Aesthetics of Vector Media

Tuesday 7 October 2025, at 10:00-12:30
Langelandsgade 139, 8000 Aarhus C
Building 1584, room 224 (entrance B)

Public but part of network meeting in the *AIsthesis – network for research on the aesthetics of AI imagery* (funded by Independent Research Foundation Denmark).

Program

10:00–10:20: Joint intro
10:20–10:40: Sebastian Rozenberg
10:40–11:00: Maja Bak Herrie
11:00–11:10: Break
11:10–11:30: Johan Malmstedt & Sebastian Breu
11:30–12:30: Joint discussion

Joint description of the session

Latent space models are becoming increasingly pervasive in content production, from text generation to visual grading. Text-2-Video models and Text-2-Image models are poised to transform aesthetics, with their integration becoming standardised in media production through Adobe Sensei and Firefly in the Adobe suite (Wilde 2023). As media generated through machine learning processes and automated pattern recognition of different kinds proliferate, latent space and similar computational architectures emerge as a mediation and medium of serious concern for aesthetics and cultural studies. This raises critical questions: does this trend signal the automatization and ultimate dehumanization of media production, or does it represent an emancipation from the formal limitations of human faculties? Both as crisis and utopia, latent space is increasingly an unavoidable process and condition (for the possibility) of digital media. As a central computational process, it now shapes technical, aesthetic, and epistemological possibilities of media, making its study essential to understanding contemporary media production. While productive, humanistic approaches to navigating this space from a media perspective are yet to be fleshed out, to move beyond isolated examples requires systematic methods that embrace complexity and actual technicalities (Offert & Dhaliwal 2024).

Uncovering latent space—through its mapping and by signaling its specific and foundational spatiality—can provide critical insights into its role in media production and aesthetics. This panel brings together three perspectives on the historical, aesthetic, and phenomenological dimensions of AI-generated media, with a focus on the concept of vector spaces. Each presentation studies how latent spaces mediate not only technical processes but also aesthetic and epistemological frameworks. By connecting computational architectures with broader historical and philosophical contexts, the panel

explores the ways latent spaces reshape understandings of similarity, spatiality, medium-specificity, and sensibility.

References:

- Offert, Fabian & Ranjodh Singh Dhaliwal. "The Method of Critical AI Studies, A Propaedeutic." arXiv preprint, 2024. https://arxiv.org/abs/2411.18833.
- Wilde, Lukas R. A. "Generative Imagery as Media Form and Research Field: Introduction to a New Paradigm." The Interdisciplinary Journal of Image Sciences 37(1), 2023: 6-33.

Paper abstracts

Sebastian Rozenberg: Recursive and Constitutive: Media Phenomenology of Vector Space
Are vector spaces in AI generative image models a medium or object in relation to consciousness, and what specific acts of reflection does it engender? This paper theorises different determinations of vector spaces as medium, emphasizing the assumed poles of empirical rational ontology and phenomenal appearance and experience. I will briefly outline three correspondances or homologies between perception/consciousness and the structures of conditional image synthesis of different kinds, multimodal image generation models.

These three correspondances are developed under the concepts of *style as format*, *instrumental similarity*, and *passive synthesis,* before returning to the generated image outputs themselves to ask what meaning or intentionality remains.

There are many correspondences and correlations between phenomenological accounts of experience and latent space, evident in concepts such as manifold, synthesis and unity. This correspondence extends to Husserl's understanding of consciousness as the filling in of perspectival adumbrations, and the topological or sequential aspects of vector space, both founded on a type of anticipation of geometrical deformations (Petitot).

More simply put, many of the structures and organisations of vector spaces are the same structures necessary for experiences to be real and realized. Different vector spaces correspond to husserlian pre-predicative structures of passive synthesis, but this paper seeks to uncover to what extent vector spaces function as a pre-predicative structure for the sensible experience of media. Problematizing both synthesis (Fazi) and sameness, my account frames the medium of vector space against Husserl's descriptions of passive synthesis and the passive data of experience, and self-sameness, formulating the tentative grounds for a media philosophy of vector space in a phenomenological key. From a perspective of recursive couplings, what constitutes the medium and mediality of a vector space? A recursive process of formatting and mediation, a dual constitution of experience and computation.

References:
- Fazi, M. Beatrice. "The Computational Search for Unity: Synthesis in Generative AI." *Journal of Continental Philosophy*, 2024. https://doi.org/10.5840/jcp202411652.

- Husserl, Edmund. *Analyses Concerning Passive and Active Synthesis: Lectures on Transcendental Logic.* Dordrecht: Springer Netherlands, 2001. https://doi.org/10.1007/978-94-010-0846-4.
- Husserl, Edmund. *Experience and Judgement.* Northwestern University Press, 1975.
- Petitot, Jean. "Sheaf Mereology and Husserl's Morphological Ontology." *International Journal of Human-Computer Studies* 43, no. 5–6 (November 1995): 741–63. https://doi.org/10.1006/ijhc.1995.1072.
- Wiesing, Lambert. *Artificial Presence: Philosophical Studies in Image Theory.* Cultural Memory in the Present. Stanford (Calif.): Stanford University press, 2010.

Maja Bak Herrie: Multimodal Collapse: Aesthetic and Medial Consequences of Image Generation Architectures

This paper examines how we can aesthetically and medially engage with the two dominant model architectures for image generation: CNNs and Vision Transformers (ViTs), and the distinct paradigms of image understanding they represent. The main argument is that model architecture – specifically its relationship to multimodality and media-specificity (Hayles) – has aesthetic consequences. While differences between model architectures are often discussed solely in terms of performance, this paper explores their aesthetic and medial implications.

Both architectures use dimensionality reduction of latent space and share a foundation in vectors and embeddings of weights. However, they also differ significantly. Whereas CNNs rely on convolutional layers to extract hierarchical features, ViTs utilize self-attention mechanisms and are based on a much more generalized architecture. Early ViTs, for instance, borrowed architectures from natural language processing tasks and applied them to images by "chopping up" input images into sequences of patches.

This paper discusses and tentatively builds an aesthetic understanding of what this new media-specificity means for the conception of "artistic mediums" (Greenberg, Lessing) and their modes of expression. For example, while CNNs and RNNs related more "intuitively" to classical artistic mediums and were primarily used for images and text respectively (Bajohr), transformer architectures are more "omnivorous," capable of processing almost any input (as long as it can be segmented and vectorized). The title of the much-cited paper "An Image is Worth 16x16 Words" (Dosovitskiy et al) underscores this "flattening" by framing images in linguistic terms, highlighting the multimodal convergence central to transformer-based approaches. This paper asks what happens *aesthetically* when such medial flattening is introduced (Krämer): what are the consequences of this multimodal "collapse" for how text and image relate to one another (Philipsen)?

References:

- Bajohr, Hannes. "Algorithmische Einfühlung: Über zwei Paradigmen digitaler generativer Literatur und die Notwendigkeit einer Kritik ästhetischer KI." *Sprache im technischen Zeitalter* 59, no. 4 (2021).
- Dosovitskiy, Alexey, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, et al. "An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale." arXiv preprint, 2020. https://arxiv.org/abs/2010.11929.

- Greenberg, Clement. "Towards a Newer Laocoön." In *Art in Theory 1900-1990: An Anthology of Changing Ideas*, edited by Charles Harrison and Paul Wood, 554–560. Oxford: Blackwell Publishers, 1992.
- Hayles, N. Katherine. "Print Is Flat, Code Is Deep: The Importance of Media-Specific Analysis." *Poetics Today* 25, no. 1 (2004): 67–90.
- Krämer, Sybille. "The 'Cultural Technique of Flattening.'" *Metode* 1 (2023): 1–18.
- Lessing, Gotthold Ephraim. *Laocoön: An Essay on the Limits of Painting and Poetry (1766).* In *Classic and Romantic German Aesthetics*, 25–130. Cambridge: Cambridge University Press, 2002.
- Philipsen, Lotte. "Prompting Aesthetic Ideologies of Generative Text-to-Image AI." *Contemporary Aesthetics*, 2025.

Johan Malmstedt and Sebastian Kawanami-Breu. The Similarity Engine: Steps Towards a Media History of Vector Modeling

This project explores the historical epistemology of vector space models, contextualizing them within the broader history of humanity's efforts to measure, quantify, and analyze "similarity" (Winkler, Goodman, Wittgenstein). Emerging from a collaborative effort, this work has developed into a nearly completed book. Our project traces vector space models back to late 19th- and early 20th-century scientific endeavors in fields such as biometrics, criminal anthropology, and computational taxonomy. These disciplines sought to grapple with multivariate complexity, employing scatterplots, multidimensional scaling, and clustering techniques to uncover familial and geometric patterns (Galton, Stanton). Confronting the limits of paper technology and the human sensorium, researchers created new visual and mathematical tools to render similarity calculable. These practices did more than organize data; they rewrote our perceptual frameworks and redefined the concept of natural order.

By situating vector space models within this historical lineage, we argue that contemporary computational approaches to similarity – such as latent space modeling – are deeply rooted in these earlier epistemic practices. Our presentation will present a case study on the moment when the cluster, a new vision of order in abstract space, became blackboxed in software for the first time. We examine how the cluster migrated from plots and simulations in the laboratory into the political realm, where it would eventually acquire ontological force in shaping publics, campaigns, and historical outcomes.

References:

- Galton, Francis. "Visualised numerals." *The Journal of the Anthropological Institute of Great Britain and Ireland* 10 (1881): 85-102.
- Goodman, Nelson. "Seven Strictures on Similarity." In: *Problems and Projects.* Indianapolis: Bobbs-Merrill, 1972.
- Stanton, Jeffrey M. "Galton, Pearson, and the peas: A brief history of linear regression for statistics instructors." Journal of Statistics Education 9.3 (2001).
- Winkler, Hartmut. *Ähnlichkeit.* Kulturverlag Kadmos, 2021.
- Wittgenstein, Ludwig. *Philosophical Investigations.* Blackwell, 1953

Bios

Sebastian Rozenberg is a PhD Candidate at the Department of Culture and Society, Linköping University, with a research focus on media aesthetics and media philosophy. He is also a Predoctoral Research Fellow at Bibliotheca Hertziana in the *Machine Visual Culture* group. He has a background in Art History and Film studies and holds an MA in Aesthetics as well as an MS in Information Science. His work integrates aesthetics, phenomenology, media theory and philosophy of technology in order to interrogate the computational basis of everyday visual appearances and experiences, developing a concept of format phenomenology.

Maja Bak Herrie is Assistant Professor in Aesthetics and Culture at the School of Communication and Culture, Aarhus University, and part of the research project *New Visions: Image Cultures in the Era of AI* (PI Lotte Philipsen). With a background in aesthetics, her work spans media theory and the philosophy of science, as she focuses on topics such as computational vision technologies, scientific imaging, photography, and artistic research. Her monograph *Thinking Through Data: How Outliers, Aggregates, and Patterns Shape Perception* was published by Stanford University Press in March 2025. Additionally, she serves as co-editor of *The Nordic Journal of Aesthetics* alongside Tobias Dias and an associate editor of *Journal of Aesthetics and Culture.*

Johan Malmstedt is a media historian and sound scholar whose work explores the intersection of signal processing and historiography. He earned his PhD in Media and Communication Studies from Umeå University with a dissertation on the stylistic evolution of Swedish public service radio (1980–1999). His research combines computational methods and critical media theory to study sound media as both data and historical artifact. Currently, he leads a postdoctoral project comparing audiovisual aesthetics in Swedish and American public service television during the 1970s and 1980s, using cross-modal analysis and deep learning.

Sebastian Kawanami-Breu is a research associate and PhD candidate at Media Studies / Humboldt University Berlin, where he also curates the *Signallabor*, the department's space for experimental and hands-on research. His research traces the role of neural network models as epistemic objects in the history of Cybernetics and the Life Sciences. He is the author of "*Phenomenotechnics and Noumenology"* (on Gaston Bachelard), and has published on the historical and political dimensions of media technology in Bloomsbury's *Thinking Media* series, *Deutsche Zeitschrift für Philosophie*, and *Gendai Shisō* (現代思想). He has also co-translated several of Mark Fisher's works into Japanese, including *Capitalist Realism* and *Acid Communism.*
https://medienwissenschaft-berlin.org/sebastian-kawanami-breu/
https://medienwissenschaft-berlin.org/signallabor/
https://www.degruyterbrill.com/document/doi/10.1515/dzph-2022-0061/html