Mapping Cultural Representations of Machine Vision

Developing Methods to Analyse Games, Art and Narratives

Jill Walker Rettberg

Department of Linguistic, Literary
and Aesthetic Studies
University of Bergen
Norway
jill.walker.rettberg@uib.no

Marianne Gunderson

Department of Linguistic, Literary
and Aesthetic Studies
University of Bergen
Norway
Marianne.gunderson@uib.no

Linda Kronman

Department of Linguistic, Literary
and Aesthetic Studies
University of Bergen
Norway
linda.kronman@uib.no

Ragnhild Solberg

Department of Linguistic, Literary and Aesthetic Studies
University of Bergen
Norway
ragnhild.solberg@uib.no

Linn Heidi Stokkedal

Department of Linguistic, Literary and Aesthetic Studies
University of Bergen
Norway
linn.stokkedal@uib.no

ABSTRACT

Machine vision technologies are increasingly ubiquitous in society and have become part of everyday life. However, the rapid adoption has led to ethical concerns relating to privacy, agency, bias and accuracy. This paper presents the methodology and preliminary results from a digital humanities project that maps and categorises references to and uses of machine vision in digital art, narratives and games in order to find patterns to help us analyse broader cultural understandings of machine vision in society. Understanding the cultural significance and valence of machine vision is crucial for developers of machine vision technologies, so that new technologies are designed to meet general needs and ethical concerns, and ultimately contribute to a better, more just society.

KEYWORDS

Machine vision, computer vision, methodology, digital humanities, network analysis, science fiction, narratives, computer games, digital art.

ACM Reference format:

Jill Walker Rettberg, Marianne Gunderson, Linda Kronman, Ragnhild Solberg, Linn Heidi Stokkedal. 2019. Mapping Cultural Representations of Machine Vision: Developing Methods to Analyse Games, Art and Narratives. In *Proceedings of ACM Conference on Hypertext and Social Media (HT '19)*, September 17–20, 2019, Hof, Germany. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3342220.3343647>

1 Utopias and Dystopias of Machine Vision

The premise of this paper is that cultural anxieties, desires and reflections about new technologies are mirrored in the

aesthetic artefacts we create, share and consume. Therefore, by analyzing the ways a technology such as machine vision is represented and simulated in art, games and narratives, we will gain a better overview of the cultural understanding of machine vision. This paper presents the methodology we are using to map cultural representations of machine vision, using a database designed to capitalize on the hypertextual structuring of knowledge.

Machine vision (defined as the registration, analysis and representation of visual data by machines and algorithms) impacts our everyday lives in myriad ways. Sometimes it is visible and allows individuals to explicitly play with and control it, as with Snapchat's selfie lenses [24] or the filters that can be added to photographs. Other times it is outside the users' control, as with the facial recognition and object recognition algorithms that sort personal photos in the "Memories" function on an iPhone, or that alert a user when their home surveillance system recognises a particular face [15, 21]. The aesthetic inference algorithms built into cameras, that adjust the quality of a photograph as it is being taken, are another example of this kind of automated assistance [23]. In other cases, machine vision is used by corporations or government agencies to monitor, target, control or assist individuals in various ways. These are typically the uses of machine vision that evoke the most anxiety in mainstream media debates. Some appear benign, as the self-driving cars that use machine vision to maneuver the streets. Others are more obviously problematic, such as commercial uses of facial recognition to target ads, or machine learning that predicts the risk of shoplifting based on visible behavior [2]. In image generation, deepfake videos have received significant attention in the media. One frequent concern is that deepfake videos presenting politicians saying something they did not actually say could impact societal trust in a democratic public sphere [11]. Law enforcement and the military are the domains which present great potential risk to individuals if mistakes are made. For instance, if facial recognition algorithms are faulty, innocent people may be suspected of a crime they did not commit, or even worse, innocent civilians could be targeted and killed by lethal, autonomous

This work is licensed under a Creative Commons Attribution-NoDerivs International 4.0 License.

HT '19, September 17–20, 2019, Hof, Germany.

© 2019 Copyright is held by the owner/author(s).

ACM ISBN 978-1-4503-6885-8/19/09. <https://doi.org/10.1145/3342220.3343647>

weapons systems. Bias in machine vision often occurs due to bias in the datasets the algorithms are trained on, as Joy Buolamwini has found in her audits of facial recognition systems from IBM, Microsoft and other companies [3]. Algorithmic bias is an important societal issue [10, 20, 29].

As part of a larger project, we use representations of machine vision in art, narratives and games to conceptually map cultural attitudes towards machine vision. We are developing a database to document works that use or depict machine vision, with a coding schema that will allow us to analyse relationships between different characteristics of the works. An example of an artwork that uses machine vision technologies is Christian Moeller's *Cheese* (2003), which examines sincerity using emotion recognition. Six actresses were asked to smile for as long as they were able to. As soon as the actress dropped her smile, an emotion recognition system would register a drop of happiness level and an alarm would go off reminding the actress to keep smiling. Other works we have documented do not use the technologies directly but portray them in different ways. For instance, in the game *Alien: Blackout* (2019), the player attempts to save a space station crew from deadly contact with an alien. In order to do this, the player has access to surveillance cameras. Maneuvering the crew to safety in the game is thus wholly dependent on surveillance camera technology.

2 Hypertextual structure for scholarly inquiry

The early scholarship on hypertext emphasised the associative organization of ideas [4, 18, 19] and authored hypertexts [30]. In his work on scholarly hypertext, David Kolb noted that informational hypertext and literary hypertext are different from hypertexts featuring scholarly inquiry, and asked “how in hypertext we might allow not just connection but assertion, self-representation, and debate about criteria” [14]. The growth of the digital humanities in the last decade has seen an increase in databases and other digital tools for developing research in the humanities, but distant reading [17] and quantitative methods in the humanities have been criticised for their potential disregard for the qualitative interpretation that is at the core of humanities methodologies [8], and for a sometimes poor use of quantitative methods [5].

Perhaps a hypertextual approach can help us escape the “digital humanities wars” [6], where the digital humanities are framed either as saviour or destroyer of the humanities. In the humanities, knowledge is understood to be situated [12] and constructed [7]. Binary oppositions and strict categorisations are seen as reductive: they obscure details and complications. How then can the digital humanities develop methods that encourage hermeneutic nuance, while using digital methods? This paper provides an example of a digital humanities project where objectivity and representivity are neither possible nor desirable. In many ways it is a continuation of early work on hypertext scholarship, which similarly emphasised relationships and nuance above binary oppositions or linear trajectories [14].

3 Methodology

The research presented in this paper is part of the larger project *Machine Vision in Everyday Life: Playful Interactions with Visual Technologies in Digital Art, Games, Narratives and Social Media*. In this larger project, the work presented in this paper will be complemented by close readings of a selection of works, by ethnographic fieldwork with users and developers of machine vision, and by a thorough survey of the history and theories of pre-digital and contemporary machine vision.

We are at an early stage in this research, which means that we are writing while deep in discussions about methodological choices. The main purpose of this paper is to describe the methodology we are developing, to share some preliminary results, and to elicit feedback. While data visualisation as a method necessitates the structuring of information in a systematic manner, we do not claim that these structures are inherent in the works included in this database. Both the creation of the coding schema, and the entering of individual works and machine vision situations into the database is an interpretative, meaning-making process. The knowledge thus produced is mediated by the analytic sensibilities, academic backgrounds, and existing knowledge brought to the table by each member of the team.

3.1 Research questions and objective

The overall objective of the *Machine Vision* project is to develop a novel theory that explains how the algorithmic machine vision of the 21st century affects the way ordinary people see themselves and understand the world. This is broken this down into three research questions:

1. Which kinds of agency develop, and which are limited in the interaction between individual users and machine vision?
2. How does machine vision lead us to experience the visual as data that can be manipulated, and does this lead us to see the world and ourselves as malleable?
3. Which values are embedded in machine vision, and which biases are introduced or supported?

Our objective in the part of the project presented in this paper is to identify patterns in how machine vision is portrayed in art, games and narratives. For instance, we may find that certain kinds of agency are more common in games than in digital art, or in dystopic representations of machine vision in contrast to more positive representations. Or, we may find that certain technologies, e.g. facial recognition, optical implants or augmented reality, are understood as neutral tools, while other machine vision technologies tend to be portrayed as hostile or alien. Identifying such patterns will then guide our qualitative research design for the later stages of the project.

Using a database built in Drupal¹, the team is developing an extensive database of interpretative metadata about games, digital art and narratives (including science fiction novels, movies, television shows and digital or online narratives) that reference or simulate machine vision. Our database architecture is an adaptation of the ELMCIP Electronic Literature Knowledge Base,

¹ The database is accessible at <http://machinevision.elmcip.net>.

and builds upon work done in the ELMCIP project [22, 25–28]. The Electronic Literature Knowledge Base documents relationships between actors and objects (e.g. authors, events, creative works and critical writing), thus making the relationships in a field of practice explicit [26]. In the machine vision database, the works documented do not have pre-existing relationships. Rather, we are creating links between the works through our interpretations and the conceptual framework that we are developing through our coding schema.

3.1 Selection of works

At the time of writing (July 2019) we have documented 136 artworks, narratives and games that thematise or use machine vision technologies. We aim to have at least 1000 entries fully coded by December 2019. We are using a strategic sampling technique aimed at documenting a wide array of both popular examples and outliers, in order to capture both mainstream representations of machine vision, and more experimental approaches. Our selection method combines expert knowledge with systematic searches of existing databases (see Table 1). The team consists of topic experts. Linda Kronman is a practicing digital artist and has a curatorial and theoretical background in the field. Linn Heidi Stokkedal is an art historian and professional photographer. Ragnhild Solberg is a games scholar. Marianne Gunderson comes from a background in gender studies and has extensive knowledge of fan fiction, creepypasta and science fiction. Jill Walker Rettberg is an expert on digital narrative, digital art and digital humanities. Each of the team members uses their existing knowledge of the field to identify as many cases as possible, and adds these to the database.

Table 1: Databases used to find examples of machine vision.

Database	Genre	Information used
Steam	Games	Tags, suggestions
Rhizome	Digital Art	Keywords in titles
Archive of Digital Art	Digital Art	Keywords, category search
Worldcat	Narrative (Novels, short stories)	Titles, book blurbs
Google Books	Narrative (Novels, short stories)	Full text search
IMDB	Narrative (Movies, TV shows)	Titles, summaries
Archive of Our Own	Narrative (Fan fiction)	Tags, keyword search
Creepypasta Wiki	Narrative (Creepypasta)	Tags, full text search
ELMCIP	Narrative (Electronic literature)	Tags, platforms, descriptions
Goodreads	Narrative (Novels, short stories)	Titles, quotes, similar literature

In addition to the works we collect from databases, we have supplemented our findings using a form of snowball sampling. For

instance, when a team member identified one interesting art work, she looked at the artist's other works and followed the artist on Instagram or other social media. This frequently led to discovering new connections and other related artists and artworks. This pleasure of moving through a database has been mentioned by others as part of the experience of analysis [13]. Actively discussing the project in social media, for instance on Tumblr, Facebook and Twitter, led to people suggesting new works we had not previously encountered. In addition, we found several works using algorithmic suggestions, such as Google's "people have also searched for" or Steam's "more like this". Using Google and other search engines to search for words such as "GAN art" or "AI art" resulted in articles describing exhibits, artists or artworks. A more direct approach was emailing new media artists and festivals for their recommendations on artworks that would be relevant to this project, in order to include artworks that are not as dominant in the media or in search engines.

Clearly it is impossible to document *all* references to machine vision in all art, narratives and games. Despite our conscious attempts to be inclusive, selection is certainly limited by the languages mastered by the team and our cultural context as Northern Europeans. To mitigate this, we plan to a) dedicate time to searching for examples outside of our cultural context and b) seek advice from experts with diverse cultural experiences.

3.2 Coding schema

Given that our goals are exploratory, we needed to design a coding system that was open-ended yet would allow us to focus on the project's primary research questions, about agency, objectivity, and bias. We began by thinking of the work (artwork, novel, movie, game etc.) as our unit of analysis, but after several months of iterative analysis, coding and discussion, we found that identifying individual *situations* in works in which machine vision occurs was much more productive. It also allowed for more nuance, as many works present different attitudes to machine vision in different situations. Each work in the database is now linked to entries on individual characters in the work (if relevant) and to entries describing situations where machine vision occurs.

The database includes games, artworks, narratives and other cultural artifacts, all with considerable internal variation in form and media, so creating a coding schema that would be applicable across all the different types of works was a challenge. In order to make sure that the codes remain relevant and applicable, the coding schema has been under continual revision and the categories listed below may still be subject to future revision. Data entry is done by team members, and individual entries are discussed by the whole group when we are in doubt and while developing the coding system.

3.2.1 Works. We record basic metadata about each work such as the title, creator (artist/author/director), the year published, the genre, and designed a set of codes to identify machine vision technologies used and/or referenced in the work, themes present in the work (e.g. climate, family, inequality, companionship), and what the general attitude towards machine vision was (e.g. alien, dangerous, exciting, helpful, hostile, overwhelming).

3.2.2 Characters. Inspired by intersectional feminist theory, we developed a system in which each character would be entered separately, in order to retain information about how specific identity categories are linked. This will enable us to analyse how intersecting identities come into play in these works. We are still discussing how to code for the representation of marginalisation and of race/ethnicity, both of which can be quite complex. We want to avoid reducing racial representation to stereotypes, but at the same time, it is obviously relevant to our research questions about bias to consider whether attitudes towards or uses of machine vision are different for characters of colour or marginalised groups in comparison to privileged or dominant groups in the fictional world or the work of art. As always, it is the *representation* of race and machine vision we are interested in, so this is less fraught than projects that attempt to codify real-life identities [9], but still potentially problematic.

Table 2: The characteristics of two of the main characters from the TV series *Neon Genesis Evangelion* (1995).

Shinji Ikari	Rey Ayanami
Customizable: Not customisable	Customizable: Not customisable
Marginalized: Not marginalised	Marginalized: Not marginalised
Race/ethnicity: Asian	Race/ethnicity: Asian
Gender: Male	Gender: Female
Species: Human	Species: Machine
Age: Young adult	Age: Young adult
Agency: Agency increased by machine vision	Agency: Agency increased by machine vision

3.2.3 Situations. Situations are specific instances of machine vision being used or referenced within a work, separated by their technologies and the sentiment of characters involved with the technology. We try to capture attitudes to machine vision at three levels: the attitudes of the person(s) operating the machine vision, of the person(s) being viewed by the machine vision, and the attitude expressed in the situation as a whole.

A work may have several situations because it presents different machine vision technologies. In the music video of Alan Walker's *Alone* (2016), we have identified three distinct machine vision situations, *Alone* (Webcam), *Alone* (Surveillance camera), and *Alone* (Augmented reality). In the Webcam situation, we see Alan Walker fans using video chat with webcams and agreeing to meet at a specific location (see Figure 1). In the database entry we describe characteristics of the operator of the machine vision and the people or things that are viewed by the machine vision. In *Alone* (Webcam) these are the same people. To describe the operators and the viewed, we either link to the main characters in the work, or when individual characters aren't the main focus, as in *Alone*, we describe characteristics of people in the group: human, white, persons of colour, female, male, young adult.

In the artwork *The Virtual Watchers* by Joanna Moll (2016) we find an example of a situation where operators and those viewed have different sentiments towards machine vision. The work presents material from a group of amateurs who use public border surveillance cameras to try to spot illegal immigrants. We coded the operators of the machine vision technology as human, white

and privileged, and noted that the operators are represented as feeling that the machine vision is exciting, helpful and protective.

Figure 1: Screenshot from Alan Walker's music video *Alone* © Alan Walker 2016 (<https://www.youtube.com/watch?v=1-xGerv5FOk>).

We coded the watched people as a targeted group, human, persons of colour and marginalised, and noted that the artwork represents them as feeling that the machine vision is hostile and oppressive.

We also note the overall attitude to machine vision as represented in each situation: oppressive for *The Virtual Watchers* (which is a very critical artwork) and empowering, helpful and pro-social for *Alone*.

We are also experimenting with ways of coding what the machine vision is doing as process, rather than emphasising a subject/object binary of viewer and viewed. For *Alone* we note that the machine vision is capturing, and at a larger societal context it is *inspiring/motivating* people, and that the relationships between humans and machine vision are portrayed as cooperative.

4 Preliminary Findings

When we had documented around 100 works, we began preliminary analysis of the coded data by exporting data into Gephi for network analysis, looking at different aspects of the data separately. Figure 2 shows a visualisation of the relationships between works and the sentiments we identified in each work. Clustering is done using the ForceAtlas2 algorithm in Gephi, which places nodes with shared connections closer to each other. We made the titles of works pale gray to emphasise the relationships between sentiments, and sized nodes by degree, so sentiment tags that have been applied to more works are larger than those that are rarely used. As Figure 2 shows, "helpful" is the most common sentiment. At this stage of research, we can't conclude that that means that machine vision is most often presented as helpful in art, games and narratives. This *may* be the case but could be caused by the small initial sample or by coding error. Since we are using network analysis not only to visualise results, but also as a strategy to understand the data, we will use these findings in continuing to refine the coding schema by examining works currently tagged as helpful.

[illegible]

5 Conclusion and further research

ACKNOWLEDGMENTS

REFERENCES

- 101