

Violaine Boutet de Monvel

Cybernetic subjectivities on a loop: From video feedback to generative AI

2023

<https://doi.org/10.25969/mediarep/21716>

Veröffentlichungsversion / published version
Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:

Boutet de Monvel, Violaine: Cybernetic subjectivities on a loop: From video feedback to generative AI. In: *NECSUS_European Journal of Media Studies*. #Cycles, Jg. 12 (2023), Nr. 2, S. 103–126. DOI: <https://doi.org/10.25969/mediarep/21716>.

Erstmalig hier erschienen / Initial publication here:

<https://necsus-ejms.org/cybernetic-subjectivities-on-a-loop-from-video-feedback-to-generative-ai/>

Nutzungsbedingungen:

Dieser Text wird unter einer Creative Commons - Namensnennung - Nicht kommerziell - Keine Bearbeitungen 4.0/ Lizenz zur Verfügung gestellt. Nähere Auskünfte zu dieser Lizenz finden Sie hier:

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Terms of use:

This document is made available under a creative commons - Attribution - Non Commercial - No Derivatives 4.0/ License. For more information see:

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Cybernetic subjectivities on a loop: From video feedback to generative AI

Violaine Boutet de Monvel

NECSUS 12 (2), Autumn 2023: 103-126

URL: <https://necsus-ejms.org/cybernetic-subjectivities-on-a-loop-from-video-feedback-to-generative-ai/>

Abstract

This paper raises an aesthetic bridge between pioneer video art and current machine learning art through the prism of instant feedback. It reinstates the former's real-time processual practices as a pertinent origin to reflect on the resolutely looped construction of cybernetic subjectivities. Given the latest developments of artificial intelligence, it infers that these now include technological agency. The article proceeds to compare the room left for human mastery in the recursive operations of both closed-circuit video setups and generative AI models. Taking into account their synthetic power, it ultimately proposes two general definitions of the artist's creative position within today's networked landscape.

Keywords

computer and machine vision, artificial neural networks, analogue and digital video synthesis, signal and data-driven art, real-time image processing

Introduction

Groundbreaking at the time of its advent, in 1965, with Sony's release of its first series of analogue recording systems intended for home use,[1] portable consumer video was notably deployed by contemporary artists to challenge human perception and self-awareness up until the mid-1970s. Among other key examples of this experiential take on the new medium, closed-circuit video installations and performances by American pioneers such as Bruce Nauman, Vito Acconci, Peter Campus, or Dan Graham typically played with the psychological drive of people to access the camera's field of vision, so as to confront their own image fed back in real time (Fig. 1). Indeed, unlike its film precursor whose viewing required prior photochemical development, analogue video offered instant feedback. The immediate display of the scanned environment onto a cabled monitor thus enabled artists

to leave the viewfinder and turn their own bodies, if not the spectators, into the live subjects of their electronic works. Consequently, early art historical accounts of such canonical use of video originally approached the notion of cybernetic subjectivities from a sole anthropological perspective, putting an emphasis on how human experience was influenced, shaped, and transformed by this particular technology. In return, it rather limited the latter's agency to mirroring that of people.[2]



Fig. 1: *Present Continuous Past(s)*, 1974, closed-circuit video installation, variable dimensions, John Gibson Gallery, New York. Photograph by Harry Shunk. Image courtesy of the Estate of Dan Graham.

This artist or viewer-centred perspective ultimately led art theorist Rosalind Krauss to institute, in 1976, narcissism as '*the* condition for the entire genre'.[3] She was then tackling and subverting altogether the modernist agenda on the wane: namely, art critic Clement Greenberg's principle of medium specificity,[4] whose pure formalism had already been confounded by the mixed-media, expandable, and life-anchored aesthetics of pop art, minimalism,[5] and the happenings[6] that emerged at the turn of the 1960s. In the late 1970s, however, video art entered a new phase. It generally shifted away from real time to embrace postproduction instead. The progressive introduction of computer-assisted editing throughout the following decade made this all the more feasible.[7] Yet, it also signed the genre's presumed death by antiquating the ontological differences between film and video, the former at this point more easily accessing the latter's image processing tools by ways of digitisation.[8] This eventually led to a whole new era: exhibited cinema, whose openly narrative, multi-projection installations rose to prominence in the early 2000s.[9]

Conversely, looking back at the inception of video art with the idea of broadening its chiefly human focus, so as to better acknowledge the original agency of any technology down the road, would certainly dig up more fortunate legacies than such professed doom in the 21st century. As art historian Yvonne Spielmann pointed out in 2005: 'It is astonishing that the debate on media theory has simultaneously declared more or less unanimously that video is obsolete by virtue of digitization.'^[10] In complete agreement with her, and with this extra objective in mind, it should be first noted that all these pioneers' closed-circuit video setups actually turned the space between the cameras and their control monitors into the arenas of a double perceptual experience: that of the participant – be it the performer, or the spectator – and that of the machine, both affecting one another's sight through live interaction. After all, the term *videre* comes from the Latin verb *videre* and means 'I see'. Alongside such canons of burgeoning media arts, concurrent investigations from artists and even scientists into strict video feedback were quick not only to acclaim the medium's dynamic quality akin to a living organism, but also to bring it – rather than our own selves – to the forefront of their processual practices.

Among others, since the late 1960s, this is notably the case of Steina and Woody Vasulka, who respectively made the exploration of both analogue and digital synthesis their lifetime work. In doing so, they distinctively paved the way onto the pendant idea that the technology itself had an agency of its own, to which they could relegate some of theirs to 'get rid of the supremacy of the human eye, the inherited modes of perception, and to reach an alternative (let's say "noncamera" or "non-human") point of view'.^[11] *mise en abyme* Briefly, there are two types of video feedback. The first consists of an optical loop by pointing a camera directly at its viewing monitor (Fig. 2), and optionally wiring them to further image processing tools (processor, keyer, synthesiser, etc.). The second involves no camera: it is a system loop that uses just a monitor connected to (at least) a mixer whose output signal is immediately fed back in (in this way, external referents are not required to create audiovisual content). Each technique has the power to synthesise an endless, intricate, always changing, and highly responsive *mise en abyme* of the outer installation and/or the setup's inner circuitry, in a more or less sustained dialogue with the maker.^[12]



Fig. 2: *Video Feedback*, ca. 1983, optical feedback setup, variable dimensions, Exploratorium, San Francisco. Photograph by Susan Schwartzberg. Image courtesy of the artist.

While the blinding developments of artificial intelligence (AI) over the last decade have now greatly marginalised human perception by arguably reversing our prosthetic relationship to technology, it only takes a theoretical leap so that the latter node be finally granted its own implicit kind of sentience in the grand cybernetic feedback loop.[13] Contemplating this possibility, cognate notions such as Refik Anadol's or Trevor Paglen's *machine hallucination*,[14] Grégory Chatonsky's *artificial imagination*,[15] and Maurice Benayoun's *art-subject*[16] have emerged from ongoing artistic experiments with generative AI models – especially the likes of GANs (Generative Adversarial Networks), first introduced by computer scientist Ian J. Goodfellow and his colleagues in 2014.[17] As part of a research that aims at bridging pioneer video art to current machine-learning art through the prism of both mediums' real-time agency and synthetic capacity, this paper thus proposes to revisit the interplay of so-connected cybernetic subjectivities from the technology's point of view, i.e. with its self-generating cycles as a starting point.

To do so, this study will confront the room left for human intervention, participation, and control in the emergence of images, as well as evaluate the sophistication of the behaviours displayed by both feedback and feed-forward systems. The argumentation will revolve around key examples of practices and their underlying concepts taken as much from the past (Steina, Woody Vasulka, James P. Crutchfield) as from the present (Refik Anadol, Trevor Paglen, Maurice Benayoun, Grégory Chatonsky). The bridge raised between yesterday's ventures into closed-circuit video and today's so-called recursive aesthetics will further reinstate the former as a pertinent origin for reflecting on the specificities of ma-

chine learning: namely, its own looped construction of subjectivity through the perpetual, although convoluted bond it maintains with our very existence, now processed as data rather than just signals. In this regard, this proposition as a whole readily follows up on art historian Ina Blom's technocratic rewriting of early video art, as she originally pondered in 2013: 'What if we could in fact imagine an opening up of the type of narrative where all agency or power of effectuation is automatically assigned to artists or works of art?'[18]

Through the looking glass: AI training out of human data

While the collection of online data began as soon as social media platforms emerged following the advent of the World Wide Web, in 1993, their exploitation became more pervasive and sophisticated with the rise of MySpace (launched in 2003) and especially Facebook (in 2004). By allowing people to create detailed profiles, share personal information, connect with others, and engage in various virtual activities, both forums started backing up a vast amount of human inputs, which they would further reemploy – or 'feed-forward', to quote media theorist Mark Hansen's dedicated term[19] – for mainly consumerist purposes such as targeted advertising and improved user experience. Indeed, it is by deducting from such recursive operations that he came up with the following theory in 2014: 'conscious experience of twenty-first-century media increasingly occurs as the result of a complex compositional process involving digital techniques of data-gathering and granular synthesis that facilitate the "feeding-forward" of multiple experiential sources into a potential future synthesis within consciousness'.[20]

Speaking of which, from a technological angle this time, this is only by the turn of the 2010s that the utilisation of so-aggregated digital contents to implement machine-learning systems progressively gained traction. One significant milestone in this area was the release of large-scale datasets like ImageNet[21] in 2009. Its labelling of millions of pictures has played a crucial role in advancing computer vision to this day – more specifically the image recognition and object detection tasks that are instrumental to training discriminative AI models via deep-learning techniques.[22] This field of computer science, which has a wide range of practical applications across industries (healthcare, automotive, manufacturing, agriculture, entertainment, etc.), focuses on enabling various types of software to understand, interpret, and extract meaningful information from visual records. It is about developing always more efficient algorithms that basically aim at replicating the human ability to perceive, analyse, and make sense of the world through the eye.[23] These discriminative capabilities are what ultimately fuelled the consecutive development of the latest generative AI models, starting with GANs in 2014, variations of which have since given rise

to many popular deepfake applications, notably to swap faces in digital photographs or videos.[24]

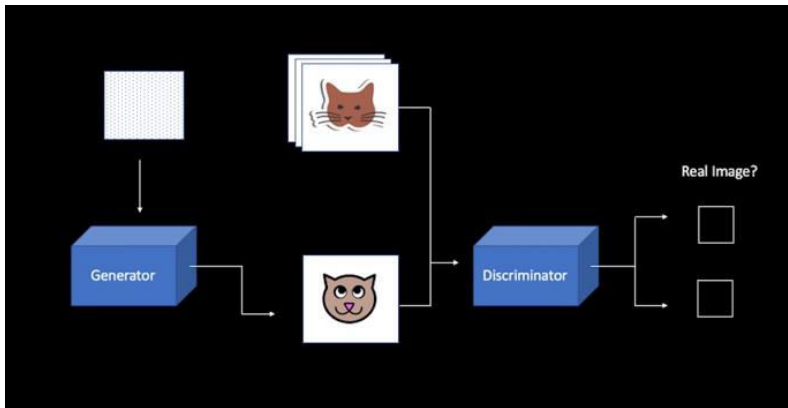


Fig. 3: GAN. Source: Lei 2019. Image courtesy of Nin Lei.

Briefly, a GAN consists of an adversarial feedback loop between two neural networks running concurrently (Fig. 3). At one end, a generator tries to create realistic data from random noise; at the other end, a discriminator tries to distinguish between these fake, synthetic outputs and the actual images coming from its training set based on one or several predetermined object categories. Iterative interaction between both algorithms keeps driving the learning process, until the generator finally passes the discriminator's test by successfully deceiving it. Their overall functioning is therefore recursive in the sense that fed-back information is repeatedly used to produce new content. Last but not least, another pivotal moment for the ever-growing integration of public data into AI training came with the widespread adoption of large language models like GPTs (Generative Pre-trained Transformers), introduced in 2018 by Open AI. This company further used them to develop text-to-image models such as DALL-E (released in 2021) and DALL-E 2 (in 2022).[25] These can generate images following user prompts, i.e. words and sentences that attempt to describe what is in fact to be statistically calculated, yet this time without the explicit adversarial setup of a GAN (a different deep-learning architecture referred to as diffusion is at play, and has since broadly superseded the former). Among other available AI models that work in similar ways, MidJourney[26] and Stable Diffusion[27] were also launched in 2022, before undergoing consecutive versions as well.

In other words, and to get back to the topic of data collection, potentially anything from simple hashtags to audiovisual content that people post online – be it their own creation or someone else's – can eventually become resources for machine-learning technologies. This

does not just attest to the mirror-like, although complex relationship that ineluctably binds the potential of generative AI, if not necessarily to the imagination of a single human agent then to the overall realm of human cultural inputs (in this case, whatever is digitised and rendered machine-readable for further algorithmic discrimination). It also implies that such recursive operations, and therefore aesthetics, can – to some extent – be curated accordingly (through carefully-adjusted datasets, or prompts, among other manageable variables). While these highly-automated systems function at speeds and scales that vastly exceed our perceptual capabilities, it is thus possible to *steer*[28] their outputs even though we cannot predict them, nor would achieve them alone at such pace anyway. Finally, this corroborates Hansen’s take on agency in the 21st century, which he urges us to reconceptualise ‘as the effect of global patterns of activity in networks, where absolutely no privilege is given to any particular individual or node’.[29]

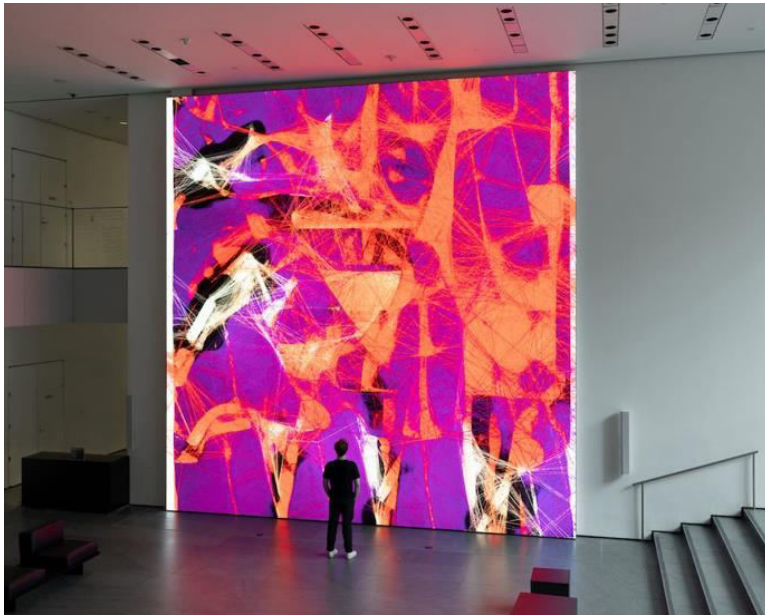


Fig. 4: *Unsupervised*, 19 November 2022 – 29 October 2023, exhibition view, The Museum of Modern Art, New York. Photograph by Robert Gerhardt. © The Museum of Modern Art.

Appropriately, along with these new tools came a renewed aesthetic discourse about the unprecedented empowerment that such computational technologies display, as well as the peculiar nature of their production. Visual culture theorist Antonio Somaini, for instance, referred to their advent as a tectonic shift,[30] further pondering in 2021: ‘What do such images represent, what kind of agency do they have, what is their temporal status, and how do they mediate our visual relation to the past, the present, and the future?’[31] Following

up on these now decade-long advances in generative AI since the introduction of GANs, many artists have begun exploring their creative potential by incorporating them into their art making through a sustained practice, often in collaboration with AI developers and engineers, subsequent exhibitions, as well as – for some – extensive writings. Refik Anadol, among others, opened his high-profile solo show *Unsupervised* at MoMA in 2022 (Fig. 4).[32] It consisted of a single large-scale projection installation, which reinterpreted in real-time the publicly available data of the museum’s comprehensive collection via the live, consecutive run of two artificial neural networks. The first served to process the entire digitised archive of MoMA through the algorithm StyleGAN2 that was introduced by NVIDIA researchers in 2020.[33] This powerful deep-learning model notably allows for high-quality image synthesis as well as fine-grained control over the generated content, which the artist had configured to remain on the abstract side. The other concurrently delved into these digital records through the Latent Space Software, a program custom-developed by his studio since 2017.[34]

In the context of machine learning, the latent notion refers to the compressed representation of previously raw data in a multi-dimensional space that retains only their most important features. Accordingly, Anadol’s dedicated software enables us to perceive what are in fact numerous categorical interpolations seamlessly drawn by computer vision, in this case, within the museum’s wide array of artworks. Following the multiple paths of their common denominators (average colour palette, dimensions, etc.), the program thus proceeded without explicit reference to any, for it was otherworldly synthesis out of them all that was ultimately at play. Finally, site-specific input captured via sensors from the immediate surroundings such as changes in light, movement, acoustics, and temperature was integrated as well to affect the resulting, continuously shifting, dream-like, hazy imagery. In doing so, this added environmental feedback discreetly included the audience, beyond its awareness, into the demiurgic looping act of the otherwise imposing AI setup (its side machinery to process everything live took up the space of a small, enclosed room).

Through the rear-view mirror: Machine vision beyond the human eye

Anadol’s *Unsupervised* is actually one of the latest exhibitions through which this pioneer of data-driven art has been prospecting the realm of *machine hallucination* since 2016, training his generative AI models on what he refers to as ‘collective visual memories’,[35] i.e. overall human cultural inputs thus turned into datasets. As a matter of fact, the DeepDream project developed by Google in 2015 played a significant role in popularising this concept over the past decade.[36] This program used CNNs (Convolutional Neural

Networks) – originally designed by computer scientist Yann LeCun in the late 1980s[37] – to find and iteratively enhance motifs in any submitted image, which would put forth surreal, psychedelic variations of it ‘presented as a dream belonging to the machine itself.’[38] It comes with no surprise then that this theme has also recurred in the quest of other early adopters and thinkers of machine-learning technologies in general, and GANs in particular, ‘in order to explore and reveal *the altered states of machine vision*’.[39]

Trevor Paglen stands among them, well at the helm, with his dedicated series titled *Adversarially Evolved Hallucination* that was exhibited in *A Study of Invisible Images* at Metro Pictures, in 2017.[40] The show consisted in the artist’s first body of works to emerge from his ongoing research into computer vision, which was realised in part with the aid and insight of AI developers during a residence at Stanford University. To make the different prints of this series, he initially created massive datasets with pictures borrowed from folk literature, psychology, poetry, and other downright unusual resources for training AI models. For each, he then carefully curated custom taxonomies such as ‘humans’, ‘omens’, ‘monsters’, or ‘dreams’,[41] overall related to power and irrational things. Finally, he set up GANs to recognise patterns from these object categories, and further synthesise images associated with their respective corpses.



Fig. 5: *Human Eyes (Corpus: The Humans)*, 2017, dye sublimation on aluminum print, 48 x 60 inches, from the series *Adversarially Evolved Hallucination*, 2017. Image courtesy of the artist; Altman Siegel, San Francisco; and Pace Gallery. © Trevor Paglen.

Among these, for instance, The Humans (Fig. 5) instructed the neural networks to detect and adversarially generate solely anthropomorphous features – although the original dataset may not have consisted of any such literal, or objective representations. Yet, what Paglen refers to as *machine hallucination* does not only come from the fact that his GANs were trained to see certain object categories where they were not necessarily, or accurately, figured from a depictive angle, at least. It also comes from the fact that he did not wait for the generators to fool their discriminators, which they would have done eventually by producing somewhat realistic outputs based on otherwise rather imaginative resources. What he did instead was select many synthetic images throughout the learning process, before it was complete, so that the resulting eerie prints of his series turned out beyond uncanny, at the threshold of abstraction, although some aspects in them may be vaguely recognisable.[42]

This notion of machine hallucination understood as the outer limits of computer vision, in this case when oddly informed from the start, is directly related to the artist's concurrent inquiry into the exact implementation of training sets composed of personal data, which, he has observed, raises major ethical concerns. He made these the topic of extensive writings, notably in collaboration with Kate Crawford, a leading scholar in the socio-political implications of machine-learning technologies.[43] As mentioned earlier, generative AI models such as GANs, GPTs, and CNNs can be partly fed out of user-generated contents, which have produced still unresolved issues regarding people's privacy, especially across social media platforms; also, the labelling required to build datasets – i.e. prior to feeding them to any machine-learning system – transpires huge human biases when it comes to describing persons in line with gender, ethnicity, and age, among other common classification factors. Far from being automated, this overall job relies on massive human labour, ranging from AI developers to clickers, the latter being responsible for manually annotating various types of data. Consequently, 'training sets reveal the historical, geographical, racial, and socio-economic positions of their trainers'[44] – and so does any ensuing synthesis, as imperceptibly as it may.

The need to fathom – in order to find a way to thwart – such seemingly almighty technologies and their convoluted, ideological ties with our very existence further led Paglen to reinvest the landscape of machine vision. Not unlike Anadol, he did so by prospecting its hallucinatory power (as opposed to its capacity for deepfake or confounding realism), yet this time wielded like an act of resistance rather than a purely aesthetic one. In the context of engineering, machine vision specifically refers to an industrial subset of computer vision

enabling automated systems to detect visual characteristics of objects to better ensure product quality, process control, reduced human error, and efficiency. In the artist's rhetoric, it encompasses the equivalent treatment of people by other people as commodities under the totalising surveillance of algorithmic setups, whether online or in real life, for both capitalist and policing schemes. Questioning the even remote possibility of hiding from such insidious objectives, instantly operated by programs way beyond our perceptual reach, made him ultimately realise that 'if we want to understand the invisible world of machine-machine visual culture, we need to unlearn how to see like humans'.[45]



Fig. 6: The VASULKAS / Steina: Machine Vision, Woody: Descriptions, 21 October – 26 November 1978, exhibition view, Albright-Knox Gallery, Buffalo. Photograph by Kevin Noble.

The latter is precisely what video art pioneer Steina set out to do 50 years prior, so as to get away from a sole anthropological perspective and investigate the true agential realm of her medium of choice instead, i.e. beyond the limitations of the eye or the viewfinder used as its strict extension.[46] In 1975 she initiated *Machine Vision*, a series of kinetic sculptures that combined cameras and mirrors. It was presented complete for the first time at the Albright-Knox Art Gallery in 1978 (Fig. 6).[47] Among other closed-circuit video setups, her entire display centred around the installation *Allvision II* (1978). It consisted of two cameras placed on the opposite ends of a horizontally-rotating rod with a small mirror orb fixed in between them, whose reflections, thus scanned from both sides, were immediately fed back on two monitors cabled nearby. The overall experiment was an attempt to demonstrate the principle of a total point of view. By embracing, on the spot, the whole space in a distorted, panoptic way, which included not only the audience but also the other

live video sculptures, it signified, according to her, 'the awareness of an intelligent, yet not human vision'.[48]

Dialogue with the tools: Artificial morphogenesis through video synthesis

Out of many pioneers who have explored the techniques of video feedback since the mid-1960s, Steina particularly stands out because she did so neither to expand the reach of human perception in a prosthetic sense, nor for the purely aesthetic motives that can inform its most minimalist – optical or system – studies. Regarding the latter, artist William Gwin actually warned others against the contemplative trap of falling into the hypnotising, mandala-like imagery easily obtained from such analogue setups, i.e. if they were to be somewhat left to their own devices. As he reported to the National Center for Experiments in Television in 1972: 'Its prettiness can be so enticing that time and energy are destroyed without leading to any serious expression or work.'[49] (Note that the same could be said of any content so-generated with AI.) In lieu of beauty then, Steina created a path for an alternative approach, wholly dedicated to uncovering the full scope of video agency by interfering with its closed-circuit possibilities, most typically as one would play an instrument (she did the violin). In this regard, she is credited for having implemented feedback loops interpolating both image and sound signals, which enabled her to interact live with intrinsic electronic faculties comparable, under these circumstances, to synaesthesia: 'That's when we realized that there didn't have to be a camera – a voltage, a frequency could create an image.'[50]

She conducted all these experiments beside her partner, Woody Vasulka, also at the forefront of this peculiar line of research that resolutely took a major step aside from any consideration towards television. Indeed, a vast majority of early adopters and thinkers of video accounted for in art history considered their practice in opposition to the mass media. They generally used instant feedback as a weapon, whether actual or metaphorical, to challenge the otherwise strictly one-way broadcast system of communication, and therefore control.[51] As this medium's forefather Nam June Paik proclaimed following the advent of portable consumer video: 'Television has been attacking us all our lives, now we can attack it back.'[52] Unconcerned with these politics, the Vasulkas got rather invested in the notion of *dialogue with the tools*, which they together used to describe their respective studies, both analogue and later on digital: 'that's how we belong with the family of people who would find images like found objects. But it is more complex, because we sometimes design the tools, and so do conceptual work as well.'[53] His turned out even more didactic, as he diligently observed different video setups deploy the entire spectrum of their synthetic capacities, just as a biologist would watch micro-organisms under the microscope.



Fig. 7: *Time/energy Structure of the Electronic Image*, 1974-75, video studies, Albright-Knox Gallery, Buffalo. Photograph by Kevin Noble. Image courtesy of Buffalo AKG Art Museum, Digital Assets Collection and Archives.

Next to Steina's *Machine Vision* at the Albright-Knox Art Gallery, he presented, for that matter, *Time/energy Structure of the Electronic Image* (1974-75),[54] a series of eleven large panels showing sequences of video stills that described basic electronic waveforms such as sine, triangle, or square (Fig. 7). These abstract geometric shapes were all realised with a Rutt/Etra scan processor. Conceived in 1973 by live video pioneer Bill Etra and engineer Steve Rutt, this early analogue raster manipulation device integrated an oscillator circuit, an amplifier, as well as an internal feedback mechanism, which could thus produce an output signal and self-sustain it on a monitor.[55] It was actually one of several synthesisers developed by artists at the time in order to gain mastery over real-time image processing.[56] Vasulka's experiments, documented in these panels, meant to comprehensively demonstrate the capacity of this specific tool to create content out of its own noise, which opened, according to him, 'a new self-generating cycle of design within consciousness and the eventual construction of new realities without the necessity of external referents as a means of control'.[57] (Note that these words perfectly foreshadow Hansen's aforementioned theory.)

The last sequence, however, distinctly displayed the picture of a hand, which was placed into the loop via the addition of a VCR (Video Cassette Recorder). As he explained, it constituted 'the first input of conventional reality into a previously self-contained system of electronically generated and processed imagery, and should be understood as part of an

electronic process'.[58] Indeed, the fact that system feedback requires no other source than itself to enable video synthesis does not mean in the least that it repudiates altogether the artist's touch (which this hand reminds us of). The latter manifests via the manipulation of knobs and sliders, prospectively adjusting the parameters of the audiovisual modules to dynamically alter the waveforms (colour, brightness, contrast, motion, speed, etc.). The same goes with optical feedback, whose overall experience is equally tactile and thus far extends the limited realm of sight.[59] While Woody Vasulka favoured the former over the latter throughout his career, stressing the construction of images without a camera, his dialogues with the tools were never about leaving all agency to them, neither controlling them, but co-creating with them by learning how to steer their self-generating cycles so that to deploy the full range of their synthetic power. In the same show he presented further works realised with the Digital Image Articulator, also known as the Vasulka Imaging System, which was designed with computer scientist Jeffrey Schier in 1976.[60] Consisting of analogue components as well as a microcomputer, it marked the transition of his live video studies towards programming.

Before feeding-forward to the current conclusion of this comparative history, a final comment must be made. As specified above, video synthesis, no matter the technique, is not only highly but essentially reactive to human interplay performed via console switches. This likens, in a way, to the variables that constitute the possibilities of carefully curating training sets, or prompts with GANs and GPTs, respectively – although the latter's tactility is obviously limited to that of typing, or other human-computer types of interaction, and thus should not sum up their entire experience. In contrast to system feedback, however, these two generative AI models that artists have abundantly appropriated over the past decade fundamentally require external referents to create new data. They simply cannot proceed purely from random noise. Whether it relies on sight, touch, and/or literacy then, this perpetuated connection to human operability leads to the point of wondering if artificial subjectivity can ever be truly detached from our own. When it comes to tackling the former's original agency, the courrant analogy drawn between self-sustaining technologies and living organisms certainly keeps on attesting to such intricate mirroring, not only in the artistic field, but also in the scientific one.

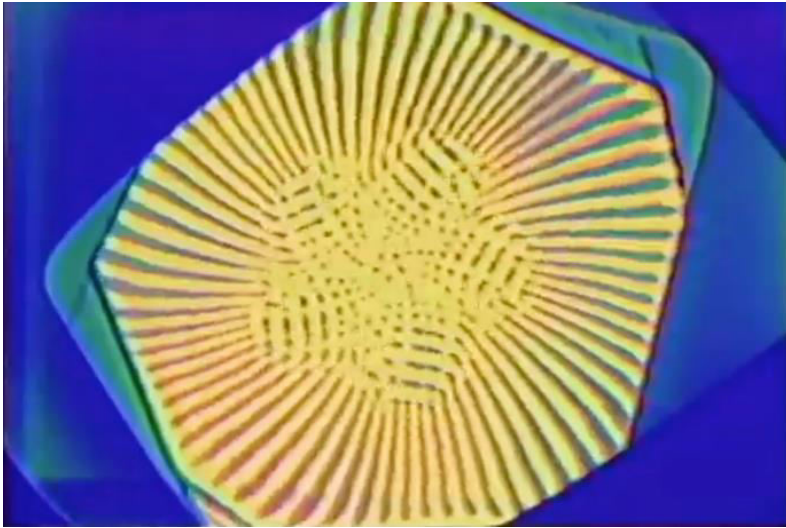


Fig. 8: *Space-Time Dynamics in Video Feedback*, 1984, U-matic video, colour, 16' Entropy Productions, Santa Cruz. Image courtesy of the artist.

Speaking of which, physicist James P. Crutchfield experimented with optical feedback within the frame of his Ph.D. regarding the emergence of order from chaos, which he defended at the University of California, Santa Cruz in 1983. Considering it 'an almost ideal test bed upon which to develop and extend our appreciation of spatial complexity and dynamical behavior',[61] he was notably interested in its potential to simulate biological morphogenesis, i.e. the self-structuration of natural forms.[62] These analogue studies resulted in a video compilation titled *Space-Time Dynamics in Video Feedback* (1984; Fig. 8),[63] which was realised by pointing a camera directly at its viewing monitor. Yet, in order to illustrate the organisational process of distinct complex systems, he had first to meticulously set up both machines at different distances and angles from one another. Far from being a completely self-sustained operation, animating the generated patterns into the desired directions also required the camera to be partially set in motion. This could be achieved optically, by playing with its zoom and focus; manually, by moving it around towards the screen; or mechanically, via kinetic devices specifically designed for such use (artist Skip Sweeney's one, illustrated in the introduction, is an example of those). Finally, the camera model that Crutchfield chose to work with (a Sony Trinitron HVC-2200)[64] – because of the sensibility of its tubes that were highly susceptible to burning if exposed to bright images for too long – typically enabled the original synthesis of blobby, amorphous forms, and further granular visual artefacts reminiscent of early cellular differentiation. Another piece of equipment with differing sensitivity would have produced alternative effects.

Towards a general redefining of artistry with generative technologies

The Vasulkas' notion of *dialogue with the tools* and its auxiliary one of co-creativity with the machine turned out to be 'a decisive criterion of the transition from analog to digital'. [65] Taking on these cues since the age of the internet, two current pioneers of data-driven art have been working towards a new understanding of artistry with 21st-century media that include, as defined by Hansen,[66] technological agency with regard to the phenomenal advances of generative AI – Maurice Benayoun is one of them. Along with architect Tobias Klein, he conceived the neuro-design prototype *Brain Factory* in 2016, enabling image synthesis via biofeedback (Fig. 9).



Fig. 9: *Brain Factory*, 2016, neuro-design installation, variable dimensions, Microwave Festival, 25 October 2018, Hong Kong. Image courtesy of the artists.

In its installed version, viewers are individually invited to wear an electroencephalography (EEG) headband to give shape to symbolic archetypes such as 'love' or 'freedom' through brain-computer interaction (BCI).[67] This allows them to tentatively sculpt – not with touch, but directly with their minds – an AI-generated matter further displayed on a screen. Its virtual, perpetually mutating, aqueous substance is synthesised in real time, based on their respective brain activity: namely, EEG signals interpreted as either positive or negative reactions to the dynamic visual abstractions concurrently converted by algorithms. Because actual control over the artificial shapes completely evades the participants, who can only think hard about their given concept, such interaction from brainwaves straight to waveforms simply leaves no room for human mastery. Instead, it

implements a dialogue with the machine that rather consists in people discriminatively gauging the former's live processing of their inner thoughts: 'We could consider that the assessment process is a form of dynamic curatorship.'^[68] Each unique outcome has ultimately been reified into a 3D-printed sculpture and documented along with every iteration of *Brain Factory*.

This experiment, which originally explored the notion of co-creativity or the role of neural networks – both human and artificial – in generating artistic outputs, led Benayoun to the concept of art-subject, i.e. art itself elevated to the status of a subject of its own. To introduce this idea, which he did with art historian Tanya Ravn Ag in 2020, he first had to define agency applied to works of art: 'Agency becomes the property of the artwork that may passively accept inputs, act by reflex, or behave according to *artificial intentionality*, empowering it with a potential behavior by design, corresponding to the artist's intention.'^[69] In other words, playing on the agential level of the technology behind a given piece of art may increase the latter's sentience, yet always in dialogue with external inputs, starting with the viewers themselves: 'We call art-subject vs. art-object the complex sentient system that makes an artwork a cognitive being able to act, react, and communicate with its public.'^[70]

Understood that within the openly immersive context of *Brain Factory*, the actual shaping has been somewhat relegated to the audience and machinery, the part left to the artist was then exclusively that of *designing* the overall setup, which would alone enable both human and artificial subjectivities to express their joint creativity. Speaking of which, net art pioneer Grégory Chatonsky also came up with such an appreciation of artistry within 21st-century networks, but in his case aside from the interactivity that his exhibitions rarely allow. In 2009 he introduced the notion of *artificial imagination*,^[71] then another step towards the acknowledgement of technological agency, yet prior to the advent of GANs and text-to-image models that he has since extensively explored. Indeed, his project *Capture*,^[72] initiated that year, is one of his earliest works delving into data synthesis, using information scraped from the Web to continuously generate electronic music and further audiovisual trivia surrounding a fictitious rock band.



Fig. 10: *The Kiss 3*, 2022, AI-generated video, 2'41. Image courtesy of the artist.

Revisiting large corpuses of human culture (such as music, in the previous example) has actually constituted a leitmotif throughout his career, starting with the seventh art, to which he has dedicated a lifelong series titled *After the Cinema* (2001-ongoing).[73] For instance, *The Kiss 3* (2022) is one of many artworks through which he has specifically tested the capacity of generative technologies to reinterpret Alfred Hitchcock's 1954 film *Rear Window* (Fig. 10). In order to synthesise this video, which consists of a highly condensed, hallucinatory, almost ubiquitous version of every single scene wrapped together in less than three minutes, he prompted the entire script of the movie to DALL-E 2.[74] Such investigations made him not only recognise the latent space as the cradle of *artificial imagination*, but also deduct that snippets of human culture constituted another one among us as well: 'There is a statistical latent space in computers, just as there is a cultural latent space in each of us.'[75] This inspired him in 2022 to venture the following statement regarding the possible role of artists in mediating both: 'Artificial imagination shouldn't be just used, but experimented in the in-between that it is, between the statistical latent space of the artificial neural networks and the cultural one of our own neurons, so that the powerful illusion of reflexivity finally crumbles.'[76]

Conclusion

On these thoughts, I would like to propose, for subsequent research, two general definitions of the artist's creative part within today's networked landscape, which opens yet intertwines the agential scope of all involved cybernetic subjectivities, be they human or non-human. The first means to differentiate artistry from engineering within the field of telecommunications, keeping in mind mathematician Claude E. Shannon's 1948 'schematic diagram of a general communication system' that describes how information passes from

transmitter to receiver, gathering internal and external noise on the way.[77] It goes as follows: *an artist strives to magnify the noise within and via any given medium, as opposed to the engineer who seeks to suppress it.* This may apply broadly (which is not to say entirely) to media arts, as well as modernism and its transitional phase towards postmodernism in the mid-20th century. Indeed, many avant-gardes from Impressionism to the emergence of the happening have attempted to introduce contingency into their works – to the point of including the audience’s random behaviour as constituent noise, in the latter case. As artist Allan Kaprow specified in 1966: ‘All the elements – people, space, the particular materials and character of the environment, time – can this way be integrated.’[78]

My second definition further considers the original sentience displayed by the latest generative AI models; I also intend its reach to potentially encompass practices beyond sole machine learning. It reads as follows: *an artist seeks to steer our senses towards an outsider perception, by orchestrating a dialogue between different networked subjectivities, inclusive of technology and nature.* This follows up on art critic Nicolas Bourriaud’s relational aesthetics that he theorised in 1998, by ‘judging artworks on the basis of the inter-human relations which they represent, produce or prompt’.[79] Although missing in this paper, the current efforts to highlight minorities is certainly implied, all the more so that social media platforms have been playing a crucial role in making their experiences more widely known. That being said, by *nature*, which I temporarily oppose to *technology*, I refer to the non-human agents that constitute the animal, vegetal, and mineral world. Because they partly inform the motifs of Grégory Chatonsky, as well as Pierre Huyghe or Hito Steyerl, among others pioneers of data-driven art, their works now call for more in-depth analysis.

Finally, a firmer media-archaeological approach may help in the future to better understand the distinction and overlap between hardware and software, which are understated in this account that bridges analogue to digital technologies more theoretically. This could also pave the way onto a wider corpus related to artistically diverting such tools, i.e. aside from their designed commercial use. Other than video feedback and generative AI, practices pertaining to glitch and creative coding first come to mind in this specific line of research.

Author

Violaine Boutet de Monvel currently teaches in the Film and Media Studies Department of Université Sorbonne Nouvelle - Paris 3, where she has undertaken a Ph.D. about the forms and legacy of feedback in video art, from cybernetics to artificial intelligence. She has de-

signed two classes approaching the topic from a media-archaeological stance: one anchored in aesthetics, the other in television history. She previously taught modern and contemporary art at Université Paris 1 Panthéon-Sorbonne. She has presented her research at various venues, most recently Università Ca' Foscari Venezia, Universitatea Politehnica din Buarești, and École Normale Supérieure. She was also the recipient of grants and residencies at NYU, Brown, and Emory University. She is the author of monographs and exhibition catalogues on several artists (Benjamin Sabatier, Grégory Chatonsky, Pierre Ardouvin, Amélie Bertrand, etc.), as well as many articles in art magazines (*ArtReview*, *Aperture*, *Frieze*, *Flash Art*, *Artpress*, etc.).

References

- Benayoun, M. and Ag, T. 'After the Tunnel: On Shifting Ontology and Ethology of the Emerging Art-Subject' in *Proceedings of the 26th International Symposium on Electronic Art*, 29-40. Montreal, 2020.
- Blom, I. 'The Autobiography of Video: Outline for a Revisionist Account of Early Video Art', *Critical Inquiry*, 39, no. 2, Winter 2013: 276-295.
- _____. *The autobiography of video: The life and times of a memory technology*. Berlin: Sternberg Press, 2016.
- Bourriaud, N. *Relational aesthetics*, translated by S. Pleasance and Woods. Les presses du réel. 1998. Reprint, Dijon, 2002.
- Boutet de Monvel, V. 'L'art vidéo pionnier sous le prisme de l'agentivité : quelle(s) icône(s) pour un médium dont la spécificité était le feedback?' in *La circulation des images en Europe*, edited by L. Saint-Raymond. Paris: Mare et Martin, 2023: 259-277.
- Branson Gill, J. 'Video: State of the Art'. New York: The Rockefeller Foundation, 1976.
- Cathcart, L. (ed.) *Vasulka: Steina, Machine Vision/Woody, Descriptions*, exhibition catalogue. Buffalo: Buffalo Fine Arts Academy, 1978.
- Crutchfield, J. 'Space-Time Dynamics in Video Feedback', *Physica*, 10, no. 1-2, January 1984: 229-245.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. 'Generative Adversarial Nets' in *Proceedings of the 27th International Conference on Neural Information Processing Systems*. Cambridge: MIT Press, 2014: 2672-2680.
- Greenberg, C. 'Modernist Painting', *Arts Yearbook*, no. 4, 1961: 103-108.
- Gwin, W. 'VIDEO FEEDBACK: How to Make It, An Artist's Comments on Its Use; A Systems Approach'. San Francisco: National Center for Experiments in Television, 1972.
- Hansen, M. *Feed-forward: On the future of twenty-first-century media*. Chicago: University of Chicago Press, 2014.
- High, K., Hocking, S., and Jimenez, M. (eds) *The emergence of video processing tools: Television becoming unglued*. Bristol: Intellect, 2014.
- Joselit, D. *Feedback: Television against democracy*. Cambridge: The MIT Press, 2007.

CYBERNETIC SUBJECTIVITIES ON A LOOP: FROM VIDEO FEEDBACK TO GENERATIVE AI

Judd, D. 'Specific Objects', *Arts Yearbook*, no. 8, 1965: 74-82.

Kaprow, A. *Assemblage, environments & happenings*. New York: Harry N. Abrams, 1966.

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., and Aila, T. 'Analyzing and Improving the Image Quality of StyleGAN' in *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020: 8110-8119.

Krauss, R. 'Video: The Aesthetics of Narcissism', *October*, 1, Spring 1976: 50-64.

LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. 'Gradient-Based Learning Applied to Document Recognition', *Proceedings of the IEEE* 86, no. 11, November 1998: 2279-2324.

Lei, N. 'Generative Adversarial Network Technology: AI Goes Mainstream', IBM, 17 September 2019: <https://www.ibm.com/blogs/systems/generative-adversarial-network-technology-ai-goes-mainstream/>.

Manovich, L. 'Computer Vision, Human Senses, and Language of Art', *AI & Society*, no. 36, 2021: 1145-1152.

McLuhan, M. *Understanding media: The extensions of man*. Cambridge: The MIT Press, 1964.

Paglen, T. 'Invisible Images (Your Pictures Are Looking at You)', *The New Inquiry*, 8 December 2016: <https://thenewinquiry.com/invisible-images-your-pictures-are-looking-at-you/>.

Paglen, T. and Crawford, K. 'Excavating AI: The Politics of Images in Machine Learning Training Sets', 2019: <https://excavating.ai/>.

Païni, D. 'Le retour du flâneur', *Artpress*, March 2000.

Rush, M. *Video art*. London: Thames and Hudson, 2003.

Shannon, C. 'A Mathematical Theory of Communication', *Bell System Technical Journal*, 27, no. 3, July 1948: 623-656.

Somaini, A. 'Film, Media, and Visual Culture Studies, and the Challenge of Machine Learning', *NECSUS European Journal of Media Studies*, no. 10, Autumn 2021: 49-57.

_____. 'On the Altered States of Machine Vision. Trevor Paglen, Hito Steyerl, Grégory Chatonsky', *AN-ICON. Studies in Environmental Images*, no. 1, 2022: 91-111.

Spielmann, Y. *Video: The reflexive medium*, translated by A. Welle and S. Jones. Cambridge: The MIT Press, 2008 (orig. in 2005).

Vasulka, W. and Nygren, S. 'Didactic Video: Organizational Models of the Electronic Image', *Afterimage*, 3, no. 4, October 1975: 9-13.

Wiener, N. *Cybernetics or control and communication in the animal and the machine*. Cambridge: The MIT Press, 1961 (orig. in 1948).

Youngblood, G. *Expanded cinema*. New York: E.P. Dutton, 1970.

Notes

- [1] Sony's inaugural 1965 domestic video system consisted of a camera (VCK-2000) and a separate reel-to-reel recorder marketed with or without a built-in monitor (CV-2000 and TCV-2010, respectively). They were both mains-powered and transportable in a hard suitcase. Two years later, in 1967, the Japanese firm released a subsequent series of battery-operated units known as the Portapak. See https://www.smecc.org/sony_cv_series_video.htm (all links accessed on 31 August 2023).
- [2] See Boutet de Monvel 2023.
- [3] Krauss 1976, p. 50.
- [4] See Greenberg 1961.
- [5] See Judd 1965.
- [6] See Kaprow 1966.
- [7] CMX Systems (later Avid Technology) actually introduced the very first non-linear video editing console as early as 1971 (CMX-600), which used computer interfaces to control the playback and editing of video. Only six were manufactured at the time for professional use in television studios. See <https://www.wci.nyc/the-first-non-linear-edit-system/>.
- [8] Michael Rush, for instance, stands among the first art historians to have conjectured the end of video art at the turn of the 2000s: 'Already Video art has become a subsection of Filmic art, a term better suited to the actual practice of most media artists today.' See Rush 2003, p. 165.
- [9] See Païni 2000.
- [10] Spielmann 2005, p. 13.
- [11] Lenka Dolanova with Woody Vasulka in High & Miller Hocking & Jimenez 2014, p. 276.
- [12] See Gwin 1972.
- [13] Cybernetics is an interdisciplinary field of study founded by mathematician Norbert Wiener in 1948. It focuses on understanding the natural principles of communication, control, and regulation via feedback loops within any living organism ('the animal') or between different ones, further implementing them into artificial systems ('the machine'). See Wiener 1948.
- [14] See Paglen 2016.
- [15] See <http://chatonsky.net/category/corpus/ami/>.
- [16] See Benayoun & Ag 2020.
- [17] See Goodfellow, Pouget-Abadie, Mirza, Xu, Warde-Farley, Ozair, Courville & Bengio 2014.
- [18] Blom 2013, p. 277. This paper served as the introduction of a subsequent book on video agency. See Blom 2016.
- [19] See Hansen 2014.
- [20] *Ibid.*, p. 45.
- [21] <https://www.image-net.org/>.
- [22] See Somaini 2021, p. 50.
- [23] See Manovich 2021, p. 1146.
- [24] See Lei 2019.
- [25] <https://openai.com/dall-e-2>.
- [26] <https://www.midjourney.com/>.
- [27] <https://stability.ai/stable-diffusion>.
- [28] This idea of steering the outputs of a given technology's self-generating cycles, which certainly compliments the etymology of the term cybernetics that comes from the Greek word *kybernetes* (meaning 'steersman' or

CYBERNETIC SUBJECTIVITIES ON A LOOP: FROM VIDEO FEEDBACK TO GENERATIVE AI

'governor'), was suggested through the consecutive sessions of two video feedback workshops, which the author took part in during the spring of 2023. One of them, Introduction to Analog Video Signals, was conducted by Andrei Jay from the artist collective Phase Space (now Phase Shift); and the other, History of Video Art Techniques, by Hunter Lombard, also known as Cable Visions. They are both warmly thanked for their insight and generosity, along with their collaborators Paloma Kop and Jonathan Sims. See <https://youtu.be/VY4tbBpCGAA?si=m9kIQ4DVMo-52aQk>.

- [29] Hansen 2014, p. 2.
- [30] See Somaini 2021, p. 49.
- [31] Ibid., p. 53.
- [32] <https://www.moma.org/calendar/exhibitions/5535>.
- [33] See Karras, Laine, Aittala, Hellsten, Lehtinen & Aila 2019.
- [34] See <https://refikanadolstudio.com/projects/unsupervised-machine-hallucinations-moma/>.
- [35] Ibid.
- [36] See <https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html>.
- [37] See LeCun, Bottou, Bengio & Haffner 1998, and LeCun 2019.
- [38] Somaini 2021, p. 52.
- [39] Somaini 2022, p. 101.
- [40] See <https://www.metropictures.com/exhibitions/trevor-paglen4>.
- [41] See <https://paglen.studio/2020/04/09/hallucinations/>.
- [42] See Somaini 2022, p. 104.
- [43] See Paglen 2016, and Paglen & Crawford 2019.
- [44] Paglen 2016.
- [45] Ibid.
- [46] See http://www.vasulka.org/Steina/Steina_MachineVision/MachineVision.html.
- [47] See <https://buffaloakg.org/blog/throwback-thursday-vasulkas-steina-machine-vision-woody-descriptions>.
- [48] Steina in Cathcart 1978, p. 9.
- [49] Gwin 1972, pp. 4-5.
- [50] Steina in Cathcart 1978, p. 23.
- [51] See Joselit 2007.
- [52] Nam June Paik in Youngblood 1970, p. 302.
- [53] Branson Gill 1976, p. 22.
- [54] Woody Vasulka in Cathcart 1978, p. 15.
- [55] See <https://www.fondation-langlois.org/html/e/page.php?NumPage=456>.
- [56] Others notably include Nam June Paik and engineer Shuya Abe's Paik/Abe Video Synthesiser designed in 1969-70, Stephen Beck's Direct Video Synthesiser in 1970-71, and Dan Sandin's Sandin Image Processor in 1971-73. See High & Miller Hocking & Jimenez 2014, pp. 117, 154, and 467, respectively.
- [57] Woody Vasulka in Vasulka & Nygren 1975, p. 9.
- [58] Ibid., p. 13.
- [59] This notion of tactility to partly inform the experience of video feedback, besides the realm of sight, echoes – to some extent – media theorist Marshall McLuhan's similar concept to describe the multisensory experience of watching television in its early low-resolution days: "The TV image requires each instant that we "close" the spaces

CYBERNETIC SUBJECTIVITIES ON A LOOP: FROM VIDEO FEEDBACK TO GENERATIVE AI

in the mesh by a convulsive sensuous participation that is profoundly kinetic and tactile, because tactility is the interplay of the senses, rather than the isolated contact of skin and object.' See McLuhan 1964, p. 314.

- [60] See <https://www.fondation-langlois.org/html/e/page.php?NumPage=457>.
- [61] Crutchfield 1984, p. 229.
- [62] See Blom 2016, p. 69.
- [63] See <https://www.youtube.com/watch?v=B4Kn3dijMCE>.
- [64] See Crutchfield 1984, p. 232.
- [65] Spielmann 2005, p. 206.
- [66] See Hansen 2014.
- [67] See <https://benayoun.com/moben/2017/07/04/the-brain-factory-trailer/>.
- [68] Benayoun & Ag 2020, p. 35.
- [69] Ibid., p. 34.
- [70] Ibid., p. 36.
- [71] See <http://chatonsky.net/category/journal/ima/>.
- [72] See <https://chatonsky.net/capture/>.
- [73] See <http://chatonsky.net/category/corpus/after-cinema/>.
- [74] See <http://chatonsky.net/kiss-3/>.
- [75] <http://chatonsky.net/postmodern/>.
- [76] <http://chatonsky.net/exp-with-artificial-imagination/> (re-translated by the author).
- [77] Shannon 1948, p. 624.
- [78] Kaprow 1966, pp. 195-196.
- [79] Bourriaud 1998, p. 51.