Visual Studies, Vol. 26, No. 1, March 2011

Routledge
Taylor & Francis Group

The algorithmic turn: photosynth, augmented reality and the changing implications of the image

WILLIAM URICCHIO

The digital turn, and with it increased use of location-aware technologies, has yielded innovative image applications and posed new questions about the status and value of the image. These applications rely on algorithmically defined relations between the viewing subject and the world viewed, offering robust alternatives to the visual economies of the past. If we take seriously Heidegger's insights regarding the Welt-bild *as a metaphor for the modern era, the algorithmic reconfiguration of subject-object relations in this emerging visual regime potentially offers insights through which we can reflect upon the current era – and a metaphoric alternative. This article uses two entry points to explore this possible reconfiguration and, with it, the question of value. Downloadable applications such as Photosynth aggregate location-tagged photographs into a near-seamless whole, and offer a way to consider such issues as collaborative authorship of the image, unstable points of view and the repositioning of subject-object relationships – all elements that fundamentally challenge western representational norms dominant in the modern era. In this new regime, the spatial referents of greatest value are points of uniqueness sought out and built upon by the program's algorithms – and not those perceived by the viewer. The viewer is in turn free to explore an extensive and dynamic image space unconstrained by (and, indeed, without access to) an authorised or 'correct' viewing position. A second case, built upon certain augmented reality applications, works by 'recognising' particular spaces and, through the use of computationally enhanced viewing screens, superimposing new images over real space. In this case, a system of virtual spatial annotation depends upon the 'correct' positioning of the viewer (and portable computing device) in the world. The two cases stand in a roughly reciprocal relationship, turning on different notions of algorithmic intermediation and subject-object relations and dynamics for the generation of meaning and value.*

Rien n'est plus ennuyeux qu'un paysage anonyme.
(Prosper Merimée)[1]

TERMS OF ENGAGEMENT

I have long been haunted by Jonathan Culler's *Semiotics of Tourism*, with its careful parsing of travel vs. tourism, its astute characterisation of behaviours such as our reflexive denigration of those more tourist-like than we, and its charting of the elusive search for authenticity (Culler 1990). Culler traces the process by which cultural attractions are marked as signs produced by an international system of signification, and responsive to and inscribed within an economic order. Yet he avoids a blanket indictment that would read this process as the mere flattening of the authentic into caricature, complicit with the demands of multinational capitalism (a charge he locates with the 'sentimental nostalgia for the organic'). His essay instead calls for an exploration of the persistent and ubiquitous semiotic mechanisms central to any social order, a task that this article takes up. It is my contention that we can see evidence of change in the dominant 'semiotic mechanisms', evidence that perhaps speaks to a deeper transformation of the social order. The goal of this article is to explore one aspect of that change: the algorithmic construction of the image. While I will not take up the grander challenge of discussing the transformation of the social order, I will address what I take to be one of its key symptoms – namely, cracks in the façade of the subject-object relationship characteristic of the modern era. Let me make clear at the outset that cracks do not a transformation make, but they give us an early warning and an important place to look for further signs of change.

My argument in a nutshell is that over the past decade or so we have had increased access to new ways of representing and seeing the world, ways dependent on algorithmic interventions between the viewing subject and the object viewed. This intervention has many manifestations, from the changed model of authorship and expertise that Wikipedia represents over and against the Enlightenment paradigm represented by Diderot's

William Uricchio is Professor and Director of Comparative Media Studies at the Massachusetts Institute of Technology (MIT) and Professor of Comparative Media History at Utrecht University in the Netherlands. His research focuses on the emergence of media technologies and practices, and recent books include *We Europeans? Media, Representations, Identity* (Chicago, 2008) and *Media Cultures* (Heidelberg, 2006).

ISSN 1472-586X printed/ISSN 1472-5878 online/11/010025-11

encyclopaedia, to the dynamic and location-aware cartographic systems that we can find on our iPhones and TomToms over and against the fixed cylindrical projection grid of Gerardus Mercator's sixteenth-century maps. In the case of the image, *grosso modo*, the long regime of three-point perspective and its reification of an underlying understanding of subject-object relations constitute the representational order that is exhibiting fractures. And those fractures – the topic of this article – appear in the form of algorithmic visualisation systems such as Microsoft's Photosynth and image recognition-based augmented reality applications.

In order better to locate the terrain and implications of this argument, we might briefly look at the larger regime into which the algorithmic represents an incursion. Its underlying premise can be found in Culler's assertion of the centrality of semiotic mechanics to the social order, a view that echoes a position developed by Heidegger in his discussion of the *Welt-bild*, where the world as picture 'does not mean a picture of the world but the world conceived and grasped as picture' (Heidegger 1977 [1938], 129). Heidegger goes on to specify that the world picture 'does not change from an earlier medieval one into a modern one, but rather the fact that the world becomes picture at all is what distinguishes the essence of the modern age' (Heidegger 1977 [1938], 130). For Heidegger, the moment at which the world becomes picture is the same moment that the human emerges as the subject in a characteristically modern subject-object relationship. It marks the birth of the modern. The modern subject-object relationship takes many forms, but can most clearly be seen in our notion of science. Heidegger says:

> We first arrive at science as research when and only when truth has been transformed into the certainty of representation. What it is to be is for the first time defined as the objectiveness of representing, and truth is first defined as the certainty of representing, in the metaphysics of Descartes. (Heidegger 1977 [1938], 127)

Heidegger here asserts the fundamental linkage of mode of representation and epistemic (and social) order. The Cartesian metaphysics to which he refers offer a widely accepted definition for the modern project that remains dominant today, even if under siege. In this order, precision and exactitude as embodied in classical mathematical physical science are essential ingredients. Heidegger states: 'Here all events, if they are to enter at all into representation as events of nature, must be defined beforehand as spatiotemporal magnitudes of motion. Such defining is accomplished through measuring, with the help of number and calculation' (Heidegger 1977 [1938], 119).

Culler's general proposition that a social order has central 'persistent and ubiquitous semiotic mechanisms' is embodied by Heidegger's essay, which argues that the modern social order can be defined through a representational system characterised by precisely defined subject-object relations (the world as picture), a metaphysics of exactitude and an underlying spatiotemporal grid. Descartes himself emblematises this order, thanks to his development of analytical geometry and the Cartesian coordinate system that followed (enabling geometric shapes to be described in algebraic form). He stands for a notion of the modern typified by classical cartography, three-point perspective and the amplified point of view manifest in authorship and the cult of expertise (Diderot). The broad space between Descartes and Heidegger, despite countless historical undulations and discoveries, turns on a consistent logic, on something we might term the *algorismic*, that is, a notion of mathematics as a language of precision, of calculability, of predictability. And it remains deeply familiar to us, pervading our lives whether through our financial and insurance systems, our notions of science or the construction of our default technologies of visual representation. While three-point perspective and clearly defined subject-object relations are not inherently algorismic, their operations are consistent with it and depend upon its logics.

The algorismic that Descartes did so much to promote as a mathematical system, an epistemic regime and even a social order, and that Heidegger and others argued as a defining characteristic of the modern era shares etymological roots with the term *algorithmic*, but the two have very different definitions and histories. In contrast to the precision, calculability and specificity of the algorismic, an algorithm refers to a process, a program with clearly defined limits, a finite instruction sequence. One is a calculable sum, whose value lies in the correctness of its result; the other a (finite) process, a formula capable of accommodating different values and yielding different results. Each term harkens back linguistically to the great ninth-century Persian mathematician Al-Khwarizmi, whose work appeared in the West by the thirteenth century, and as mathematical models, each can be traced back to the ancient Greeks. Yet their histories differ considerably. The algorismic predominated, enjoying a particular boost in use in the sixteenth and seventeenth centuries. By contrast, although Euclid demonstrated their operation *c.*300 BC, algorithms remained a marginal form of calculation

until the second half of the twentieth century, when the introduction of computers enabled the efficient processing of programs.[2] The mathematical advances of the twentieth century owe less to the arithmetic of the algorismic than to the accelerated development of algorithmic instruction sequences, a process that continues to grow exponentially thanks to the ongoing validity of Moore's Law. The chips in our phones, cameras and bank cards, the supermarket 'loyalty tags' scanned as we check out, and the cookies and behaviour tracking that companies such as Amazon and Google use to predict and assess our interests all attest to the pervasiveness of an infrastructure designed to generate data whose meanings can only be deciphered algorithmically. And it is this algorithmic layer that stands between the calculating subject and the object calculated, and that refracts the subject-centred world charted by Descartes, that merits closer inspection. Its implications, if we take thinkers like Heidegger and Culler seriously, can be profound.

MARKING VALUE

One of the semiotic strategies that Culler discusses is the 'marker'. Drawing on Dean MacCannell's *The Tourist,* Culler defines a marker as any kind of information or representation

> that constitutes a sight as a sight: by giving information about it, representing it, making it recognisable. Some are 'on-site' markers, such as plaques telling that 'George Washington slept here' or that this vial of dust comes from the moon. Some are mobile markers, such as pamphlets and brochures designed to draw people to the site, give information at the site, and serve as souvenirs or representations of the site . . .The proliferation of markers frames something as a sight for tourists. The existence of reproductions is what makes something an original, authentic, the real thing – the original of which the souvenirs, postcards, statues etc. are reproductions – and by surrounding ourselves with reproductions we represent to ourselves, as MacCannell astutely suggests, the possibility of authentic experiences in other times and in other places. (Culler 1990, 5)

Culler's essay and MacCannell's insights go a long way towards explaining how value is created by collapsing temporalities, overlaying a site or object in the present with a reference to its provenance or past, or even evoking the authentic through the act of reproduction. A crumbling bit of rock in one's hand transforms when we learn that it is a piece of one of Giza's pyramids or a fossilised piece of Diplodocus bone. The process of generating value through the assignment of meaning is a familiar if complicated one, one with its own spatialities and temporalities: the artefacts within the retail outlet, the antique shop and the second-hand store are distinguished by these dynamics rather than by any intrinsic worth. Mere purchase and ownership, no matter how brief, can recast an object from retail to second hand; and today's second hand in one market could possibly be revalued as a valuable antique in another.

The image, however, has an even more multivalent status. As an artefact, say in the form of a photograph, the image is collectable, subject to the same ebbs and flows in valuation as other things, particularly those privileging symbolic and cultural attributes. Tastes, both professional (art markets) and vernacular (eBay), help to set these values. And, like souvenirs, the image has the ability to link us back to the people, places and events it represents. Its very presence attests to 'the possibility of authentic experiences in other times and in other places'.

But as Culler suggests, the image can also be complicit with the valuation of its referent, the place or thing represented. In this case, the relationship between the image and that which it represents can be discursive, amounting to little more than a claim; or it can work through resemblance, something that we simply recognise or take as recognising. The space between the two is slippery and sometimes fraught. Just as the claim that a piece of stone is from Cheops' pyramid changes the stone's status, so too the (discursive) claim that a photograph is of a particular place constructs its (representational) meaning. Claims are not always accurate; film and television directors routinely move between these two modes of assigning meaning, establishing a location through resemblance – iconic shots of a particular place such as Manhattan – only to cut to shots taken in less expensive locations such as Toronto or Pittsburgh as if the same location. This technique of using discursive strategies to evoke representational assumptions was explored by early Soviet filmmaker Lev Kuleshov, who argued that at least in the sequential image flow of film, meaning is constructed by the relationship of one shot to the next rather than fully residing within any one image.

The assignment – and destabilisation – of meaning can occur within the realm of the image, as Kuleshov and his followers argued; and theorists such as Culler have extended the argument, turning it on a different sort of framing, but nevertheless undercutting the assumed stability of imagistic representation and meaning by looking outside it, at the broader signifying context. And

more recently, theorists such as Tim Ingold and Sarah Pink have sharpened the critique by addressing factors such as multisensory settings, movement and place, reminding us that there is more to the image than meets the eye. The idea of the image as 'becoming' through movement – movement of the world, movement of the (photographing) subject and movement even of the subsequent viewer – is taken up and developed in Sarah Pink's contribution to this issue, and will resonate with the algorithmic applications to which we will shortly turn.

The paradox of the image as a site of meaning and value turns on its reciprocal relationship to the object or place photographed. As discussed, sometimes the location can give the image its significance, and sometimes, as Culler notes, the image can give the location its meaning. Functioning somewhat differently than a replica – say, a miniature Eiffel Tower, which can be seen from any angle and still maintain the legibility of its link to the real thing – the image carries with it a more constrained sense of representation. An extreme close-up of one of the tower's beams would not necessarily convey the meaning 'Eiffel Tower'; it could only do this by relying exclusively on discursive claims. Moreover, even at the representational level, where what we would see is a beam or metal surface, assuming that either was recognisable, we would still have to rely on a discursive claim to distinguish *this* beam or surface from any other metal beams or similarly surfaced objects. And such an image, no matter how much one proclaimed its indexical status, would have a difficult time conveying value back to the object. At least when its representational capacities are being called upon (as distinct from its cultural value as art), the value and meaning of a photographic image are constrained by the viewer's ability to recognise and make visual correlations between the two-dimensional image and the thing represented. And in this particular game, convention and point of view matter, strategies that have been finely honed and codified in the course of the modern era, and that are part and parcel of a particular way of seeing and being in the world. Photosynth and certain augmented reality applications offer some intriguing variations on this process, and in so doing, raise some fundamental questions about our visual regime.

MOVEMENT WITHIN THE IMAGE: PHOTOSYNTH

Photosynth (www.photosynth.net), a software application developed in part by Noah Snavley (PhotoTourism) at the University of Washington and acquired and further developed by Microsoft, essentially analyses digital images for sites of uniqueness, generates a three-dimensional point cloud of the represented space and reassembles the images into a near-seamless composite. Viewers can then explore the assembled three-dimensional photographic space in almost any direction (including in depth), with Seadragon technology enabling instant high-definition access to whatever portion of the whole is on the screen. This both precludes any feeling of zooming into an existing picture (a situation in which granularity increases and the authoritative status of the original composition is reinforced) and facilitates the transition from one image-space to another. The overlays and transitions from one image to the next are discernible but not obtrusive, and as a viewing option, the contours of the original photographs can be marked with white lines ('quads') for navigational purposes (viewers can also see all of the individual photographs laid out side by side if they would like, as well as different angles of the point cloud). Users can build their own 'synths' by surfing the web or particular sites such as Flickr, and tagging images with a particular location ('Piazza San Marco in Venice') or event ('Obama's inauguration'). Synths can be constructed with between dozens of, and even several hundred, photographs; for the moment, Microsoft recommends staying below a 600-photograph limit since the strain on the system's processing power visibly interferes with the experience. Available to the public since 2008, Photosynth's capacities seem to be developing steadily, with major improvements in the synthing of different images introduced in 2010.

In the Photosynth website's words (http://photosynth.net/about.aspx):

> Photosynth takes your photos, mashes them together and recreates a 3D scene out of them that anyone can view and move around in. Different than static photos and video, Photosynth allows you to explore details of places, objects, and events unlike any other media. You can't stop video, move around and zoom in to check out the smallest details, but with Photosynth you can. And you can't look at a photo gallery and immediately see the spatial relation between the photos, but with Photosynth you can.

Very much a work in progress, Photosynth's status as a medium also seems to be in formation, and this description suggests its sites of slippage relative to existing media forms. It manages to wed place and movement through its 'point cloud' – a three-dimensional configuration of sites of visual uniqueness upon which the images are arranged and through which movement is facilitated – a logic with implications that might best be approached by

considering Photosynth's limit cases. The system is predicated upon the visual contours of place. Like a jigsaw puzzle, it works from known or given points, and extends outwards, building an ever-changing whole by aligning the points of different images and suturing them together. Unlike a puzzle, however, Photosynth addresses a third dimension, and therefore must confront the issue of significant overlap among its images. Since the synths are composited from digital photographs, some of their challenges regard such formative elements as image contrast ratio, composition (angle, proximity, focus) and lens characteristics, all of which are capable of rendering points of uniqueness in ways that the system cannot recognise. If an extreme close-up of several mosaic pieces from San Marco's basilica is tagged and fed into a synth, it may or may not be accepted, depending upon the certainty of its location as extrapolated through the point cloud. If its pattern is distinctive, allowing it to fit uniquely within other shots of the same area, then it will probably be accepted, allowing the synth's user to traverse the plaza even to the point of a close-up of the mosaics. Shots that, because of extreme angles or distorting lenses or lighting configurations, are unrecognised are simply not successfully synthed. These tolerances express the operations of the underlying algorithm, reflect the current state of the navigational interface and attest to particular conventions regarding imagistic correctness – and they seem quite dynamic, since the capacities of the system grow ever more refined.

As might be expected, Photosynth does particularly well with architectural environments: stability is a plus and facilitates the construction of point clouds. However, it also yields impressive results in nature. Rather more dynamic than the relative stasis of architecture, nature raises the larger issue of duration. What happens to the points of uniqueness in a forest when blowing wind, or plant growth, or seasonal changes alter the configuration of leaves and branches? One can find quite impressive synths of forests on Photosynth's site, but most share tranquil conditions and a small temporal window, and presumably make good use of this relative stasis. But we might put the question to architectural spaces as well. Mixing images taken during the day and at night will result in a high percentage of rejected images, but what about shots taken before and after, say, a war or an earthquake? New chips on a surface or distorted corners of buildings, windows or bricks complicate the construction of a point cloud, exposing the system's durational tolerances and rendering its embrace of historicity ambivalent at best.

At the other end of the spectrum, what about the 'transient, the fleeting, the contingent' that Baudelaire ascribes to the modern? The most widely seen example of a Photosynth application that addresses this problem is CNN's appropriately titled 'The Moment', a synth of 628 user-submitted photographs of the moment that President Obama took his oath of office. Although many of the photographs are of the inauguration itself, large portions of the synthed image include crowds of onlookers, dignitaries and even musicians. 'The Moment' actually reveals quite a bit of temporal slippage – changes in bodily position, different configurations of the flags in the wind and so on. The constants, the point clouds, seem to be grounded in architectural detail and the configuration of the podium more than anything else, and the more one examines close shots of the crowd, the more one is reminded of early-nineteenth-century photography, where long time exposures often resulted in ghost figures. These fugitive images emerged from the gaze locked within three-point perspective and subject to the limits of early photochemical emulsions. With Photosynth, a different mechanism is at play: the tolerances of algorithmic reassembly; but the ephemeral, nevertheless, seem to pose a very real challenge to the system.

Photosynth permits one to 'wander' through the depicted space, moving deeper into (or farther back from) the frame and exploring off-screen space, whether to the left or right, or the top or bottom of the initial frame. Indeed, the frame (of the computer screen) is rendered arbitrary, like the lens of a magnifying glass in relationship to a newspaper or the viewfinder of a telescope to the night sky. As the Photosynth advertisement notes, the sense of mobility and user-generated interactivity within the image-space challenges our inherited definitional assumptions of photography. But it does something more, restoring an experience of movement to the image and the depiction process, vividly illustrating the arbitrary constraints of any one image within a larger reality. Much like Sarah Pink's discussion of Google Street Views in this issue (also an algorithmically enabled application), Photosynth assembles points and enables movement within and among them, offering us different ways of displaying and engaging with visual information. But beyond Photosynth's ability to do this dynamically and in three dimensions, the application distinguishes itself through its subverting of any particular point of view, in the process disrupting the modern configuration of the subject-object relationship, as well as its reliance on collaborative authorship. Let us take these factors in turn.

Consider Canaletto's painting *Piazza San Marco with the Basilica* (1730). An impressive display of three-point perspective, the basilica is at the centre of the frame, and the lines provided by the Procuratie Vecchio and

FIGURE 1. *Piazza San Marco with the Basilica*, by Canaletto, 1730. Fogg Art Museum, Cambridge, MA.

Procuratie Nuovo, as well as the gutters in the plaza, literally map out Canaletto's sight lines. The representation is completely stable, the geometries a testament to the algorismic regime of fixity and precision. There is one and only one position from which the image's geometries work, and the viewer and artist share that point of view. Photosynth, by contrast, draws on some 600 photographs taken from different positions, angles and distances, and with an array of lenses. The user can navigate within and among these, in the process discovering many well-composed images and new ways of seeing the plaza and its buildings. But there is no correct or authorised viewing position, no 'master shot' within which everything else is a recomposition. Instead, there is simply a three-dimensional space made up of many textures and granularities, and the means to move within it. Whereas in Canaletto's case, the subject-object relationship is geometrically fixed, in Photosynth's dynamic assemblage the relationship is unstable, subject to the whims of the user and the capacities of the algorithmic. We can move within and across particular points of view, confront radical shifts in angle and inhabit a visual cacophony bound together by a point cloud that enables our viewing position. While it is certainly the case that the individual images that make up the synth each largely embody the conventions of three-point perspective – something inscribed in the lenses that most cameras come equipped with, and something basic to our cultural notions of visual legibility – I wish to call attention to the fabric of radical disjunctures in viewing position and the fundamentally unstable nature of the composite. In Photosynth, many points of view are called upon, and inscribed into an algorithmically dynamic mix. Like Wikipedia entries, the whole only works because of individual contributions – acts of authorship, bits of information, points of view – that are algorithmically reworked into an ever-changing mix. Rather than relying on the expertise and reputation of one known individual, the reader must take a more active role in making sense of the ensuing composite of anonymous voices, in assessing it, in moving across its links to pursue additional information. The work of individuals – whether in the Wikipedia entry or in the photographic contribution to a particular synth – is theoretically traceable, but the larger mix supersedes and largely effaces those traces, enabling something like a collective point of view in which the user's agency and actions are paramount.

The organisational logic of the synth depends neither on any one person's perspective (*à la* Canaletto), nor on multiple people's perspectives (even though these are the imagistic building blocks of synth), nor even on the navigational prowess of the end user. Instead, they depend on an algorithmic intervention. The selection of which images are in or out, of how they are spatially arranged, of which points offer access to which image all reflect the interplay of algorithmically determined points of uniqueness, on the one hand 'skinned' with photographic information, and on the other 'activated' by the navigational work of the user. This intervention stands in sharp contrast to Canaletto's painting, which embodies the subject-object relationship emblematic of the modern era, the relationship charted by Descartes and manifest in Heidegger's *Welt-bild*. The viewing subject inscribes her position, her subjectivity, in the calculus through which the object is apprehended and represented. Three-point perspective is the technique by which this relationship is visually expressed; it serves as the language of this relationship and the guarantor of its stability. Photosynth, by contrast, is enabled through an algorithmic intervention in the subject-and-object relationship. A process of neither the subject's nor the object's making determines point of access, sights seen,

FIGURE 2. Photosynth image of Piazza San Marco, Venice. Available online at http://photosynth.net/view.aspx?cid=c8c6203d-c361-4700-a343-2a5cd0c060f3.

connections made, experience gained. This is done, in a manner of speaking, in collaboration with both the subject and the object. But the effect undermines the unity of the viewing position hard-wired into Canaletto's painting, in which we see the Piazza from Canaletto's perspective. In Photosynth's case, while we are aware of seeing through many complementary sets of eyes, the enabler of those viewpoints and of the larger composite remains unseen and out of reach. It is a program layer that changes, that is redefined, that offers different affordances – in addition to being the single conduit through which we can access the image – while being completely outside the control of the user.

Authorship, in this context, is both problematic and pluriform. Although mostly effaced, it is the author of the individual photographs; largely enacted, it is the author of the experience – that is, the navigating user; fundamentally enabling, it is the author of the algorithm; and in terms of what we actually see and select from, it is the algorithm as author. Descartes' triumphant subject and the *Ich* implied in Heidegger's *Welt-bild* are not eradicated, for their traces remain abundant. Rather, they are fundamentally repositioned by the algorithmic regimes that now stand between subject and object.

IMAGES IN MOTION: AUGMENTED REALITY

Algorithms take many forms, and no digital camera or computer-based viewing system would be complete without them. But I would like to turn to a particular application that contrasts in important ways with Photosynth: location-based augmented reality (AR). AR systems overlay existing physical reality with an additional (augmented) layer, making visible information that can only be seen through a lens or on a screen (viewing technologies are quickly developing, and although camera-equipped cell phones and handhelds currently predominate, head-mounted displays – glasses – are already on the market, and working prototypes of retinal displays are now in the laboratories). The informational overlay is tagged to particular places in the world, and at the moment, three different systems may be used to link the real and the virtual. First, *fiduciary markers* are the most basic, consisting of graphically coded tags that physically attached to the object for which an overlay is sought. The device's recognition of the code triggers the appearance of the data layer in its screen, like the other systems, treating the information or image as spatially substantial (one may, in effect, walk around a virtual object as if it physically exists). Second, currently the dominant application, *digital compass tracking* systems essentially offer a kind of triangulation process using geo-positioned data and GPS (Global Positioning System), a compass and an accelerometer. Wikitude, for example, calculates users' positions by using these elements, linking them together with the Wikitude data set to provide geographic information (e.g. longitude and latitude), history and information regarding points of interest. Finally, *natural feature tracking* systems represent a fast-emerging image technology that assigns data to location by making visual correlations between physical places (i.e. 'recognising' them) and the information to be appended. An image-recognition system, it requires the user to position sights within a viewfinder, which the algorithms then process to find any correlations with the stored database. The system's search for unique identity points is conceptually related to Photosynth's, except that in this case, the user is in the physical world attempting to

correlate real and virtual data in order to trigger a virtual graphic overlay. For the moment, natural feature tracking's intensive processing demands have limited its use to powerful handhelds and relatively iconic locations, but among industry insiders, it seems to be the application of choice.

Augmented reality applications such as Layar (http://www.layar.com/) and Wikitude (http://www.wikitude.org/), even though they use digital compass and not natural feature tracking, illustrate the uses of AR. They can overlay dynamically responsive information on the screen of the handheld; point out the locations of any particular sort of business, displaying them as overlays in the device's screen, linking those locations to web-based information repositories and providing travel instructions; give visual access to a world of 3D user-generated fantasy characters that co-inhabit real space; and offer data overlays regarding underground infrastructure such as water, sewage and electrical conduits. The AR sector has experienced exponential investment growth over the past year or two, since most smart phones now have the requisite navigation devices and AR applications turn on simple (and free) software downloads. This growth has also been driven by the advertising industry, which sees AR applications as 'the next big thing'. However, more creative and interventionist applications have also been on the rise, giving users, for example, the ability to overlay commercial billboards with art projects or critical commentary, or to serve as tools for game play. These latter, more dynamic and even spontaneous applications point to the reason that natural feature tracking is preferred over other systems. If an art project seeks to overlay something relatively ephemeral such as Burger King's billboard advertisements and signs with a different graphic and logo (one of Julian Oliver's *artvertising* endeavours), the AR system must be able simply to recognise the graphic form rather than relying upon previously annotated GPS coordinates (something unlikely in this scenario).

My interest in natural feature tracking AR systems turns both on their temporal characteristics (their potential to embrace the ephemeral), and on their spatial requirements (the device-equipped human must be physically positioned in a particular way in order to get particular information). The requirements for physical positioning are quite specific – point the camera at the wrong location, use an unusual angle or an inappropriate time of day, and you will get other-than-expected results. In this, natural feature tracking AR systems share something with Photosynth's image requirements, where the assignment and alignment of points of uniqueness can only work if certain imaging rules are followed, assuming that the AR user has specific goals in mind and is not 'just exploring'. Unlike Photosynth, however, even if the user is 'just exploring', the resulting information in AR systems is quite specific, with meaning assigned to particular locations and information provided that can serve as incentives to act (in the case of retail businesses, coupons might appear, or contact information, store hours and location instructions).

The act of mediated looking in AR systems – that is, looking through the camera of the handheld computing device – is also always an algorithmically enabled navigational act. Through the lens of the AR-activated device, there is no such thing as an innocent gaze: the act of gazing and the views consequently seen are transformed into a process of signification as images are laden with particular meanings. Paul Virilio has described a related, if far more extreme, phenomenon in his characterisation of sight-activated missile guidance systems, where seeing is synonymous with targeting (Virilio 1989). The innocence of uninformed exploration – something hypothetically possible with Photosynth – is transformed into an encounter with the always-already-marked and significant. The user sets the parameters of signification, be they social (are any of my friends around?), cultural (tourist information once found in travel guides is now location-bound) or commercial (where's the nearest Starbucks?).

The differences between the two systems are striking. In the case of Photosynth, we move within a virtual image-space generated by the algorithmic assemblage of many points of view, many authors and even many times. And we do so presumably from a position of relative bodily stasis, sitting behind our computers. In the case of the AR systems under consideration, we move in the physical world, our movements and positioning tracked and algorithmically processed. The handheld device, our interface to the world, is both the agency of that tracking and the means of the augmented world's appearance. Point of view is embodied, constant, and synonymous with the viewing subject. Rather than wandering through a virtual image-space made up of many different points of view, we bridge the physical world with a virtual image-space that embodies one point of view: the algorithmically modulated viewing subject. Both Photosynth and the AR applications at hand require that we take the dynamic character of the image as given. Whether lacking a dominant point of view (Photosynth) or embodied in a particular point of view (AR), movement is an expected experiential dimension: movement within the image (Photosynth) or the image in motion (AR). Movement in Photosynth is

limited to the virtual world, but opens up the possibility of 'seeing through' many different points of view, through other authors' eyes. By contrast, movement in the AR domain is sited within the physical world, embedded in a multisensory environmental mix of sounds, smells and presences, but point of view is ultimately limited to that of the viewing subject.

The construction of meaning, too, functions quite differently between the two systems. Photosynths enter the world framed by meaning in the sense that they bear descriptions, are reinforced and built by image tagging, and, depending on their topic, bring with them the cultural baggage of the ensemble – whether Piazza San Marco or Obama's inauguration. That framing gives way to an experiential domain in which countless other vistas and associations may be generated, since the system gives access to up to 600 or so points of view, plus the interpretive frames and interests that the user brings to bear in interacting with them. The 'steering' mechanism is simply the interaction among the user, the coded source material and the enabling algorithm in the intervening data organisation and navigation system. In the case of location-based AR applications, meanings are as precise as the viewing position. Places are dynamically tagged with particular data almost simultaneously with the act of viewing, and these tags can in turn be linked to deeper data repositories (websites) and even other links (telephone connections), reaffirming the assignment of meaning to place. AR systems effectively overlay the viewer's access to the physical world with specific (and selectable) grids of signification.

Yet for all of their differences, applications like Photosynth and augmented reality share a fundamental realignment of subject-object relations thanks to their algorithmic processing layer. As stated earlier, this does not preclude human agency, nor the stubborn fixities of the world viewed. Rather, it resituates them, defining their parameters and enabling their interactions. And although difficult to 'see' (after all, we attend to the images before us and not to the underlying selection and organisation process), the algorithmic domain ultimately determines *what* we see, and even *how* we see it. There is a great temptation to rearticulate this as a form of 'old school' ideology, a distorting screen between us and the world, but this would be incorrect. For starters, it would imply the existence of an ideology-free past, somehow claiming that when subjects and objects met directly, unencumbered by the unseen hand of the algorithm, the world was a transparent place and its rule set visible to all. The long modern era of the algorismic was no less ideologically complicit. The flattening of the algorithmic and the ideological in the case of repositioned subject-object relations would also miss Heidegger's point about the *Welt-bild*. The fundamental reworking of our position as subjects vis-à-vis the world in this new algorithmically enabled era, with all of its many affordances and challenges, may well in itself suggest aspects of an emerging new order, one in which processual intermediation repositions the old certainties. We are moving from the *Welt-bild* emblematised by Canaletto's *Piazza San Marco* to the very different logics and subject positions of Photosynth, Google Street View, augmented reality and their ilk.

PERSPECTIVE

Let us return to Culler's notion of the marker – 'any kind of information or representation that constitutes a sight as a sight, giving information about it, representing it, making it recognisable'. We might also say that it constitutes a site as a sight, transforming space into place. And it is a site of value. In the two rather divergent algorithmic applications that we have considered, we have seen different approaches to the economy of the image. In one case, linguistic tags ('Obama's inauguration') begin a sorting process that is carried on by algorithmic analysis of the image, and its dissection and reassembly as a navigable point cloud. Human tagging and algorithmic processing both mark and construct a coherent image-space, with Photosynth aggregating different points of view, and enabling the user to wander among them. In the other case, locations in the physical world are virtually marked and given significance. Augmented reality systems create a window into this rich overlay of data and meaning, revealing the locations of banks and pharmacies, apartments for rent and T-mobile hot spots, virtual artworks and location-based historical data. The algorithmically enabled interplay between the viewer's position in the physical world and this virtual information layer is transformative, creating sites of meaning and enabling action.

One is tempted to call upon the distinctions drawn between travel and tourism that Culler charts – one an open and unstructured experience, with signification arising from new and unexpected ways of seeing; the other, carefully charted and well-labelled, fixed in semiotic import, with the path from one sight to another well-marked, and commodification often lurking nearby. But Culler is too intelligent a critic to settle for such essential definitions. He complicates these distinctions, pointing instead to shifting regimes of definitional referents, to an ages-old dynamic that elevates one experience at the expense of the other, destabilising any intrinsic meaning and pointing instead to a larger

dynamic. Both modes of engagement have historically struggled over the notion of the authentic. Deployment patterns of the term suggest that its meanings are highly relative, with the pot calling the kettle black when it comes to authentic encounters vs. tourism. Definitional relativism aside, times have changed. The notion of the authentic in the age of the modern, where subjects and objects encountered one another in relative peace, should probably be reconceptualised in an age where algorithmic interventions have modified the agency and place of both. The project of postmodernism might be read as the beginnings of such an endeavour. Culler reminds us that the authentic is not something unmarked or undifferentiated, but rather, that authenticity is a sign relation. And in so reminding us, he offers a possible way out of this dilemma. Culler says:

> A semiotic perspective advances the study of tourism by preventing one from thinking of signs and sign relations as corruptions of what ought to be a direct experience of reality and thus of saving one from the simplistic fulminations against tourists and tourism that are symptoms of the touristic system rather than pertinent analyses. Tourism, in turn, enriches semiotics in its demonstration that salient features of the social and natural world are articulated by what Percy calls 'symbolic complexes' and its revelation of the modern quest for experience as a quest for an experience of signs. Its illustration of the structural incompleteness of experience, its dependency on markers, helps us understand something of the nature of semiotic structures. (Culler 1990, 9)

Perhaps the algorithmic can benefit from a similar insight. Although of a different order than the clearly defined subject-object binary that characterised the modern era for the last few hundred years, the algorithmic turn remains rooted in human experiential and semiotic practices. On one hand, that element of human continuity might override newly emerging disruptive potentials, helping us instead to locate and assess the alterations to our established modes of interaction and being in the world, and make creative use of them. At a moment of transition, it is difficult to tell whether this is simply a default mode – the momentum of the past – or if it reveals a level of adaptive insight – part of the same fabric of imagination that creates and uses these new affordances. On the other hand, perhaps we are indeed approaching a significant turn in our organisation and use of data, and these new applications can both demonstrate and enable an underlying collaborative symbolic and experiential domain outside the tradition of the transcendent subject. The implications of either stance for our understanding of the visual, its relations to the subject and the parameters of its valuation and significance remain profound.

If the algorithmic reworking of subject-object relationships were simply limited to Photosynth and AR, the response would be simple. But the structural similarities in terms of algorithmic intermediation, a reconfiguration of subject-object relations, and new dynamics for the generation of meaning and value that can be found in such diverse applications as Wikipedia, dynamic cartographic interfaces, taste recommendation systems (Amazon), Google AdSense, even automatic stock market trading all suggest that something larger is at stake. These developments are but fissures in the still-robust ramparts of the modern. The stable subject position inscribed into three-point perspective and emblematised by Heidegger as the *Welt-bild* retain their cultural taken-for-grantedness. The arithmetic precision of the algorismic remains the basis of our currencies and economic lives. But as we explore the new affordances of the algorithmic, and as our capacities to deploy them grow in tandem with the progression of Moore's Law, we might also begin to reflect more critically about the differences in emerging modes of representation. If nothing more, these differences offer compelling heuristic entry points to interrogate our assumptions regarding representation and to reflect upon the implications of the alternatives.

NOTES

[1] 'Nothing is more boring than an unnamed landscape' (Culler 1990, 6).

[2] A long and interesting history of efforts to work with algorithms pre-dates the computer era. The work of Leibnitz (1956) and Harsdörffer (1971), among others, greatly complicates the simple bifurcation I have heuristically asserted in this article. For a compelling discussion, see Trettien 2009.

REFERENCES

CNN. 2009 The moment. http://edition.cnn.com/SPECIALS/2009/44.president/inauguration/themoment/.

Culler, J. 1990. *Framing the sign: Criticism and its institutions.* Norman: University of Oklahoma Press. http://www.scribd.com/doc/36984937/culler1.

Harsdörffer, Georg Philipp. 1971. *Poetischer Trichter; Die Teutsche Dicht – Und Reimkunst, Ohne Behuf Der Lateinischen Sprache, in VI. Stunden Einzugiessen* [A poetic funnel for infusing the art of German poetry and rhyme in six hours, without benefit of Latin language]. Hildesheim and New York: G. Olms.

Heidegger, M. 1977 [1938]. The age of the world picture. In *The Question Concerning Technology and other essays*, translated and edited by William Lovitt. New York: Harper Torchbooks.

Leibniz, G. W. 1956. On the general characteristic. In *Philosophical papers and letters*, 2nd ed., edited and translated by Leroy E. Loemker. Boston: D. Holland.

MacCannell, D. 1976. *The tourist*. New York: Schocken.

Trettien, W. 2009. Computers, cut-ups, & combinatory volvelles: An archaeology of text-generating mechanisms. MSc thesis, Massachusetts Institute of Technology. http://whitneyannetrettien.com/thesis/.

Virilio, P. 1989. *War and cinema: The logistics of perception*. London: Verso.