Check for updates

RESEARCH ARTICLE

REVISED Machine vision situations: Tracing distributed agency

[version 2; peer review: 3 approved]

Marianne Gunderson[1], Ragnhild Solberg[2], Linda Kronman[1], Gabriele De Seta[1], Jill Walker Rettberg[1]

[1]Department of Linguistic, Literary and Aesthetic Studies, University of Bergen, Bergen, 5020, Norway
[2]Department of Art and Media Studies, Norwegian University of Science and Technology, Trondheim, 7491, Norway

v2 **First published:** 31 Aug 2023, **3**:132
https://doi.org/10.12688/openreseurope.16112.1
Latest published: 15 Mar 2024, **3**:132
https://doi.org/10.12688/openreseurope.16112.2

Abstract

This article proposes a new method for tracing and examining agency in heterogeneous assemblages, focusing on the role of machine vision technologies in creative works. We introduce the concept of the "machine vision situation" and define it as the moment in which machine vision technologies come into play and make a difference to the course of events. By taking situations as the unit of analysis, we identify moments at which machine vision technologies take part in actions without reducing them to either tools or protagonists, instead allowing for more complex agential entanglements between human and non-human actors. Grounded on an interdisciplinary theoretical framework, this article demonstrates how an analytical unit such as the machine vision situation is a valuable method for tracing how agency is distributed. We illustrate this through three examples by applying the method to creative works – narratives, digital games, and artworks – revealing key aspects of distributed agency and calling attention to the excess, complications, and messy entanglements that might otherwise be overlooked in analyses of agential assemblages. The machine vision situation is shown to be a flexible unit of analysis that can be productively incorporated in both quantitative and qualitative studies and applied to other contexts in which human and non-human agencies interact.

Plain language summary

Machine vision – the ability of machines to "see" and interpret visual information – has advanced significantly in recent years, with applications ranging from self-driving cars to medical diagnosis. However, there is a growing recognition that this technological advancement does not simply power a wide variety of new tools and systems, but also results in new distributions of agency alongside (and, at times, against) human decision-making.

Our article explores this idea in depth, examining how machine vision

Open Peer Review

Approval Status ✓ ✓ ✓

	1	2	3
version 2 (revision) 15 Mar 2024	✓ view	✓ view	✓ view
version 1 31 Aug 2023	? view	? view	

1. **Susana Tosca**, University of Southern Denmark, Odense, Denmark
2. **Anastasia Salter**, University of Central Florida, Orlando, USA
3. **Kristin Veel**, University of Copenhagen, Copenhagen, Denmark

Any reports and responses or comments on the article can be found at the end of the article.

technologies and human beings are represented as agents in works of narrative such as games, art, and fiction. Analysing the representation of machine vision in artistic works reveals how these technologies are experienced and imagined in different contexts. We introduce the framework of “machine vision situations” to analyse the complex and dynamic relationships between humans and machines in both fictional and real-world contexts.

A machine vision situation is a moment in which a machine vision technology is seen or represented as making a difference to the course of events. This situation can be analysed by identifying the actors involved and making a list of verbs that describe each of their contributions to the event. This method results in a dataset that can be analysed quantitatively, but it also generates a starting point for a qualitative analysis of distributed agency between human and non-human agents in both fictional and real-world situations.

Keywords
machine vision, digital humanities, agency, science fiction, games, art,



This article is included in the Horizon 2020 gateway.



This article is included in the European Research Council (ERC) gateway.



This article is included in the Artificial Intelligence and the Social Sciences and Humanities collection.

Corresponding author: Marianne Gunderson (marianne.gunderson@uib.no)

Author roles: Gunderson M: Conceptualization, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Solberg R**: Conceptualization, Methodology, Writing – Review & Editing; **Kronman L**: Conceptualization, Methodology, Writing – Review & Editing; **De Seta G**: Conceptualization, Supervision, Writing – Review & Editing; **Rettberg JW**: Conceptualization, Funding Acquisition, Methodology, Project Administration, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 771800, Machine Vision in Everyday Life: Playful Interactions with Visual Technologies in Digital Art, Games, Narratives and Social Media).
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2024 Gunderson M *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Gunderson M, Solberg R, Kronman L *et al.* **Machine vision situations: Tracing distributed agency [version 2; peer review: 3 approved]** Open Research Europe 2024, **3**:132 https://doi.org/10.12688/openreseurope.16112.2

First published: 31 Aug 2023, **3**:132 https://doi.org/10.12688/openreseurope.16112.1

REVISED **Amendments from Version 1**

This version of the paper has been revised to address the comments and suggestions from the reviewers. Specifically, we have:

- elaborated on how the analysis of fiction, art, and games can enrich our understanding of machine vision technologies.

- contextualized the machine vision situation method in relation to Clarke *et al.*'s situational analysis and emphasised the connection between our work and Haraway's concept of situated knowledges.

- shortened the paragraph on Bennett's approach to agency, and moved the paragraph on Hayles' concept of nonconscious cognition into the 'Myriad and Mosaic Virus' section.

- added a short discussion about the agency of the player/user to the 'Myriad and Mosaic Virus' and the 'Detroit: Become Human' section, explaining our criteria for including the player as an actor in relation to our project focus and aim and pointing to how the inclusion of the player/user could bring out different insights.

- moved the paragraph on the database findings to the concluding section and added a paragraph highlighting our main findings from the database.

In addition, we have re-edited the paper for readability and conciseness. We are grateful to both reviewers for their thorough reviews, encouraging comments and insightful remarks, which contributed significantly to the improvement of the manuscript.

Any further responses from the reviewers can be found at the end of the article

Introduction

Machine vision, understood as "the registration, analysis, and representation of visual data by machines and algorithms" (Rettberg *et al.*, 2019, 1), brings us sights (the inside of one's colon, the travel patterns of a population under lockdown, or a live feed of owl hatchlings huddled in a bird box) which we would never have been able to observe without technological mediation. The history of machine vision can be traced back to the invention and popularisation of the camera and can arguably be extended to include the history of glasses, binoculars, microscopes, and other noteworthy early examples of technologies that have enabled us to surpass the limitations of embodied human vision. Developments in algorithmic automation and machine learning dramatically broaden the extent to which machine vision reaches into our lives. Technologies that fall under the umbrella of machine vision encompass widely different objects and uses ranging from surveillance cameras and webcams to facial recognition and cancer detection. The computational processes underpinning much of machine vision also make possible the creation of synthetic images such as deepfakes, multimodal machine learning models that generate images from written prompts, and augmented reality applications such as amusing smartphone camera filters.

The proliferation of machine vision agents and processes raises new questions about agency. Who creates the image when you type a prompt into an image generator like DALL-E or Midjourney? Is it you, or the machine learning model? Who is responsible when a person is wrongfully arrested after being misidentified by a facial recognition algorithm? Is it the police, the algorithm, or its programmers? The police, the algorithm, or its programmers? Advancements in machine learning are driving the development of increasingly autonomous technological agents: self-driving cars are built with the ability to act upon information in real time with no or minimal human input, and generative models reconfigure the relationship between human and non-human creativity and labor. How can we understand the distributions of agency between human and technological actors such as machine vision systems? How can we identify the contributions of each agent without losing track of their connectedness? How are human and non-human actors affected or changed by their interactions with machine vision? In other words, how do we understand agency and responsibility when automated entities interpret, act upon and even create visual data?

Our answer, which we develop throughout this article, is based on the premise that agency is not an exclusively human capacity. The notion of non-human agency has been developed across a wide range of theories and approaches that cluster around the fields of posthuman theory, new materialism, affect theory, and science and technology studies (Bennett, 2010; Hayles, 2017; Latour, 2005; Massumi, 2021). To understand agency in this broader sense, we need to take account not only of human choices and actions, but also of those of other lifeforms such as plants, animals, and microorganisms, as well as of non-living matter such as technical objects, weather phenomena, and other "things" (Hayles, 2017). Acknowledging the agency of the non-human world undermines the position of the (hu)man as the sovereign actor who controls and transforms nature through the deliberate use of tools (Marchand, 2018). What used to be a passive backdrop for human heroism comes to life, and the stage becomes populated by a proliferation of agents, allowing new dramas to play out. By focusing on a specific domain of technology and representation – the role of machine vision in cultural production - this article proposes a method for bringing these dramas to light.

When studying the distribution of agency among human and non-human actors, it is essential to develop research methods that do not rely on a binary opposition between human and non-human actors nor operate with the assumption that agency is a zero-sum game in which the agency of one actor automatically diminishes or cancels out the agency of others. Drawing on Lauren Berlant's (2011) use of the situation as an analytical tool, we propose the concept of *machine vision situations,* defined as the moment at which machine vision technologies come into play and are seen to make a difference in what is taking place. Based on the concept of machine vision situations, this paper develops a method for tracing the distribution of agency between human and non-human actors, discusses its relevance to different approaches to non-human agency, and demonstrates its potential for both quantitative and qualitative analysis of distributed agency in other contexts.

Technological change is always accompanied by public debate. We argue that fiction, art, games and cultural expression more broadly are modes of discourse that allow societies to access responses to technology that are not always expressed in rational debates (Rettberg, 2023:13–15). Many existing technologies were first described in science fiction or myths: technological development is often driven by fiction and imagination (de Seta, 2022). Playing games or watching a movie allows people to become familiar with technologies before they have direct access to them (Solberg, 2022a). Technologies are critiqued and challenged in art games and narratives, since art, in the broadest sense of the word, can express emotional responses to technological change as well as exploring what could go wrong or how technologies could be used differently (Gunderson, 2021; Kronman, 2023).

The rapid development of AI-powered machine vision technologies like facial and object recognition, drones, image generation, deepfakes and surveillance has led to governments and organisations around the world implementing guidelines, policies and regulations addressing the challenges they bring. The machine vision situations method brings out insights from fiction, art, games and other aesthetic modes of expression that are often underexamined in academic and political deliberations, and can thus add to the knowledge used by policy makers and developers to guide the development and application of artificial intelligence and machine vision technologies.

In the following sections, we begin by reviewing theoretical approaches including Bruno Latour's actor-network theory (Latour, 2005), N.K. Hayles' concept of nonconscious cognition (Hayles, 2017), and Lauren Berlant's (2011) and Brian Massumi's approaches to affect (Massumi, 2021). We then trace how we conceptualized the machine vision situation as an analytical unit, outline the method we developed around it, and discuss how it informs and is informed by multiple approaches to non-human agency. This discussion is followed by three examples of machine vision situations from Annalee Newitz's short story *Drones Don't Kill People* (2014), the digital game *Detroit: Become Human* (2018), and Anna Ridler's digital artworks *Myriad (Tulips)* (2018) and *Mosaic Virus* (2019a), illustrating how this tool can be used to analyse narrative works, games, and artworks. Finally, we discuss how this tool may be used beyond the context of creative works and point towards their further development.

Machine vision situations

We developed the concept of the machine vision situation as we worked on the *Database of Machine Vision in Art, Games, and Narratives*. Completed in 2021, the database documents 77 digital games, 190 digital artworks, and 233 movies, novels, and other narratives that use or represent machine vision technologies (Rettberg *et al.*, 2022a; Rettberg *et al.*, 2022b). Within a total corpus of 500 works, we have identified and analysed 874 specific situations where machine vision is central (Rettberg *et al.*, 2022a; Rettberg *et al.*, 2022b). Based on the definition of machine vision mentioned above, and an initial survey of creative works, we developed a list of 26 different technologies that formed the basis for identifying machine vision situations in the works.[1] Our goal in developing the database was to draw on the computational powers of the digital humanities to conceptually map commonalities and tendencies in representations of machine vision across a wide range of works. This is a common approach in the digital humanities (see, for instance, Sinclair & Rockwell, 2015, 288). Computational methods of analysis have often been placed in opposition to "the individuated and situated practices of human reading and interpretation" (Drucker, 2017, 631), but as Drucker points out, this is a false dichotomy. As Jill Walker Rettberg argues, "data is always partial and situated" (Rettberg, 2020, 4), and designing a database or a text analysis program is an interpretative act (Drucker, 2017), something the lively discussions within our research group can attest to. Designing a system in which to characterise often complex works through a limited number of keywords is a process of making careful trade-offs between specificity and generalisability. While many digital humanities projects rely on machine reading and interpretation of a corpus, our project relied on "manually" selecting the works to include in the corpus. The database structure functioned as a lens through which we constructed a body of data that would then be analysed using digital tools. However, the actual data resulted from each researcher's reading and interpretation of the works. The process thus combines classical humanistic, interpretative approaches to text with digital, quantitative methods.

In order for works to be registered in a database, there must first be a conceptual infrastructure (Feinberg, 2017) within which to register them. This process involved deciding on the categories under which works would be classified. To see patterns across genres and technologies, we needed a classification schema that allowed us to find shared characteristics in a very diverse dataset. The challenge was finding a format that allows for direct and easy comparison between the works required the design of a structure through which we could extract comparable meanings from, for example, a digital art installation, a digital game, a movie, a novel, and a piece of digital fiction. In our case, this meant negotiating and discussing which aspects of the works had relevance for the project and deciding on the basic units of meaning through which to build a corpus of data that would lend itself to network analysis by making some features of the text itself explicit, with the goal of making them processable by some computer application (Renear, 2004 in Pierazzo, 2015:307). This infrastructure, for us, is built around the situation.

A situation, according to Lauren Berlant, is "a state of things in which something that will perhaps matter is unfolding amid the usual activity of life" (Berlant, 2011:5) and "a genre of living that one knows one's in but that one has to find out about, a circumstance embedded in life but not in one's

[1] For a complete list of these technologies, together with the definitions we used for the purpose of this project, see Rettberg *et al.*, 2022a.

control" (2011:195). This concept of the situation as a moment of disturbance in everyday life that "forces one to [...] become interested in potential changes to ordinariness" (2011:195) is a very fruitful concept with which to analyse the cultural imagination of emerging technologies, such as machine vision. The task is to identify the situations in which machine vision technologies come to matter in the action of what is unfolding within or through the work. A machine vision situation may be an excerpt from a novel, a scene in a movie or a sub-plot in a game, or it may encompass a whole work, as may be the case with short stories or artworks – what we were looking for was the moments at which machine vision technologies were taking part in the action, so to speak. Each work may contain one or more of these situations in which one or more machine vision technologies come into play.

The focus on the specific situation is also inspired by Donna Haraway's concept of situated knowledges (Haraway, 2007), and our research is grounded in the premise that knowledge about sociotechnical processes will always be rooted in specific contexts. Instead of aiming for the objectivity of a God-like view from nowhere (Haraway, 2007, 115), our aim was to develop a method committed to the particular, contextual, and diverse practices and meanings that emerge in specific circumstances. This commitment may seem contradictory to a database-based partly quantitative approach, in which information is extracted from its original context and made to fit into a pre-determined structure. However, this tension turned out to be a generative force that nudged us to circle back to the data and developing close readings of individual machine vision situations.

By defining machine vision situations as a core unit of the database, we identified moments where machine vision technologies became what Bruno Latour would call "matters of concern" (Latour, 2004). In *Reassembling the Social* (2005), Latour argues that accounts of social agency are limited by their reliance on a conception of the social which precludes the inclusion of anything, or any *thing*, that is not composed, collectively or individually, of people. This social realm is seen as separate and clearly distinguishable from other realms of reality, such as physics, biology, or geology. In actor-network theory, which Latour developed in the 1980s and 1990s with Madeleine Akrich, John Law, Michel Callon and others (Akrich, 2023), the social is redefined as "a type of connection between things that are not themselves social" (Latour, 2005:5) and the task of the social scientist is redefined as "the tracing of associations between heterogeneous elements" (Latour, 2005:5). Action, in the actor-network approach, is no longer limited to that which is intentional or meaningful human behaviour, but may just as well rest in "the domain of 'material' 'causal' relations" (Latour, 2005:71). If something, or *some thing* "makes a difference in the course of some other agent's action" (Latour, 2005:71), that thing has agency within the context of which it is seen to matter. Building on this, the idea of machine vision situations is that looking at the doings of humans and technologies together allows us to see the difference each agent makes in the situation and to understand the specificity of their interactions. While actor-network theory has developed its own methodological tools for studying the interrelations of human and non-human agency, for our purpose we needed to develop a system that would allow us to generate data that could be studied quantitatively and which was suited to analyse works of fiction, games, and art. In order to capture these interactions within the various works, we set out to establish a model with which to encode the actions and actors involved that was simple enough to use as the structure of our database.

Another major challenge was figuring out how to capture meaningful interactions across a wide range of machine vision technologies. Describing the agency expressed by a specific machine vision technology within a work was complicated by our realisation that machine vision technologies are often presented as doing many different things that cannot be easily reduced to a single entry. It was clear that we could not form our analysis of machine vision simply around the act of seeing. Some actions might more easily be described as generating, classifying, or identifying, but even that seemed insufficient to capture what these devices were doing in our lives on a meaningful level. The same technology may do different things in different contexts or different works. However, it is easy to take the agency of non-organic objects as given by their known functions and uses. As Massumi states, "a thing is when it is not doing" (2021:7), meaning that when an object is not doing, it is fixed, known, static. However, we are interested in the doing, the undetermined potentiality for action, and their ability to cause something to happen, without assuming to know in advance the difference they would make. Instead of looking at machine vision technologies in their "thingliness", as objects with known uses and fixed functions, we set out to study their manifold actions as embodied and/or articulated in a wide range of works. As a result, we decided to describe what technologies and other actors are doing in each situation through an open vocabulary of verbs. In the case of the machine vision database, this 'verbosity' may be an apt method to trace the difference these technologies make as linguistic markers of our vibrant relationship with that which we are used to thinking about as matter, resources, or tools. By attributing verbs to machine vision technologies, they become visible as agents within these situations.

Following Latour (2005), in order to understand the unfolding of social (and non-social) events, the contributions of each element in the chain, or network, of action must be taken into account: the researcher should trace the connections that enable social agency, through a network composed, as it may be, of objects, groups, documents, and people. We landed on a system in which we registered one or several machine vision technologies and any characters and/or entities that were active in each machine vision situation.[2] We used the term character as it is used in narratology: "a text- or media-based figure in a storyworld, usually human or human-like (..) in contrast to

[2] For a detailed explanation of how we defined characters, entities and the technologies, see Rettberg *et al.*, 2022a.

'persons' as individuals in the real world" (Jannidis, 2014). In our dataset characters were often human, but many were also robots, sentient AIs, or as in one of the example analyses later in this paper, sentient drones. We registered information about how characters in each situation were represented, such as their gender[3], age, etc. Entities are objects, institutions, or generic categories such as "users" or "images" for which gender or similar characteristics are not relevant. Each agent (a technology, character, or entity) was then assigned a verb to describe their interaction with the other agents in the situation. This enabled us to capture some of their liveliness, while remaining open to the possibility that technologies, people, and other entities may be seen to be doing several things at once, without predetermining the types of actions that may be taken by any of the actors involved.

We made a deliberate choice to position technology, people, and institutions on the same level in the machine vision situation. Our analytical model has no predetermined structure that pre-defines technologies as passive, or as tools or objects, or that positions humans as users, creators, or subjects. This even-handed treatment allows both machines and people to be assigned agency in a given situation. Suddenly, their active participation in a chain of events comes to the foreground. The analyses of machine vision situations later in this paper demonstrate how this works in practice.

Studying humans, technologies, and other entities as potentially lively "bodies with the power to affect and be affected" (Massumi, 2021:16) in unforeclosed ways also required a system for capturing the effects other actors have on each entity involved in a situation. We used the present continuous tense (e.g. *watching*) for actions undertaken by a character, technology, or entity, and the past participle as used in the passive tense (e.g. *watched*) for actions that happen to a character, technology, or entity. This enabled us to identify the various activities that machine vision technologies are involved in, as well as their effects on the various actors involved. This inclusion of passive verbs reveals how the various agents in the situation are affected by each other's actions, bringing out how they are both being changed and effecting change on other bodies within the situation.

In the database, however, the technologies (alongside the characters and entities) appear as separate, predefined categories. This is a departure from the actor-network-theory approach of allowing the subjects to appear as you trace their interactions. By designating them with controlled vocabularies, the actors appear more fixed and stable than they might have any reason to be. The limitations of the understanding of agency imposed by the database structure are particularly apparent when seen through the lens offered by Jane Bennett in her book *Vibrant Matter*. Building on Latour, she theorises non-human agency through the Deleuzian concept of assemblages, "ad hoc groupings of diverse elements" (Bennett, 2010:23) through which agency is effectuated and unfolds. Agency, for Bennett, is not controlled from one central node; rather, the ability to make something happen may arise from the emergent properties of assemblages of matter, bodies, and forces. The machine vision situation is our attempt to map some of these assemblages and trace the agency as it emerges between and among the diverse bodies involved. The verbs attached to the various agents highlight the distribution of agency throughout the assemblage, bringing out their diverse contributions of causes and effects while demonstrating their interconnectedness.

Although our conceptualization of situation emerged from the process of designing a research database, it shares several aspects with the method of situational analysis developed by Adele Clarke and several other scholars. Clarke positions situational analysis as "an extension of grounded theory (GT) method of analysis for qualitative research" (2016:1) in which "the situation itself becomes the unit of analysis, and the researcher maps the situation to analyze it" (1), and which can be applied to multiple kinds of data, including interview transcripts, ethnographic notes, visual content, and so on. Situational analysis is underpinned by a wide range of theoretical approaches across disciplines, including interactionist sociology, feminist embodiment, and infrastructural ecology (Clarke *et al.*, 2022:25), as its unit of analysis – the situation – is conceived in ecological terms as being a composition of co-constitutive elements (Clarke, 2019:15). While sharing with situational analysis this centrality of the situation as an analytical unit, as well many theoretical interlocutors (including Massumi, Haraway, and Latour), the approach we propose in this article is designed to operate on a smaller scale. Rather than conceptualizing entire social arenas or research projects as a situation to be mapped, we identify situations in the scenes, events, or moments of interaction that are depicted in narrative works such as artworks, digital games or novels. This difference in scale reflects a broader methodological choice: rather than departing from grounded theory, our project applied a targeted approach to cultural representation through which we sought to identify the role of a specific technology (machine vision) in a heterogeneous corpus of texts. Despite the different scale at which our approach to situations operates, the method we propose can easily complement situational analysis by allowing researchers to identify pivotal moments or events and zoom into the details of what is happening in these more limited situations within a larger context.

To sum up – in order to study what is being done with, by, and to machine vision technologies, we begin by identifying situations in which these technologies are seen to make a difference. We then identify the agents involved in the situation – the specific technologies, characters, and other entities

[3] According to the SAGER guidelines, Heidari *et al.* (2016) recommend reporting on both sex and gender in research design. However, since most of the individuals that are registered in our dataset are characters in fictional or artistic works, sex cannot realistically be determined, and gender has been determined by the researchers through interpreting the social, visual and linguistic context cues within the fictional works, if a character's gender is not explicitly stated in the texts. For more information, see Rettberg *et al.* (2022b).

participating in the action. The next step is to look closely at the situation, and attribute verbs to each of the agents involved, based on what they are seen to be doing, using passive verbs to describe the effects of actions, if appropriate. The verbs should be allowed to emerge from the text(ure) of the situation, and the same agent may have several different, even contradictory, verbs attached to it. Relatedly, the verbs are not limited to the supposed or assumed uses or activities in which a body or object of its kind is assumed to take part.

In what follows, we will go through machine vision situations from three different works: a short story, a digital game, and an artwork, to demonstrate how this analytical model can be applied to explore the representation of agency in a variety of genres.

Example analyses: Assemblages in the works

Drones Don't Kill People

The result of this analytical model is a snapshot of actions attached to each actor, which can, on their own, form remarkably compelling narratives. For example, in the short story *Drones Don't Kill People* by Annalee Newitz (2014), narrated from the perspective of an artificially intelligent surveillance drone, one machine vision situation occurs when a group of largely autonomous AI drones, subcontracted to a corporation working for the Turkish government, are set to monitor a professor and his family and to pass on potentially relevant information. One night the drones observe the unnamed mother answering a question posed by the daughter. This is what the incident looks like in the database:

> *Professor's Daughter: is Questioning*
>
> *Professor's Wife is: Answering, Explaining, Revealing*
>
> *Drones are: Spying, Recording, Interpreting, Selecting, Debating*
>
> *Corporation is: Subcontracting, Employing, Surveilling, Deciding*

Here we see that although the drones perform expected machine vision actions such as recording, they are also partaking in other more cognitive and collaborative activities involved in the same process. The inclusion of "spying" furthermore emphasises that what they are doing has normative implications in a larger social context. The mother and daughter's actions are linked to the drones' activity through the inclusion of "revealing". Finally, the corporation's actions are included in the assemblage, highlighting that the drones are not acting in a vacuum. The corporation is the agent attributed with the action of "surveilling" since they, not the drones, decide what to do with the information which is collected. Together, they enter into what Bennett (2010) might describe as an assemblage of machine vision-mediated activity. However, this does not mean that agency is equally distributed among the actors, which will become evident in the next entry from the same short story.

Based on the information collected from this incident, the professor and his family are identified as political dissenters and activists, and the drones are ordered to assassinate them. In the database, the entry for this second situation looks like this:

> *Corporation is: Subcontracting, Employing, Deciding, Ordering*
>
> *Drones are: Obeying, Deciding, Killing, Watching, Recording.*
>
> *Professor is: Killed*
>
> *Professor's Wife is: Killed*
>
> *Professor's Son is: Killed*
>
> *Professor's Daughter is: Evading, Watching, Screaming, Killed, Recorded.*

Both the drones and the corporation take part in the act of killing the professor and his family. The unequal distribution of power between the drones and the corporation is made explicit through the inclusion of the verbs "ordering" and "obeying," while the inclusion of "deciding" emphasises the drones' active role. Here, the act of killing is attributed to the drones, although this will be contested later in the story, as foreshadowed by the title. The assassination went according to plan, although it did include one "statistically anomalous event" when the drones missed the daughter on the first shot. Her screams of terror as she watched them kill her family were recorded in their distributed memory. The way this situation is written, the family, with the exception of the daughter, have very little agency, as their lives are extinguished before they can register what is happening or respond – as Bennett would say, while the drones are acting, they are "suffering action" (2010:21). The daughter suffers the same fate, but her final actions reverberate throughout the story – through the recording of the incident, her screams will continue to have the power to affect the drones.

The act of assigning the actors and verbs involved in a situation is inevitably interpretative, and revisiting these entries inevitably brings up potential edits and additions that could have been made. For instance, rereading the short story revealed a third situation, overlooked at the time of data entry, but readily apparent in the light of this paper: after the contract with the Turkish company expires, the drones are subcontracted to the Uyghur Republic[4] government to monitor a desert highway in contested territory bordering on China. Little of relevance is happening, and for the first time, the drones are left alone with nothing to do, so they look about on the web for stuff to analyse. As drones, they are programmed to recognise faces, so when they come across images of drones with faces painted on them, they begin investigating their own identity.

> *Government is: Ordering, Surveilling, Ignoring*
>
> *Drones are: Bored, Analysing, Sharing, Recognising, Changing*
>
> *Images are: Fascinating, Recognised*

[4] In this near-future science fiction story the Uyghur population in Northwest China has been able to gain independence and establish their own country.

In this situation, images appear as an agent for their capacity to capture the drones' attention and elicit a moment of recognition and interest in themselves as drones. Here, something unexpected occurs: the drones take the first step towards self-awareness and, thus, consciousness. As a result, they become interested in themselves as agents in the world. This is the true turning point in the story, taking the drones down the path of forming a political movement fighting for the rights of AIs. The images are clearly making a difference to the drones' course of action in the Latourian (Latour, 2005) sense, and in this assemblage, they are the pivotal actors.

In all three of these situations, something new emerges in the machine vision assemblage. In the first one, the ambiguity of the information makes the drones debate whether or not to include it in their report. In the second situation, the recording of the screaming daughter will reverberate in the drones' collective memory. In the third situation, the moment of self-recognition prompts the emergence of self-awareness, opening the drones up to change and their emergence as political subjects. The assigned verbs do not necessarily capture these becomings as they are matters of affect, not action – they are what takes form beyond the actions taken by each of the entities of the situation, changes in positionings, and the seeds of future events (Massumi, 2021). Nevertheless, tracking these changes can point to moments where action and affect come together to make something new. This would not become apparent in a quantitative analysis of a great number of machine vision situations like the ones collected for the database. However, each situation is an invitation to look closer, follow the connections and interactions that become apparent through this method, and look for what emerges beyond the verbs.

Registering incidents in a short story using this framework may be relatively straightforward. However open to interpretation and shifting meanings, a text is still a relatively static object in the sense that the words appear in the same order every time you encounter them, and the narrative structure easily lends itself to the process of identifying actors and actions. In what follows, we will show how machine vision situations work when considering other genres: first, a game, and then two works of art.

Detroit: Become Human

The doing can take on various forms across media. Audiovisual media like films and digital games often present actions visually, and in most games, this visualisation will also require some sort of haptic action from its user. Using situations to trace agency in games shows these various entanglements of action and bodies across levels of virtuality.

Consider the game *Detroit: Become Human*, published in 2018 by Quantic Dream, which is a branching narrative game that follows androids as they navigate their emerging sentience in a society that places them in positions of servitude and submission. One of these androids is the police inspector Connor, whose job is to find and capture rogue sentient androids. Due to Connor's android body, the presence of various machine vision technologies embedded in his body is unquestioned. Connor's augmented vision includes an augmented reality overlay of the world, combined with object recognition technology, vision beyond the human spectrum, and even reconstructive/generative image software that can visualise past events leading up to the crime scenes before him. These technologies perform various actions, but are combined in the character of Connor, played by the player. Connor's actions are also actions on behalf of the game software, the video game console, the player, and various machine vision technologies, showing the interdependence and permeable boundaries between various agents in the assemblage (Bennett, 2010).

In one situation, two guards escort Connor into a corporate elevator to present him to their boss. Once inside the elevator, a perceptive Connor/player can identify a surveillance camera in the top corner, hack into it, kill the two guards, and escape without repercussions. The situation entry looks like this:

Connor is: Hacking, Fighting, Cloaking, Killing
Surveillance camera is: Hacked, Blinded
Law enforcement is: Fighting, Killed
Corporation is: Blinded

Here, we see that the entry creates a directional relation between Connor's hacking and the surveillance camera that is hacked. In turn, Connor's hacking not only blinds the surveillance camera, but also the corporation to which it belongs. The surveillance camera becomes a Latourian (Latour, 2005) "thing" that influences another agent's action. In assigning the same passive action of "blinded" to both surveillance camera and corporation, a temporal relation is created between the two. This relation emphasises how the surveillance cameras are the prosthetic eyes of the corporation in this situation.

Many games offer diverging and possibly mutually exclusive narrative paths, and *Detroit: Become Human* is no exception, as it is structured around presenting influential narrative choices to the player. If Connor does nothing in the elevator, he will be shot and killed when the doors open again. If Connor attacks the guards in the elevator without disabling the surveillance camera, he will have to fight more guards upon exiting, because he will have been spotted acting deviantly, i.e. more sentient than the humans prefer. Connor can also fail at fighting the guards. Following this pattern of diverging paths, the elevator situation will not even occur in some playthroughs because it depends on previous choices.

The science fiction world of *Detroit: Become Human* exemplifies the characterisation of technologies. In this situation, Connor is the agent, not the various processes that combine into him. The player's bodily labour of choosing and possibly failing at performing actions is hidden in this particular situation

analysis. This reveals how human and non-human agencies combine or work in tandem beyond the fictional world. Even if the player is not an explicit agent in this situation, they are implicit in the presence of the player character's actions. In a different context, the situation could be expanded to include the player, as well as the console, or even the company or programmers who made the game, and so forth – one could follow these connections forever.

This method also highlights the challenge of determining which actors are registered and are thus made visible. The choices of what to include as an actor would inevitably affect the dynamics in a given situation. For instance, by including the player in the above situation, we would not only introduce a set of new actions, but also new tensions, meanings and potential readings of the situation. In the case of *Detroit: Become Human*, this inclusion could bring out issues of race and structural oppression, since in this game, androids are treated like an oppressed minority in ways that mirrors the struggles of racialized minorities in the U.S (Dehnert & Leach, 2021; Leach & Dehnert, 2021; Schubert, 2021), and the inclusion of the player as an actor could bring up interesting tensions and contradictions between the player's identity and experiences and that of the player character. However, in this project, we decided to solely include the actors represented within the work, which means that the database does not register the player or the console here, but the combined action of distributed agents, as represented by Connor. For our purposes the player is included only when they directly interact with machine vision, for instance with the augmented reality display of *Pokémon Go* (Niantic, 2016). This decision arose out of the need to have situations be comparable across genres and types as work, and our focus on the representation of technologies within the contents of the works themselves. Although this means that we missed out on some potential insights related to the embodiment of the player, this was a necessary trade-off given the thematic scope of this project. A project with different research focus and priorities would most likely make different decisions about which actors to include or exclude, as the boundaries of the assemblage must be drawn with concern for the specific purpose of the knowledge project of which it is a part.

Myriad and *Mosaic Virus*

In art, situations allow multiple layers of analysis. Digital artworks can put the viewer in a situation where they interact with machine vision technologies. They can represent fictional or actual machine vision situations in the same way as games and narratives do. Digital artworks also often highlight their own creation, that is, the situation in which the artwork was created. Anna Ridler's diptych *Myriad (Tulips)* (2018) and *Mosaic Virus* (2019a) comprises two of many artworks where artists use machine learning-based AI image generation to create art. *Myriad (Tulips)* exhibits a dataset of hand-labelled polaroids classifying – according to colour, type, and stripe – a myriad of tulips, that is, ten thousand tulips. AI image generation often appears to be an automated process, but as Ridler describes it, dataset curation involves an "insane amount of work and it is usually work that is hidden" (Ridler, 2019c). To create the dataset, Ridler selected the tulips at the market, photographed them, and sorted and labelled them. Thereafter, Generative Adversarial Networks (GANs), a type of machine learning model, were trained with the dataset. What the network learned from the ten thousand images of tulips was to hallucinate "botanical impossibilities" (Ridler, 2019b). In the second part of the diptych, *Mosaic Virus*, these generated images are displayed on screens that show how the AI-generated tulips evolve. As Ridler explains in her artist statement, the "tulips are controlled by the price of bitcoin". The aesthetic choice of bringing AI, bitcoin, and tulips together combines the historical speculative bubble created by the 17th-century Tulipmania with contemporary speculative investments in AI cryptocurrencies.

The identified machine vision situation in the database looks like this:

> *Creator (Artist) is: Classifying, Labeling, Selecting, Speculating*
>
> *Machine learning, Image generation is: Learning, Generating, Hallucinating, Co-creating*
>
> *Images are: Classified, Generated*

Little of the aesthetics of the artwork remain in the reduced situation entry; however, what is implied is a distribution of cognitive labour between the artist and the machine. The artist creates categories and then interprets photographs classifying the tulips for the datasets. In both Ridler's classification of tulips and our analysis of this as a machine vision situation, "interpretive flexibility" occurs, despite rigid classification protocols (Feinberg, 2017). Ridler notes that even "something as simple as a tulip is difficult to put into discrete categories – is it white or pale pink, is it orange or yellow". In turn, the image-generating machine learning model learns to recognise patterns in the 10,000 examples in the dataset, in order to generate new images of tulips.

While human agency is still crucial to create an AI artwork like *Mosaic Virus,* some cognitive functions like discerning patterns and drawing inferences are externalised to machine vision. In *Unthought*, N. K. Hayles proposes a definition of cognition that allows us to understand technological objects as cognisers: "Cognition is a process that interprets information within contexts that connect it with meaning" (Hayles, 2017:22). She argues that while consciousness may be attributable primarily to humans, cognition is a capacity that we, according to Hayles (2017), share with other animals, perhaps even plants, as well as with technical devices. Hayles describes computational media, and especially AI, as "quintessentially cognitive technologies" (Hayles, 2017:41) with the ability to process information, identify patterns, and make inferences. When we use these technologies, we effectively enter into what Hayles terms "cognitive assemblages," which, for the example of a cell phone, would include "relay towers and network infrastructures, including switches, fibre optic

cables, and/or wireless routers, as well as other components" (Hayles, 2017:8).

The creation of AI art emerges out of cognitive assemblages involving both human and machine interpretation. However, when AI art was popularised by image generators like DALL-E and Midjourney producing new images from text prompts, it became easy to forget the human labour involved. The outcry from artists whose labour has been scraped into internet-sized datasets to train AI image generators highlights AI's utter dependence on data created by humans (Benzine, 2022). We can bring the human back into AI creativity through Hayles' notion of "punctuated agency" that "operates within regimes of uneven activity, longer periods when human agency is crucial, and shorter intervals when the systems are set in motion and proceed on their own without direct human intervention" (Hayles, 2017:32).

In the situation from Anna Ridler's *Mosaic Virus*, the concept of "punctuated agency" helps us understand how the cognitive creation of an artwork is distributed between the artist and the machine: To create *Mosaic Virus*, the artist makes conscious aesthetic decisions when collecting and classifying a dataset, choosing a model, and adjusting parameters. The machine learning model then operates within these parameters. Hayles discusses how human non-conscious processes "feed forward intuitions to conscious awareness" (2017:41); likewise, we can understand the technical non-conscious of machine learning models as generating a type of "technical intuitions" (Kronman, 2020). In *Mosaic Virus,* technical intuitions are expressed as visual hallucinations of impossible tulips. Technical intuitions are then fed forward to conscious cognition when Ridler curates the generated images as an artwork and when the audiences of *Mosaic Virus* make sense of what the machines hallucinated. This demonstrates the back-and-forth activity of punctuated agency between human and machinic agents within the same assemblage. The specificity of punctuated agency between the actors is not explicitly registered within the machine vision situation structure. Again, however, the format functions as an entryway or invitation to look closer and uncover the connections and processes behind the activities brought out by the verbs.

While our research captures how creative works represent machine vision, this method also opens for the possibility to address the agencies of users interacting with machine vision in creative works. In the *Database of Machine Vision in Art, Games and Narratives*, there are several examples of digital games and artworks that have the entity "user" as an agent. Some artworks facilitate some type of interaction between the viewers and machine vision technologies, in those cases we would include the viewer of the artwork as a "user agent" in our situation. Research interested in analysing user interaction would include the user as an agent in each machine vision situation, which shows again how the analysis model is a process of trade-offs between specificity and generalisability.

Tracing agency

The example analyses above demonstrate how the machine vision situations model can help in qualitative analyses of individual works, even though the model was initially designed to support quantitative analysis in the context of the *Database of Machine Vision in Art, Games and Narratives*. Taking the machine vision situation as the core unit of analysis allowed us to create structured data about the distribution of agency in machine vision interactions in a form that could be analysed quantitatively. We have begun to publish findings from our analyses of this dataset, identifying what the most common technologies are doing in the works in which they are used or represented, and what is most commonly being done to them (Rettberg, 2022b). For instance, drones are most commonly represented as recording, killing, transmitting and targeting and are represented as being controlled by human beings (Rettberg, forthcoming). Based on a machine learning analysis of the machine vision situations dataset, Rettberg (2022a) developed the method of 'algorithmic failure' to identify particularly salient cases for further study. The data collected on digital games formed the basis for three studies: one on the use of surveillance cameras as an interface in digital games, proposing the term 'cyborg vision' to account for the experience of embodied surveillance that these games offer to the player (Solberg, 2022a); a second on how holograms mediate between human and non-human actors in games (Solberg, 2021); and a third on enhanced vision in games and its relation to ideas of domination and power (Solberg, 2022b). The data collected on artworks formed the basis for Kronman's (2023) analysis of different approaches to hacking machine vision in art, revealing how art is used to expose bias in machine vision, as well as papers on non-conscious cognition and agency (Kronman, 2020), and aerial perspective and prediction in machine vision assemblages (Kronman, 2019). Gunderson (2021) drew on the database to analyse the representation of augmented reality in popular culture. On the level of the dataset as a whole, the information about each character's represented gender, race, age, species and sexuality in connection with the assigned verbs can also be used to study racial bias and other biases in how machine vision technologies are imagined across many creative works. The dataset has been deposited in the UiB Open Research Data repository and is available for futher research under a Creative Commons licence (Rettberg *et al.*, 2022a).

The main insight to come out of this cumulative process is just how many different things machine vision technologies can be seen to do when they become involved in events. When we started out, we naively assumed that these technologies' contribution to unfolding actions would mainly involve actions such as "seeing", "representing", or "recording", but it soon became clear that this was not the case. As shown by the examples above, these visually oriented actions are present, but they are often accompanied by actions that would not initially be ascribed to these types of technologies. Some of these actions have historically been thought of as fundamentally human capacities, such as imagining, interpreting, or deciding.

In other cases, the agential contributions of machine vision are inseparable from acts such as killing, creating, or tricking. What a specific technology can be said to do, whether in fiction or reality, can only be determined by attending to the actual situation in which it can be said to make a difference. Although we may imagine that we are in control and that these technologies are mere tools for human intentions, this idea is undermined by our own depictions of them in games, art, and narratives.

While each of our papers grew out of the quantitative work that went into making the database, several have turned to qualitative, close readings of works and situations as their main analytical method (Gunderson, 2021; Kronman, 2020; Kronman, 2023; Solberg, 2021; Solberg, 2022a; Solberg, 2022b). Based on this experience, this paper demonstrates how the concept of machine vision situation, and the formal data structure we developed for it, can serve as a tool to identify the compelling and relevant moments for analysis within a work, and function as a framework for not just quantitative, but also qualitative analysis of distributed agency. As such, it offers the theoretical basis for and a description of the machine vision situation as an analytical model, a model we argue is broadly applicable well beyond studies of machine vision in cultural works.

Although initially designed for the purpose of quantitative analysis across works, in the above three examples, the machine vision situation functions as the entry point for qualitative analysis of human-non-human assemblages. Simply identifying the actors and their doings in a situation produces new insights into how humans, technologies and other entities affect and are affected by each other in everyday interactions. This structure creates a rich foundation for closer analysis, inviting the researcher to trace the connections between the agents, understand the processes behind the verbs, and see what emerges from the situation as a whole. In *Drones Don't Kill People* the situations helped us understand the distribution of power between the actors. Furthermore, they brought out the pivotal moments through which something new emerged out of human/non-human assemblages. In *Detroit: Become Huma*n, we saw how the agency as represented within the work calls attention to its entanglement with the agency that was not included in the situation – that of the player and the console and the game itself as a technical/digital object. With the artworks *Myriad* and *Mosaic Virus*, analysing the machine vision situation reveals how cognition is distributed between human and technical actors in machine learning processes, while also calling attention to what does not fit into the situation structure – the process of punctuated agency between them. In each of these works, the machine vision situation reveals core aspects of the distribution of agency. Furthermore, the situation functions as a provocation, an incitement to look closer at what does not fit the structure, the excess, the complications – and that is often where a more profound understanding is to be found. By following a set structure, messy entanglements that might otherwise have been overlooked call attention to themselves and become obvious. A close reading of one or more related situations allows the researchers to disentangle complex distributions of agency in individual moments when machine vision technologies make a difference, while a large dataset of situations creates the foundations for a quantitative analysis of agency across different contexts. In so doing, the machine vision situation framework straddles the gap between quantitative and qualitative research methods.

While our project collected and analysed data about creative works, we are confident that the machine vision situation can be productively applied to other contexts in which human and non-human actors interact, including everyday encounters with machine vision and the discursive imaginaries through which people make sense of these technologies. Agency is distributed through action, interaction and narrativisation, and we expect that the framework of the machine vision situation can be applied similarly across disciplines and methodological approaches. Although conceptualised through the study of machine vision technologies, there is no reason why this method should be limited to machine vision technologies. As a method, the situation structure should be generalisable to any interaction between human and non-human actors in which agency is distributed in novel ways; hence, further theoretical work could expand it into a more general concept of a "sociotechnical situation".

By describing what both technologies and humans do in each situation and attaching actions to each of them, we produce information on how they are represented as actors and how actions, or doings, are distributed between them. As a result, we are able to trace the interpretative, meaning-making and communicative contributions of machine vision technologies, as well as their material functioning. This means that we as researchers also constitute an assemblage with the database; the database emerging through the parameters we give it, and the database subsequently directing our actions, interpretations and readings. The concepts, theories and ideas that arise from it cannot be traced back to a single researcher or even be limited to our efforts as a team; the database itself has an active role in their creation. Our thoughts and ideas, and future articles, will contain traces of the effects of this assemblage.

The knowledge we produce is conveyed with and through the structure we created to produce it, but that does not mean that it is contained by it. In the process of creating a quantitative dataset, discussing the structure, reading, playing, interacting with works, and agonising over which verbs to assign, a method emerges and takes on a life of its own. Quantitative data unfolds and reveals itself as a qualitative wellspring.

Ethical approval

Ethical approval and consent were not required.

Data availability

The underlying data has been deposited in the UiB Open Research Data repository DataverseNO, https://doi.org/10.18710/2G0XKN (Rettberg *et al.*, 2022a). This dataset contains all the entries on narratives, games, and artworks in the Machine Vision

Database. See Rettberg *et al.* (2022b) for a detailed description of the data.

This project contains the following underlying data:

- 00_README.txt
- 01_codebook.csv
- 02_technologies_sentiments_topics_definitions.csv"
- creativeworks.csv
- situations.csv
- characters.csv
- narrativegenres.csv
- situation_description.csv
- situations_visual.csv
- creators.csv
- worksinfo.csv
- machinevisionscripts.R

Data is available under the terms of the Creative Commons Zero "No rights reserved" data waiver (CC0 1.0 Public domain dedication).

Author contributions

Gunderson was the lead writer, developed the theoretical framework and wrote most of the paper. Rettberg acquired funding, was the principal investigator for the Machine Vision project, and had the initial idea of analysing agency using a database. Gunderson and Solberg first conceptualised the situations model by combining agents and verbs. Rettberg, Kronman, Gunderson and Solberg collectively developed the analytical model from there in discussion and by analysing works for the database. De Seta joined the project after the database was established and contributed to the database and further development of the ideas. Kronman wrote the section about digital art, Solberg wrote the section about games, Rettberg paragraphs about digital humanities methods and the project overall, and de Seta contributed to the overall framing and conclusion. All authors participated in revisions and editing, with Gunderson coordinating the process.

Acknowledgements

The authors would like to thank Annette Markham, who participated in discussions at the early stages of developing the situations model, and the many contributors to the Machine Vision database, especially Stein Magne Bjørklund, Linn Heidi Stokkedal and Kurdin Jacob. For a full list of contributors to the database, see Rettberg *et al.* (2022b).

References

Akrich M: **Actor Network Theory, Bruno Latour, and the CSI.** *Soc Stud Sci.* 2023; **53**(2): 169–173.
PubMed Abstract | **Publisher Full Text**

Bennett J: **Vibrant Matter: A Political Ecology of Things.** London: Duke University Press, 2010; 200.
Reference Source

Benzine V: **A.I. Should Exclude Living Artists From Its Database, Says One Painter Whose Works Were Used to Fuel Image Generators.** *ArtNet News.* 2022.
Reference Source

Berlant L: **Cruel Optimism.** Durham London: Duke University Press, 2011.
Publisher Full Text

Clarke AE: **Situational Analysis.** In: *The Blackwell Encyclopedia of Sociology.* edited by George Ritzer, 1st ed., Wiley, 2016; 1–2.
Publisher Full Text

Clarke AE: **Situating Grounded Theory and Situational Analysis in Interpretive Qualitative Inquiry**. In: *The Sage Handbook of Current Developments in Grounded Theory.* edited by Antony Bryant and Kathy Charmaz. Sage Handbooks. Los Angeles London New Delhi Singapore Washington DC Melbourne: SAGE reference, 2019.
Reference Source

Clarke AE, Washburn R, Friese C: **Situational Analysis in Practice: Mapping Relationalities Across Disciplines**. 2nd ed. New York: Routledge, 2022.
Publisher Full Text

Dehnert M, Leach R: **Becoming Human? Ableism and Control in Detroit: Become Human and the Implications for Human-Machine Communication.** *Human-Machine Communication.* 2021; **2**: 137–52.
Publisher Full Text

De Seta G: **Technologies of Clairvoyance: Chinese Lineages and Mythologies of Machine Vision.** In: *Machine Decision Is Not Final: China and the History and Future of AI.* edited by Benjamin Bratton, Anna Greenspan, and Bogna Konior. S.l.: MIT Press, 2022.

Dream Q: **Detroit: Become Human.** 2018.
Reference Source

Drucker J: **Why Distant Reading Isn't.** *PMLA.* 2017; **132**(3): 628–635.
Publisher Full Text

Feinberg M: **A Design Perspective on Data.** In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.* Denver Colorado USA: ACM, 2017; 2952–2963.
Publisher Full Text

Gunderson M: **Populærkulturelle Forestillinger Av Utvidet Virkelighet: Makt Og (u)Leselige Identiteter Når Verden Blir En Skjerm.** *Tidsskrift for Kjønnsforskning.* 2021; **45**(02–03): 89–104.
Publisher Full Text

Haraway D: **Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective**. In: *Technoscience: The Politics of Interventions.* edited by Kristin Asdal, Brita Brenna, and Ingunn Moser, Oslo: Unipub, 2007; 109–31.

Hayles NK: **Unthought: The Power of the Cognitive Nonconscious.** Chicago: University of Chicago Press, 2017.
Reference Source

Heidari S, Babor TF, De Castro P, *et al.*: **Sex and Gender Equity in Research: Rationale for the SAGER Guidelines and Recommended Use.** *Res Integr Peer Rev.* 2016; **1**(1): 2.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Jannidis F: **Character**. In: *Handbook of Narratology.* edited by Peter Hühn, Matías Martinez, John Pier, and Wolf Schmid, 2. ed., De Gruyter Handbook. Berlin: De Gruyter, 2014; 14–29.

Kronman L: **The Deception of an Infinite View - Exploring Machine Vision in Digital Art.** In: *Proceedings of the Politics of the Machine, Beirut.* 2019.
Publisher Full Text

Kronman L: **Intuition Machines: Cognizers in Complex Human-Technical Assemblages.** *A Peer-Reviewed Journal About.* 2020; **9**(1): 54–68.
Reference Source

Kronman L: **Hacking Surveillance Cameras, Tricking AI and Disputing Biases: Artistic Critiques of Machine Vision.** *Open Library of Humanities Journal*. 2023; **9**(2).

Latour B: **Why Has Critique Run out of Steam? From Matters of Fact to**

Matters of Concern. *Crit Inquiry.* 2004; **30**(2): 225–248.
Publisher Full Text

Latour B: **Reassembling the Social: An Introduction to Actor-Network-Theory.** Clarendon Lectures in Management Studies. Oxford; New York: Oxford University Press, 2005; 301.
Reference Source

Leach R, Dehnert M: **Becoming the Other: Examining Race, Gender, and Sexuality in *Detroit: Become Human.*** *Review of Communication.* 2021; **21**(1): 23–32.
Publisher Full Text

Marchand JS: **Non-Human Agency**. In: *Posthuman Glossary.* edited by Rosi Braidotti and Maria Hlavajova, New York: Bloomsbury Academic, 2018; 292–95.

Massumi B: **Parables for the Virtual: Movement, Affect, Sensation**. *Twentieth Anniversary edition with a new preface. Post-Contemporary Interventions*. Durham: Duke University Press, 2021; 408.
Reference Source

Newitz A: **Drones Don't Kill People.** *Lightspeed Magazine.* 2014; (54).
Reference Source

Niantic: **Pókemon Go.** Android, 2016.

Pierazzo E: **Textual Scholarship and Text Encoding.** In: *A New Companion to Digital Humanities.* edited by Susan Schreibman, Ray Siemens, and John Unsworth, John Wiley & Sons, 2015; 307–21.
Publisher Full Text

Rettberg JW: **Situated Data Analysis: A New Method for Analysing Encoded Power Relationships in Social Media Platforms and Apps.** *Humanit Soc Sci Commun.* 2020; **7**(1): 5.
Publisher Full Text

Rettberg JW: **Algorithmic Failure as a Humanities Methodology: Machine Learning's Mispredictions Identify Rich Cases for Qualitative Analysis.** *Big Data & Society.* 2022a; **1**(6).
Publisher Full Text

Rettberg JW: **What do fictional characters do with machine vision in games, novels and movies?** Code notebook, 2022b.
Reference Source

Rettberg JW: **Machine Vision: How Algorithms Are Changing the Way We See the World**. Cambridge: Polity Press, 2023.
Reference Source

Rettberg JW: **Controlling Drones in Contemporary Science Fiction.** In: *Drones in Society: New Visual Aesthetics*, edited by Elisa Serafinelli. Palgrave, Forthcoming.
Reference Source

Rettberg JW, Gunderson M, Kronman L, *et al.*: **Mapping Cultural Representations of Machine Vision: Developing Methods to Analyse Games, Art and Narratives.** In: *Proceedings of the 30th ACM Conference on Hypertext and Social Media*, Hof Germany: ACM, 2019; 97–101.
Publisher Full Text

Rettberg JW, Kronman L, Solberg R, *et al.*: **A Dataset Documenting Representations of Machine Vision Technologies in Artworks, Games and Narratives.** DataverseNO, V2, [dataset], 2022a.
http://www.doi.org/10.18710/2g0xkn

Rettberg JW, Kronman L, Solberg R, *et al.*: **Representations of Machine Vision Technologies in Artworks, Games and Narratives: A Dataset.** *Data in Brief.* 2022b; **42**: 108319.
http://www.doi.org/10.1016/j.dib.2022.108319

Ridler A: ***Myriad (Tulips)***. 2018.
Reference Source

Ridler A: **Mosaic Virus.** 2019a.
Reference Source

Ridler A: **Using AI to Produce 'Impossible' Tulips.** Interview by Elaine Ayers, Hyperallergic Magazine, 2019b.
Reference Source

Ridler A: **An Interview with Anna Ridler, the Artist behind Mosaic Virus and Myriad (Tulips).** Interview by Syon Lee, Felix Online, 2019c.
Reference Source

Schubert S: **'Liberty for Androids!': Player Choice, Politics, and Populism in Detroit: Become Human.** *European Journal of American Studies.* 2021; **16**(3): 1–18.
Publisher Full Text

Sinclair S, Rockwell G: **Text Analysis and Visualization: Making Meaning Count.** In: *A New Companion to Digital Humanities.* edited by Susan Schreibman, Ray Siemens, and John Unsworth, John Wiley & Sons, 2015; 274–90.
Publisher Full Text

Solberg R: **Hologrammer i grenseland: Ikke-menneskelige aktørers tilstedeværelse og handlingsrom i spill.** *Norsk medietidsskrift.* 2021; **28** (4): 1–20.
Publisher Full Text

Solberg R: **(Always) Playing the Camera: Cyborg Vision and Embodied Surveillance in Digital Games.** *Surveillance & Society.* 2022a; **20**(2): 142–56.
Publisher Full Text

Solberg R: **"Too Easy" or "Too Much"? (Re)imagining Protagonistic Enhancement through Machine Vision in Video Games.** *Przegląd Kulturoznawczy.* 2022b; **4**(54): 548–69.
Publisher Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔ ✔

Version 2

Reviewer Report 22 April 2024

https://doi.org/10.21956/openreseurope.18792.r38965

© **2024 Salter A.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Anastasia Salter**
University of Central Florida, Orlando, Florida, USA

I assessed and approve the revisions.

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 22 April 2024

https://doi.org/10.21956/openreseurope.18792.r39522

© **2024 Veel K.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Kristin Veel**
University of Copenhagen, Copenhagen, Denmark

It has been a great pleasure to read this article which I think provides an important methodological contribution and presents a series of interesting analyses that will be of great value to further work in the field. The article aptly outlines “machine vision situation” as a theoretical concept and demonstrates its usefulness as an analytical tool which allows for the integration of distant and close reading methods in fruitful ways. It is of particular interest to see how the method works across literature, games and art. This scope allows for valuable considerations regarding how to address the presence of machine vision technologies in a text/artwork/game and focusing on its implications for the narrative/story world/aesthetics across different fictional genres.

The authors have addressed the concerns raised by the previous reviewers and the article now stands as a thorough study that is innovative and thought-provoking both methodologically, theoretically and analytically.

One minor typo:
"The police, the algorithm, or its programmers?" is repeated twice in the introduction.

Is the work original in terms of material and argument?
Yes

Does it sufficiently engage with relevant methodologies and secondary literature on the topic?
Yes

Is the work clearly and cogently presented?
Yes

Is the argument persuasive and supported by evidence?
Yes

If any, are all the source data and materials underlying the results available?
Yes

Does the research article contribute to the cultural, historical, social understanding of the field?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: My research engages with the cultural implications of technologies in art, literature and film and I have for instance worked specifically on drones as well as machine vision aesthetics.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 05 April 2024

https://doi.org/10.21956/openreseurope.18792.r38964

© 2024 Tosca S. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Susana Tosca
University of Southern Denmark, Odense, Denmark

I have read the improved version of the article and I think that it satisfactorily addresses the concerns raised in my review and more. The authors have added essential context so that their contribution shines in their own right.

Is the work original in terms of material and argument?
Yes

Does it sufficiently engage with relevant methodologies and secondary literature on the topic?
Yes

Is the work clearly and cogently presented?
Yes

Is the argument persuasive and supported by evidence?
Yes

If any, are all the source data and materials underlying the results available?
Yes

Does the research article contribute to the cultural, historical, social understanding of the field?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: I am a media researcher, with speciality in reception, audience studies, entertainment media.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 03 November 2023

https://doi.org/10.21956/openreseurope.17395.r34881

© 2023 Salter A. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Anastasia Salter
[1] University of Central Florida, Orlando, Florida, USA
[2] University of Central Florida, Orlando, Florida, USA

I enjoyed reading this article and found the framing in humanities discourse (and particularly creative works) unusual for this type of work. This is a timely and compelling study of representations of machine vision technologies, with a particularly useful investigation of the agency given to non-human actors in the context of a range of media. The wide range of works included in the underlying dataset allows the authors to interrogate those representations with attention to how we categorize and understand technologies such as facial recognition, satellites, drones, etc in the context of "human-non-human assemblages.

Given the current popular attention to machine vision related work (and particularly the intense reactions to image generators), this article's close analysis of three works will be of particular use to those guiding conversations and mediating reactionary discourse. As an educator, I certainly will be drawing on this useful analysis of "Drones Don't Kill People" for my own AI humanities classes, and I appreciate its inclusion here.

From a game studies perspective, the only section that raised some concern for me was the discussion of Detroit: Become Human. This game's representation of Black masculine identity and use of parallels between android segregation and US history specifically has been the subject of some significant work already (I'd particularly recommend Rebecca Leach and Marco Dehnert's article from 2021). This analysis decontextualizes the human (and non-human) stakes of the agents by not engaging with these important aspects of embodiment, which is particularly critical to the author's discussion of "a society that places them [androids] in positions of servitude and submission." As Susana Tosca has already noted in the previous reviewer report, there's also a positioning of the interactor here that I concur warrants further discussion.

References

1. Dehnert M, Leach R: Becoming Human? Ableism and Control in Detroit: Become Human and the Implications for Human-Machine Communication. *Human-Machine Communication*. 2021; **2**: 137-152 Publisher Full Text

Is the work original in terms of material and argument?
Yes

Does it sufficiently engage with relevant methodologies and secondary literature on the topic?
Yes

Is the work clearly and cogently presented?
Yes

Is the argument persuasive and supported by evidence?
Yes

If any, are all the source data and materials underlying the results available?

Yes

Does the research article contribute to the cultural, historical, social understanding of the field?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Digital humanities, electronic literature, game studies

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 15 Mar 2024

Marianne Gunderson

Dear Prof. Salter,

Thank you for your constructive and helpful review. We appreciate your thoughtful feedback and found it very useful for our revision process. We have made the following changes based on your comments:

As mentioned in our response to Tosca's review, we agree that much more can be said about player or user agency in games and art; their absence in our analyses was an unforeseen result of our choice of examples. We have added a paragraph to both the art and games sections emphasising this and indicating how the user or player could come into play, as well as when and why we chose to include or exclude the player or user as an actor in our analyses. Here, we also point towards how the boundaries for which actors to include must be redrawn according to the priorities and limitations of the specific research project and indicate what this may bring out in the resulting analysis. Per your suggestion, we also brought the context of race and embodiment in relation to Detroit:Become Human into this section, drawing on the article you recommended as well as a couple of others.

With these revisions, we believe that we have addressed all your major objections and hope that this version will meet your expectations.

Competing Interests: No competing interests were disclosed.

Reviewer Report 28 September 2023

https://doi.org/10.21956/openreseurope.17395.r34883

© 2023 Tosca S. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Susana Tosca
[1] University of Southern Denmark, Odense, Denmark
[2] University of Southern Denmark, Odense, Denmark

I really enjoyed reading this article. It provides a valuable contribution to understanding the topic of machine vision and the way we consider media-human configurations in general (which they by the end call "sociotechnical situations"), so it also has methodological interest. The "Machine vision situation" as a productive idea that allows for more nuanced analysis than for instance a thematic analysis, it refer to instances not just where machine vision appears but where it becomes an issue, there are tensions, breakdowns, etc. The article provides an alternative way of articulating and analysing non human agency, something that STS literature not always achieves. It was useful and inspiring.

I found the paper particularly strong in the way it shows that a combination of digital methods and qualitative analysis, can produce interesting results, and it fully convinces in its claim that designing a database is an interpretive art. The three analyses cases, although brief, also do a good job of show how machines seeing can be broken up in a lot more agentic verbs like generating, classifying, identifying...The examples are strong and well argued. I must admit that after reading the first example I thought: is this all there is to analyse? But then it struck me that the power of this method is not in the individual semantic depth, but the sum of all these agency oriented analyses in a huge database that can yield interesting results. Perhaps this cumulative aspect could be made more clear in the paper. In short, to expand/explain their claim that “Taking the machine vision situation as the core unit of analysis allowed us to create structured data about the distribution of agency in machine vision interactions in a form that could be analysed quantitatively”. I would also like to know more about how the analysis of fictional material can enrich our understanding of machine vision "in real life", as it is just mostly assumed it does, without an actual argument.

I have also indicated above that there are possibilities for improvement, which I would recommend the authors address to review the paper. The first and most important criticism in a paper centered around analysing situations is that the method of situational analysis is not even mentioned. Adele Clarke has worked on this method (inspired by grounded theory) for many years now, starting a whole school of thought around it, with countless examples of analyses and applications. This omission is even more glaring given that Clarke is also inspired by STS and including technology and non human agency in her way of approaching situations. Some of the situation theorists she bases her method on are also absent here, like C. Wright Mill or Donna Haraway.

The second important issue in my view is that in the discussion of computer games and artworks, the agency of the user/interactor is not a part of the analysis, which is limited to the, shall we call it, diegetic aspects of machine vision. But unlike in literature, users have to activate/manipulate machine vision in these assemblages. Can this be addressed somehow in the discussion?

As a minor issue, the Introduction section is a bit repetitive. I think I counted three times of similar expressions to "we came up with the concept of the machine vision situation". It can be tightened up. Also, the "previous literature" part seems to be a bit less well integrated: the Jean Bennett section reads like a detour, and the Hayles section is also unclear in its contribution to this specific paper. They need to either be unpacked more or maybe go altogether if the authors agree that

situational analysis is a relevant background to be incorporated (which would require quite some space).

My final remark is that I would not call this "method agnostic" as the authors do several times throughout the paper. True, the situational approach can be applied to different kinds of situations, but it has a very clear ANT inspiration that would certainly not be considered neutral by most people in the Humanities.

Is the work original in terms of material and argument?
Yes

Does it sufficiently engage with relevant methodologies and secondary literature on the topic?
Partly

Is the work clearly and cogently presented?
Yes

Is the argument persuasive and supported by evidence?
Partly

If any, are all the source data and materials underlying the results available?
No source data required

Does the research article contribute to the cultural, historical, social understanding of the field?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: I am a media researcher, with speciality in reception, audience studies, entertainment media.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 15 Mar 2024

Marianne Gunderson

Dear Prof. Tosca,

Thank you for this thorough reading and detailed review of our paper. We found your comments to be of great aid in our revisions and we believe that they have significantly contributed to the improvement of this text. We have made the following changes based on your feedback:

In response to your comment on the cumulative aspects of the method, we have restructured some sections so that the quantitative results are highlighted in the conclusion and added a paragraph about the main findings from the database. We further clarified the relationship between the qualitative and quantitative contributions of this method by emphasising how the method also serves as an entry point to qualitative analysis of distributed agency.

Regarding your query about Clarke's situational analysis, we agree that this was a glaring omission. We have added a paragraph discussing how the machine vision situation method relates to situational analysis, as well as how it could potentially supplement it. We have also included a paragraph emphasising the connection between our work and Haraway's concept of situated knowledges, which has been part of the theoretical backdrop for our work. Also, per your suggestion, we have shortened and integrated the Bennett section and moved the Hayles paragraph into the Mosaic Virus section, where its relevance should be more apparent.

We agree that much more can be said about player or user agency in games and art, their absence in our analyses were an unforeseen result of our choice of examples. We have added a paragraph to both the art and games sections emphasising this and indicating how the user or player could come into play, as well as when and why we chose to include or exclude the player or user as an actor in our analyses. Here, we also point towards how the boundaries for which actors to include must be redrawn according to the priorities and limitations of the specific research project and indicate what this may bring out in the resulting analysis.

The term 'method agnostic' was used to indicate that the method can be used both as a quantitative and a qualitative tool, but we have removed this term to avoid ambiguity.

Finally, in response to your observations, we have sharpened the introduction by removing superfluous repetitions.

We believe that we have addressed all your major objections and hope that this version will meet your expectations.

Competing Interests: No competing interests were disclosed.