Languages of Representation

written by Jeremy Huggett | 11/12/2020



Adapted from original by Pete Linforth from Pixabay

Shawn Graham recently pointed me (and a number of colleagues!) to a new paper entitled 'Computer vision, human senses, and language of art' by Lev Manovich (2020) in a tweet in which he asked what we made of it … so, challenge accepted!

Lev Manovich is, of course, a professor of computer science and a prolific author, focusing on cultural analytics, artificial intelligence, and media theory, amongst other things. In this particular paper, he proposes that numbers, and their associated analytical methods, offer a new language for describing cultural artefacts. The idea that this is novel may be news to those who have been engaged in quantitative analyses across the humanities since before the introduction of the computer, but aspects of his argument go further than this. The actual context of the paper is as yet unclear since it is online first and not yet assigned to a volume. That said, a number of other open access online first papers in *AI & Society* seem to address similar themes, so one might imagine it to be a contribution to a collection of digital humanities-related papers concerning images and computer vision.

It's an interesting paper, not least since – as Manovich says himself (p2) – it presents the perspective of an outside observer writing about the application of technological methods within the humanities. Consequently it can be tempting to grump about how he "doesn't understand" or "doesn't appreciate" what is already done within the humanities, but it's perhaps best to resist that temptation as far as possible.

Manovich asks what the intellectual consequences of adopting computer vision methods in humanities research might be (p1), although his answer doesn't really extend beyond the use of numbers to better capture information and provide better, more precise descriptions. However, in posing the question, he maintains a largely clear divide between a perception of humanities researchers reliant on linguistic description and categorisation, and computer vision researchers

using numerical description instead. He suggests that although the use of computational methods has become popular in areas such as literary studies and history (i.e. primarily textual humanities) there is far less interest in areas which analyse visual culture (p2). He therefore concludes that "humanists that study the visual have been slow to make use of computers" (p2) and supports this claim by reference to the pages of a number of digital humanities journals and conference programmes (p3). Archaeologists might rightly bristle at this claim, but it becomes apparent that Manovich's definition of the humanities is one which largely excludes archaeology (it is referenced only once throughout the paper). This tends to reinforce the perception of a relatively remote relationship existing between digital archaeology and digital humanities (e.g. Huggett 2012) so perhaps it is unfair to draw too many parallels with archaeology in this paper.

Manovich argues that the application of new techniques and tools in the visual digital humanities are used to answer existing questions and generate new questions, but don't change our research in fundamental ways (p2), unlike in fields such as computer vision. He suggests that the use of machine learning to refit fresco fragments, for instance, is useful but doesn't lead to big new ideas for the field (p2). It's an argument I've made myself on several occasions – that a key objective of digital archaeology should be to advance archaeological theory and practice. His archaeological example isn't perhaps the best one, however, and in a way it actually demonstrates the limitations of apparently archaeological applications undertaken by digital researchers who aren't archaeologists or who don't have archaeologists on the team. This isn't an unusual situation – the history of digital archaeology has plenty of examples where archaeology is used as a challenging testbed for tools and techniques but where relatively little substantial is contributed to the field by way of return, and this continues to happen in archaeology as well as in other areas of the humanities. For instance, Mateusz Fafinski (2020) has recently written of the problems when historical data is used by data scientists in an un-nuanced fashion. Coincidentally, the paper critiqued by Fafinski concerns the use of numerical techniques to analyse images (as emphasised in Manovich's paper), and is now flagged by its publisher as under further editorial review due to unspecified problems (Safra et al. 2020). It would be interesting to see whether, if Manovich looked at one of the number of applications of numerical techniques and machine learning by archaeologists – recently summarised by Davis (2020), for example – he would draw the same conclusions about archaeology. Perhaps he would: for instance, many of the applications of automated classification seem primarily aimed at reproducing practice rather than necessarily leading to new ideas as such.

Manovich proposes that digital numeric representations of cultural artefacts provide a better, more precise descriptive language than human natural languages, and that this is in part because numerical representations are closer to how our human senses perceive and encode external stimuli (p2). From a human perspective, the methods we use to capture sensorial, cognitive, and emotional aspects are limited, not least by the language(s) with which we communicate what we see, hear, feel, smell or taste. He argues that the use of numerical data to capture text, shapes, audio and images is beneficial because it is better than natural language at capturing smaller variations that may be beyond linguistic description, that natural languages are not capable of representing the nuances and differences that our senses can detect and appreciate:

> … if we can accurately and exhaustively "put into words" an aesthetic experience, it is likely that this experience is an inferior one. In contrast, using numerical features instead of

> linguistic categories allows us to *[sic]* much better aspects of an analog experience." (p6)

So he suggests that, for example, we may not be able to perceive a 1% difference in brightness or a miniscule difference between two objects, whereas a computer can.

That may be true, but it rather presupposes that such fine differences are useful or relevant in the first place. As archaeologists, we are certainly accustomed to digital recording devices such as survey instruments recording data to several decimal places, whereas once we're beyond a tenth or so of a centimetre it becomes an arguably spurious level of detail. The vagaries of the material record mean that recording features to within a literal hair's-breadth will frequently have little value or worse, mislead. We also have plenty of experience of taking data and recategorizing it, reducing it to a smaller range of options because the larger alternative is not considered useful, not simply because the tools we use or our own cognition are unable to handle it. In other words, the level of precision is only significant up to a point and may be entirely bogus given the circumstances surrounding data capture and processing. At one level then, Manovich's argument is a classic instance of more is better, and is commonly a feature of big data arguments, for instance. In contrast, it is becoming increasingly appreciated that, for example, a neural network with significantly more layers than others does not inevitably perform more reliably. To an extent, therefore, his argument risks falling into the trap that numeric necessarily equals objective, authoritative, accurate, precise (and he specifically refers to precision on numerous occasions). That said, while he argues that a data representation using numerical values can capture information with more precision than a linguistic description, he does observe that natural language includes the use of metaphor, intonation, rhythm etc. that describe perception in different ways, which is certainly resonant with, for example, the challenges of capturing tacit knowledge or other aspects of human perception that aren't well-suited to digitalisation. So although he doesn't explicitly say so, this perhaps implies that the use of digital and natural languages can best be seen as complementary.

His argument that a benefit of numerical representations is that they more closely relate to how our senses operate seems something of a red herring to me. Our sensory organs translate stimuli into electrical impulses within nerve cells and trigger chemical neural transmitters which transfer information between neighbouring neurons, but the numerical capture of this information is a translation of the electrical energy detected by sensors into numerical data to create traces, and so in that respect little different from geophysical probes sensing archaeological features below the ground, or satellite imagery operating beyond the visible range. So there is nothing natural or ineffable about numerical representations: they are simply an invented human language, and consequently as limited – albeit in different ways and to different degrees – as 'natural' human languages.

In the end, Manovich concludes that

> we can use digital computers to capture analog dimensions of artifacts and our aesthetic experiences as numbers. This *[sic]* numbers can use continuous scales that allows us to capture tiny differences between artifacts and details of artifacts with as much precision as we want. And we do can this for arbitrary large numbers of artistic and cultural artifacts. (p7)

and that this ability to describe phenomena more precisely than before is the first step in expanding our knowledge of a domain (p8). It's interesting that this observation coincides with increasing cases coming to the fore where, for example, large-scale numerical approaches can be shown to break down in the face of inbuilt but unrecognised bias, or fails to recognise the limitations of the numerical representations of otherwise analog data. This is something Manovich himself refers to in an earlier discussion of data (Manovich 2019, 64ff), specifically that data are constrained by their representation in a computational (numerical) environment in much the same ways as they are in an analog one. It's also interesting that other papers in this (presumed) collection also see datasets such as used in machine learning as problematic in various ways (e.g. Chávez Heras and Blanke 2020, Malevé 2020, Offert and Bell 2020) and some seek to address the question of the intellectual consequences of using computational vision methods (e.g. Emsley 2020). Those papers also show that, far from digital humanities being backward and primarily in receipt of computer vision and numerical methodologies from the computational and cognitive sciences, they can be a net contributor to those same fields.

References

Chávez Heras, D., and Blanke, T. (2020). On machine vision and photographic imagination. *AI & Society*. https://doi.org/10.1007/s00146-020-01091-y

Davis, D. S. (2020). Defining what we study: The contribution of machine automation in archaeological research. *Digital Applications in Archaeology and Cultural Heritage*, 18, e00152. https://doi.org/10.1016/j.daach.2020.e00152

Emsley, I. (2020). Causality, poetics, and grammatology: The role of computation in machine seeing. *AI & Society*. https://doi.org/10.1007/s00146-020-01061-4

Fafinski, M. (2020). Historical data. A portrait. *History in Translation* (blog), 29 Sep 2020, https://mfafinski.github.io/Historical_data/

Huggett, J. (2012). Core or Periphery? Digital Humanities from an Archaeological Perspective. *Historical Social Research/Historische Sozialforschung*, 37(3), 86–105. Final author version available at: http://www.cceh.uni-koeln.de/files/Huggett.pdf

Huggett, J. (2018). Who watches the digital? *Introspective Digital Archaeology* (blog), 26 Mar 2018. https://introspectivedigitalarchaeology.com/2018/03/26/who-watches-the-digital/

Malevé, N. (2020). On the data set's ruins. *AI & Society*. https://doi.org/10.1007/s00146-020-01093-w

Manovich, L. (2020). Computer vision, human senses, and language of art. *AI & Society*. https://doi.org/10.1007/s00146-020-01094-9

Manovich, L. (2019). Data. In H. Paul (Ed.), *Critical Terms in Futures Studies* (pp. 61–66). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-28987-4_10

Offert, F., and Bell, P. (2020). Perceptual bias and technical metapictures: Critical machine vision as

a humanities challenge. *AI & Society*. https://doi.org/10.1007/s00146-020-01058-z

Safra, L., Chevallier, C., Grèzes, J. and Baumard, N. (2020). Tracking historical changes in trustworthiness using machine learning analyses of facial cues in paintings. *Nature Communications,* 11, 4728 (2020). https://doi.org/10.1038/s41467-020-18566-7