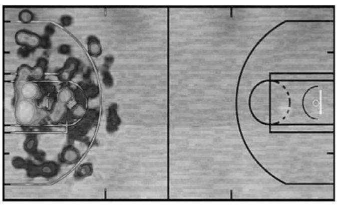
Introduction

In *The Imaginary Signifier* (1982), Christian Metz coined the term "Scopic Regime"



A visualisation of Kevin Durant's shots created by SportVU, a company whose player tracking technology originates in an IDF rocket programme.

»to describe how certain technologies, media and historical conditions might prescribe how we view the world.[1] Metz wished to problematise the notion of a single, universal vision. A decade later, Martin Jay further elaborated this notion by claiming that not only is vision subject to change from one age to another, but, during a given age, several "competing" regimes might exist in parallel. For instance, in parallel to "Cartesian perspectivalism" — which originated in the southern Renaissance and which Jay, among many others, identifies with modernity — there existed several other, alternative, modes of vision.[2] Thus the concept of a "Scopic Regime" enables us to understand vision not as neutral, or given, but rather as a contested and dynamic sphere, prone to influence from a variety of elements — philosophical, technological, political or economic. In this sense the shift from analogue photography to digital imaging technologies, and more generally to what I shall term computational vision, is a change from one regime to another. This shift, its origins, characteristics and implications, will be the subject of the following essay.

During the past five decades we've witnessed the rise of a new "Scopic Regime." The logic of this is ubiquitous: it defines the way the online world is designed and structured, how the buildings we live in are built and what insurance we are allowed. In some cases, primarily when put to military and governmental use, this logic has a decisive influence upon ethical and political questions. Just as during the Renaissance a certain technology of linear perspective redefined the way we view the world, influencing painting but also warfare and governance, so today different technologies establish a new mode of "vision," playing a determining role within both the cultural and the political realms.

As is well known, photography was crucial in redefining the contemporary modes of vision and representation. Photography's automated nature made it seem concrete and grounded, it embodied a unique combination between elements considered to be subjective, for instance, an artist's hand, a selection or "framing" process, while simultaneously deploying those elements which contribute to a sense of objectivity, a mechanical process and a chemical reaction.[3] Thus, thanks to the development of an automatic, seemingly objective, apparatus, photography solidified the fundamental, widespread belief in the image's legibility, in its mimetic credibility or "adequacy." As Lorrain Daston and Peter Galison write in their formidable study of objectivity:

> What the photograph … offered was a path to truthful depiction of a different sort, one that led not by precision but by automation — by exclusion of the scientist's will from the field of discourse.[4]

Roland Barthes, the major proponent of this conceptualisation of analogue photography, wrote extensively on this medium's unique quality as testimony. According to Barthes, the photographic image will necessarily refer to "what has been"; the photo is "real" simply because unlike a painting, which might fabricate representation from pure imagination, the photo demands a physical object so it might re-present it. Barthes went so far as to claim that analogue photography is akin to a “message without a code,” that is, as “raw,” or purely legible information.[5]

The digital image, on the other hand, is all code. It is often described as a numeric representation (normally binary) of a two-dimensional image, but behind such a definition lies a much more ambivalent reality. In fact, the raw data collected by the electro-optic sensor isn’t a “representation” of an image. On the contrary, the image is just one possible “output” converted from the raw data. An electro-optic-sensor (such as a standard CCD) consists of a thin silicon wafer divided into millions of light sensitive squares (photosites). Each photosite corresponds to an individual pixel in the final image. The sensor turns light (photons) into electrons (charge). Once exposed, the electrons at each photosite are passed to a charge-sensing node, amplified, and relayed to read out electronics to be digitised and sent to the computer. The output voltage is consequently converted to a digital signal — rendered into binary code. This code is then, finally, converted into an image.

To the same extent one could have programmed a device that would render optical input (light) into sonic or textual output via code. In other words if we must decide upon the primacy of a medium, the principal mode which lies at the heart of the electro-optic logic, it would be code, or numeric data, which would gain prominence. The digital image is a two-dimensional representation of numeric data, rather than the other way around. In this sense the digital image is symptomatic of the current “Scopic Regime” for within this entire regime the image is subjected to the logic of code and computation.

The prehistory of the digital image,or the story of project 5980

In December 1940 the Rockefeller Foundation awarded Norbert Wiener a two thousand dollar grant to pursue a research project. “Project 5980,” dubbed the “Debomber,” was an attempt to develop an efficient aerial defence system, which had become a crucial concern following the German airborne attacks on Britain. Wiener, together with Julian Bigelow, intended:

> to place the analysis of the problem of prediction upon a purely statistical basis … to model the behaviour of the airplane within the frame of reference belonging to the airplane, rather than referring it to that of the observer on the ground.[6]

The task of seeing and targeting the bomber was transferred to an automated algorithmic process, one which was based on statistical patterns extracted from a database, a history. The “Debomber” was never completed, it was premature: digital memory of sufficient capacity was not yet invented, and computers too were barely more than an idea. But its historical implications have been extensive both in the field of cybernetics, a science founded by Wiener, and in the development of the first computers, at Princeton University, where Bigelow joined the Electronic Computer Project. Vision would never be the same.

Several key concepts first developed by Wiener and Bigelow laid the foundation for "computer vision," artificial intelligence and more generally for the field of cybernetics. Above all else, Wiener and Bigelow's identification between noise, entropy and information and the negation of entropy, or negentropy, would be key to developing a method for separating or filtering signal from noise by way of a statistical analysis of a large database.[7] Furthermore, in parallel to his futile work on the anti-aircraft gunner, Wiener *did* succeed in developing a noise-reducing filter. This filter, aptly named the "Wiener filter," is used to this day to reduce sonic or visual noise. Ever since, numerous filters and formats have been invented, like the codecs and mediums of compression, such as JPEG or the MP3, which are all based upon the same principle: the statistical separation of signal from noise.

Here we first encounter the principle, which is to be so essential to the current "Scopic Regime": it is a regime that transcends the visual realm — in the beginning, before the image, was code. In other words, what distinguishes our current dominant mode of vision is the way by which the image is subjected to an exterior logic; a logic which one could translate into a myriad of modes of expression.

As Peter Galison has pointed out, the development of the "Debomber" introduced an "ontology of the enemy." Wiener argued that "human behavior could be mathematically modeled and predicted, particularly under stress; thereby articulating a new belief that both machines and humans could speak the same language of mathematics."[8] According to Galison, the servomechanical enemy would become for cybernetic vision "the prototype for human physiology and, ultimately, for all of human nature."[9] Thus, Wiener would go so far as to claim that "as objects of scientific inquiry, humans do not differ from machines."[10] But this moment also engendered a new understanding of vision itself, becoming, as Orit Halpern suggests: "a material artefact — an algorithm — capable of actions and decisions such as identifying a 'prey' or an 'enemy.'"[11] programme.

Un Coup de Dés Jamais N'Abolira Le Hasard (A Throw of the Dice will Never Abolish Chance)

During interviews I conducted together with Ryan Jeffery with a selection of entrepreneurs, data analysts, engineers and ad agency employees, a recurring motif we encountered was the use of the so-called "Monte Carlo" method.[12] "Monte Carlo" is the name given to "a broad class of computational algorithms that rely on repeated random sampling to obtain numerical results."[13] Stanislaw Ulam, a Jewish-Polish mathematician who immigrated to the United States during the Second World War and played an important role in the development of the "Manhattan Project," first thought of the system while recovering from a long period of illness hoping to improve his chances of winning in the game of Solitaire. Ulam understood that instead of relying upon theoretical combinatorial calculations, one could simply churn through a large set of random simulations one after the other and observe the results (wins, in the case of Solitaire). The larger the set of simulations, the more accurate the inducible probabilistic distribution becomes. Of course crunching numbers in magnitudes of this order was impossible without the invention of the computer. Ulam was one of the first people to realise the dramatic implications of exponentially accelerating computational power. But the "Monte Carlo" method foregrounds another important aspect of "computer vision," namely, its indifference towards the distinction between stimulatory and sensory data, that is between data that is *collected*, and simulations that are *generated*.

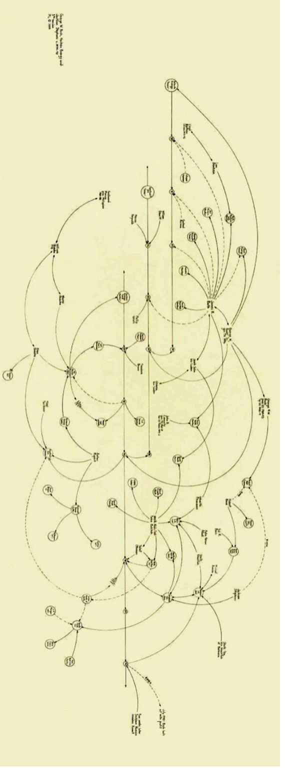
Benoit Mandelbrot, an important pioneer in the use of computer generated imagery (CGI) has written about his contribution to the development of mathematical thinking through image analysis:

> it was near-universally believed among pure mathematicians around 1980 that a picture can lead only to another, and never to fresh mathematical thinking. … A striking innovation that helped thoroughly destroy this belief resided in my work's heavy reliance on detailed pictures, in contrast to schematic diagrams. Incidentally, a picture is like a reading of a scientific instrument … More precisely, my discoveries of new mathematical conjectures relied greatly on the quality of visual analysis and little on the quality of the pictures … [14]

Today these so called "fractal landscapes" — virtual surfaces generated using random probabilistic algorithms designed to produce "fractal" behaviour which mimics the appearance of natural terrain are used in military simulations and forensic investigations, whilst also populating a growing part of our visual everyday; occupying the worlds of film, advertising and stock imagery. Moreover, CGI, whether generated using a probabilistic model, animated laboriously, or rendered on top of an existing image, plays a major role in the current shift towards "post-production." As Hito Steyerl describes:

> In recent years, postproduction has begun to take over production wholesale. In newer mainstream productions, especially in 3-D or animation, postproduction is more or less equivalent to the production of the film itself. Compositing, animation, and modeling now belong to postproduction. Fewer and fewer components actually need to be shot, because they are partially or wholly created in postproduction. Paradoxically, production increasingly starts to take place within postproduction. Production transforms into an aftereffect.[15]

In other words, computational vision is, to a great extent, a vision that constantly spills from virtual to the real; both are perpetually re-touched or re-modeled to fit each other, creating a constant feedback loop. For instance a fashion model's makeup is picked to match the High Definition (HD) resolution they'll be filmed in. Similarly, insurgents devise ways to hide as visual noise by reducing their thermal signature, whereas hotels name themselves "hotel in" to improve Google search results. All these examples demonstrate how computer-generated models, whether visual or merely theoretical, inform and augment reality, and "reality" is in turn understood in terms of these models.



Mark Lombardi, George W. Bush, Harken Energy, Jackson Stephens, 1979-1990.

Several weeks after 9/11 the Whitney museum received a phone call from the FBI asking whether they would be able to visit works by Mark Lombardi then on display. The artist, who committed suicide in 2000, created for years intricate diagrams mapping power relations between public and private interests.

Things you can do with metadata

In the days following the publication of the National Security Agency (NSA) documents leaked by Edward Snowden, many in the United States senior administration, including President Obama himself, claimed to their defence that though the NSA gathered warrantless data in bulk it never actually eavesdropped, or listened, to any of its citizen's conversations.[16] But metadata, or data about data might be a computers' preferred type of data. Content, whether it is a website's content, a phone conversation or an image, is by contrast much more difficult for a computer to "digest." Any form of eavesdropping on phone conversations would necessitate at least partial human analysis and would therefore be nearly impossible to conduct on such a large scale. But within the current "Scopic Regime" content is rendered nearly superfluous; topological analysis of metadata is more than enough. In the case of phone calls, the people you called, the length of the conversation, its geo-loco data and the decibels registered might all be considered metadata.[17] These variables have sufficient incriminating potential; even a simple printout of one's call history can lead to a range of conclusions. By analysing metadata one can map entire communities, either with the aim of understanding the chain of command of a guerrilla organisation, or in an attempt to classify which users, or customers, are more lucrative or influential, and thus to produce personalised advertising campaigns. This is not hypothetical, as General Michael Hayden, former director of the NSA, has asserted: the United States military has killed people based on an analysis of meta data, rather than content. Suspects have been targeted according to Activity Based Intelligence (ABI) by analysing a range of activities within a given area — Subscriber Identity Module (SIM) card operations, transactions and movements — rather than by listening to the content of actual conversations.[18]

Shibboleth

The most common test used today to distinguish between a bot — a software application that runs automated tasks over the Internet — and a human user is the Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA). Unlike the original Turing test — invented by Alan Turing with the aim of discerning whether a machine has reached human-like levels of artificial intelligence — a CAPTCHA, rather ironically, relies upon a machine to judge whether the user is a human. The way this is done is rather simple, and one which we encounter on a daily basis: the computer generates an image which contains a set of blurry, smudgy characters (usually a combination of letters and numbers) and we, as users, are set the task of recognising or deciphering the vague message. For humans, the inverted Turing test, as inconvenient as it may be, is still quite a simple task. For your average "bot" on the other hand, solving such a riddle would amount to a great intellectual feat. Human vision and recognition is here used as a "shibboleth" of sorts, a means of differentiation between man and machine. In an ironic turn of events, the machine thus becomes the measure of all things. The images that appear in CAPTCHA tests are produced in such a way as to sabotage the possibility of automated decryption, or Optical Character Recognition (OCR). To do so one need only blur or colour the background, smudge the text so that it isn't horizontal, or distort the characters and their spacing randomly. Apparently even though computers are capable of executing billions of calculations in a matter of seconds, they're quite limited when it comes to "seeing."

And yet, this gap,between visual understanding and data analysis is slowly closing, not because supercomputers have suddenly reached human levels of intelligence and are able to recognise images and semantically categorise them, but rather because during the past five decades we have consistently made an effort to approach the world of computers, to make our world more comprehensible to them. Computers are essentially blind to images, they "see" only thanks to external devices, like data about data. Consequently a computer is able to visualise information in the same way that a blind person navigates the city with a white cane. In other words we might say that the computer, is also able to see *blindly*.

Social networks enable computers to translate and map human intersubjective relationships, to classify facial characteristics and quantify our wants and opinions. So called "smart devices" render our everyday life into troves of data: medical, social and geographic. During the next decades we can expect more and more sensors to *mine* sources of data from archives which are currently growing exponentially. This is the logic behind the vision of an "internet of things," which advertising companies across the globe are fiercely promoting. The more parts in our life we make intelligible and quantifiable to computational machines, the more we are prone to subject ourselves to a logic and epistemology designed to their benefit.

It's essential we keep in mind that computer vision too is just a perspective, one which stems out of a concrete historical context, with its own set of biases, preconceptions and blind spots. During this essay several of these were identified — first was the preference towards quantity over quality, which is based upon the probabilistic nature of data analysis. This, as we have seen, is manifested in many different fields, from the underlying logic of noise-reducing algorithms, to the architecture of surveillance programs. The primacy of code over all other modes of expression, and consequently, the negation indexicality, is another such bias. The rise of post-production, computer generated imagery and physical modeling, all have economic as well as aesthetic and epistemological implications. Finally, the pervasiveness of metadata informs the way knowledge is mapped and extracted and people surveilled. As these different examples all demonstrate, what is at stake is essentially a new mode of governance and control, as well as a novel way of seeing, and consequently, representing.

Endnotes

1. Bruno Latour, "Step Toward the Writing of a Compositionist Manifesto," in New Literary History, vol. 41 (2010); 471–90.
2. Martin Jay, "The Scopic Regimes of Modernity," Vision and Visuality, ed. Hal Foster (Seattle: Bay Press, 1988). See also, Erwin Panofsky, Perspective as Symbolic Form (New York: Zone Books, 1997).
3. Lorraine Daston and Peter Galison, "The Image of Objectivity", in Representations (1992); 81–128.
4. Daston and Galison, "The Image of Objectivity," 117.
5. Roland Barthes, Camera Lucida (London: Vintage, 2000); 76–77.
6. George Dyson, Turing's Cathedral: The Origins of the Digital Universe (New York: Pantheon Books, 2012). See also, Peter Galison "The Ontology of the Enemy: Norbert Wiener and the Cybernetic Vision," Critical Inquiry 21, Autumn (Chicago: University of Chicago Press, 1994).
7. Norbert Wiener, The Human Use of Human Beings: Cybernetics and Society (New York: Doubleday Anchor, 1954); 7–12. See also, David Arthur Bell, Information Theory and its Engineering Applications (London: Isaac Pitman & Sons, 1956); 8–12.
8. Orit Halpern, Beautiful Data (Durham: Duke University Press, forthcoming).
9. Galison, "The Ontology of the Enemy," 233.
10. Arturo Rosenblueth and Norbert Wiener, "Purposeful and Non-Purposeful Behavior," Philosophy of Science (1950); 326.
11. Halpern, Beautiful Data, forthcoming.
12. The interviews were conducted as part of our work on an ongoing artistic research project and film titled All that is Solid Melts into Data.
13. See: http://en.wikipedia.org/wiki/Monte_Carlo_method (accessed 1 November, 2014).
14. Benoit B. Mandelbrot Fractals and Chaos The Mandelbrot Set and Beyond (New York: Springer-Verlag, 2004).
15. One of the clearest examples of this phenomena from recent years was Avatar which used "35,000 processor cores with 104 terabytes of RAM and three petabytes of network area storage," making it one of the most powerful supercomputers in the world. See, Hito Steyerl, The Wretched of the Screen (Berlin: Sternberg Press, 2012); 182.
16. See: http://www.nybooks.com/blogs/nyrblog/2014/may/10/we-kill-people-based-metadata/?insrc=wbll (accessed 1 November, 2014).
17. The "smarter" the phone, the more diverse a set of metadata it can provide about its users. The latest iPhone 5s is home to the so-called M7 motion sensor, and the new HTC has two backside camera lenses that are capable of collecting three-dimensional spatial coordinates.
18. See: https://firstlook.org/theintercept/2014/02/10/the-nsas-secret-role/ (accessed 1 November, 2014).