REVIEW (IDEAS)

ARTEFACT LOST - AI AESTHETICS IN VIDEO PRODUCTION

J. Rosenbaum, RMIT University

> *"Whatever you now find weird, ugly, uncomfortable, and nasty about a new medium will surely become its signature. CD distortion, the jitteriness of digital video, the crap sound of 8-bit - all of these will be cherished and emulated as soon as they can be avoided. It's the sound of failure: so much modern art is the sound of things going out of control, of a medium pushing to its limits and breaking apart. The distorted guitar sound is the sound of something too loud for the medium supposed to carry it. The blues singer with the cracked voice is the sound of an emotional cry too powerful for the throat that releases it. The excitement of grainy film, of bleached-out black and white, is the excitement of witnessing events too momentous for the medium assigned to record them."* (Eno 1996)

Considering Eno's words as AI development in images and video creation continues to advance, we cannot help but draw parallels between the discomfort of early glitchy machine learning creations to the homogeneity of recent generative AI works. In the early days of Deep Dream (Mordvintsev, Olah & Tyka 2016), Generative Adversarial Networks (Goodfellow et al. 2014) and Style Transfer (Gatys, Ecker & Bethge 2015), AI had artefacts and unsettling aesthetics that could discomfit people as much as amaze them. The quirkiness of AI aesthetics was a part of its algorithmic proof of creation by something inhuman. As AI systems have become more refined, we see fewer artefacts, glitches and computational signatures that lend AI images and videos their charm. With the push towards generative AI the focus is on realism, on accuracy and removing that signature influence. In return while the quality of the generations has improved, a signature of the medium being pushed to its limits is lost.

The checkerboard artefacts that beset GAN images and videos through deconvolutional overlap (Odena, Dumoulin & Olah 2016), the puppyslugs (Lu Linvega 2015) of Deep Dream and the low framerate flickers of generative video were all signatures of machine learning images. Systems such as ATTNGAN (Xu et al. 2017) and VQGAN+Clip (Crowson & Murdock 2021) which introduced text to image while continuing to produce alien imagery that maintained the glitch aesthetic of the machine. This changed as generative AI systems such as DALLE (Ramesh et al. 2022), Midjourney (Holz 2022) and Stable Diffusion (Rombach et al. 2022) emerged and started to work towards photorealism with erasing the hand of the machine as the goal. Rather than make images that appear to be made by a machine, these systems attempt to emulate humans, bringing about dialogues surrounding creative jobs, the role of the artist and copyright debates and discussions. It has been an ongoing goal for developers to make increasingly realistic work, but in turn, they have surrendered the unique signature of the AI aesthetic.

Figure 1. Akten, M. (2015) Journey through the Layers Mind

In video works artefacts are magnified, the motion creating a captivating dance of glitches and artifacts that emphasize the role of the AI in the creation. This is seen in works such as Memo Akten's *Journey Through the Layers of The Mind* (2015) which uses Deep Dream, and Anna Ridler's *Mosaic Virus* (2019) which uses GANs. My own work using GANs through my PhD embraces the glitches as part of the learning of the AI as I disrupt its data to change its bias in *Set in Stone* (Rosenbaum 2022). The recent text to video works using systems like Gen 1 and 2 by Runway (Esser et al. 2023) continue to embrace the move towards realism and clarity of image. The goal is to make videos as true to life as possible, a video that looks like it was made by human hands, created entirely by machine. *Echoes of Sand And Stone* (2023) seeks to be the "most realistic AI trailer" with its use of Midjourney and Gen2, and while it is visually attractive, the quick cuts and post-work obscure most algorithmic glitches and signatures. The goal of this work is to create something that looks like it is made by human hands. The faces and figures look very similar, the palette and styles all echoing a homogenization of data as well as the bias of the AI.

Figure 2. Argyropolous. (2023) Echoes of Sand and Stone

I maintain that the lack of control that Eno describes is part of the stress in the machine, the signature of an alien generative voice in the hands of a human creator. It is the symbol of the collaboration with the machine and the exploration of the unknown. As we move towards cleaner and clearer imagery, we gain clarity and realism, but lose originality, suffused in a homogenization of data until all images generated by AI look very similar. This similarity robs us of our imaginative inferences with AI generated works (Rosenbaum 2020) and magnifies the biases inherent in the large data systems. Why try to make something that could be made by a human, when the very concept of working with a machine is working with an alien influence. Why seek to emulate human filmmakers when we can create new aesthetics and artistic movements? That signature, that mechanical stress and influence, will become the signature that we miss.

Figure 3: Rosenbaum. (2022), Set in Stone

REFERENCES

Akten, M 2015, ‘Journey through the layers of the mind (2015)’.

Argyropoulos, E 2023, *MOST REALISTIC AI TRAILER: Echoes of Sand and Stone - MIDJOURNEY & RUNWAY GEN-2*.

Crowson, K & Murdock, R 2021, ‘S2 VQGAN+CLIP Classic.ipynb - Colaboratory’, *Google Colaboratory*, viewed 18 February 2022, <https://colab.research.google.com/drive/1_4Jl0a7WIJeqy5LTjPJfZOwMZopG5C-W?usp=sharing#scrollTo=g7EDme5RYCrt>.

Eno, B 1996, *A year with swollen appendices*, Faber and Faber, London.

Esser, P, Chiu, J, Atighehchian, P, Granskog, J & Germanidis, A 2023, ‘Structure and Content-Guided Video Synthesis with Diffusion Models’.

Gatys, LA, Ecker, AS & Bethge, M 2015, ‘A Neural Algorithm of Artistic Style’.

Goodfellow, IJ, Pouget-Abadie, J, Mirza, M, Xu, B, Warde-Farley, D, Ozair, S, Courville, A & Bengio, Y 2014, 'Generative adversarial nets', *Advances in Neural Information Processing Systems*, vol. 3, no. January, pp. 2672–2680.

Holz, D 2022, 'Midjourney. Website and service.'.

Lu Linvega, D 2015, 'Puppyslug all the way down. #deepdream http://t.co/Dro7ErdYdw', *Twitter*.

Mordvintsev, A, Olah, C & Tyka, M 2016, 'Inceptionism: Going Deeper into Neural Networks'.

Odena, A, Dumoulin, V & Olah, C 2016, 'Deconvolution and Checkerboard Artifacts', *Distill*, vol. 1, no. 10, p. e3.

Ramesh, A, Dhariwal, P, Nichol, A, Chu, C & Chen, M 2022, 'Hierarchical Text-Conditional Image Generation with CLIP Latents'.

Ridler, A 2019, 'Mosaic Virus', <http://annaridler.com/mosaic-virus>.

Rombach, R, Blattmann, A, Lorenz, D, Esser, P & Ommer, B 2022, 'High-Resolution Image Synthesis with Latent Diffusion Models'.

Rosenbaum, J 2020, 'Hidden Worlds: Missing histories affecting our digital future', *Idea Journal*, vol. 17, no. 02, pp. 275–288.

Rosenbaum, J 2022, *AI perceptions of gender*.

Xu, T, Zhang, P, Huang, Q, Zhang, H, Gan, Z, Huang, X & He, X 2017, 'AttnGAN: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks', *arXiv:1711.10485*.