — its specific *properties, constraints, and logic*. We are familiar with such analysis for photography, painting, film, or digital media — but it barely exists for the AI medium. That is what I want to do here. What are its affordances, its constraints, and its limitations? What artistic strategies can we develop to work with — and against — its unique properties? How to exploit its properties in the most interesting ways? These are the questions that matter if we want to understand not just what AI means for culture in general, but what it means for the practice of making new cultural artifacts.

Because I am a practicing artist with 50 years of experience, it is natural for me to approach this question by comparing this new

medium to all other media I used in my art practice — painting, drawing, etching, analog and digital photography, and programming (i.e., generative art). And as a theorist and historian of media, I am equally interested in what is truly unprecedented about AI, and what continues earlier historical trajectories.

For the art world and the media art field, "AI artists" usually do not include people who use popular AI tools without modification. Only if an artist trains their own model on particular data, or fine-tunes an existing model, are they recognized. It is also expected today that such artists express a critical attitude toward AI technology; often the content of the work is AI itself.

In contrast, I am interested in normal AI tools — both multimodal ones and those focused on image and video generation, music generation, and so on — used today by millions of members of the creative industry (photographers, designers, architects, filmmakers and film editors) and tens of millions of content creators.

Understanding a new medium requires studying how it is *actually used at scale*, not only its exceptional or critical deployments - — such as interactive installations, real-time performance systems, and custom-trained artistic models. (Imagine a general film history or theory that disregards all narrative films and is only concerned with experimental cinema.) Approached in this way, the 'AI medium' consists of trained models and the apps and web services that structure users' interactions with models through particular interfaces and controls.

It is tempting to describe AI Medium using our existing concepts such as remix, bricolage or post-modern. The last term seems to fit quite well: art that consists of quotes and references to earlier historical art; combining parts and elements of already existing culture; looking back rather than forward. We can also recall the 19th century eclectic paradigm in art and architecture that deliberately revived and combined historical styles.

All these concepts share a common assumption: that culture moves forward by recycling, recombining, and referencing what already exists. In short — "*Make it Old*" rather than modernist "*Make it New*." AI content seems to confirm this perfectly — the models are literally trained on everything that *came before*.

Yet we should resist using only these concepts — because they can make us miss what is new about the AI medium. In my talk, I want to explore some of the key properties of this medium that such concepts cannot capture.

Some of the properties I will describe and analyze are structural — they follow from how these models fundamentally work. Others may diminish as models improve. Both are worth understanding, because even temporary properties of a medium shape the culture produced with it during that period.

1. The Trade-off Between Variability and Control

AI gives us unprecedented ease of creation and variability — but at the expense of precise control over details. This is how this medium fundamentally works: based on extracting patterns from large collections of media artifacts, and then generating new artifacts through prediction — through probabilistic actions.

The lack of control is therefore not a temporary limitation waiting to be engineered away. It is a constitutive property of the medium. Probabilistic generation and total authorial control are in tension by definition.

What does this mean in practice for culture? For amateurs and casual users, the lack of precise control over details is more often than not irrelevant, as long as generated artifacts look good. But what about creative professionals? Although GenAI produces high-quality results, designers do not have complete control over every detail — the model makes countless decisions autonomously. Consider such cases as, for example, generating an advertising photo for a product, or making a book cover or another graphic design where exact position and size of every graphic element really matters. In such cases, this lack of full control remains a fundamental challenge.

There is a big part of creative universe where the lack of precise control and inability to edit results is acceptable — but probably much bigger part where it is not.

This perspective also gives a possible new way to categorize the art, media and design universe. Normally we use categories like media types, or mass produced vs unique, or 2D vs 3D vs 4D, or static vs interactive, and others. But now the arrival of GenAI brings another important boundary: all cases where high precision, control over minute details and *editability* are crucial, vs. all cases where it's not.

2. Unprecedented Degrees of Freedom

Continuing this theme, I want to describe one aspect of AI image generation which is probably most important and interesting to me among all its other aspects. Modern artists used a variety of methods and tools to give up control and add randomness and indeterminacy into their artistic process. But usually, all such methods and systems only operate in one or at maximum a few dimensions. For example, take a particle system — yes, its parameters (particle size, transparency, speed, etc.) can be controlled by random functions — but in using such particle systems, we can still imagine beforehand the range of possible animated visuals we may see. While particles can be rendered in a variety of ways, all these possible visuals will have a definite family

resemblance. The degrees of freedom — the range of possible we may expect — is limited.

Although GAI can be used in this way (limiting its degree of freedom, and this is how it is often used for particular practical media tasks), it also enables a different mode. The artifacts it can generate in response to user requests can have more degrees of freedom, more unpredictability, and cover a much larger area of creative universe of everything which is possible, but invisible to us.

Modern experimental arts (including visual arts, music, literature, and performance) developed a variety of methods for introducing indeterminacy into artistic works. These methods and also ways to think about different types of indeterminacy are very useful for understanding AI medium.

Possible examples of large unpredictability and flexibility and many degrees of freedom in earlier art movements include of course Jazz. I am also thinking of many people who performed at Judson Dance Theatre in New York the 1960s (Yvonne Rainer, Trisha Brown, and others) — performances where some structures were fixed and other space was deliberately left for performers to improvise. And also happenings. And also more uses of random procedures in modern music by Cage and others. But — at least in terms of image generation, I don't think we ever had a medium which offered so many degrees of freedom as AI

medium.

3. The Anti-Avant-Garde

AI learns the logic of our world, thinking and behaviors — the everyday expectations embedded in trillions of web pages and billions of images. It learns how the normal world is organized, what entities it has, and how its entities normally behave.

It also absorbs how things normally look — the visual and logical conventions of everyday photography and media. At the same time, AI models often learn idealized visualized representations. If a training dataset only includes images with high aesthetic ratings, the model absorbs not how things look, but how they look at their best. So the world learned by AI models has entities with idealized appearances that otherwise behave in entirely normal, familiar ways.

This becomes visible the moment you try to violate that logic. Ask AI to generate a cat chasing a dog, and it resists — defaulting back to the world as it expects it to be. The medium pushes back.

This might seem to contradict the fact that many users successfully generate surrealist imagery with AI tools. But surrealist prompts work precisely because they combine elements that are individually normal and well-represented in training data. "Melting clock in a desert" works because GenAI knows clocks perfectly,

knows deserts perfectly, and can combine them. It is still operating within the logic of known visual categories.

The resistance appears when you try to violate relational logic — causal, behavioral, physical. A cat chasing a dog breaks an expectation about how animals behave. That is different from combining two normal things in an unusual spatial arrangement.

In short: AI is good at visual surrealism but resists logical surrealism. Dalí-style imagery — yes. Actually inverting the rules of how the world operates — much harder.

It privileges familiar reality and its logic. This makes GenAI structurally the opposite of what avant-garde art historically tried to do — which was precisely to defamiliarize the normal, to make the expected strange. Artists who want to use it for genuinely experimental work are constantly fighting the medium's own tendencies.

Aesthetic Conservatism

The AI medium is also visually conservative. AI models privilege an aesthetic that blends 19th-century realism and academism (i.e., the aesthetics of Paris Salon painting): idealized lighting, anatomically consistent figures and faces, recognizable objects, polished surfaces, legible spatial depth, and clearly defined contours.

Many of these conventions were established in academic painting and later became standard in photography and commercial imagery. Photography also introduced additional conventions that dominate modern visual culture, such as centrally positioned subjects and images rendered in standard one-point perspective. As a result, these visual structures appear throughout contemporary imagery — from professional photography to advertising, stock images, and social media.

Such images are far more common on the web and therefore dominate the datasets used to train AI models. Even when other image types are included — such as illustrations that receive high aesthetic ratings — they typically share similar characteristics: clear, coherent, centered, and visually pleasing. The statistical dominance of this aesthetic biases models toward producing similar visual structures.

As a result, AI image models rarely introduce radically new visual structures. Instead, they recombine familiar visual patterns within the *statistical boundaries* learned during training. Their outputs exemplify what creativity researchers have earlier called “bounded creativity” — the production of variation within existing constraints rather than the invention of entirely new forms.¹

¹ Margaret A. Boden, *The Creative Mind: Myths and Mechanisms* (1990; revised ed. 2004).

Even when prompts request unusual or inconsistent imagery, generation tends to drift toward familiar visual structures learned from the training data. Thus models treat anomalies as errors rather than possible artistic innovations. In other words, they privilege conformity over experimentation.

One may wonder if we actually need any more visual innovation. 20th-century artists already expanded visual repertoire in all possible directions. The development of computer graphics in the last decades of that century added many more possibilities. So if real innovation is still to come, it will probably have to do not with what and how images can represent, but with how they are generated (for example, via neural interfaces) and how we interact with them — through new interfaces that go beyond screens, projections, the web and social media platforms, and VR/AR. It is strange that we expect AI to produce radically new visual techniques or entirely new types of art. After all, we no longer expect this from art itself. After more than a century of constant innovation and endless artists' manifestos — from Impressionism to Net Art — came a period of pluralist coexistence, where all agendas, styles, and media are considered equally legitimate. Contemporary art has quietly accepted that the era of successive revolutions is over. So why should we expect AI — a medium trained entirely on existing culture — to restart it?

Most likely, our desire to see AI innovate aesthetically is simply the result of our *intellectual and cultural inertia*. For more than a

century, modern culture has equated artistic value with the production of radical novelty. Although we may think that today we are far from modernist values and ideas, we have actually inherited its “make it new” mandate.

4. A Medium That Thinks

The conservatism I just described is only one consequence of a deeper property. McLuhan argued that communication media shape how people think. With GenAI, we encounter something fundamentally different. *This is a medium that itself thinks* — or more precisely, performs cognitive tasks as well as, and often better than, most humans, including domain experts.

Its cognitive abilities span multiple tasks. A single model can generate all kinds of images, understand them, edit them, reason about them. It can discuss any artwork with a precision and richness that often exceeds what most viewers can articulate. And it can look at your own work and suggest improvements — identifying compositional issues, visual hierarchies, color relationships — pointing to visual problems that even experienced artists might intuitively sense but struggle to put into words.

But what makes this truly unprecedented is not cognition itself. Earlier algorithmic media — manually written computer programs for media generation used in the arts since the late 1950s — could

also have functionality similar to some of the properties of human thinking. For example, if/then conditionals in generative programs can be said to emulate human reasoning. However, this functionality had to be explicitly put into the programs by the author. [And it was task-specific: a program written to generate particular visual compositions, such as Georg Nees's *Schotter* (1968) or Frieder Nake's *Hommage à Paul Klee* (1965), could only do that.]

GenAI is the first artistic medium whose cognitive capacity is built in rather than externally programmed. And this is perhaps its most important defining property.

But this characteristic also has another side. AI cognition and world knowledge were acquired by absorbing how things normally are — by extracting patterns from the entire existing media universe. This is why it gravitates toward the normal: Its thinking defaults to the familiar. But visually, what it generates may look either ordinary or idealized, depending on how the training data was selected and the user instructions. Either way, the medium's cognition and its conservatism are connected — both are products of the same learning process.

5. A Medium That Understands All Other Media

GenAI has built-in knowledge of art history, contemporary visual culture, artistic techniques and media. No artistic medium before it has ever known so much about all the others.

Consider the list of 400 "artistic techniques" from the Midlibrary web resource. Adding any name from this list to a prompt tells Midjourney (or other AI image tools) to simulate the appearance of a given medium. The range is truly encyclopedic — from the familiar to the entirely esoteric. Sashiko, Fabergé egg, Parquetry, Decoupage, Calotype print, Stained glass, Halftone print, data visualization, Chinese ink brush, Gyotaku... the list goes on and on. AI seems to have mastered them all.

Along with this vast knowledge of artistic techniques and media, AI can simulate visual styles of thousands of fine artists, manga artists, photographers, fashion designers, and creatives in other visual fields. The same Midlibrary site lists "1546 painters styles" tested with Midjourney. From Frans Hals to Hilma af Klint, from Hokusai to Zaha Hadid, from Vivian Maier to Hayao Miyazaki.

But this is more than a catalogue. AI has learned the conventions and *patterns* of the media universe — not just what media aesthetics, materials' effects and artists' styles look like, but *how their elements work together*. An AI video tool can generate appropriate sound effects for a clip showing an outdoor scene, or

choose the right type of virtual lens to fit the content of a generated shot. We can call this "**media cognition**" — the practical know-how embedded in countless media artifacts, now absorbed and internalized by the model.

How well can I actually use its encyclopedic knowledge of artists and styles in practice? It can create new images in the manner of thousands of both famous and lesser-known creators. But does this ability extend to both the content and visual styles of these creators? Can we, for example, use their styles separately, applying them to new content? Turns out that it's not so easy.

6. The Entanglement of Style and Content

When you specify the name of a particular artist in a prompt, we often see a "bleed" effect: along with applying the style of an artist to the generated image, AI also transfers the content of this artist's works. The artist's characteristic subject matter seeps into the generated image. For example, specifying 'Hokusai' often generates waves or Mount Fuji regardless of the prompt, 'Mondrian' almost inevitably produces grids with primary colors even when you describe a landscape, and 'Frida Kahlo' tends to generate self-portraits with flowers and foliage even when the prompt describes something else entirely.

This happens because style and content are structurally entangled in how these models encode knowledge — not as a temporary

flaw, but as a consequence of how learning from images works. Sometimes extracting a style works reasonably well; often it does not.

Additional techniques can to some extent reduce this problem - such as fine-tuning models with LoRA. Another way to reduce this entanglement is to bypass the artist's name altogether.

Midjourney's style reference feature (--sref) allows users to supply a reference image instead of a name. The model then extracts visual qualities — color palette, texture, line work, composition — directly from the image, without activating the web of content associations that a name triggers. Prompt "a modern city street" with a Hokusai painting as a style reference, and you are more likely to get a city street rendered with Hokusai's visual qualities rather than waves and Mount Fuji. The results are not perfect, but they are significantly better.

The fact that bypassing the artist's name improves results is revealing. An artist's name is a concept with many dimensions: style, subject matter, period, iconography, all entangled. A painting is a visual artifact with extractable properties. The model handles them differently because they are fundamentally different kinds of input.

But even if the models could reliably extract style and disregard content, there are more fundamental issues that make me wonder if such an operation is even possible anyway. What does it mean to "extract style" from a given artist? Do artists actually have singular

coherent styles visible across all their works? Usually, no — because usually an artist's style (or if we want to use a better term — visual language) evolves during their career. (To take just one example, van Gogh's style changes between his Paris period, Arles period, and Auvers-sur-Oise period). So even when we find that an AI model can successfully imitate a given artist, this is either something which is averaged across these different styles — or a particular style associated with the artist's most famous works.

If we consider pre-modern European and East Asia, this gets even more interesting. Pre-modern paintings often contain multiple "styles" inside the same image. This is not stylistic inconsistency but intentional hierarchical rendering. Faces, hands, or sacred figures might be painted with high precision, while clothing, landscape, or secondary objects are treated schematically. Such paintings intentionally do not have a single style, and AI usually struggles to reproduce the hierarchical distribution of styles within a single image. The model captures average properties of images - not intentional stylistic hierarchies.²

² As a result, as of today (Spring 2026) AI often produces images that have uniform brushwork everywhere, consistent texture across the whole image, similar level of detail across objects. This is why AI paintings often feel slightly too homogeneous.

Conclusion: The Challenges of AI Medium

We have explored *six properties of Generative AI as a creative medium*. It is a medium that thinks with extraordinary breadth — and at the same time gravitates toward the normal. It has encyclopedic knowledge of art and culture — but struggles to separate style from content. It allows you to quickly explore many ideas and possibilities — but not to have full control over how they are visualized. And it replaces the unique object with endless variations. (Of course, Generative AI has many other important properties: some of them I already explored in *Artificial Aesthetics book* and my other recent writing.)

The most interesting question remains open: what kind of art does this medium make possible that was not possible before?

But maybe this is the wrong question entirely — if computers have always been designed to simulate existing media and existing social and cognitive operations (counting, indexing, summarizing, searching), why should we expect AI to be any different?

And yet, it is inevitable for us to ask: not what Generative AI can imitate, recombine and interpolate between — but what it can genuinely enable that no previous medium could? That is the question I find myself returning to, both as an artist and as a theorist.

[January 2025 - March 2026]