

Google Street View: From Spatiotemporal Construction to Algorithmic Intervention

Tiancheng Cao

s1255916

leo.t.cao@gmail.com

Research MA Thesis Arts and Culture

Leiden University

Supervisor: Helen Westgeest

Second Reader: Eric de Bruyn

August 2017

24,910 words

Table of Contents

Introduction	1
Chapter One Towards a Notion of Non-Simultaneity	5
Instantaneity and Simultaneity: The Recounting of Filmic and Photographic Time	6
The Construction of an Intra-Frame Temporality	11
Not Now but Almost Here: Street with a Narrative	20
Chapter Two The Technologized Virtual (Museum) Space	28
From the Mobilized Virtual Gaze to the Panoramic Perception	28
Museum without Walls: Google Arts & Culture	38
The Late Capitalist Museum and the Technologized Subjects	50
Chapter Three The Algorithm, The Screenshot, The Photograph	58
From the Photographic Paradigm to the Algorithmic Turn	58
Performative Cartography and the Screenshot Aesthetic	65
The Screenshot versus the Photograph	73
Conclusion	82
List of Illustrations	84
Bibliography	86

Introduction

“Searching for Art Just Got Better. Where Will You Start?”¹ On May 31, 2017, exactly ten years after the launch of Street View, Google published the blog post announcing a deeper integration between two of its major products: Google Maps and Google Arts & Culture. In a promotional video, we start with a search in Google Maps for the Tokyo National Museum of Western Art. On the resulting cartographic view showing the rough floor plan of the building, a yellow human icon (Pegman) is dragged over to hover above a labyrinth of blue lines. As the Pegman lands on the map, our view switches to the museum interior, where several paintings are on display in the Street View mode, some with a circle superimposed at the bottom right corner. A click on the circle reveals more information about the work and the artist, and another click on “view more” takes us to the Arts & Culture webpage, where the view is zooming in and out on a high-definition reproduction of one of Claude Monet’s *Water Lilies*.

The whole video is only about forty seconds long, but it is not a condensed compilation showcasing new features added to the two products. Instead, it is a genuine depiction of how an actual user would operate in the product interface and transition between different views (from cartographic to photographic, from bird-eye to spherical). While the process of searching for art starts with Maps and ends with Arts & Culture, the connective tissue that links the two is provided by the technologies behind Street View. Speaking in terms of user experience, we may have already been used to navigating through a digital map or browsing a webpage with illustrative images, but the exploration of a three-dimensional virtual space that is constructed entirely from photographic images remains for many a relatively new experience that I believe deserves more scholarly attention.

If Google’s search engine has managed to establish itself as by far the most popular tool with which to access information, its mapping services are now trying to replicate the success by creating an ever-growing virtual representation of the physical world. In terms of image production, one of the most intriguing aspects of Street View is the fact that the entire virtual space is stitched from countless individual digital photographs. To collect

¹ Niccolai, 2017.

these images as raw materials for post-production stitching, Google dispatches its vehicles all over the world, each equipped with a specially designed camera set that simultaneously captures up to fifteen images in different directions (Figures 0.1 and 0.2).² The number of cameras on this rosette set helps to minimize the stitching artifacts and create a more seamless 360-degree experience for users. To view these images, a user can access Street View either from Google Maps, as described in the promotional video, or through Google Earth, in which case the view transitions from satellite images to aerial photographs before “landing” on the street. In the street-level spherical view, the arrow indicates the direction in which a user can proceed, and the cross sign marks the spot where the next spherical view is available (Figure 0.3).



Figures 0.1 and 0.2 A rosette camera set with fifteen cameras is installed on all Street View vehicles.



Figure 0.3 Screenshot from Google Street View, Witte Singel, Leiden, retrieved June 2017.

² The so-called Google’s Fleet includes Street View cars, trekkers, trolleys, snowmobiles, and trikes, intended for various indoor and outdoor circumstances. See <https://goo.gl/oaPs9X>, all URLs accessed July 15, 2017, unless otherwise specified.

From a theoretical perspective, the relevance of a research project dedicated to Street View also stems from its rooted connection with photographic images. In a simplified manner, we can think of this connection as the assembly and arrangement of many single-frame photographs, which immediately brings the filmic medium into the discussion, since film can also be seen as one established way of assembling and arranging single-frame images. This triangular relationship thus not only introduces to Street View the various topical debates surrounding the photographic medium, but also propels us to extend the comparative analysis between photography and film onto this new form of visual representation. While time and space are both recurring themes in the discourse of lens-based media and seem always to structure the discourse around still and moving images alike, Street View signifies the intrusion of a new factor: the algorithm. The unprecedented and still expanding scope of Street View makes it impossible for the program to be created and maintained only with human labor. The image stitching process, for instance, is completed by pre-programmed algorithms that are in effect free from direct human intervention. While this observation is by no means to suggest the absence of human agency within Street View, it is nonetheless important that we take into account the role played by the algorithm and the implications it may carry for our increasingly media-saturated visual culture.

Accordingly, motivated by the three theoretical issues of time, space, and algorithm, in this research I aim to find out how we can understand Street View as a new medium in terms of its spatiotemporal characteristics as well as the invisible algorithmic interventions. Specifically, I will attempt to address the following questions: how does the construction of a virtual space in Street View with single-frame photographs necessarily invoke a parallel presentation of time? What are the spatial parameters that define the relationship between the viewer and the virtual space? And, to what extent do the algorithmic interventions in Street View signify a paradigmatic change in the visual representation of physical reality?

To answer these questions, I will conduct an interdisciplinary research that draws on existing literature mainly from film and photographic studies, new media studies, as well as museum studies, with a particular focus on intermediality, the interconnectedness between Street View and other existing media forms. The research will be divided into three parts. Chapter One “Towards a Notion of Non-Simultaneity” attempts to establish

the temporal scheme of Street View images based on a comparative review of photographic and filmic time, as well as an outline of the various types of movement involved in the production and perception of the images. Also, based on Robin Hewlett and Ben Kinsley's *Street with a View* (2008), I will examine how the new temporal scheme has endowed the images with a strong narrative potential. Chapter Two "The Technologized Virtual (Museum) Space" considers how the relationship between the virtual space and the viewer is structured by the "mobilized virtual gaze" and the "panoramic perception." Meanwhile, as technologies behind Street View are being used to depict museum space, an institutional layer is added to the virtual environment, to which I will respond with a new media approach and, based on the Arts & Culture project, attempt to account for its increasing popularity. Chapter Three "The Algorithm, The Screenshot, The Photograph" looks beyond the spatiotemporal construction and examines the mechanism beneath the image surface. I will consider how the "algorithmic turn" has superseded the "photographic paradigm" and in so doing evoked a burgeoning "screenshot aesthetic." Then, Clement Valla's *Postcards from Google Earth* (2010-ongoing) will serve as the case study, revealing how this new practice has resulted again in the focus on single-frame images. Finally, I will conclude the chapter with a comparative analysis between these screenshot images and the original photographic images used to construct the virtual space in the first place.

Searching for art may have just got better, but with rapid technological innovations in both media and cartography we may well expect Street View to become even better in the near future. Underlying this relentless betterment are not only the challenges it poses for scholarly research to comprehend and, quite literally, catch up with all the changes, but it also accentuates the very urgent need for scholars to establish and constantly revise the theoretical link between Street View and other representational schemes that precede its inception. By addressing the above issues surrounding this new media platform, I hope to position this research within the larger interdisciplinary debate on time, space, and algorithm, shedding new light not only on the link between Street View and the established tradition of visual representation but also on its possible future position within the increasingly technologized visual culture.

Chapter One Towards a Notion of Non-Simultaneity

The fact that Google's vehicles collect single-frame digital photographs as raw materials for the construction of a virtual space positions Street View in close intersection with film and photography. On the one hand, the display of filmic images, whether analog or digital, still depends on the progression of individual photographic frames; on the other hand, the way in which single-frame images are organized in films (one after another) is diametrically opposed to that in Street View (one next to another). Therefore, such an observation necessitates an in-depth examination into the relationship between Street View images and the more traditional lens-based photographic and filmic images.

From which perspective should we approach this examination? It is important for us to first realize that with the digitization of visual images comes a steady convergence between photography and film, since both mediums now increasingly share the same technological nature (algorithmic image), distribution system (the internet), and contexts of production and reception (mobile and computer screens).³ Nonetheless, the taxonomic distinction between a photograph and a film remains persistently visible, one important aspect of which is the association of movement with the former and that of stillness with the latter. It is for this reason that this chapter focuses on temporal aspects of Street View images and addresses the following question: how does the construction of a virtual space in Street View with single-frame photographs necessarily invoke a parallel presentation of time? To answer the question, I will start with a comparative recounting of the temporal schemes of filmic and photographic images. Then, I will examine the various types of movement involved in the navigation of Street View, with special attention paid to the role played by what I call the "intra-frame temporality." Finally, based on the case study of Robin Hewlett and Ben Kinsley's participatory project *Street with a View* (2008), I will examine how the proposed temporal scheme in Street View has endowed the images with a strong narrative potential.

³ Cohen and Streitberger, 2016, 8.

Instantaneity and Simultaneity: The Recounting of Filmic and Photographic Time

To unravel the temporality of Street View images, the temporal aspects of both filmic and photographic images require a comprehensive examination. Even though the invention of photography predated and was a necessary prerequisite for that of film, I will nonetheless start with the latter, whose temporal dimension is in comparison less complicated. For philosopher Stanley Cavell, the material basis for the medium of film is “a succession of automatic world projections,” in which “succession” includes the motion depicted on screen, the motion of successive frames, and the juxtapositions of cutting.⁴ Pertinent to our discussion here is the second movement, the continuous replacement of static frames that, according to film historian Thomas Elsaesser, creates a mere “illusion” of motion in the eye of the beholder.⁵ This classical interpretation of film movement also resonates with film theorist Peter Wollen, who in the 1980s questioned the difference in semantic structure between still and moving images and argues instead that it is not “movement but sequencing (editing, *découpage*) which made the main difference by determining duration differently.”⁶ Despite the nuances in phrasing, what Cavell, Elsaesser, and Wollen have in common is the emphasis on static frames and how their temporal succession purportedly establishes the material basis of film movement. Indeed, the process of frame progression is so intrinsic to our understanding of moving images that we continue, even in the digital era, to base the production and perception of them on the concept of the frame. Thus, the temporal scheme of filmic images is based on what I call an *inter-frame temporality*.

The recurrent emphasis on the role played by the individual frame naturally leads our discussion back to photographic images, where the capability of conveying the passage of time is seemingly at odds with the very apparatus of the medium. Perhaps for this reason, the classical view on photographic time seems to be centered on an irresistible charm of instantaneity. For film critic André Bazin writing in the 1950s, the photographic image is an act to preserve, “to snatch it from the flow of time, to stow it away neatly ... in the hold of life.”⁷ For film theorist Christian Metz writing in the 1970s, it is the representation of “a

⁴ Cavell, 1979, 96-97.

⁵ Elsaesser, 2011, 117.

⁶ Wollen, 2007, 110.

⁷ Bazin, 2005, 9.

point in time that has been frozen.”⁸ For Wollen, photography is ice and film is fire, one with “the cryogenic power to preserve objects through time without decay,” while the other “all light and shadow, incessant motion, transience, flicker, a source of Bachelardian reverie like the flames in the grate.”⁹ These depictions of photographic time, however, are by no means comprehensive. Already in the 1870s, British photographer Eadweard Muybridge’s photographic studies of animal locomotion aimed to amalgamate discrete segments of instantaneity so that human vision could perceive a continuous motion flow. Into the second half of the twentieth century, photography’s association with instantaneity continued to be challenged. Among the various artistic attempts, Japanese photographer Hiroshi Sugimoto’s *Theaters* series, which captures the entire duration of a movie in one single frame, is often seen, “in a literal way, [as] the embodiment of temporal duration.”¹⁰

The photograph as paradox, as art historian Thierry de Duve puts it, between “time exposure” and “snapshot” seems to be the structuring condition for the debate concerning photographic time,¹¹ but time exposure remains so ambiguous a concept that a distinction should be made between what I call *simultaneous duration* and *non-simultaneous duration*. In his reflection upon the experience of time in still photography, Belgian photographer Maarten Vanvolsem claims that in a single-frame photograph exposure time “is the same for every single spot within the frame. Not only is the amount, the quantity (the exposure time, duration) the same but also the time when the exposure took place. There is simultaneity.”¹² This claim is not a medium-specific observation, but is rather in line with the broader developments in the tradition of pictorial art since the Renaissance. By the seventeenth century, as art historian Lew Andrews notes, “[p]ictorial space had in effect congealed, had become unyielding and unchanging, and had acquired a unity and simultaneity that it had never before possessed.”¹³ This change in pictorial space, among other things, led writer and philosopher Gotthold Ephraim Lessing to his distinction between arts of space (such as painting) and arts of time (such as poetry).¹⁴ Commenting

⁸ Metz, 1991, 19.

⁹ Wollen, 2007, 110.

¹⁰ Green, 2006, 9.

¹¹ Duve, 1978, 113.

¹² Vanvolsem, 2005, 54.

¹³ Andrews, 1998, 104-105.

¹⁴ Lessing, 1836, 150-151.

on Lessing's dichotomy, artist and writer Victor Burgin argues that such a distinction "underwrites the categorical separation of the still and the moving image on the basis of a supposed absolute difference between simultaneity and succession."¹⁵ Based on these arguments we can see the extent to which the notion of simultaneity has structured the debates within the discipline of visual culture.

For many contemporary visual artists, however, the gravitation to simultaneity fails to maintain and has in fact become an overly generalized assumption that they actively aspire to challenge. For instance, in British artist Sam Taylor-Johnson's *Five Revolutionary Seconds* (1995-2000), a series of 360-degree panoramic images are made each with a five-second exposure, depicting interior environment and its human inhabitants with a continuous flow that records multiple temporal instances (Figure 1.1). The same applies to Vanvolsem's strip photography (Figure 1.2), which is "a chronological compilation of the individual line images continuously recorded by the camera during an event."¹⁶ In both cases, a panoramic image format is used in combination with an exposure process that is certainly durative but by no means simultaneous. In Vanvolsem's words, "two different spots of the image will have been exposed at a different time. ... They will have different time coordinates when compared to a time line."¹⁷

Indeed, the "time coordinate" of Sugimoto's *Theaters* would be the entire duration of the movie, and the temporal markings would be the same across the image frame. In the images of Taylor-Johnson and Vanvolsem, in contrast, the difference in time coordinates inevitably leaves a temporal trace that helps the viewer reconstruct the direction, circular or linear, in which the camera has moved. If Sugimoto's photographs *accumulate* time onto a contemplative façade, then Taylor-Johnson's and Vanvolsem's images *distribute* time into an expanded physical space. This difference between simultaneous and non-simultaneous duration is also implied in the comparative case study by art historians Hilde Van Gelder and Helen Westgeest, who characterize Vanvolsem's strip images as "a

¹⁵ Burgin, 2009, 302.

¹⁶ Petersen and Davidhazy, 2013, 616.

¹⁷ Vanvolsem, 2005, 54.

multiplication of processes in time” whereas Sugimoto’s photographs “a concentration of processes in time.”¹⁸



Figure 1.1 Sam Taylor-Johnson, *Five Revolutionary Seconds IX*, 1997.

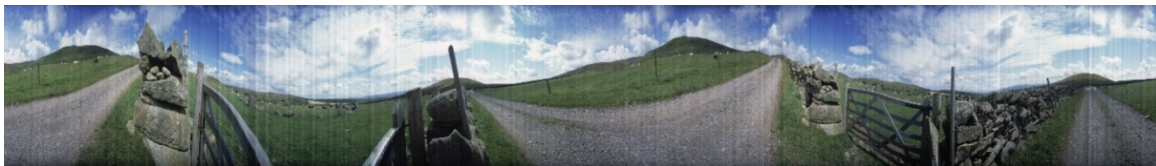


Figure 1.2 Maarten Vansoleem, *Specious-present 2 (Seas)*, 2006.

The difference can also be explained in a graphic format, as is demonstrated Figure 1.3, in which the temporal axis is “cut through” by the photographic frame. In Figure 1.3a, the “plane of simultaneity” (shaded square) is the visual equivalent of an instantaneous snapshot, neatly stowed away, in Bazin’s words, from the flow of time. Next to it, as the plain of simultaneity begins to accumulate itself in time, a depth in time (shaded cube) is added onto the frame, so the resulting image, one of time exposure, becomes the *concentration* of such a temporal depth. In Figure 1.3b, however, only a slit of space is open on the frame, but as the exposure process proceeds, the temporal *multiplication* (shaded rectangle) is continuously transferred onto, and indeed translated into, a spatial expansion supported by the material base of an unwinding film, which is not unlike how a seismograph continuously translates the duration and intensity of the Earth’s vibrations and movements into a graphic record.

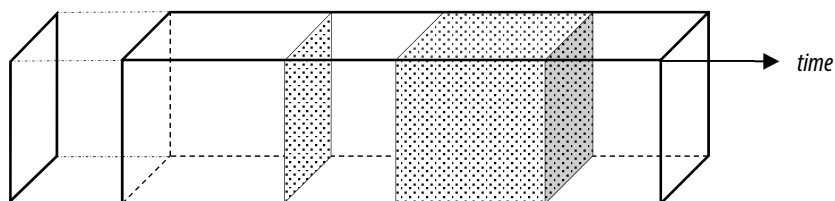


Figure 1.3a Demonstration of Snapshot and the “concentration” of time in Sugimoto’s works

¹⁸ Van Gelder and Westgeest, 2011, 84.

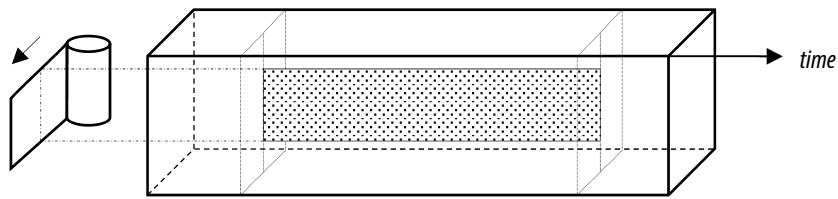


Figure 1.3b Demonstration of the “multiplication” of time in Vanvolsem’s works

	Simultaneity	Non-Simultaneity
Instantaneity	<i>Snapshot</i>	?
Duration	<i>Sugimoto</i>	<i>Taylor-Johnson & Vanvolsem</i>

Table 1.1 Temporal schemes of photography based on (non-)instantaneity and (non-)simultaneity

This comparison also completes the review of the temporal scheme of photographic images. Although the paradox between snapshot and time exposure has been a recurring theme in the debate concerning the medium’s relation to time, the notion of duration, as we have demonstrated in this section, is anything but homogenous or unified. Instead, we should always approach the axis of instantaneity and duration while taking into account the notion of (non-)simultaneity at the same time. Table 1.1 recapitulates the two axes in a matrix format, along with four possible combinations of temporal schemes. If the snapshot indeed represents a “point” frozen in time, it will undoubtedly occupy the intersection between instantaneity and simultaneity. In contrast, Sugimoto’s *Theaters* series can be located at the conjuncture of duration and simultaneity. While the innovative techniques of Taylor-Johnson and camera modifications of Vanvolsem successfully integrate duration with non-simultaneity, we are still faced with a seemingly impossible combination of instantaneity and non-simultaneity. Although we cannot simply attribute this combination to Street View images, the two temporal axes, including the question-marked part left unexamined, will nonetheless provide the theoretical basis on which we approach the temporal scheme of these images in the next section.

The Construction of an Intra-Frame Temporality

With the establishment of a more complete depiction of filmic and photographic time, we can now proceed to the investigation of the temporal aspects of Street View images. Unlike live stream videos or surveillance footages, these images are not updated in real time and definitely appear static and frozen. Still, navigating through the virtual space invokes an emphatic sense of time and movement in several important aspects.

First of all, time is an inherent aspect of the image collection process, since the Google's Street View vehicles (or satellites and aircrafts in the case of Google Earth) are constantly on the move when they capture the images.¹⁹ In addition, the viewer is more or less aware of the process in which the images are produced, and this awareness may contribute to a critical reflection upon the relationship between still image and passing time. This aspect of movement is comparable to the perception of Sugimoto's photographs, since knowledge about the production process, that an entire movie is exposed in front of the camera, plays a key role in shaping the interpretation of the series. Similarly, we can compare this aspect of image collection with American artist Edward Ruscha's *Every Building on the Sunset Strip* (1966) (Figure 1.4), in which slight variations in perspective and exposure among consecutive image blocks that are assembled into one sequence reveal to us not only the production process, but also a confrontation between the images that are still and the time that passes regardless.



Figure 1.4 Edward Ruscha, *Every Building on the Sunset Strip*, 1966.

The second aspect of movement is the physical movement on the part of the viewer. Unlike the cinematic scenario, where the succession of frame is mostly beyond the control of the viewer, programs such as Street View, by virtue of its inherent interactivity, require

¹⁹ For satellite and aerial images in Earth, Google works with different image providers, with provider name and date of acquisition displayed at the bottom of each view.

constant user commands to perform their proper functions. Starting the program, typing an address, and navigating through the space, all these user commands, conducted with either keyboard input or voice-control, are necessary components for the program to display anything accordingly. At the current stage, the required movement on the part of the viewer is similar to that in computer and mobile gaming: fixed in position with hand-based controls. However, with the introduction of virtual reality in Street View, the viewer would be able to move around in real-life space to initiate a corresponding movement in the virtual space. In this case, bodily movement would involve physical displacement, not unlike the experience of someone viewing a nineteenth century panorama, where walking around is a necessary component of the viewing process.

The last piece of the puzzle is that which connects the movement of production and that of perception: the image itself. Clearly, Street View images do not present movement as do filmic images; in fact, when individual images are combined into a composite format, inter-frame temporality becomes a conceptual impossibility, since the concealing of frames in order to create a seamless space is the very aim of image stitching. Nonetheless, frame-based movement is but one possible way of presenting time. In Table 1.1, there remains a conceptual conjunction left unexamined, the seemingly impossible conjunction of instantaneity and non-simultaneity.

Indeed, we might ask, how can an instantaneous situation also be non-simultaneous? If a photograph represents, as Metz puts it, a “point” in time, it would constitute, in linguistic terms, a *punctual situation*. According to linguist Bernard Comrie, punctuality, as opposed to durativity, is “the quality of a situation that does not last in time (is not conceived of as lasting in time), one that takes place momentarily. ... A punctual situation, by definition, has no internal structure.”²⁰ This, however, clearly contradicts our basic understanding of the photographic medium, especially its mechanism of image production. No matter how short a predetermined exposure time is, it remains durative from a linguistic perspective and will therefore never become truly instantaneous. From a phenomenological approach, art historian George Baker also concedes that the reduction of photographic images to a “purely visual stasis” is a condition that modernist photography never succeeded in

²⁰ Comrie, 1976, 42.

achieving.²¹ The question then becomes: what is the “internal structure” of the snapshot as a pseudo-instantaneous process? How can the internal temporal constituency of a photographic exposure reconcile the apparent contradiction between instantaneity and non-simultaneity?

“You wait and wait, and then finally you press the button – and you depart with the feeling (though you don’t know why) that you’ve really got something.”²² Yet, French photographer Henri Cartier-Bresson knows exactly why: the release of the camera shutter at a “decisive moment.” Nevertheless, instantaneity is such an expedient term that what happens to the camera shutter within that fraction of a second is usually taken for granted. “You Press the Button, We Do the Rest” was not only a catchy slogan that marked the commercial success of the Kodak camera at the dawn of the twentieth century, it also ushered in a new era when the history, if not definition, of the medium transmuted from being technologically focused to being culturally oriented. Google does not have the luxury of simply pressing the button, and the niceties of shutter mechanism remain a technical challenge for its researchers. In a 2010 report specifying these challenges involved in capturing and presenting street-level images, the research team notes:

We tried mechanical shutters in R3 and R4 but settled on CMOS sensors with an electronic rolling shutter for R5 through R7. A key problem in these later designs was to minimize the distortion inherent in shooting from a moving vehicle while exposing each row of the image at a different time.²³

Several terms require explanation. R3, R4, and R5 are all previous versions of the rosette camera set (Figure 0.1). Both CMOS and CCD are commonly used technologies for the production of image sensors, but most digital cameras today, including those built in mobile devices, use a CMOS sensor.²⁴ There are several similarities between the two types, but one major distinction is the readout mode, the way in which each sensor reads and transfers signal charges accumulated on the pixels. A CCD sensor typically uses a *global*

²¹ Baker, 2005, 126.

²² Cartier-Bresson, 1952, n.p.

²³ Anguelov et al., 2010, 34.

²⁴ CMOS: Complementary Metal Oxide Semiconductor; CCD: Charge-Coupled Device.

shutter mode, which reads and exposes all the pixels simultaneously. With a CMOS sensor, in contrast, pixels are read row by row, creating a time delay between each row's exposure. This case, as described in the report, is often referred to as a *rolling shutter* mode.

Strictly speaking, both readout modes pertain only to digital imaging, but the concept of a non-simultaneous exposure sequence predated the digital era. In a typical mechanical focal-plane shutter, two metal strips, known as front and rear curtains, are positioned in front of the film.²⁵ The exposure process is regulated by the front curtain moving across the frame followed by the rear curtain, leaving a slit of space in between through which film is exposed to light (Figure 1.5). The width of the slit is in direct proportion to the length of the exposure, so the slit can be regarded as the mechanical counterpart of one row of pixels. Expectedly, image distortion caused by this progressive exposure also has its analogue predecessor, the best-known example of which is perhaps French photographer Jacques-Henri Lartigue's 1913 photograph *Le Grand Prix A.C.F.* (Figure 1.6). In the image, we see people and electricity poles in the background lean to the left, while the race car, especially the wheels, lean to the right. Lartigue was panning the camera so as to capture a sharp image, but it was not fast enough, so during the short yet not instantaneous exposure process, the car was still moving forwards, whereas everything in the background was, in relative terms, moving backwards. In the meanwhile, the narrow slit in the shutter was scanning through the exposure frame from top to bottom. Figure 1.7 demonstrates how image distortion can be created when fast moving objects are being photographed with a rolling shutter.²⁶

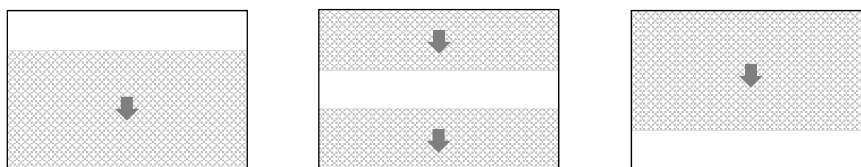


Figure 1.5 Demonstration of the exposure process with a rolling shutter.

²⁵ For a discussion about non-focal-plane mechanical shutters, see Vanvolsem, 2011, 43. He also notes that the Dutch translation for focal-plane shutter is *spleetshutter*, which literally means “slit shutter.”

²⁶ Images projected through the lens onto the film or the sensor are upside-down, so the downwards moving slit will in fact expose the bottom of the image first, thus the upwards arrow in Figure 1.7.



Figure 1.6 Jacques-Henri Lartigue, *Le Grand Prix A.C.F.*, 1913.

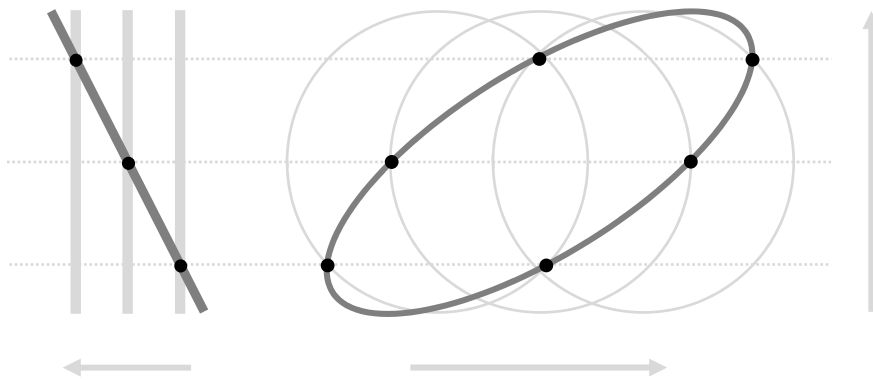


Figure 1.7 Demonstration of image distortion in photographing fast moving objects with a rolling shutter.

The minimization of focal-plane distortion remains a practical concern for the research team, but its very presence carries theoretical implications regarding the temporal aspects of the photographic snapshot, whose supposedly unqualified instantaneity can be refuted from both a linguistic and a phenomenological perspective. Instead, the internal structure of the snapshot taken with a rolling shutter, whether mechanical or electronic, is one of *non-simultaneous duration*, which necessitates that various parts of the same photograph be exposed at different points in time and bear different temporal imprints, or, in Vanvolsem's words, time coordinates. In Figure 1.7, for example, all the dots along the same horizontal axis will have the same time coordinates, but across the entire frame, there will be, to quote Van Gelder and Westgeest, a *multiplication* rather than *concentration* of coordinates in time. In terms of the graphic model that we construct in Figure 1.3a, this

would mean that the plane of simultaneity never really exists, unless a global shutter mode is used in the exposure process.

If we are to continue the mathematical metaphor, then two different coordinates will be sufficient to form a *vector*, which, unlike a *scalar*, is a quantity with both magnitude and direction. For instance, the duration of a flight is a scalar, but the displacement of an airplane from X to Y is a vector. Similarly, a two-second exposure, or an exposure of any duration for that matter, is a scalar, but a non-simultaneous exposure sequence that scans through the frame should be seen as a vector. Now, along the vector that connects all the different time coordinates, we could draw another axis, which would not only indicate the inverse direction in which the slit moves,²⁷ but also become, quite literally, a segment of the “timeline.” Looking back at Figure 1.7, we can now add another layer of meaning to the upwards arrow. Suppressed under the split second, this invisible yet integral timeline is the internal temporal constituency of the snapshot that I have tried to foreground. Comparing with the *inter-frame temporality* that we have attributed to film, I call this new temporal scheme in photographic snapshot *intra-frame temporality*.

This newfound notion of intra-frame temporality is clearly predicated upon the logic of non-simultaneity, which brings the photographic snapshot to a closer alliance with the images by Vanvolsem and Taylor-Johnson. What both artists strive to underscore with either a modified camera design or an alternative photographic technique has been a crucial constituent of the medium from the very outset. To some extent, the association of photography with stillness can be regarded as an act of retroactive construction, as writer and artist David Company contends:

Stillness in images only became apparent, understandable and truly desirable in the presence of the moving image. ... Cinema, we could say, wasn't just the invention of the moving image, it was also the invention of stillness as a sort of by-product. In the era of cinema, the frozenness of the snapshot ... came to be understood as the essence of the photographic.²⁸

²⁷ Again, the direction is inverse because the image projected through the lens is upside-down.

²⁸ Company, 2007, 189.

In a similar vein, Burgin argues that “[t]o equate movement with film and stasis with photography is to confuse the representation with its material support.”²⁹ likewise, new media theorists Ingrid Hoelzl and Rémi Marie contend that this transmedial association is “the result of a technological and conceptual standardization.”³⁰ Therefore, it could be argued that what the notion of intra-frame temporality helps to (re)consolidate is the medium’s intrinsic capability to present time and movement. When we compare inter- and intra-frame temporality, the role of the frame, or lack thereof, becomes an important consideration. When the images are stitched together, frames that otherwise mark the beginnings and endings of the intra-frame timelines will also disappear. This would have resulted in an inexplicable temporal jumble, but Google’s solution to the aforementioned image distortion issue provides a convenient alternative. As the researchers realize, “[t]he cameras must be in portrait orientation so that the exposure window’s movement is roughly parallel to vehicle motion.”³¹ In this modified configuration, the pixels are in effect read column by column, thus rotating the vertical intra-frame temporal axis into a horizontal position. When the images are stitched together in this way, each and every intra-frame temporal axis will also be weaved into one single, unified timeline (Figure 1.8). The change in camera orientation also brings the production of Street View images closer to that of Vanvolsem’s strip images, since in his modified camera, it is through a fixed vertical slit (Figure 1.3b) that the rolling film is continuously exposed to the light source. If a strip photograph is “a chronological compilation of the individual line images” and “display(s) time as a visual component,”³² then Street View images are the chronological compilation of the individual frame(less) images, which in themselves involve time as an invisible but inherent component.

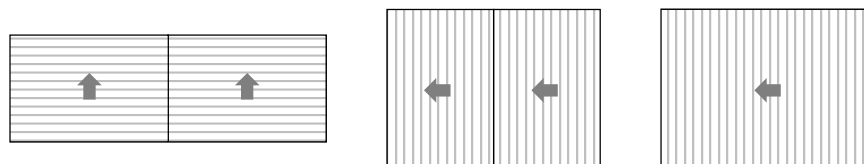


Figure 1.8 Demonstration of intra-frame temporal axes weaved into one single timeline.

²⁹ Burgin, 2009, 302.

³⁰ Hoelzl and Marie, 2015, 4.

³¹ Anguelov et al., 2010, 34.

³² Petersen and Davidhazy, 2013, 616.

The constructed timeline is the third and perhaps the most important dimension of movement in Street View images, though it is not “movement” in the literal sense of the word, as in the collection and perception of the images. Rather, it suggests a temporal potential (the accumulation of intra-frame temporality) within the image, something inherent and persistent irrespective of in what sequence user navigation proceeds. It is also of note that all three types of movement – that of production, that of perception, and that within image itself – can be aligned by a simple visual annotation. In both Maps mode (Figure 1.9a) and Earth mode, a grid of blue lines will indicate the availability of Street View images, with tiny arrows pointing to the direction in which the vehicles have traveled; In Street View mode (Figure 1.9b), a similar yellow line displays the trace of vehicle movement as well. In addition to indicating the traces of image production, these lines also track movement of perception, functioning as limitations within which we as virtual spectators are able to locate the view. Without the presence of the blue line,³³ the Pegman cannot drop and will land instead in the legend section (bottom right in Figure 1.9a). Finally, these lines also serve as the embodiment of the horizontal timeline pieced together by the individual intra-frame temporal axes. In other words, the annotated lines add a palpable form to the passage of time, which has become possible in the first place exactly as a result of the convergence of three types of movement.

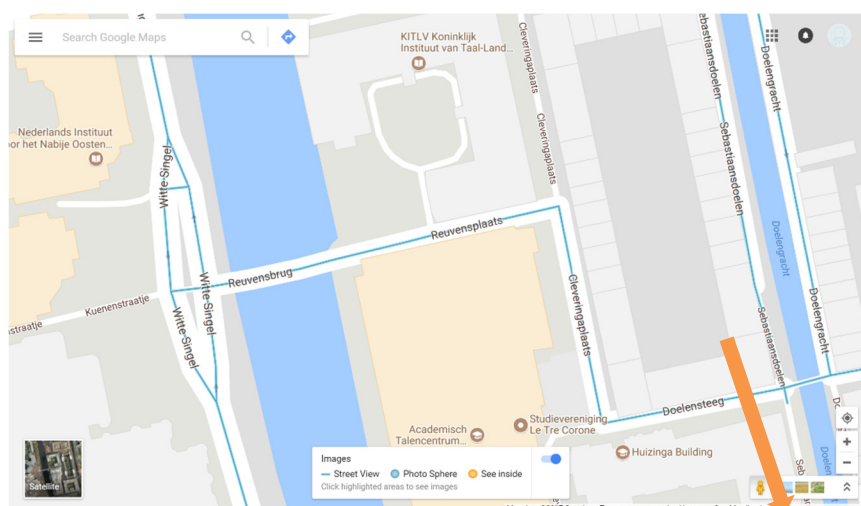


Figure 1.9a Screenshot from Google Maps, Reuvenplaats, Leiden, retrieved June 2017.

³³ When only a single 360-degree panoramic image, known as Photo Sphere, is available, a blue circle will be displayed and the viewer can click on or drag the Pegman to the circle to view the image.



Figure 1.9b Screenshot from Google Earth, Witte Singel, Leiden, retrieved June 2017.

This passage of time was further heightened in 2014, when Google decided to release “historical imagery from past Street View collections dating back to 2007 to create this digital time capsule of the world,”³⁴ adding yet another layer to the temporal displacement of the images. The historical images are accessible via a clock icon, and the viewer can move the slider to see images from a designated year and, if available, month. Brooklyn-based artist Justin Blinder incorporates cashed images from Street View and sources from local government of Manhattan and Brooklyn in his project *Vacated* (2013-ongoing), featuring animated GIF images of urban streets dramatically morphing back and forth between a previous date and a later period, often with a newly gentrified façade. The artist’s implied social agenda aside, the visual effect of contrast is in part attributable to the potential of Street View images to present time in a multifaceted manner. To summarize, it is now clear that Google’s construction of the virtual space in Street View has indeed managed to invoke a parallel presentation of time: the intra-frame temporalities of individual photographs are combined to form a continuous timeline, which, along with physical movement involved in the processes of image production and perception, becomes the basis for the invisible yet inevitable presentation of time. In other words, based on the temporal scheme that we construct in this section, we might say that navigation in Street View is as much a trip in space as a journey through time.

³⁴ Official Google Blog, 2014.

Not Now but Almost Here: Street with a Narrative

As is the case with film, the establishment of a unique temporal scheme in Street View immediately propels us to consider the relationship between this new temporality and narrative. In his discussion about the viewer's psychological response to photography and its space-time construction, de Duve makes the following comments:

For an image to be read requires that language be applied to the image. And this in turn demands that the perceived space be receptive to an unfolding into some sort of narrative. Now, a *point* is not subject to any description, nor is it able to generate a narration. Language fails to operate in front of the pin-pointed space of the photograph, and the onlooker is left momentarily aphasic.³⁵

Indeed, a point in time, as Comrie would argue, is a punctual situation with no internal structure, hence the incapability for any narrative to unfold. The space of a single-frame photograph may continue to be “pin-pointed,” but its temporal nature should nevertheless be regarded as a line segment rather than a point, at least when a rolling shutter is used. For Street View images, neither space nor time can be fixed onto a single point. Therefore, with the coexistence of spatial expansion and temporal extension, in this section I will examine the narrative potential in Street View images. Robin Hewlett and Ben Kinsley's *Street with a View* (2008) is one such case, in which staged performance is blended in with an otherwise nondescript residential community. In collaboration with local residents from Pittsburgh's Northside, the two artists staged a number of scenes along Sampsonia Way, “ranging from a parade and a marathon, to a garage band practice, a seventeenth century sword fight, a heroic rescue and much more.”³⁶ The scenes were then captured by a Street View car that was, based on a collaboration with the two artists, driving around the neighborhood. In the end, the project becomes part of the image archive, and now it only can be accessed through the Street View program (Figure 1.10).

³⁵ Duve, 1978, 119, emphasis in original.

³⁶ Kinsley, 2008.



Figure 1.10 Robin Hewlett and Ben Kinsley, *Street with a View*, 2008, installation shot.

Nonetheless, the representation of staged performance does not automatically qualify the project as narrative. In his 1966 essay “Notes Toward a Phenomenology of the Narrative,” Metz defines the narrative as a “closed discourse that proceeds by unrealizing a temporal sequence of events.”³⁷ Before reaching this definition at the very end of the essay, Metz has meticulously examined its five structural elements: *enclosure*, *temporal sequence*, *discourse*, *unrealization*, and *events*. Now, a similar analytical process can help us better understand Hewlett and Kinsley’s participatory project in terms of its narrative potential.

The basic requirement of narrative *enclosure*, according to Metz, entails a beginning as well as an ending. The former can be easily recognized in the image where two participants (Figure 1.11a) are standing at the crossroad, not only guiding the vehicle to the performing crowds, but also beckoning to the viewer, as if saying “Come! The story begins here!” As the performance continues, however, the viewer is eventually left wondering where the show ends. No one is dressed in uniform suggesting an end, so the viewer proceeds slowly and carefully lest they miss anything curious out of the corner, until the previously overcast day morphs into bright sunlight all of a sudden between 243 and 237 Sampsonia Way (Figures 1.11b and 1.11c). Most of the Street View images are captured in a bright sunny day, so the unusual gray day, along with scattered rain drops left on the camera lens, already contributes to an inadvertent touch that adds to the narrative milieu at the very

³⁷ Metz, 1991, 28.

beginning. One can only speculate that the pre-arranged date of recording happened to be under less desirable weather conditions, but whatever the case, the sudden change of weather certainly marks the clear beginning and ending of this performative sequence.

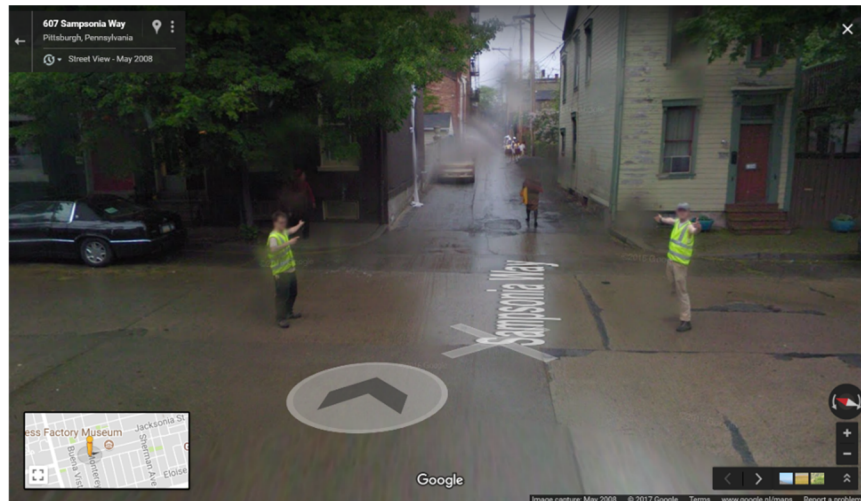


Figure 1.11a Screenshot from Google Street View, 607 Sampsonia Way, Pittsburgh, retrieved June 2017.



Figure 1.11b Screenshot from Google Street View, 243 Sampsonia Way, Pittsburgh, retrieved June 2017.



Figure 1.11c Screenshot from Google Street View, 237 Sampsonia Way, Pittsburgh, retrieved June 2017.

With the construction of an invisible timeline in the previous section, the second structural element of a narrative, *temporal sequence*, is now fairly self-evident. As is noted by Metz, one of the basic functions of narrative is “to invent one time scheme in terms of another time scheme.”³⁸ But in *Street with a View*, one could argue that the time schemes of the moving vehicle (production), the virtual spectator (perception), and the temporal axis (image) are all intricately interwoven. A temporal sequence can be found in all three time schemes, but one major difference is the irreversibility of production time and image time. Perceptual time, in comparison, is more flexible and constantly “scrambled” by the viewer, as they move, stop, continue again before turning around and going back, despite road signs that remind them that Sampsonia is a one-way street. The relationship between different time schemes also leads to Metz’s distinction among narrative, description, and image, which he illustrates with the following example:

A motionless and isolated shot of a stretch of desert is an image (space-significate-space-signifier); several partial and successive shots of this desert waste make up a description (space-significate-time-signifier); several successive shots of a caravan moving across the desert constitute a narrative (time-significate-time-signifier).³⁹

³⁸ Metz, 1991, 18.

³⁹ Metz, 1991, 18.

If we replace desert with street and caravan with, for example, marching band, this short example by Metz then becomes perfectly pertinent to the analysis of *Street with a View*, but some caution is needed here. Firstly, Metz's definition of an image (space-significate-space-signifier) expectedly follows the Renaissance convention in which "pictorial space gradually lost its temporal resonance."⁴⁰ In other words, this is a relatively confined definition where the temporal structure of an image is not only instantaneous but also simultaneous, or, simply put, pin-pointed. Secondly, the distinction in the construction of a cinematic narrative as opposed to one within Street View will largely depend on how the word "successive" is interpreted. Strictly speaking, a series of successive shots entails the presence of one discrete shot, framed and frozen, to be succeeded by the next. So, unless succession is understood as a generalized continuation of certain process, it will always betray its inter-frame nature that deprives Street View images of any potential for narrative unfolding. In any case, it is important to remember that the temporal sequence within Street View is based on the notion of intra-frame temporality and an image stitching process that weaves the individual segments into an integrated timeline.

The question of *discourse* is less complicated. Being essentially a statement, a discourse necessitates some sort of subject, a narrator behind the texts, visual or literary. Obviously, we know that this is the project by two artists, but in terms of a "narrative process," as Metz puts it, we can also argue that the images perceived by the viewer are both selected and arranged, like "an album of predetermined pictures." In a sense, and again this is comparable to a cinematic instance, the sequence itself becomes the first and foremost "grand image-maker."⁴¹

Finally, the element of *events* can be understood along with its inevitable *unrealization*. The parade, the marathon, and the sword fight are all events that have been ordered into a temporal sequence, but what does unrealisation entail? And, why is the unrealisation of events inevitable as they become part of a narrative? To address these questions, we can start with the question of what constitutes reality, on which Metz comments:

⁴⁰ Andrews, 1998, 105.

⁴¹ Metz, 1991, 20-21.

Reality assumes *presence*, which has a privileged position along two parameters, space and time; only the *here* and *now* are completely real. By its very existence, the narrative suppresses the *now* (accounts of current life) or the *here* (live television coverage), and most frequently the two together (newsreels, historical accounts, etc.).⁴²

In other words, for Metz it is the coexistence of *here* and *now* that constitutes reality, which suggests that the suppression of at least one of the two elements would result in an event's unrealization. In *Street with a View*, we may argue that the parameter of *now* is indeed missing. First of all, images in Street View were all collected sometime in the past, negating the possibility of a *now*. Then, not only did the project take place further back in time, but more importantly, every time the viewer decides to (re)visit the place, they will always have to first access archive images in Street View and set the destination for May 2008. Just as the events are buried in real-time history, so too are the records of events concealed in the image archive. The required click on the clock icon to go back in time thus turns itself into a symbolic gesture that validates the image's status as what we might call *not now*.

On the other hand, the suppression of *here* is more complicated, not least because our cultural understanding of *presence* has constantly been redefined by the advancement of media technology. For Bazin in the 1950s, the charm of family albums derives from a sense of presence, something that is so intrinsic to photography that, after its invention in the 1830s, plastic arts are finally freed from the obsession with realism.⁴³ In other words, photography's ability to depict reality with a higher level of fidelity has rendered obsolete the sense of proximity established by the painter between the viewed and the viewer. The same could be argued about cinema, when its invention in the 1890s largely displaced the role of painted panorama, which delivered an illusory spatial totality that once fascinated the nineteenth century. Bazin pushes this logic of displacement one step further, declaring that "cinema has not yet been invented," since every technological development will necessarily lead the medium closer to its origin, an origin that is "in complete imitation of nature."⁴⁴ That original, yet-to-be-invented cinema may or may not ever arrive, but along

⁴² Metz, 1991, 22, emphasis in original.

⁴³ Bazin, 2005, 14.

⁴⁴ Bazin, 2005, 21.

the path approaching its possible realization, new media technology – at the moment it seems to be virtual reality – will always play a key part in reshaping what it means for us to be in the presence of an event. Therefore, *here* should first and foremost be recognized as dynamic concept subject to changing cultural influences.

Accordingly, the viewing condition of Street View is also constantly evolving. From the computer screen, to the mobile phone screen, and now to the head-mounted virtual reality display, an unparalleled sense of immersion is leading us one step closer to the creation of a total illusion, and in so doing expanding the notion of *presence* and, by extension, *here*. Perhaps no one will actually confuse virtual presence with physically being there, but the fact that we may involuntarily move our physical body to, for instance, dodge an incoming object in the virtual environment already indicates the extent to which the human body may perceive and respond to the virtuality as the physical. The reaction itself is anything but new, as can be seen in film viewers or video game players, but the fidelity of simulation achieved with virtual reality technology would make a convincing case for an expanded definition of *presence*: one perceives the virtual environment as if physically being there. Of course, at present Street View is rather rudimentary in this respect, since the bodily sensorium is reduced to a mere visuality. Nonetheless, the capability of virtual presence in Street View is something that we could term *almost here*.

The combination of *not now* and *almost here* in Street View can be further understood through a comparison with cinema. According to media theorist Pepita Hesselberth, the *presence-effect* in cinema can be described as an intensified experience of *here*, *now*, and *me*.⁴⁵ As we have discussed earlier, the *now* in *Street with a View* is suppressed and restructured into a *not now*. While the same process might equally apply to the cinematic *now*, the cohesion of cinematic narrative and the use of continuity editing have by and large contributed to a suspension of disbelief (denial of *not now*) on the part of the viewer. For Hewlett and Kinsley, however, the narrative elements are both anachronistic and amorphous, and the pace of narrative is beyond their control and rests instead on the choice of the viewer. Furthermore, cinema has the potential to anthropomorphize the camera view, creating a personified, albeit sometimes unidentified, entity whose diegetic

⁴⁵ Hesselberth, 2012, 244.

position can be assumed by the viewer so as to elicit a sense of *me*. In contrast, the camera view in *Street with a View* fails to be anthropomorphized and instead remains that of the apparatus: the perhaps unintended rain drops are a constant reminder; the crowds need to make way for the vehicle to pass through; and the by-standers are returning their gaze towards the vehicle, only to find later that their faces have been blurred. Therefore, despite the immersive nature of Street View images that provides the viewer with an emphatic sense of *not now* and *almost here*, in Hewlett and Kinsley's project the invisible yet clearly perceptible camera view results in a certain narrative distance between the viewer and the scene that precludes an intensified experience of *me*.

With the construction of *not now* and *almost here*, the process of unrealization, the last structural element in Metz's definition of a narrative, is now complete. We can therefore confirm that Hewlett and Kinsley's work has managed to harness the inherent temporality within Street View images to present this bizarre narrative sequence. With the pin-pointed space of photography morphing into a more continuous spatiotemporal sequence, the viewer that has been left momentarily aphasic is eventually reinstated their capability of language, the capability to read the story unfolding when navigating inside the virtual space of Street View.

This observation also concludes the chapter, since we have come to realize that the presentation of time in Street View, apart from the various types of movement involved in the program itself, can also be amplified with a properly premeditated artistic intervention. Finally, it is important for us to realize that *Street with a View* is by no means the only case in which artists construct a narrative with Street View images. Many artists have taken screenshots from Street View and used these collected images to build their own narratives (consider ongoing projects such as Michael Wolf's *A Series of Unfortunate Events* and Jon Rafman's *Nine Eyes*). Still, what remains special about *Street with a View* is the fact that artistic intervention takes place in the *pre-production* stage. That is, choices made by the two artists will have a lasting effect on appearance of Street View (archive) images, whereas the practice of collecting screenshots (to be discussed in Chapter Three) should be regarded as a *post-production* intervention in which Street View images constitute a reservoir of found archive.

Chapter Two The Technologized Virtual (Museum) Space

In the previous chapter, I have attempted to establish the temporal scheme and how it reveals the inherent narrative potential within Street View. Although the depiction of time inevitably runs in conjunction with the spatial construction, the space itself has hitherto been regarded as a preexisting backdrop, whose precise parameters are yet to be examined. Also, despite a brief discussion about the necessity of user input, the nature of interaction between the viewer and the virtual space still requires an in-depth inquiry. In addition, the fact that Street View does not necessarily present views of the street suggests its versatility as a representational platform that has increasingly been used to showcase interior space.

Based on the above considerations, this chapter departs from the temporality of Street View and takes a spatial turn, with special attention paid to those spatial aspects that may have considerable spectatorial implications. What are the spatial parameters that define the relationship between the viewer and the virtual space? And, how can we account for the rising popularity of the representation of museum space in the Street View configuration? To address these questions, I will begin with notions of the “mobilized virtual gaze” and the “panoramic perception” as a gateway to understanding Street View space as one of navigation. Then, using Google’s Arts & Culture project as a case study, I will examine the construction of the virtual museum space and demonstrate how this new representational form has been rendered as a new media interface. Finally, I will extend the discussion to some museological, art historical, and mass cultural developments in the twentieth century that may account for the growing popularity of this emerging virtual (museum) space.

From the Mobilized Virtual Gaze to the Panoramic Perception

For all the contemporary technologies involved in Street View, the construction of a virtual space is anything but new. Panoramic rotundas popularized in nineteenth century Europe already constitute an early instance of virtual space in which the visitor could perambulate. Looking into the twenty-first century, we can think of Microsoft Photosynth as one rival technology of Street View that also builds a virtual space based on digital photographs of

the physical world.⁴⁶ Whether as a corporeal entity or through a sense of virtual presence, the viewer of a virtual space is always endowed with the ability to move around, but, as new media theorist Lev Manovich reminds us, we should “take into account the new way in which space functions in computer culture: as something traversed by a subject, as a *trajectory* rather than an *area*.”⁴⁷ In other words, computer-generated virtual space ceases to be an equally accessible physical expansion, and instead becomes one of navigation, as instructed by signs and functions provided in the program’s interface.

The immediate predecessor of Street View, often acknowledged as the first publicly exhibited virtual navigable space, is *Aspen Movie Map* (1978-1981) designed by the Architecture Machine Group at MIT, which is essentially a simulation of real-life driving experience through the city of Aspen, Colorado.⁴⁸ Using a joystick and a touchscreen, the viewer can choose a direction at each intersection, and the “movie” proceeds frame by frame unless they push the stop button (Figure 2.1). Unlike most simulators, *Aspen Movie Map* is constructed entirely from photographic images. Yet, these images are not stitched together as in Street View, and instead they follow the cinematic logic of single-frame replacement, hence the name “movie map.” “Each shot was logged in a computer database. When the eventual user wanted to ‘move’ a specific direction, the computer would call up the appropriate shot from the laserdisc player.”⁴⁹



Figure 2.1 Computer History Museum, *Aspen Interactive Movie Map*, 1978-1981, screenshot from video.

⁴⁶ Photosynth was released to the public in 2008, but was shut down by Microsoft in February 2017.

⁴⁷ Manovich, 2001, 279, emphasis added.

⁴⁸ For a detailed recount of the project, see Naimark, 2006.

⁴⁹ Weber, 2012.

Indeed, the many similarities between *Aspen Movie Map* and Street View attest to their common nature as being a virtual navigable space modeled on an existing physical world. Furthermore, the reconceptualization of Street View space as one of navigation is also consistent with the temporal scheme laid out in the previous chapter, since the trajectories followed by the viewer are precisely the routes along which the image-collecting vehicles have travelled. The position of the viewer is therefore confined to the blue/yellow lines or the cross signs superimposed onto the photographic images, while the vast expansion of the off-trajectory “area” remains beyond their reach. This navigational logic then draws an analogy between the space of Street View and that of a typical first-person shooter game. As a matter of fact, an Amsterdam-based advertising agency actually turned the program interface into a shooting game, dubbed *Google Shoot View*.⁵⁰ The gamespace analogy also extends to the fact that, in both Street View and most shooting games, a first-person point of view is always accompanied by a two-dimensional map usually at the lower corner, indicating one’s position within a larger cartographic grid.

As to the specific modes of navigation, Photosynth provides a useful model that we can apply to the Street View space. Before Photosynth users upload images to create their customized “synth” (a stand-alone navigable space), they need to choose from one of the four modes: spin, walk, wall, and panorama (Figure 2.2). Based on this classification, it is obvious that Street View navigation is the combination of “walk” and “panorama.” If the act of walking can represent human mobility in the most primal sense, and panoramic structures symbolize one of the earliest human attempts to create a virtual reality, then the unification of mobility and virtuality, the resulting *virtual mobility*, necessarily invokes what film historian Anne Friedberg calls the “mobilized virtual gaze.” According to Friedberg, the *virtual gaze* is “a *received* perception mediated through representation,” while the *mobilized gaze* has been an inherent part of all “cultural activities that involve walking and travel.” The compound term is introduced “to describe a gaze that travels in an imaginary *flânerie* through an imaginary elsewhere and an imaginary elsewhere.” For

⁵⁰ The image of an assault rifle, along with a simulated telescopic sight, was superimposed onto the Street View interface. Although the images did not respond in any way to the shooting, Google still decided to block the agency’s access to the Street View API (application program interface) only four days after the game’s release in December 2011. See also Albanesius, 2011.

Friedberg, the effect of virtual visibility was produced most dramatically with the invention of photography, and the mobilized virtual gaze logically culminated with the popularity of cinema.⁵¹

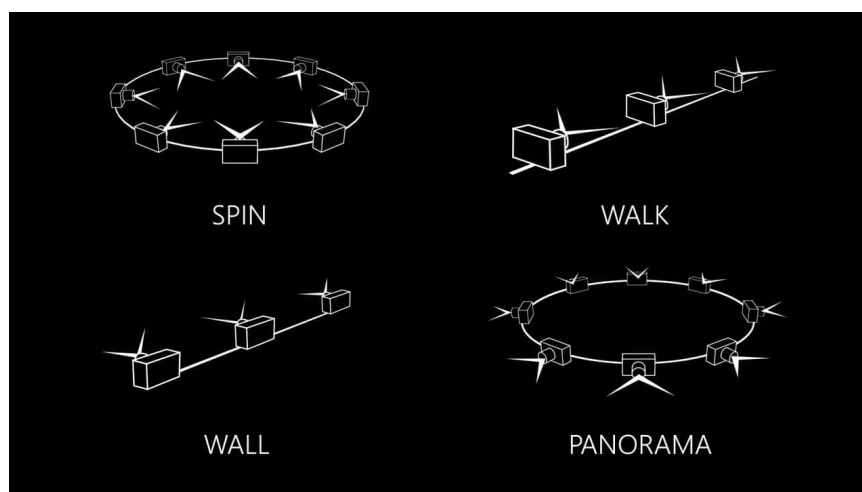


Figure 2.2 Four navigational modes in Photosynth, 2013, screenshot from video.

Already we can establish an association between Friedberg’s “imaginary elsewhere” and “imaginary elsewhere” with our *almost here* and *not now* outlined in the previous chapter. Indeed, both theoretical schemes aim to depict a sense of virtual presence provided by various visual representations. Although Friedberg, writing in 1994, mainly focuses on cinema and television and stops short of extending the genealogy of the mobilized virtual gaze to more contemporary visual technologies,⁵² her theoretical framework that compares and distinguishes cinematic and televisual spectatorship remains relevant to our discussion of the Street View space. Friedberg’s comparative scheme is summarized in Table 2.1.

It can be expected that many cinematic spectatorial principles that have already been challenged by televisual viewing will remain contested in Street View: the screen, whether on a computer or a mobile device, indeed functions only as light source; repeated viewings are possible as long as an internet connection is available, but it is very likely that repeated viewings may not necessarily result in the same repeated view due to database updates (the

⁵¹ Friedberg, 1994, 2-3.

⁵² Friedberg devotes a short section to virtual reality, which she argues would further challenge the principles of cinematic spectatorship by promoting participatory and interactive users, but she approaches the medium as “an almost contentless means of communication, looking for a marketable purpose” and therefore does not delve into its spectatorial implications. See Friedberg, 1994, 143-147.

fact that *Street with a View* is buried in the image archive is a case in point); screen size becomes even smaller compared to that of an average household television, and in the case of head-mounted display, the view appears “larger than life” only because the device is placed so close to the eyes; this viewing mode also break with the two-dimensional tradition, since side-by-side stereoscopic images can already achieve a convincing degree of three-dimensional illusion. The above is a quick list of how Street View viewing follows or disobeys some of the basic spectatorial principles outlined by Friedberg. Deserving special attention are the two aspects in Table 2.1 concerning the viewer’s mobility and interactivity, or lack thereof, not least because they are the direct results of Street View being virtual navigable space.

Cinematic Spectatorship	Televisual Spectatorship
<i>dark room with projected luminous images</i>	<i>TV as light source rather than projection</i>
<i>immobility of spectator</i>	<i>modicum of mobility</i>
<i>single viewing</i>	<i>reruns, time-shifting functions, rentals</i>
<i>noninteractive relation between viewer and image</i>	<i>channel alternatives</i>
<i>“larger than life” framed image</i>	<i>home-sized image scale</i>
<i>two-dimensional screen surface</i>	<i>two-dimensionality (except for 3D TV)</i>

Table 2.1 Principles of cinematic vs. televisual spectatorship.

In an attempt to establish the relationship between the screen and the body of the viewer in what he calls “screen-based representational apparatus,” Manovich reviews the historical conditions of Alberti’s window, Dürer’s perspectival machines, camera obscura, photography, and ultimately cinema, in which the subjects all need to remain immobile in order to see the images correctly, if not at all.⁵³ If the physical imprisonment of the body is the cost of virtual mobility, then Friedberg observes a more nuanced tradeoff between mobility and virtuality:

⁵³ Manovich, 2001, 103-109.

as the “mobility” of the gaze became more “virtual” – as techniques were developed to paint (and then to photograph) realistic images, as mobility was implied by changes in lighting (and then cinematography) – the observer became more immobile, passive, ready to receive the constructions of a virtual reality placed in front of his or her unmoving body.⁵⁴

In contrast, predicated on the assumption of continuous user command, navigable space is interactive by definition, and interaction itself consists of an action/reaction feedback loop between a human user and a computer program. The action on the part of the user then materializes through its movement, which can be as “micro” as using a remote control, clicking a button, and touching a screen, or as “macro” as actual bodily displacement in the physical space in order to initiate a corresponding movement in the virtual world. Such a distinction is important because the *micro-movement* seems to resemble what Friedberg means by “modicum of mobility” in televisual spectatorship, but what is being enacted by such moderate mobility is actually a “distracted gaze” in which such activities as “ironing, laundry, and childcare become rhythmic components of viewing.”⁵⁵ Televisual spectatorship thus constitutes the *possibility* of mobility. But in a navigable space like Street View, user mobility becomes a *necessity*, if not prerequisite, of the viewing experience. It is precisely when user movement is interrupted that the gaze becomes distracted. Based on this distinction, we could also argue that *Aspen Movie Map* has a lower degree of interactivity (dependence on user input) in that without any user command, the view proceeds uninterrupted and literally becomes a “road movie,” whereas the dynamic Street View space would “freeze” into a photographic snapshot.

The necessity of viewer mobility therefore becomes what fundamentally distinguishes the traditional spectatorship and that of more recent interactive mediums, despite their common role played in the gradual centrality of the mobilized virtual gaze. For Friedberg, the nineteenth century *flâneur* is the ardent champion of the mobilized gaze, whose experience of the world was constantly virtualized by arcades, panoramas, and dioramas. The same can be said about a user navigating through the virtual space, a *virtual flâneur*,

⁵⁴ Friedberg, 1994, 28.

⁵⁵ Friedberg, 1994, 136.

who “is happiest *on the move*, clicking from one object to another, traversing room after room, level after level, data volume after data volume.”⁵⁶

Still, one thing of note here is the role played by virtual reality, which again complicates our conclusion drawn above. Manovich mentions how the earliest virtual reality systems continued the tradition of bodily imprisonment and sometimes reduced the human body to a giant joystick,⁵⁷ but such imprisonment is transmuting into a form of *attachment*, thanks to developments made in device miniaturization. Such wearable devices as head-mounted display or retina display no longer imprison the viewer like a painter’s camera obscura or a movie-goer’s seat. Rather, they are attached to the human body with varying degrees of invasiveness,⁵⁸ and it is this decreasing level of invasiveness that allows for increased user mobility. Manovich’s optimistic claim that virtual reality “probably represents the last act in the long history of the body’s imprisonment”⁵⁹ is not technically wrong, but I hope to nonetheless emphasize a persistent continuation rather than a radical termination. Imprisonment may have morphed into attachment, but it would be naively mistaken to therefore equate mobility with freedom, or to assume that the viewer has finally ceased to hold the passive role, because the previous impossibility of movement and the present necessity to move are both manifestations of the mandate of a particular representational apparatus. That we are not allowed to move or talk in the movie theater and that we have to keep moving when geared up in virtual reality attest to the very same condition in which a self-mutating power regime operates: we may not be locked in any more, for now we carry the prison with us.

The relevance of this power-based interpretation will resurface in the following section when we delve into the institutional role of the museum, but for now it suffices to say that the advancement of technology has always been a crucial factor in shaping our perception of the physical as well as the virtual world. If the discussion so far concerning an extended mobilized virtual gaze is an indication of the interactive nature of navigable space, I want

⁵⁶ Manovich, 2001, 274, emphasis added.

⁵⁷ Manovich, 2001, 110.

⁵⁸ Carrozzino and Bergamasco have classified virtual reality systems based on the level of immersion, which is inversely related to the level of invasiveness. External devices that are not attached to the body thus possess a higher level of immersion compared to wearable devices. See Carrozzino and Bergamasco, 2010, 454.

⁵⁹ Manovich, 2001, 114.

to highlight a parallel technological development of the nineteenth century that, albeit unwittingly, revolutionized human perception in a no less significant manner. Just as “machines of virtual transport ... extended the virtual gaze of photography to provide virtual mobility,”⁶⁰ so did the growing popularity of railroad travels produce what cultural historian Wolfgang Schivelbusch terms “panoramic perception” by generating a new experience of distance and time.

In his account of the social transformation effected by the advent of railroad travel, Schivelbusch notes that the increased speed and “mathematical directness” fundamentally changed the previously close relationship between the traveler and the traveled space.⁶¹ Schivelbusch’s metaphor of the train being a projectile shot through the landscape effectively characterizes the traveled space as a navigable space. Not unlike a Street View navigator whose virtual position is limited to predetermined routes, a railway passenger possesses a vantage point that similarly confined within the carriage, which in turn runs only along preconstructed railroad tracks. Therefore, compared with *Aspen Movie Map*, the view from the train would become a fast-forward movie map without the pause button. In addition, if we use the four navigational modes presented in Figure 2.2, then the view in *Aspen Movie Map* can be said to follow the “walk” mode, whereas the railroad in effect provides the passenger with a view based on the “wall” mode.

The speed of railway journey, according to Schivelbusch, also transformed the traveled landscape itself, which becomes what phenomenologist Erwin Straus regards as “geological space,” a systematized structure in which every place is “determined by its position with respect to the whole and ultimately by its relation to the *null point of the coordinate system* by which this space obtains its order.”⁶² The accelerated journeys shortened the perceived distance between two places, and it is precisely this *perceived* distance that is, in a much more radical manner, being redefined within Street View. Virtual space, with its instant access, essentially invalidates the correlation between physical distance and time required to reach a particular location. In Street View, or any digital maps for that matter, the time it takes to jump from Paris to New York has nothing to do with how far the two cities are

⁶⁰ Friedberg, 1994, 4.

⁶¹ Schivelbusch, 2014, 53.

⁶² Straus, quoted in Schivelbusch, 2014, 53, emphasis added.

geographically separated. Rather, it depends exclusively upon the speed of database access, which is in turn determined by the speed of internet connection and the specifications of the device on which the program operates. In this sense, we can regard the database where images are stored as the de-spatialized “null point” of the software program by which the systemized virtual space obtains its order.

In the meanwhile, Schivelbusch also outlines a nineteenth century tendency whereby the velocity of railway journeys stimulated what he calls “panoramic perception,” “the ability to perceive the discrete, as it rolls past the window, indiscriminately.”⁶³ This new mode of perception results from the passenger’s gaze at successive scenes rolling past the window, and results in, as described by philosopher Dolf Sternberger, some sort of panoramization of the visual fields: “the views from European windows have lost their depth, becoming part and parcel of the same panorama world surrounding them and constituting a painted surface everywhere.”⁶⁴

The observation of an almost homogenized panoramic field is of significant relevance, because a similar principle can be said to govern the virtual navigable space as well. In his comments on navigable space in relation to modern painting, Manovich notes several modern painters who belong to what he calls the “space-medium” tradition, where space is understood “as a homogeneous, dense field, where everything is made from the same ‘stuff.’”⁶⁵ Following the tradition of space-medium, these painters rejected the dichotomy between a distinct object and an enveloping space, and instead depicted “a dense field that occasionally hardens into something which we can read as an object.”⁶⁶ Such a monistic ontology is equally applicable to Street View space, where the façade of a building is not exactly an independent structure separable from an invisible Cartesian coordinate system, but merely a dense field of pixels that we *read* as an architectural entity.⁶⁷ It is precisely these pixels that function as the “part and parcel” in Sternberger’s panoramization and

⁶³ Schivelbusch, 2014, 60-61.

⁶⁴ Sternberger and Neugroschel, 1977, 15.

⁶⁵ Manovich, 2001, 265.

⁶⁶ Manovich, 2001, 255.

⁶⁷ Technically, this argument holds true for photographic images in general, but in the flattened space of a photograph, such homogenization is consistent with the lack of interactivity and the fixity of perspective, whereas in navigable space, this monistic ontology stands in contrast to the multiplication of perspectives and the possibility of movement.

“the same stuff” in Manovich’s space-medium. The only difference is the respective origin of such homogenization, since the nineteenth century scenario was occasioned by the velocity of public transportation, whereas the twenty-first century instance was realized through a massive project of post-production image stitching.

If the cost of nineteenth century virtual mobility is the gradual imprisonment of the body, then panoramic perception is achieved at the expense of a close connection with the traveler and the immediate foreground. Not only are the passengers forever separated from the traveled space by the window glass, but they also fail to perceive the proximity as they do the distant backdrop due to the increased speed of the train.⁶⁸ As Schivelbusch notes:

[the preindustrial traveler] saw himself as part of the foreground, and that perception *joined* him to the landscape, included him in it, regardless of all further distant views that the landscape presented. Now velocity dissolves the foreground, and the traveler loses that aspect. He is removed from that “total space” which combines proximity and distance.⁶⁹

The disassociation of proximity and distance and the removal of the traveler from the total space are important considerations not because they likewise apply to Street View space, but because the exact opposite is true. Placing the viewer in a total space and offering them a sense of presence are precisely what immersive mediums strive for. Spherical panoramas in Street View enable the viewer to look around, making it possible to reestablish the link between foreground and background. And, in a virtual reality setting, the synchronization of head movement and perspectival changes strengthens the restored link even further. It is for the same reason that Schivelbusch interprets the rise of photography as an effort to resurrect “the intensive experience of the sensuous world, terminated by the industrial revolution,” since the loss of foreground in panoramic perception is compensated in the new medium with its enriching close-ups.⁷⁰ Thus, Street View can be seen as a continued

⁶⁸ For the same reason, some of the late nineteenth century moving panoramas modelled on railway carriages featured multiple layers of landscape that moved at different speeds, essentially creating a parallax effect for the viewer that later became a common practice in animation techniques and video game design.

⁶⁹ Schivelbusch, 2014, 63, emphasis in original.

⁷⁰ Schivelbusch, 2014, 63.

effort to resurrect, after photography and cinema, and further intensify our experience of the sensuous world, with an extra emphasis on the sense of immersion.

The parallel between nineteenth century railway travels and Street View navigation is evident in their dependence on machines of transportation traversing the physical space. Indeed, the “mobility of vision ... became a prerequisite for the ‘normality’ of panoramic vision.”⁷¹ Although railway passengers see through a transparent glass window, their gaze is nevertheless mediated and virtualized by the unprecedented speed they experience. Therefore, at the end of the section, we come to realize a possible convergence between the mobilized virtual gaze and the panoramic perception: in response to the first research question raised at the beginning of the chapter, we now realize that it is precisely the two notions that combine to constitute the core principle behind a user’s navigation through Street View space, defining not only the mode of perception but also the relationship between the viewer and the surrounding virtual environment.

Museum without Walls: Google Arts & Culture

In his book about television as a cultural form, cultural critic Raymond Williams regards communication technologies like radio and television as “systems primarily devised for transmission and reception as abstract processes, with little or no definition of preceding content.”⁷² Although it is difficult to draw the same conclusion for Street View, the separability between means of communication and any preceding content suggests that Street View may well be a slight misnomer and in fact belies its representational versatility. Indeed, from the project’s very inception, its developers have envisioned its application in the representation of both exterior and interior space. In the following two sections, I will focus on the use of Street View technology in the construction of a different type of navigable space: the virtual museum space.

Before we begin the inquiry, it is important to take a look at the organizational scheme behind the construction of such an interior virtual space. At the very heart of the scheme is Google Cultural Institute (Figure 2.3), a not-for-profit initiative founded in 2011 “that

⁷¹ Schivelbusch, 2014, 64.

⁷² Williams, 2003, 18.

partners with cultural organizations to bring the world’s cultural heritage online.”⁷³ For the “culturally curious” audience, Google Arts & Culture (GA&C), formerly named Google Art Project, is the digital platform that provides audience with this online access. The Institute’s digitization tools consist mainly of two parts: Art Camera and Museum View. The former is an image capturing device (Figure 2.4) that creates ultra-high-resolution photographic reproductions of the museum collection, while the latter uses a Street View trolley (Figure 2.5) that navigates through museum space, taking photographs in a manner (as described in the introduction) similar to its street counterpart. The Cultural Institute also provides a free collection management system and storytelling tools for participating partners. In addition, The Lab located in Paris is an independent department dedicated to the bridging of new technology and its application in visual culture.

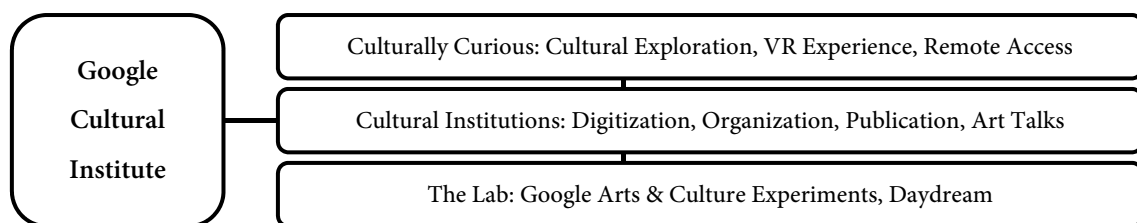


Figure 2.3 Google Cultural Institute and its major components.



Figure 2.4 The Art Camera; Figure 2.5 A Street View trolley.

Based on the structural overview, we can see that the presentation of a virtual museum space revolves around the process of digitization, whose two components, Art Camera and Museum View, both indicate the primacy of photographic images, be they single-frame

⁷³ Google Cultural Institute, 2015.

reproductions or stitched panoramas. In the webpage showing an individual reproduction, the Pegman icon can be found below the image that redirects the viewer to, if available, its corresponding Museum View. Similarly, a thumbnail image (or a superimposed circle) will appear in Museum View to inform the viewer of an available photographic reproduction (Figure 2.6a). The high-resolution reproductions allow users to zoom in and observe extremely minute details (Figure 2.6b) at a level that is not only unprecedented but also unconceivable in an actual museum setting, usually due to the limited dimension of the artworks, the required distance to be maintained between the viewer and the objects, and undesirable viewing conditions, such as direct lighting that causes reflections or crowded gallery space with no possible vantage point. Therefore, the advantage of viewing objects as paintings, manuscripts, and photographs in this configuration is fairly self-evident, but it also raises the question about viewing three-dimensional objects. Sculptures, for instance, are mostly photographed from one particular angle,⁷⁴ and currently the number of possible perspectives (the cross signs) offered in Museum View is rather limited. Consequently, an object that is otherwise placed at the center of the room and to be viewed from multiple angles may end up registering a very different façade in its digital incarnation.

But this is hardly surprising, since the inadequacy is a direct result of the limited modes of navigation introduced at the beginning of this chapter. Street View and Museum View both follow the “walk” and “panorama” modes, while the “spin” (inverse panorama) and “wall” also included in Photosynth are not yet available (Figure 2.2). While this may not be a serious problem for Street View, since streets constitute a grid that compartmentalizes the geological space, in Museum View the viewer is placed *inside* one compartment within a larger spatial grid (as depicted by the floor plan). Thus, the fact that most paintings and photographs are hanging on the wall while sculptures are usually placed at the center of the gallery demonstrates the very urgent need for a “spin” view and a “wall” view in Museum View. Technically, adding two extra navigational modes is not impossible. As former Photosynth group manager David Gedye suggests, to create a smooth “panorama” view or “spin” view, the recommended number of photographs taken is between twenty-five and

⁷⁴ While it is possible that multiple images have been taken from different perspectives, GA&C only selects one image for online publication, usually with a frontal angle that tends to flatten the three-dimensionality of the represented objects.

forty (one photograph taken every ten to fifteen degrees).⁷⁵ While this would require a much higher, to borrow a statistical term, sampling frequency than currently provided in Museum View, a navigational experience tailored for the virtual museum space is already technically achievable. Indeed, Google is making progress in this aspect, as can be seen in the latest version of Google Earth, where smoothly rotatable three-dimensional images are available for landmark sites. If, or perhaps when, the “spin” view becomes available in Museum View, experience of the virtual museum space would probably resemble that of an actual visit even more.

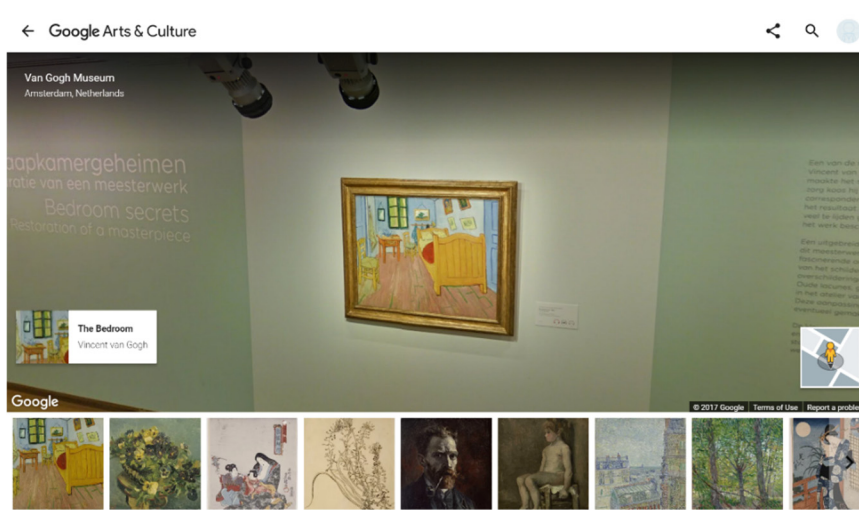


Figure 2.6a Museum View of Van Gogh Museum, screenshot from Arts & Culture, retrieved June 2017.

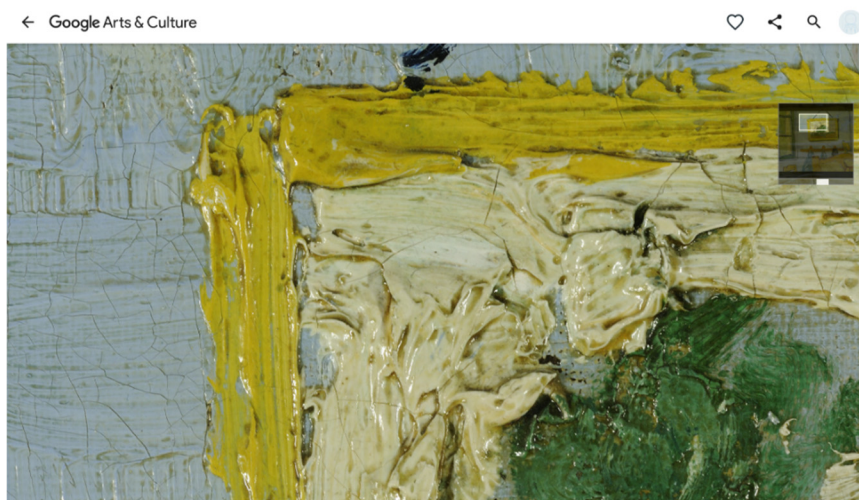


Figure 2.6b Detailed view of van Gogh's *The Bedroom*, 1888, screenshot from Arts & Culture, retrieved June 2017.

⁷⁵ Larsen, 2013. In the demonstration video, Gedye also suggests that the photographs be taken with a constant zoom level as well as a considerable amount of overlapping among each other.

However, the question we may need to ask is concerned precisely with this possible resemblance. Why does Google choose to digitize numerous museum collections so that an actual visit can be to some extent circumvented, while at the same time trying to reconstruct a virtual space that so closely resembles and so convincingly reminisces of its physical counterpart? In other words, how can we reconcile the two components of the digitization process that seem to operate under two contrasting sets of principles? One possible response to this paradox may come from the observation of museum visitors, whose passion for a personal encounter with the artworks is sometimes coupled with an obsession to in the meanwhile take a photograph of the objects, a virtual representation of the very thing that they seek to physically perceive. Still, a more systematic understanding of the paradox may derive from a new media approach, in which the dynamic between “database” and “interface” is playing a key part.

Manovich has outlined some of the basic principles regarding new media, which may shed some light on the variety of operations inside the Cultural Institute. The construction of a virtual museum space, for instance, operates in accordance with the principle of numerical representation, that not only can the images be described with a mathematical function, but they also become programmable and subject to algorithmic manipulations. The process of numerical representation, or digitization, is usually achieved through sampling and quantization, which convert continuous data into discrete units.⁷⁶ In the case of Museum View, a continuous walk through the museum space is converted into several discrete spherical panoramas, where the captured images are stored and transferred as numerical data, and later stitched together with an image processing algorithm. Sampling frequency, as described earlier, may affect how much Museum View resembles the physical environment, since the denser the sampling becomes, the more continuous the experience will turn out to be.

Perhaps of greater relevance is the principle of variability, which suggests that it is now “possible to separate the levels of ‘content’ (data) and interface. A number of different interfaces can be created to the same data.”⁷⁷ On the one hand, Art Camera can be seen as

⁷⁶ Manovich, 2001, 27.

⁷⁷ Manovich, 2001, 37.

the armed version of the museum registrar, cataloguing one data entry after another into the database. On the other hand, Museum View can be seen as one of the interfaces created from such a database. It just so happens that the organizing scheme of this particular interface, something that we may call *indexical interface*, is the very architectural layout of the interior space where sampling takes place. The logical question that follows would be whether there exist what we may call *unindexical interfaces*, through which we can access virtual spaces that are not modelled on any existing architectural interiors. They are, in other words, not indexically linked to any physical space.

The answer is a definite yes, and in fact the presentation of unindexical interfaces is a major part of GA&C,⁷⁸ where Featured Projects and Featured Stories are basically themed exhibitions. They are usually organized according to their institutional origins, but every now and then some participating institutions would collaborate on a larger project. Either way, these projects focus on the digitized exhibits themselves, with little or no regard for the original spatial configuration in which they are exhibited or stored. The exhibition interface usually includes a horizontal view (Figure 2.7a) of successive multimedia panels,⁷⁹ which can be said to follow the “wall” mode of navigation. Single spherical panoramas are sometimes included, but most of the panels feature texts and single-frame images only. An earlier version of the exhibition interface displayed thumbnail images with hyperlinks and a slider that functioned as an interactive virtual floor plan (Figure 2.7b), but since then it has been redesigned to embrace its current style.

While the act of navigation is clearly visible in both cases, the virtual space represented here is nonetheless flattened, two-dimensional, and almost with an eBook style.⁸⁰ In contrast, The Lab with its GA&C Experiments⁸¹ has created a different kind of unindexical interface that provides access to a more three-dimensional virtual space. In the Experiment named Curator Table, for example, thumbnail images of the entire database are laid out as though on an imaginary table (Figure 2.8a), presumably simulating the scenario I which

⁷⁸ The homepage of the GA&C website usually includes the following sections: Daily Digest, Featured Project, Featured Stories, Explore with Street View, Zoom In, Explore by Time and Color, Artists, and Partners. See <https://goo.gl/OhRNvZ>.

⁷⁹ Depending on the size of the screen, some mobile devices may use a vertical view as in a typical web page.

⁸⁰ Most of the featured projects and stories are labeled with an eBook icon. See Figure 2.7a.

⁸¹ GA&C Experiments currently include the following sections: Tags, X Degrees of Separation, Free Fall, Curator Table, and t-SNE Map. See <https://goo.gl/MJLUpK>.

curators work on the available museum collections. Zoom in on one section of the Table will reveal images grouped by key words (Figure 2.8b), and when the viewer scrolls the close-up view horizontally, images in the foreground will actually move faster than those in the background, thus adding to the visual effect of three-dimensionality.

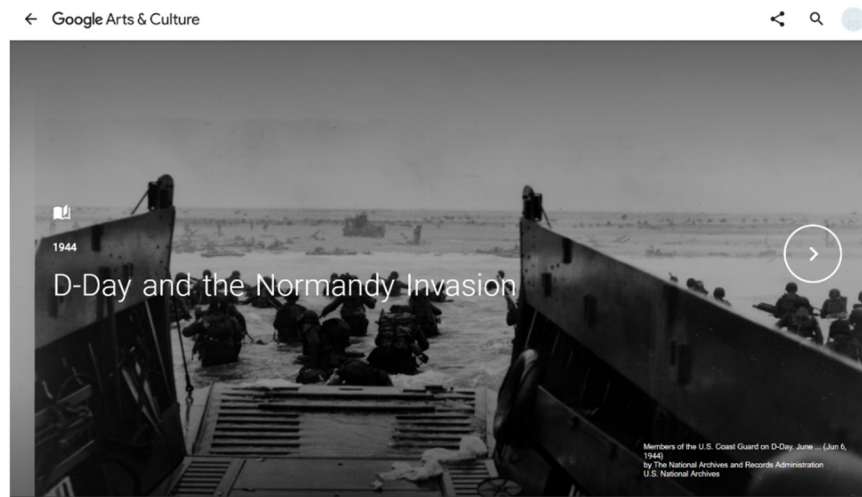


Figure 2.7a Cover panel of an online exhibition, screenshot from Arts & Culture, retrieved June 2017.

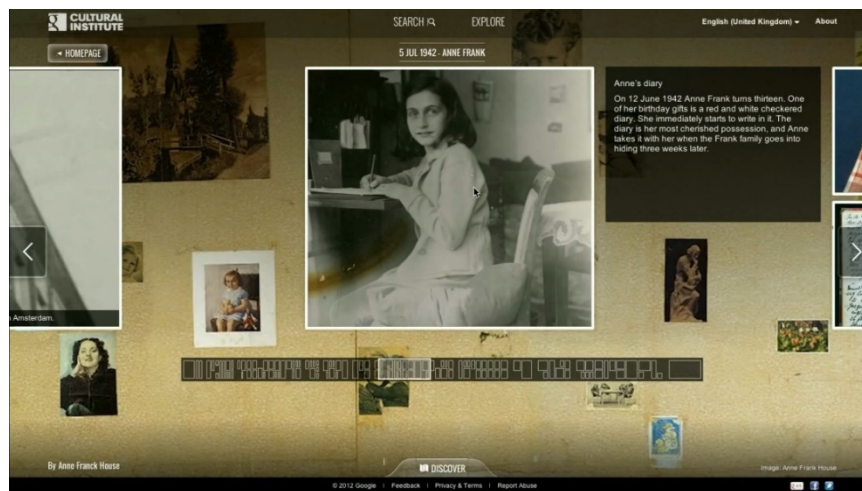


Figure 2.7b Exhibition interface from a previous version, screenshot from video, retrieved June 2017.

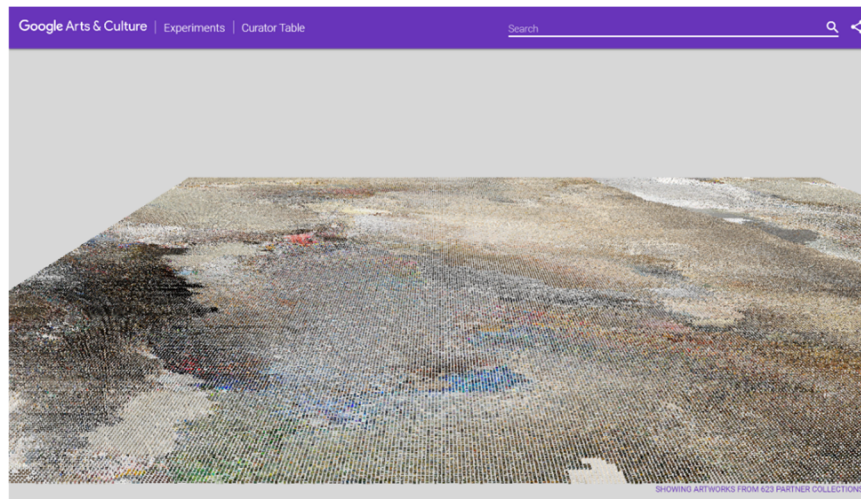


Figure 2.8a Curator Table, screenshot from Arts & Culture, retrieved June 2017.

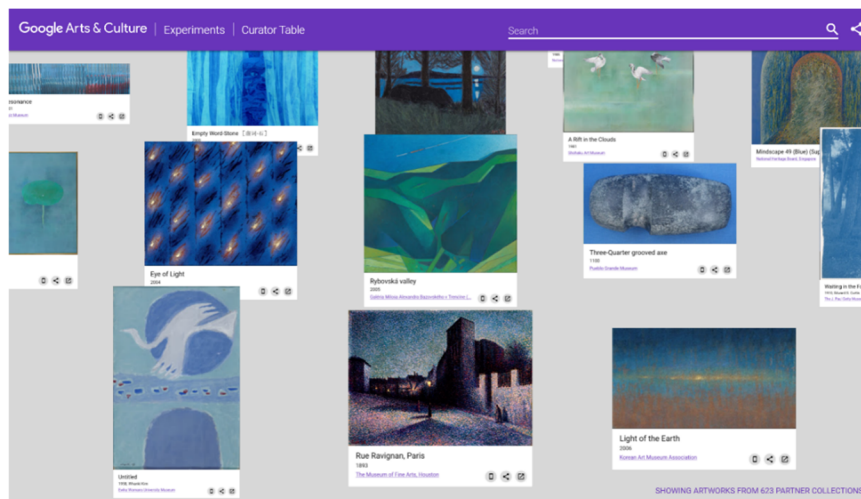


Figure 2.8b Curator Table, screenshot from Arts & Culture, retrieved June 2017.

Whether it is virtual space modelled on an actual museum building, or multimedia virtual interface with an eBook style, or imaginary three-dimensional virtual space with all the images available at the same time, these indexical/unindexical interfaces are a perfect demonstration of the variability principle, that multiple interfaces can be created from the same database. In all the above cases, “we are no longer interfacing to a computer but to culture encoded in digital form.” Manovich therefore uses the term “cultural interface” to describe “the human-computer-culture interface: the ways in which computers present and allow us to interact with cultural data.”⁸² The central role played by cultural interfaces

⁸² Manovich, 2001, 70.

in our interaction with cultural data also addresses the digitization paradox mentioned earlier: Art Camera contributes to the building of an ever-expanding database, whereas Museum View is simply one of the possible interfacial variations.

As I have tried to demonstrate so far, the new media approach and particularly the notion of cultural interface can help us establish a more methodical understanding of the Cultural Institute and GA&C, but it is critical that we caution against one potential misconception, that the human-computer-culture interface has somehow eliminated the process of human intervention or institutional framing, or that these technologized cultural interfaces simply serve as a transparent window through which cultural contents can be accessed with a touch of mythical neutrality that is somehow long overdue. Just as “the [museum] space between viewer and canvas is controlled, institutionalized, and policed as a special, real kind of space,”⁸³ so is the virtual space infused with, so to speak, the ideology of the white cube. Similarly, just as the cessation of bodily imprisonment in virtual reality does not necessarily result in the freedom on the part of the viewer, neither does the insertion of the computer in the human-culture relation prevent the institutional frame from continuing conferring upon the exhibits its own ideological connotations. If anything has changed, it is the messaging mechanism instead of the message itself. Masked in a digital form beneath the cultural interface, the mechanism now exhibits even more invisibility. This argument can also be illustrated with what Manovich refers to as “non-transparency of the code,” which postulates that computer interface, as a code carrying cultural contents, necessarily brings its own intended messages.⁸⁴

What, then, is the message that remains unchanged? Before we answer the question, it is important that we realize the digitization of museum collections started long before the establishment of Google Cultural Institute: individual museums have been promoting their institutions and collections with dedicated websites since the early 1990s; the same decade also saw the proliferation of CD-ROM-based virtual museums, but these early instances “rarely attempted to simulate in 3D the physical space of the museum.”⁸⁵ In new media terms, they created an unindexical interface from a relatively small cultural database.

⁸³ Bolter and Grusin, 2000, 59.

⁸⁴ Manovich, 2001, 64-65.

⁸⁵ Huhtamo, 2010, 122.

Further back in time, we can even argue that the conception of virtual museum based on photographic reproductions in fact predated the digital era. Among the pioneer scholars who have attempted to understand the implications of photographic reproductions on the museum institution is art historian André Malraux, who observes:

[We] have far more great works available to refresh our memories than those which even the greatest of museums could bring together. For a “Museum without Walls” is coming into being, and (now that the plastic arts have invented their own printing-press) it will carry infinitely farther that revelation of the world of art, limited perforce, which the “real” museums offer us within their walls.⁸⁶

The “museum without walls” envisioned by Malraux is an idealized, imaginary museum of art that we carry around in our own minds. If the sheer quantity of artworks was the main reason we needed such an imaginary museum, then photography, the then newly invented “printing-press” of plastic arts, was that which made it possible. Malraux’s museum thus suggests the instrumental primacy of photographic images, but we should not simply take it at face value and equate this notion with a randomly selected collection of photographic reproductions. Rather, in Malraux’s conceived museum without walls,

picture, fresco, miniature and stained-glass window seem of one and the same family. For all alike ... have become “colorplates.” In the process they have lost their properties as *objects*; but, by the same token, they have gained something: the utmost significance as to *style* that they can possibly acquire.”⁸⁷

If this argument sounds all too familiar, that is because it involves yet another instance of homogenization that is in line with Sternberger’s panoramization and Manovich’s space-medium. As in the indiscriminate train window and the homogenous dense field, a “rather *specious unity* [is] imposed by photographic reproduction on a multiplicity of objects,”⁸⁸

⁸⁶ Malraux, 1974, 16.

⁸⁷ Malraux, 1974, 44, emphasis in original.

⁸⁸ Malraux, 1974, 46, emphasis added.

whose differences in the original properties are now reduced to distinctions in mere styles. However, it is precisely this “specious unity” that allows us to juxtapose works of art and compare stylistic elements that used to be incomparable. From a disciplinary perspective, “[f]or the last hundred years ... art history has been the history of that which can be photographed.”⁸⁹ It was thus from the attempts to organize photographic reproductions that the discipline of art history forged and preserved its identity. Yet, this attempt to juxtapose and compare is, according to Malraux, a modern phenomenon that remained unconceivable before the establishment of modern art museum. In other words, “the art museum functions as facilitator of a dialogue ... that lies at the heart of our modern response to art,”⁹⁰ and Malraux’s museum without walls can therefore be regarded as an imaginary extension of the same dialogue initiated within the museum’s institutional frame and governed by the self-perpetuating principles of art history.

With another Experiment called X Degrees of Separation, the Cultural Institute pushes this logic of image juxtaposition one step further. In this experiment, a user can select two images from the database, and a matching algorithm will search for a connecting pathway, a chain of artworks, between the two images based entirely on visual feature analyses. The resulting associative sequence (Figure 2.9) is reminiscent of art historian Aby Warburg’s last project, *Mnemosyne Atlas*, an unfinished attempt to map the “afterlife of antiquity.” Warburg himself created a series of thematically organized black panels (Figure 2.10) on which he juxtaposed photographic reproductions of artworks and artefacts, which he saw as “images of great symbolic, intellectual, and emotional power [that] emerge in Western antiquity and then reappear and are reanimated in the art and cosmology of later times and places.”⁹¹ Compared to Warburg’s black panels, X Degrees of Separation clearly lacks the metonymic, intuitive logic and the overarching art historical ambition, and relies only on the formal elements automatically detected by the algorithm, but this is in fact closer to Malraux’s vision, since in his imaginary museum objects as “colorplates” are different only in terms of their nuances in style.

⁸⁹ Malraux, 1974, 30.

⁹⁰ Allan, 2010.

⁹¹ Johnson, 2013. Researchers from The Warburg Institute and Cornell University Library have discerned nine major thematic sequences from the otherwise “non-discursive” and “frequently digressive” *Atlas*. See <https://goo.gl/9AYH9m>.

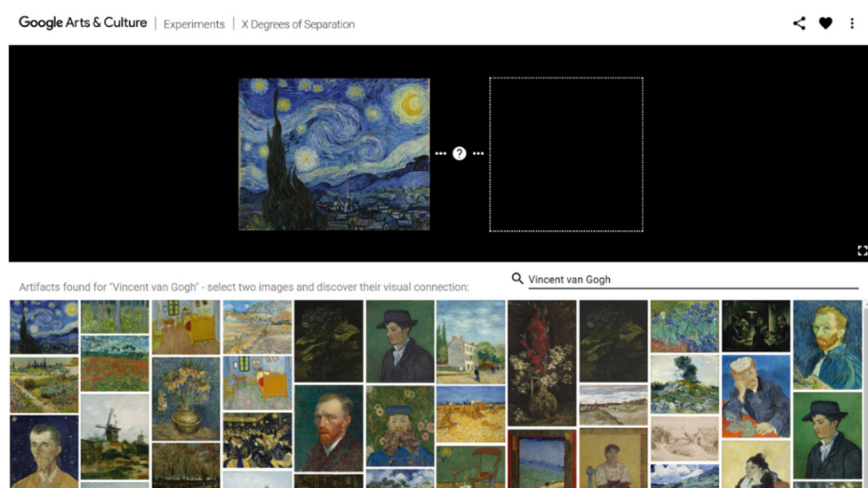


Figure 2.9a X Degrees of Separation, selection view, screenshot from Arts & Culture, retrieved June 2017.

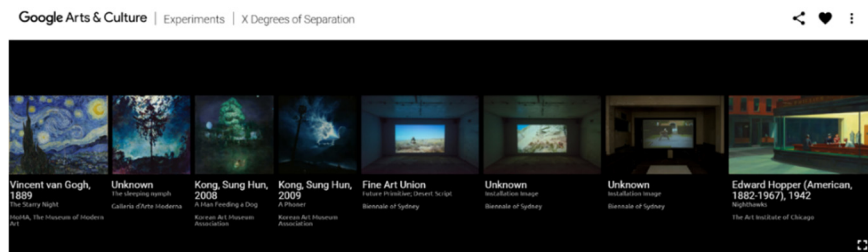


Figure 2.9b X Degrees of Separation, result view, screenshot from Arts & Culture, retrieved June 2017.



Figure 2.10 Panel 32 from Warburg's *Mnemosyne Atlas*.

If, as can be seen from Malraux and Warburg, we can access and even to some extent “activate” a particular strand of art history through thematically organized photographic reproductions, then cultural interfaces provided by the Cultural Institute can be regarded as a more technologized descendant of Malraux’s imaginary museum and Warburg’s

memory atlas. Despite their differences, all three address the issue of access to our cultural history with the help of photographic reproductions and in so doing inevitably reinforce the regulating narrative and organizational scheme that we now know as art history. This, to answer our previous question, is the message that remains unchanged: the white cube, the black panels, or the cultural interfaces, they are all manifestations of an abstract process that converts a social history of cultural objects and practices into an idealized history of art.

Based on this historical continuum, we can now conclude with much certainty that the virtual museum is not a preordained substitute for its moribund physical counterpart, but instead operates under the same reigning force that not only takes advantage of but also passes its torch onto a more technologized representational platform. Museum View, that navigable virtual museum space that brings us into this discussion in the first place, can be thought of as the product of an ongoing power struggle between the lingering dominance of traditional museum institutions and the emerging ascendancy of multimedia platforms like GA&C. As long as the former is preferred to the latter by the majority of the culturally curious, these platforms will have to continue perfecting their representational dexterity in the name of an “unmediated” museum experience; otherwise, we might witness a gradual waning of indexical interfaces in favor of unindexical ones. The myriad of socioeconomic, technological and cultural factors shaping this power struggle is perhaps beyond the scope of our discussion here, but we could expect a continued coexistence – cooperation as well as competition – between the two kinds of institutions in the foreseeable near future.

The Late Capitalist Museum and the Technologized Subjects

Although it may be too soon at this point to jump into any conclusion as to the end result of the intricate relationship between traditional museum institutions and newly emerged, technologized cultural institutions, as it is, however, we can still trace some of the recent developments in museology that inform our understanding of this power dynamic.

If recent efforts of Google Cultural Institute, especially the inclusion of hyperlinks that connect Museum View with individual reproductions, have foregrounded the symbiotic relationship between cultural database and cultural interface, a similar trend is certainly

not unheard of in museum practices of the twentieth century. For instance, cultural historian Michelle Henning has traced the blurring distinction between storage and archive to Mundaneum (originally named Palais Mondial, or World Palace) created in 1910. According to Henning, this breakdown of distinction continues in present-day open storage display, which “threatens to banish the exhibition as something which frames and stages and re-presents, but in practice, the exhibition becomes the means through which the storage system or archive is navigated.”⁹² Here, we can also observe how the new media dichotomy of interface and database has structured Henning’s interpretations. In a similar vein, media theorist Wolfgang Ernst has previously charted the correspondence between contemporary open-stack systems and the structure of new media at greater length:

The museum is no longer the terminal for parcel post from history, art, and culture; instead the institution becomes a flow-through and transformer station. Its demand now is mobilizing, unfreezing the accumulation of objects and images in its repositories, making them accessible to the public by displaying the stacks or recycling them into the exhibition area.⁹³

Again, the dichotomy of interface (flow-through and transformer station) and database (accumulation and repositories) is clearly visible, but with the new demand for mobilized and unfrozen accumulation, the museum is, according to Ernst, “no longer concerned with memory in temporal terms” and adopts instead the “synchronic standard operations of machine memory based on reversible time.”⁹⁴ This museological shift from diachrony to synchrony would in turn echo observations art critic Rosalind Krauss made earlier in 1990:

The encyclopedic museum is intent on telling a story, by arraying before its visitor a particular version of the history of art. The synchronic museum – if we can call it that – would forego history in the name of a kind of intensity of experience, an aesthetic charge that is not so much temporal (historical) as it is now radically spatial.⁹⁵

⁹² Henning, 2007, 40.

⁹³ Ernst, 2000, 25-26.

⁹⁴ Ernst, 2000, 26.

⁹⁵ Krauss, 1990, 7.

For Krauss, the model for such a shift towards synchrony is Minimalism, which, according to Minimalist sculptor Robert Morris, is predicated upon a stronger “awareness of oneself existing in the same space as the work,” which in turn “takes relationships out of the work and makes them a function of space, light, and the viewer’s field of vision.”⁹⁶ As we can see from these remarks, the so-called synchronous museum has an inherent affinity with space and is closely related to the variability of the viewer’s position within this spatial context. This again evokes Friedberg’s mobilized virtual gaze and its many implications for the viewer. Indeed, as Krauss continues, a crucial part of Minimalism’s logic is to reform the notion of the viewing subject, “a subject radically contingent on the conditions of the spatial field, a subject who coheres, but only provisionally and moment-by-moment, in the act of perception.”⁹⁷

In essence, Krauss has underscored the primacy of what she calls the “lived bodily perspective” in Minimalist art, which becomes relevant to our discussion in that it in part accounts for the continued appeal of Museum View, the persistent popularity of an indexical interface for all its imperfections. “For it is the immersion of the body in the world, the fact that it has a front and a back, a left and a right side, that establishes ... a kind of internal horizon which serves as the precondition of the meaningfulness of the perceptual world.”⁹⁸ This is particularly the case in an immersive virtual reality setting, which, as we have demonstrated previously, returns the preindustrial foreground dissolved in panoramic perception to the viewer so that they can truly have “a front and a back, a left and a right side,” or at least the sense thereof. The immersed virtual subject, to paraphrase Krauss, thus coheres provisionally and click-by-click in the act of navigation. In contrast, unindexical interfaces will very likely relapse into “the decorporealized and therefore bloodless, algebraicized condition of abstract painting in which a visuality [is] cut loose from the rest of the bodily sensorium.”⁹⁹ Thus, Museum View’s interactive and immersive nature can be said to have recorporealized the viewer in a gesture of compensation, and in

⁹⁶ Morris, 1968, 232.

⁹⁷ Krauss, 1990, 9.

⁹⁸ Krauss, 1990, 9.

⁹⁹ Krauss, 1990, 9.

so doing deprioritizing the sense of visuality and to some extent reconnecting it to the rest of the bodily sensorium.

From the mobilized virtual gaze to the lived bodily perspective, our discussion seems to finally come full circle. Still, Minimalism's ambition to radically restructure the conditions of viewer/artwork encounter in resistance to the banality of consumer culture is, as Krauss suggests, coupled with its simultaneous "participation in a culture of seriality, of multiples without originals – a culture, that is, of commodity production."¹⁰⁰ Now, such an internal contradiction can also be mapped onto Google's construction of the virtual space. On the one hand, immersivity, sense of presence, and instant access have been the prime goals of Street View and Museum View, but on the other hand, these virtual spaces remain highly mediated and deliver a received perception. The supposed seamlessness of the total space is constantly being stretched or ruptured by mostly unintended "glitches" that operate in a self-reflexive manner, revealing the otherwise invisible means by which the virtual space is created. In Street View, blurred faces and license plates remind us of the algorithms that operate beneath the image surface; stitching artifacts and image distortions reiterate the fact that the virtual space is literally a patchwork; the invasive presence of Street View Car is ubiquitous, as is best exemplified in a "selfie" captured in Japan in 2015 (Figure 2.11).

Certainly, these can all be seen as mere technological imperfections that are temporary and can be corrected in future updates, but media theorists Jay David Bolter and Richard Grusin would suggest that this apparent contradiction is in fact a necessary condition of all new media. Bolter and Grusin's double logic of remediation postulates the coexistence of "transparent immediacy" and "hypermediacy," in which the former "leads one either to erase or to render automatic the act of representation," while the latter "acknowledges multiple acts of representation and makes them visible."¹⁰¹ Following this double logic, we may forever approach Street View and Museum View *in the name* of an unmediated experience, only to find ourselves repeatedly led back to the world of simulacra.

¹⁰⁰ Krauss, 1990, 8.

¹⁰¹ Bolter and Grusin, 2000, 33.



Figure 2.11 Screenshot from Google Street View, Nagiso, Nagano Prefecture, retrieved June 2017.

Up to now, our investigation of Museum View has mostly been centered on the view itself, whereas the museum's institutional voice of the has rarely been discussed. While it is hard to imagine any consensus among the vast cornucopia of museum practitioners, in a 2015 interview with *Financial Times*, Sree Sreenivasan, then chief digital officer at the Metropolitan Museum of Art in New York, addressed a common concern across cultural institutions that “the virtual would trump the physical,” when there are too many photographic reproductions of the artworks available to the museum-going public. His response was a “virtuous circle,” in which a positive virtual experience might lead more potential audience to visit the museum in person, who would take more pictures and disseminate these images to even more people through various social networks.¹⁰² For how long can such a circle be sustained is open to question, but the fact that more and more institutions have chosen to collaborate with Google while extending their own social network presence has confirmed the symbiotic relationship between the virtual and the physical.

If Sreenivasan's optimistic comments can be read as a welcoming attitude towards new technology, it certainly is not something new on the part of the museums. Since the early 1990s, technological innovations in video projection have paved the way for the relocation of moving images within the museum space on an unprecedented scale, which, according to film theorist Erika Balsom, displaced the role of monitor in pursuit of greater spatial

¹⁰² Sreenivasan, 2015.

arrangements.¹⁰³ In other words, this shift liberated the viewer from the monitor's limited scale and relatively confined viewing position, and can be seen as to stand in line with the synchronic museum's predilection for a spatially contingent aesthetic charge. When more installation works tend to "favor immersive spectacle, overwhelming the viewer through large-scale projections of a high sensory intensity,"¹⁰⁴ the relocated moving images, Balsom suggests, end up serving "as a central component of an increasing *spectacularization* of the museum space that brings it into a closer proximity to mass culture than ever before."¹⁰⁵ This welcoming attitude towards innovations in media technology since the 1990s is the basis on which Balsom disproves the assertion that the museum space is safeguarding the moving images against the erosion of commercial cinema. In fact, Balsom argues:

the crisis of the black box is equally the crisis of the white cube – for both institutions are implicated in the predicament of what happens to the public culture of the twentieth century within the digital mobility of the twenty-first.¹⁰⁶

This "digital mobility" of again brings us back to the mobilized virtual gaze, and from such an institutional perspective we may now reconsider Museum View, an obsessively spatially charged indexical interface, as part of the larger scheme that equally seeks for an increasing "spectacularization" of the museum space. As Krauss comments:

The industrialized museum has a need for the technologized subject, the subject in search not of affect but of intensities, the subject who experiences its fragmentation as euphoria, the subject whose field of experience is no longer history, but space itself.¹⁰⁷

If spectacularization indeed affects the museum institution's "proximity to mass culture," the digital mobility, or the mobilized access to digital contents, has also induced significant changes with regard to the position of arts and culture within the public and the private

¹⁰³ Balsom, 2013, 34.

¹⁰⁴ Balsom, 2013, 55.

¹⁰⁵ Balsom, 2013, 61, emphasis added.

¹⁰⁶ Balsom, 2013, 49-50.

¹⁰⁷ Krauss, 1990, 17.

domains. If televisual spectatorship “served an at-once mobile and home-centered way of living: a form of *mobile privatization*,”¹⁰⁸ then the museum’s now closer proximity to mass culture will very likely continue this form and allocate even more cultural resources to be consumed in a mobile and home-centered manner.

Remarkably, the idea of a “domestic pinacotheca” was already envisioned by avant-gardist László Moholy-Nagy in the 1920s, which included a filing system for lithographs and collotypes, a “radio picture service” as an image distribution system that we might regard as the distant ancestor of television or the internet, and even a hologram-like three-dimensional imaging system¹⁰⁹. Like Malraux’s museum without walls, Moholy-Nagy’s domestic picture gallery underscores the primacy of photographic reproductions, but, unfortunately, most of these ideas never really materialized in his lifetime. This should not come as a surprise, since, after all, it was the maturation of projection technology that triggered the institutional relocation of moving images, and it was the ubiquity of networked computers and mobile devices that prompted projects like GA&C. Still, this is in no way meant as some sort of technological determinism; instead, if “what makes the museum obsolete as a nineteenth-century accumulative institution is the mass media,”¹¹⁰ the reason Moholy-Nagy failed to deliver his ideas in the 1920s (or perhaps we should say the 1920s failed Moholy-Nagy’s vision) was precisely that the media technological basis of his picture gallery had yet to truly become “mass.” Similarly, we can understand Moholy-Nagy’s apparently anachronistic vision against Friedberg’s theoretical framework, and attribute his “failure” to a mobilized virtual gaze that was, in the 1920s, yet to be fully realized:

As this mobilized virtual gaze has become a fundamental feature of everyday life, experiences that produce such subjective fluidity are now as much a part of the public sphere ... as they are a part of the private.¹¹¹

¹⁰⁸ Williams, 2003, 19, emphasis in original.

¹⁰⁹ Moholy-Nagy, 1969, 25-26, emphasis in original.

¹¹⁰ Colomina, 1996, 213.

¹¹¹ Friedberg, 1994, 147.

Therefore, we may need to reevaluate the position of traditional museum institutions in their competition and cooperation with the newly established, technologized cultural platforms like Google Cultural Institute. Rather than passively responding to technological innovations and reluctantly internalizing the digital mobility of the twenty-first century, the industrialized museums find themselves in an inevitable position – amid the crisis looming the black box and the white cube – in which they need to actively accommodate the changing demand of contemporary mass culture.

Finally, to return to the questions posed at the beginning of this chapter, already we have arrived at the conclusion that both the mobilized virtual gaze and the panoramic perception are what define and regulate the relationship between the viewer and the virtual space in Street View. In the particular case of Museum View, while the general conclusion still holds, we must also take into account the ideology of the physical white cube that has now become an integrated part of the virtual museum space. As for the growing popularity of Museum View, we now come to realize that the recent museological penchant for open storage display, the Minimalist formula for a spatially charged aesthetic experience, and the mandate of an increasingly spectacularizing mass culture have all played a significant role.

Chapter Three The Algorithm, The Screenshot, The Photograph

Up to this point, I have tried to outline the spatiotemporal parameters that define Street View as a representational system. In the previous chapters, our focus has shifted from the system's temporal aspects to its virtual spatiality. This turn to spatiality is completed with a new media approach, in which the dichotomy of database and interface has structured our arguments. In this chapter, I will reconsider some aspects of the new media approach that are not directly related to the construction of virtual space, but they are nonetheless instrumental in revealing both the working mechanism of Street View as a multimedia platform, and the resonance it evokes with existing mediums such as photography.

When we previously consider the new media principle of numerical representation, the focus is dedicated more to the digitization process than to its implication that now digital images are increasingly subject to algorithmic manipulations. For this reason, the main questions to be addressed in this final chapter are: to what extent do the algorithmic interventions in Street View signify a paradigmatic change in the representation of physical reality? And, what are the artistic responses evoked by this change of paradigm? To answer these questions, I will begin with the so-called “algorithmic turn” in recent cultural discourses and how it has destabilized what Ingrid Hoelzl and Rémi Marie refer to as the “photographic paradigm of the image.” Then, I will turn to performative mapping practice and what I seek to foreground as the *screenshot aesthetic* in the forming of resistance against the dominant algorithmic paradigm. In particular, I will demonstrate how the act of capturing screenshot images in fact resonates with both the photographic medium and the photographic paradigm. Finally, given such resonance, I will conduct a comparative analysis between the screenshot images and the photographic images with which Google stitches the virtual space of Street View in the first place.

From the Photographic Paradigm to the Algorithmic Turn

Without a doubt, manipulation of photographic images is almost as old as the medium itself, but the degree of malleability, “the ease, speed and quality of digital manipulation

represent an important new stage in the technology of imagery.”¹¹² In addition, when new media objects become programmable and subject to algorithmic manipulation, the very nature of manipulation has also changed. With analogue images, the manipulation process is manually achieved based on a preset aim (say, to remove dust), but in the digital era, various “presets” (now, a dust removal command) are preprogrammed with human labor, while the executions are usually performed by different algorithms, which are essentially instructions to be followed by computer operations.¹¹³ Thus, algorithmic manipulation operates on a level that is more systematic and more automatic. In addition, not only has the algorithmic turn radically transformed the conditions of image (re)production and dissemination, for media theorist William Uricchio it has also established “algorithmically defined relations between the viewing subject and the world viewed, offering robust alternatives to the visual economies of the past.”¹¹⁴

What exactly, we might ask, are the “visual economies of the past”? Or, if we are indeed experiencing an algorithmic turn, from what does it constitute the turn? Uricchio’s answer can be recapitulated as that from the *algorismic* to the algorithmic, in which the former represents “a calculable sum, whose value lies in the correctness of its result,” whereas the latter represents “a (finite) process, a formula capable of accommodating different values and yielding different results.”¹¹⁵ Although the algorismic can serve as the underlying logic of a mathematical system, an epistemic regime, and even a social order, Uricchio limits his argument to the realm of visual representation, in which the entrenched tradition of three-point perspective and its stable geometry have become “a testament to the algorismic regime of fixity and precision.”¹¹⁶ The case in point for Uricchio is Photosynth, in which unstable compositions and perambulating viewing positions (Figure 3.1) mark a radical break away from the geometrically fixed subject-object relationship governed by the algorismic.

¹¹² Gunning, 2008, 40.

¹¹³ According to the Oxford English Dictionary, an algorithm is “a process or set of rules to be followed in calculations or other problem-solving operations, especially by a computer.” See <https://goo.gl/WsBuFC>.

¹¹⁴ Uricchio, 2011, 25.

¹¹⁵ Uricchio, 2011, 26.

¹¹⁶ Uricchio, 2011, 30.

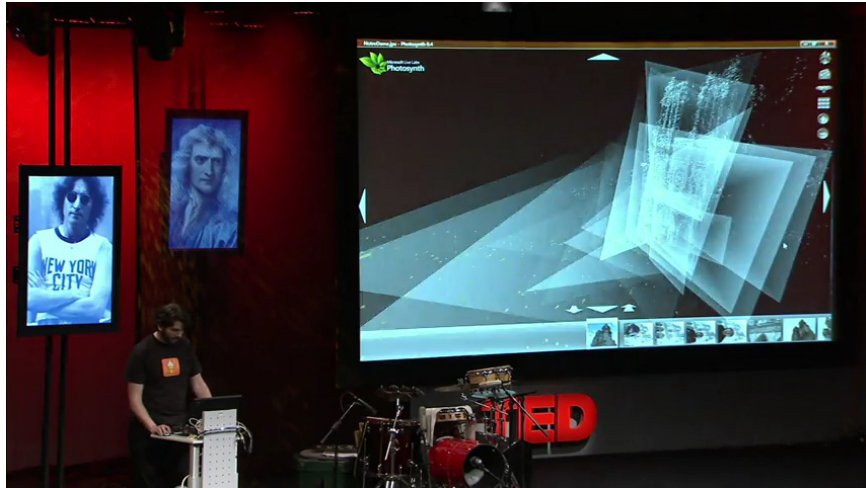


Figure 3.1 Perspectives in Photosynth, screenshot from video, retrieved June 2017.

Across such an algorithm/algorithm spectrum, Street View can then be regarded as a less radical version of visual representation compared to Photosynth. While the viewing position is not fixed (hence the mobilized gaze), it does have to follow a strictly defined pattern (the lines and the cross signs), and the carefully stitched spherical panoramas also aim to provide a frameless spatial totality that stands in stark contrast with Photosynth's fragmentary image windows. One important reason for this discrepancy may be the fact that Photosynth assembles images uploaded by users, whereas Street View collects its own raw materials. Although it is also possible for users to publish their own spherical images in Street View, they will have to follow the same specification requirements.¹¹⁷ In other words, both Photosynth and Street View include algorithmically constructed images, but they operate under the instructions of different algorithms, with the former working more in line with the logic of hypermediacy (framed and windowed views), whereas the latter in pursuit of a higher degree of transparent immediacy (seamless stitching).

With the algorithmic turn, it is also possible for us to expand some of the conclusions reached in the first chapter, particularly with regard to *Street with a View* and its inherent narrative potential. As a reminder, Christian Metz suggests that visual narrative necessarily entails a visual narrator as a "grand image-maker" that operates behind the texts. Now, instead of equating the image sequence itself with the storyteller, we can attribute this role

¹¹⁷ With the Street View mobile app, users can find the nearest places to them where a Photo Sphere is needed. Google also provides a range of imaging devices with which users can capture and publish their images. For general content policies and publishing acceptance criteria, see <https://goo.gl/7QnO2Z>.

to the operating algorithms behind Street View. Uricchio does not address narrative in his inquiry into Photosynth, but he nevertheless comments on the issue of authorship, which he views as 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.¹¹⁸

What Street View and Photosynth both have are the “enabling” author of the algorithm and the “enacted” author of the experience, but in Street View there never is an individual author to be “effaced” in the first place. As the copyright symbol constantly reminds us, the images are created by and always belong to Google. Even in the case of user contributed images, a user account name and sometimes a profile photo are provided to acknowledge the contributor, an act of accentuation that is the very opposite of effacing.

Apart from the algorismic, another case of the visual economies of the past with which we can understand the algorithmic turn is what Hoelzl and Marie term the “photographic paradigm of the image,” something that in fact predates the invention of the photographic medium. In a dedicated book chapter about Street View images, the two authors trace the “geometric convergence” of the world and the image to fifteenth-century mathematician Leon Battista Alberti’s perspectival drawing. Underlying the convergence are the rules of geometric projection, which are the operating logic behind the congruence of vision and representation as well as the alignment of the physical and the virtual. This convergence, according to Hoelzl and Marie, was “naturalized” over the ensuing centuries. Thus, the reason we tend to think of photographic images as resembling how things look in reality is that vision and representation are “set and understood *as* commensurable and thus convergent from the very beginning.”¹¹⁹ In other words, to see the world as perspectival images is an acquired ability no matter how inherited it may seem, and after human beings have over the centuries internalized this perceptual mode, photographic representations

¹¹⁸ Uricchio, 2011, 31.

¹¹⁹ Hoelzl and Marie, 2015, 94, emphasis in original.

will unsurprisingly look “natural” to us. Indeed, as filmmaker Hollis Frampton puts it, “opticians *assume* the Renaissance perspective when they grind their lenses.”¹²⁰

For Hoelzl and Marie, the condition for the alignment of the virtual and the physical is that they are perceived within a “common reference system” and are governed by the same rules of geometric projection. This common reference system is what the authors mean by the photographic paradigm of the image. With digitization, however, the old photographic paradigm remains intact only “on the visual level,” since the algorithm has taken control “on the computational level.”¹²¹ If we can assume that the visual level means what images resemble and the computational level means how images come to resemble, then, again with the new media approach, we can equate the two levels, respectively, with an interface and a database. Indeed, the interface still follows the photographic paradigm, since the rules of geometric projection remain applicable, albeit not always seamlessly, within individual images, but the organizing system, the database from which images are selected and then stitched, reads and operates nothing but pixels. Such an interface/database reinterpretation not only explains the “split” of the photographic paradigm into one visual level and one computational level (the separability between database and interface), it also accounts in part for the “perfect synergy” in which the photographic and the algorithmic paradigms now function:

While the photographic paradigm is augmented with hitherto unprecedented possibilities of multi-vision, tele-vision, navigability and real-time adaptivity, the algorithmic paradigm has evolved in order to operate in continuity with the first.¹²²

If the algorithmic paradigm has indeed “evolved in order to operate in continuity with” the photographic paradigm so as to claim as the rightful heir in the established tradition of visual representation, then the question we need to ask is what exactly incentivizes this evolvement. One reason may be the fact that Street View is yet another case in which the photographic merges with the cartographic. This mergence, according to Hoelzl and

¹²⁰ Frampton, 1983, 189, emphasis in original.

¹²¹ Hoelzl and Marie, 2015, 95-96.

¹²² Hoelzl and Marie, 2015, 96.

Marie, dates as far back as to Alberti, who used the same mathematical tools to create his perspectival drawings as well as to carry out cartographic measurements in Rome.¹²³ More recently, media theorist Tristan Thielmann has observed “parallel developments of a spatial turn in media studies and a media turn in geography,” where the former leads to the proliferation of “locative media” while the latter results in “mediated localities.” As Thielmann suggests, a proper umbrella term for these parallel developments would be “geomedia.”¹²⁴ Based on his observation, we may regard Maps, Earth, and Street View as one bundled collection of geomedia tools, and, Thielmann continues, “while such a tool continues to provide instructions for navigating physical reality, at the same time, it also always provides a platform on which all data can be plotted.”¹²⁵

In a similar vein, media theorist Francesco Lapenta argues that geomedia as a collective system of data search and organization “transform the geolocation of their users, their geosphere, into data, and connect these data to existing information that describe users’ online activities and identities (and their infosphere).”¹²⁶ According to Thielmann and Lapenta, we may argue that the ultimate ambition of algorithmically-based geomedia is to tap into users’ infosphere, to be able to collect the vast amount of user data. Furthermore, Lapenta pushes the prominence of data collection to the level of a new economy of space, and suggests:

Geomedia can not only be interpreted as the evolution and response to the matured need for new organizational criteria to coordinate and link *mediated interactions*, but also as the attempt of new actors and the same *economic and political forces* that regulate physical time and space to organize and regulate the global placeless flow of information into locally controlled and physically contextualized information systems.¹²⁷

¹²³ Hoelzl and Marie, 2015, 93-94. The two authors also point out that a more recent case would be the use of aerial photography for military and civic mapping in the nineteenth century.

¹²⁴ Thielmann, 2010, 5.

¹²⁵ Thielmann, 2010, 8.

¹²⁶ Lapenta, 2011, 20.

¹²⁷ Lapenta, 2011, 21, emphasis added.

To be sure, the “mediated interactions” would refer to the act of user navigation through the virtual space, with the addition of user posted of texts and images via highly integrated social network services, but the presence of “economic and political forces” is perhaps best exemplified in the virtual museum space. As we have discussed earlier, the ideology of the white cube pervades the physical and the virtual museum space. In a physical setting, it regulates “physical time and space,” while in a virtual configuration, it organizes and regulates instead the “placeless flow of information.” Moreover, Lapenta notes a constant tension between the “flexibility and the autonomy-enhancing qualities” provided by geomedial and “an opposite function that sees them as new organizational and regulatory systems,”¹²⁸ but this apparent contradiction can be reconciled if we place it against the larger societal shift from philosopher Michel Foucault’s disciplinary society to philosopher Gilles Deleuze’s society of control.

For Foucault, “discipline proceeds from the distribution of individuals in space,”¹²⁹ which is precisely what maps are intended for. The, so to speak, “art of distributions” and the techniques employed – enclosure, partitioning, and functional sites – pertain perfectly to the intensely surveilled and strictly regulated museum space. In contrast, what we might call the “art of dispersion” has choreographed the mobilizing of virtual museum-goers, who are no longer subject to institutional regulations such as no eating or talking, but will now become part of the capitalization process in which the production and exchange of user data are turned into commodities. Likewise, the same process also applies to the comparison between an urban navigator and its virtual counterpart. In Lapenta’s words, “users themselves (and their surrounding space) are transformed into information – a commodified image.”¹³⁰ All in all, the “dispersed and ‘flexible’ modes of authority, ubiquitous networks, and the relentless branding and marketing of even the most ‘inner’ aspects of subjective experience” are all essential elements that characterize Deleuze’s society of control.¹³¹

¹²⁸ Lapenta, 2011, 21.

¹²⁹ Foucault, 1995, 141.

¹³⁰ Lapenta, 2011, 21.

¹³¹ Shavero, 2010, 6. In June 2017, YouTube, now a Google product, introduced a Heatmaps in VR that enable 360-degree video creators “to see exactly what parts of your video are catching a viewer’s attention and how long they’re looking at a specific part of the video.” Although this eye-tracking function is not yet

Therefore, if we map this societal transformation onto the paradigmatic turn in visual representation, we realize that the algorismic regime, the photographic paradigm, and the disciplinary society all embrace the strictly defined geometric fixity, while the algorithmic turn, coupled with the new society of control, effects the logic of flexibility based on the endless stream of data. Integrating the shift with visual representation, we may conclude that visual economies of the past dictate that the viewer be fixed in a predetermined position in order to successfully perceive the representation of physical reality, but the algorithmic turn, as one visual economy of the future, has allowed us access to the virtual world through programs like Street View in a manner of greater autonomy.

Performative Cartography and the Screenshot Aesthetic

If the algorithmic has indeed replaced the photographic to become the new mechanism of power through which geomedia asserts its cultural and socioeconomic dominance, how then should we respond to this new powerhouse? In response to the arrival of a control society, Deleuze suggests that “[t]here is no need to fear or hope, but only to look for new weapons,”¹³² yet he stops short of revealing to us what exactly the “new weapons” could be. It may be assumed that the proper response will depend on the specific form of power exertion, so when we look at the film *The Truman Show* (1998), for instance, the protagonist eventually decides to sail off the fabricated island set in defiance against his equally constructed life on stage. Similarly, in *The Matrix* (1999), Neo and his crew choose to unplug from the all-encompassing neural-interactive simulation and form an army of resistance against the omnipresent matrix. Incidentally, the two films can to some extent be seen, respectively, as the embodiment of a disciplinary society and a control society, as Truman is constantly being “distributed” within the enclosed, partitioned functional sites, whereas the matrix amounts to a virtual reality interface that essentially comprises of computer codes. Yet, even with the cinematic elements of fiction and drama stripped off,

available in Street View, it nonetheless demonstrates what Shaviro means by the branding and marketing of one’s subjective experience. See <https://goo.gl/8nqAzu>.

¹³² Deleuze, 1992, 4.

most of us cannot afford to simply unplug from the media-saturated and algorithmically operating contemporary culture.

Nevertheless, what is achievable is a more critical stance through which we can engage with geomedia platforms as the powers that be. One important aspect of such engagement, as writer and artist Taien Ng-Chan suggests, is the use of map itself. Most of the mapping services today allow users to “annotate” their products (writing comments, posting images) and more sophisticated tools such as APIs even encourage “web developers to incorporate maps into their sites and to make map mash-ups,”¹³³ of which, as we have seen, *Google Shoot View* is an example. Although corporations like Google and Microsoft can maintain, if not further strengthen, their dominance with the provision of customizable mapping tools, they nonetheless “allow stories, subjectivities, and personal engagement into the digital map [and] seem to increase the democratization of cartography through widening access and amateur software tools.”¹³⁴

Similarly, media theorist Jason Farman recounts at greater length the presence of a “participatory culture” among Google Earth users and its instrumental role in the deconstruction of the digital map. For Farman, this participatory culture is comprised of two main components: an online social network and the use of “overlay,” which functions “as a way for users to augment the map by offering a different visual representation of a specific area.” The integration of overlays with the social network would then constitute a bricolage that recontextualizes and subverts “master representations” of the digital map:

The social network as a community is positioned to enact agency. This potential for agency comes through the implementation of the very tools that limit them through a repurposing, reimagining, and reconfiguring of master representations in conjunction with user-generated content.¹³⁵

¹³³ Ng-Chan, 2016, 67. Google is no exception and offers a list of mapping tools at the disposal of the user. See <https://goo.gl/JQsjNG>.

¹³⁴ Ng-Chan 2016, 67.

¹³⁵ Farman, 2010, 879-882.

With the above arguments in mind, we may revisit Hewlett and Kinsley's *Street with a View* discussed in the first chapter. The premeditated inclusion of human and narrative elements can be seen as an alternative form of pre-production map overlay, and the project itself amounts to some user-generated contents, accomplished and accessed through the implementation of Google's own mapping tools to repurpose, reimagine, and reconfigure the purportedly neutral and objective master representations. For this reason, we may contend that their artistic intervention not only confronts against the static nature of Street View images, but it is also imbued with the participatory culture that seeks to deconstruct and therefore resist against Google's ascendant geomeia platform.

The performance of *Street with a View* was implemented simultaneously as the Street View car scrolled past the scene, so in this particular case, according to new media theorist Federica Timeto, space was "figured and practiced at the same time," in which mobile annotation made "the data and the practices of space converge."¹³⁶ Such performative mapping practices therefore construct an act of active resistance in the *pre-production* process. While such pre-production interventions function as one critical tool "that ultimately reimagines the status of the map presented by Google and the viewer's relationship to that map,"¹³⁷ we have also observed numerous instances of resistance during the *post-production* process. That is, mapping images first become available and are then approached as existing archive images. What they have in common is what we might call the *screenshot aesthetic*, a peculiar predilection for artists and amateurs alike to capture the visual façade of the media interface.

Before we delve into the aesthetic, it is of note that the definition of screenshot – "an image of the data displayed on the screen of a computer or mobile device"¹³⁸ – is based on several key assumptions. First, a screenshot is the exclusive product of new media, as the parallel structure of data(base) and screen (interface) clearly suggests. Also, computers and mobile devices not only serve as the locus (both hardware and software) of image display,

¹³⁶ Timeto, 2015, 100.

¹³⁷ Farman, 2010, 884.

¹³⁸ Oxford English Dictionary. See <https://goo.gl/1hHT7E>.

they also provide the tools (built-in screenshot programs) and platforms (through third-party screenshot applications) so that screenshot capturing becomes possible.¹³⁹

How then, we may ask, can the mere capturing of screenshots, regardless of what is actually being captured, be seen as an act of resistance? To answer this question, we can devise a quick list of what exactly changes when a screenshot is taken from Street View: a screenshot image intercepts the interlocked temporal axis; the otherwise mobilized gaze is fixed at one particular point of view; the static and two-dimensional view is no longer interactive or immersive. In other words, screenshot images cease to be part of the ongoing operations within geomeia that are regulated by the algorithmic paradigm. Instead, they are extracted from the program's data flow and become image files stored locally on the device that are, albeit still programmable with another software, no longer subject to *the* algorithmic manipulations from within Street View.

Therefore, we may indeed be able to assimilate the act of taking screenshots to one of the central themes behind *The Truman Show* and *The Matrix*, since the three all entail the extraction or emancipation of a single entity from an overwhelmingly dominant system of power that constantly exerts control over its subjects. If such an analogy holds, and if the screenshot aesthetic does pose an active resistance against the algorithmic paradigm, the question follows: to what extent does this new aesthetic hearken back to the photographic paradigm that has supposedly been superseded? To begin with, we may argue that taking screenshots still follows the principle of geometric projection in the broadest sense, since the screen functions as a two-dimensional plane where some sort of "digital contact print" is being produced. Moreover, the single important thing that needs to be considered when we take a screenshot is the position, size, and shape of the rectangular frame. It is, in other words, an act of cropping, which, as American photographer John Szarkowski argued in the 1960s, is an indispensable aspect of the photographic medium:

¹³⁹ Before built-in or third-party screenshot programs became widely available, a user had to take "screen shots" with a separate camera. The resulting image essentially became a reproduction rather than a screenshot. Some computer magazines have published tutorial articles about how to take such screen shots. See, for example, Bobo, 1986, 40-45.

The photographer looked at the world as though it was a scroll painting, unrolled from hand to hand, exhibiting an infinite number of croppings – of compositions – as the frame moved onwards.¹⁴⁰

In addition, as the scroll painting metaphor reminds us, capturing screenshots also evokes the photographic in terms of its temporal scheme. In Chapter One, we have meticulously distinguished simultaneous and non-simultaneous duration, and how the use of a rolling shutter determines that a photographic exposure is always non-simultaneous and thus can never truly become instantaneous. However, the screenshot aesthetic, the cameraless form of image making, complicates the claim: while it will always take some time to capture a screenshot (hence the act itself should be deemed durative), this duration is *undefined* and becomes *irrelevant* to the resulting image. It is undefined because the time it actually takes to capture a given screenshot is always unknown, unlike a digital photograph in which the Exif (exchangeable image file format) data usually keep a record of the image's exposure time. It is irrelevant because for the same image to be captured on multiple devices (or multiple times on the same device), the operation may take different amounts of time depending on nothing but the device's specifications and conditions (which may be why the time is not kept on record). In other words, the undefined temporal parameter is both externalized and trivialized in the process of capturing screenshots. Furthermore, without the presence of a rolling shutter, this process achieves global simultaneity in that all pixels in the image will have exactly the same "temporal coordinates." Therefore, we could argue that this *undefined, irrelevant, and simultaneous duration* is the temporal scheme of screenshot images. If, for Bazin, the analog photographic image is an act to preserve, "to snatch it from the flow of time, to stow it away neatly ... in the hold of life,"¹⁴¹ we may paraphrase his comments from more than six decades ago and argue that the capturing of screenshot images also amounts to an act of preservation, only this time it is snatching from the flow of data and stowing away in the hold of digital life.

To briefly recap, the continued adherence to the principle of geometric projection, the act of cropping, and the temporal scheme that aims to create a visual preservation are all

¹⁴⁰ Szarkowski, 2009, 9.

¹⁴¹ Bazin, 2005, 9.

essential aspects of the screenshot aesthetic that render the practice as photographic as the medium of photography itself, without relying on a camera device or an exposure process. To further illustrate the screenshot practice's return to the photographic, we may look at the ongoing project *Postcards from Google Earth* (2010-) by Brooklyn-based artist Clement Valla. In this collection of bizarre images, the otherwise seamless representation within Google Earth twists and turns into some sort of surrealistic illusion (Figure 3.2). The main reason for this distortion is the coexistence of two separate data inputs: three-dimensional topographic models and two-dimensional satellite and aerial photographs mapped onto them. We can think of the latter as a database of texture images or “skins” that are matched with, stretched across, and pasted onto the topographic models, a process known as the Universal Texture. As the artist explains:

When these photographs get distorted and stretched across the 3D topography of the earth, we are both looking at the distorted picture plane, and through the same picture plane at the space depicted in the texture. In other words, we are looking at two spaces simultaneously.¹⁴²



Figure 3.2 Clement Valla, *Postcards from Google Earth*, 2010-ongoing.

The doubling of spaces is one of the basic tenets of Google Earth, but most of the time it is concealed by an adequate degree of matching between cartographic and photographic

¹⁴² Valla, 2013, 146.

data. For this reason, we can think of the series as another demonstration of the medium platform's self-reflexivity through situations of unconceivable anomaly. For Italian graphic designers Davide Giorgetta and Valerio Nicoletti, the "sudden interruption of established and conventional operations makes the error as impetuous as fascinating" so that users are encouraged to "immortalize" such happenings.¹⁴³ Indeed, these digital moments are as ephemeral as they are perishable, and their metaphorical deaths are the preordained results of the program's constant updates and the perfection of its operating algorithms. This would again bring us back to Bazin, since the screenshot image "helps us to remember the subject and to preserve [it] from a second spiritual death."¹⁴⁴ Such an undertone of screenshot as memento mori is also intimated in the title of Valla's series:

I also call these images postcards to cast myself as a tourist in the temporal and virtual space – a space that *exists digitally for a moment*, and may perhaps never be reconstituted again by any computer.¹⁴⁵

Based on this momentary digital existence, we can almost deduce, in a true Barthesian fashion, a "that-has-digitally-been" as the conceptual core behind Valla's screenshot series, or the "immortalization" of to-be-updated error images in general. In his contemplative reflection upon the *noeme* of photography, philosopher Roland Barthes comments:

I call "photographic referent" not the *optionally* real thing to which an image or a sign refers but the *necessarily* real thing which has been placed before the lens, without which there would be no photograph.¹⁴⁶

Similarly, we may call "screenshot referent" not the *geometrically* real thing to which an image or a sign refers but the *algorithmically* real thing which has been placed before the screen, without which there would be no screenshot. If we can equate the interpretation of Barthes with the semiotic notion of "index," which dictates that meaning is constructed

¹⁴³ Giorgetta and Nicoletti, 2016, 16.

¹⁴⁴ Bazin, 2005, 10.

¹⁴⁵ Valla, 2013, 149, emphasis added.

¹⁴⁶ Barthes, 1982, 76.

based on a physical relationship between the image and its referent, “that-has-digitally-been” would then entail something of a *virtual index*, suggesting “that there is a precisely specified object alive in a virtual space, with computational malleability but no definite real-world presence.”¹⁴⁷ To be sure, no one would really believe that real-world bridges could be curled and bent as we see in Valla’s collection, but any physical presence, or the belief thereof, is essentially irrelevant, since the meaning of a virtual index is established only along the axis of a virtual relationship to the referent. So, when Valla builds the collection “to make sure there is a record that this image was produced by the Universal Texture at a *particular time and place*,”¹⁴⁸ he is not talking about the spatiotemporal parameters that define the physical bridge, but those that define the moment of his on-screen interaction with the program, the moment instantly after which the captured images are no longer subject to Google’s algorithmic manipulations.

To sum up, we have now attributed several key aspects of the photographic paradigm to the practice of capturing screenshot images, which, I hope, suffices to demonstrate the inherently subversive nature of the screenshot aesthetic as well as its potential as a post-production “new weapon” against the increasingly ubiquitous algorithmic paradigm. Yet, we might caution against overt optimism. While the screenshot aesthetic brings a new hope, the algorithmic empire strikes back. For instance, mobile messaging apps such as Snapchat will notify the sender of the message (which usually disappears after a designated amount of time), if the receiver takes a screenshot of the content. A more extreme case is banking apps, in which the screenshot function is usually disabled for security reasons. The main takeaway from this power struggle is thus the dynamic nature of the algorithmic expansion and that we should never approach it without considering the underlying technological changes beneath any transmedial wax and wane.

As we have already seen, the screenshot aesthetic extends itself far beyond the realm of geomedial, and the issue of preserving obsolete and otherwise irretrievable media interface was touched upon in a recent GA&C blogpost about the archiving of digital art. While it is hard to imagine that image files saved in JPG or PNG formats would someday become

¹⁴⁷ Allen, 2016, 272.

¹⁴⁸ Valla, 2013, 148, emphasis added.

inaccessible, other digital art or digitized art relying on more complex software systems are already at the risk of disappearing:

Obsolescence is a real threat to digital art. ... Just as preservationists have identified ways to extend the life of pigment, canvas and stone, solutions must be found to assure the longevity of digital works or they may prove to be even less resilient than their physical counterparts.¹⁴⁹

Obsolescence, unfortunately, is a necessary result of technological innovations. Photosynth is a case in point: since the discontinuation of the service in early 2017, webpages that are embedded with Synth views simply display an error message instead, whereas most archive images are either captured screenshots or recorded videos.¹⁵⁰ A more general challenge is the preservation of average webpages, which are subject to the changes in both database and interface design. “This leaves archivists with few options. They could download a limited set of data directly from the platform, or create a derivative version, such as a video recording.”¹⁵¹ In this sense, screenshot images might also be categorized as a “derivative version” of the original interface, but what such a categorization fails to recognize is the inherent rivalry between the original and the derivative. It is the anticipated demise of the original that necessitates the proper function of the derivative. In other words, to tie in with our main argument about the screenshot practice facilitating an act of resistance, we might say that it is the original which only “exists digitally for a moment” that establishes the derivative as a new form of aesthetic that operates in a relatively permanent manner.

The Screenshot versus the Photograph

¹⁴⁹ Cerf, 2017.

¹⁵⁰ All Photosynth-based URLs are now redirected to the its shutdown page. See <https://goo.gl/QS7HNC>. When Microsoft first announced the shutdown on November 4, 2016, it provided an offline Photosynth viewer so that users could export and access their own Synths, but the backup had to be finished before the shutdown date, after which Microsoft would cease to maintain the functionality of the viewer. See <https://goo.gl/EdJpVG>.

¹⁵¹ Rhizome, n.d. For similar projects on GA&C, see The Theresa Duncan CD-ROMs: Putting interactive classics online with Emulation as a Service, <https://goo.gl/h9hWXc>, and Form Art: Preserving the browser with Oldweb.today, <https://goo.gl/vxmjyv>.

The attempt hitherto to establish a transition from the algorithmic turn to the screenshot aesthetic has seemingly led us to come full circle: first, individual photographs are collected by Google from around the world, then, the images are processed to create the Street View interface, and finally, the capturing of screenshots extracts single-frames images from the algorithm-based media platform. Not only does the act of capturing screenshot resonate with the photographic paradigm, but we also, quite literally, start with a massive collection of JPGs and now end up with another growing reservoir of JPGs. These observations beg the seemingly self-evident question: in exactly what ways, if any, are the screenshot images different from the original photographic images? After all, their visual resemblance is such that sometimes one could be mistaken for the other. In addressing this question, I hope as well to reconnect with my arguments made in the two preceding chapters, not only to shed new light on the screenshot aesthetic, but also to reaffirm our earlier observations about Street View images in general.

In Chapter One, we are concerned with the temporality of Street View images, which is the first aspect in which we can distinguish photographic from screenshot images. As a reminder, individual photographs collected by Google are all taken with a durative and *non-simultaneous* exposure sequence, resulting in the fact that different parts of the same photograph bear different temporal imprints, and therefore the creation of an intra-frame temporality. In contrast, as we have established in the previous section, the temporal scheme of screenshot images is an undefined, irrelevant, and *simultaneous* duration. In terms of the graphic presentation created in Figure 1.3a, the screenshot images are in fact parallel with the countless “planes of simultaneity” along the temporal axis, whereas the photographic images will always intersect with these planes. Thus, despite their similar metaphorical stoppage of time, and despite their common failure to achieve instantaneity in a truly “point-in-time” sense, the photographic non-simultaneity and the screenshot simultaneity is their first and foremost difference.

In Chapter Two, we take a spatial turn and predicate our analysis of the virtual space on the new media dichotomy of database and interface, which offers us another possible perspective through which to better understand the difference between the two kinds of images. If, as argued by Manovich, the principle of variability suggests the new possibility

to separate database from interface in new media objects,¹⁵² then we might expect certain re-convergence between database and interface when a new media object “regress” back into an, so to speak, old media object. This regression from a newer to an older media form is exactly what happens when a screenshot is taken in Street View, since from the otherwise algorithm-dominated media platform we manage to extract a single, algorithm-free media file.

Then, we might ask the question: do the photographic and screenshot images function as part of the database or the interface? For individual photographic images, the answer is clear: they operate as part and parcel of the image database with which the Street View interface is created. For screenshot images, however, things are not as straightforward. As they regress into an old media object, the supposed separation between database and interface is lost, and instead both concepts converge and amalgamate at the same JPG file. Although individual image files still consist of even more numerical codes, in terms of visual representation it is already the smallest unit. In other words, there is no invisible database separate from but connected to the visible interface; rather, the screenshot image itself is a disconnected, independent visual interface. This, compared with photographic images being used only to build the database, is another aspect in which the two fundamentally differ.

Where do these independent screenshots wind up? Many are perhaps stored locally, but more presumably circulate on the internet and, like other digital images, someday may be rehashed as image data for a different interface. The screenshot as a conscious aesthetic, in contrast, aims somewhere else: the physical museum space. Escaping the regulations of algorithms, a screenshot image immediately becomes printable, which turns the image into a physical object so that it can be processed in a standardized institutional procedure and end up hanging on the museum wall. Obviously, the photographic images as raw materials are equally printable, but their intended and destined use as database construction renders meaningless the act of printing, since, once printed, these images are no longer stitchable. In the meanwhile, at the other end of the screenshot image’s escaping route is the museum wall, which Krauss deems as the key puzzle piece of the “space of exhibition” and the

¹⁵² Manovich, 2001, 37.

aesthetic discourse that operates within it: “given its function as the physical vehicle of exhibition, the gallery wall became the signifier of inclusion and, thus, can be seen as constituting in itself a representation of what could be called *exhibitionality*.”¹⁵³

This “exhibitionality,” a notion that entails both printability of image and attachability to the wall, is at the very core of the screenshot aesthetic that attempts to align itself with the established media forms such as painting and photography. Indeed, through the capturing of screenshots, Valla manages to claim for his *Postcards* series a rightful spot inside the white cube, in a way that most photographic images are exhibited (Figure 3.3a), or bending the print to match the exhibition’s spatial configuration (Figure 3.3b), or, true to its title, in standard postcard format (Figure 3.3c). Despite the different exhibiting formats, the *Postcards* series can be seen as one instance “in which aesthetic discourse resolves itself around a representation of the very space that grounds it institutionally.”¹⁵⁴



Figure 3.3a Clement Valla, *Postcards from Google Earth*, 2010-ongoing, installation shot.

¹⁵³ Krauss, 1982, 312, emphasis in original.

¹⁵⁴ Krauss, 1982, 312-313.



Figure 3.3b Clement Valla, *The Universal Texture*, 2012-ongoing, installation shot.



Figure 3.3c Clement Valla, *Postcards from Google Earth*, 2010-ongoing, installation shot.

It is worth mentioning that what Krauss means by “space of exhibition” is not only the literal museum space, but also a “discursive space,” a domain of culture with its own user expectations as well as its own kind of knowledge.¹⁵⁵ In this regard, the exhibitionality of printed screenshots can be said to have relocated the images in the discursive space of contemporary art. Then, we need to ask, in precisely what discursive space do the original photographic images operate? Or, if the acquisition of exhibitionality is the ultimate goal behind the screenshot aesthetic, what are the unexhibited photographic images trying to achieve?

¹⁵⁵ Krauss, 1982, 311.

Just as the screenshot images arrive at their intended discursive space of art through a standardized institutional process, so too do the photographic images undergo a formulaic processing process. In fact, Google explicates the publishing of Street View images in a four-step procedure. The individual photographs (Figure 3.4a) as raw materials are to be aligned and stitched (Figure 3.4b), with the resulting image being a spherical view (Figure 3.4c). It is important that we realize that the 360-degree image at this stage is not yet the complete interface, but it is only when the images are annotated (Figure 3.4d) and become navigable that they can finally qualify as the media platform we know as Street View. What is intriguing about this publishing process is that as long as this paragraph is being read on printed paper, there is no way to actually demonstrate the image's navigability. In fact, we are categorically certain that Figure 3.4d is a screenshot image, whereas Figure 3.4c may well be *the* stitched image rather than its screenshot copy. Also, this four-step process of turning digital photographs into an annotated and navigable interface mirrors the “media turn in geography” observed by Thielmann, since after all what the integration of mediated reality in the form of photographic images promotes is the dissemination of geographical information. Consequently, to answer the above questions, the discursive space in which the photographic images operate is that of geomedia (or, to be precise, mediated localities) rather than contemporary art, and the end objective of Google's post-production process is the image's ultimate navigability, rather than the screenshot image's exhibitionality.



Figure 3.4a Google's publishing process: collecting imagery.



Figure 3.4b Google's publishing process: aligning imagery



Figure 3.4c Google's publishing process: turning photos into 360 photos



Figure 3.4d Google's publishing process: showing the right image

The difference in discursive space and end objective is perhaps the most significant aspect that distinguishes the screenshot images from the original photographic images. From here we can also derive some related comparison. For instance, if the vision behind mediated localities can be “described as if there were nothing more in the territory than what is on the map,”¹⁵⁶ Google would naturally hope that the content of the photographic images be as comprehensive as possible, minimizing the gap between the territory and the map. Therefore, whether it is a bizarre group performance, or just another mundane city block, the cameras on Google's Fleet are always indiscriminate in the framing of the physical world. In contrast, the urge for artists to immortalize certain digital situations is often associated with the appearance of errors and anomalies, whose ephemerality makes the capturing a decidedly conscious and highly selective act.

Also, what Google seeks to comprehend with its rosette camera system is the façade of the physical reality, while what we select through the capturing of screenshot images is the semblance of the virtual world, hence the difference between the index and the virtual index. The photographic and the screenshot images may occasionally look the same, but beneath their apparent resemblance are actually mediations of different “orders.” If we can

¹⁵⁶ Thielman, 2010, 12.

agree that a photographic image is the mediated representation of physical reality (thus first-order), a screenshot image of the stitched and annotated photographic images would then constitute a *second-order mediation*, namely, the mediated representation of another mediated representation. With this in mind, the resemblance between photographic and screenshot images can then be accounted for in a non-phenomenological manner: just as in the first-order mediation it is possible for vision (physical) and representation (virtual) to align and converge, so too can the photographic images (virtual) and the screenshot images (virtual on a higher order) converge in the process of second-order mediation.

	Photographic Images	Screenshot Images
Temporal Scheme	<i>Non-simultaneity</i>	<i>Simultaneity</i>
New Media Dichotomy	<i>Database</i>	<i>Interface</i>
Discursive Space	<i>Geomedia</i>	<i>Art</i>
Objective	<i>Navigability</i>	<i>Exhibitionality</i>
Content	<i>Comprehensivity</i>	<i>Selectivity</i>
Index	<i>Physicality</i>	<i>Virtuality</i>
Mediation	<i>First-order</i>	<i>Second-order</i>

Table 3.1 A comparative analysis between photographic images and screenshot images.

To summarize, the variety of issues we have discussed up to now can be simplistically expressed in the *photograph-interface-screenshot* process, in which “interface” is the part of the cartographic platform that Street View users directly encounter, the part with its own unique spatiotemporal parameters, and the part that is interactive, immersive, location-aware, and ultimately data-driven. Dominating the *photograph-interface* conjuncture is the intervention of algorithm, behind which is the broader algorithmic turn that our contemporary visual culture has yet to fully assimilate. In the meanwhile, operating at the *interface-screenshot* nexus is the screenshot aesthetic that we delineate in this chapter. Both the starting point and the end point in this photograph-interface-screenshot process are single-frame images, but they are, as is recapped in Table 3.1, ontologically different.

Therefore, to address the main questions of the chapter, despite the paradigmatic shift in visual representation engendered by the algorithmic intervention, the practice of capturing screenshot images does not, and is never intended to, take us back to square one. Instead, the resistance against algorithmic manipulation facilitated by the screenshot aesthetic is leading us into uncharted new territory, a territory that not only addresses old issues but also raises new critical questions.

Conclusion

The things that make virtual reality compelling have been with us for tens and hundreds of thousands of years. At its core, it's about storytelling. Everything along the way is a step towards trying to get more immersed in the experience and more immersed in the story. But, you know, we're still putting something between you and the world. What we really wanted to do was to make something feel real.¹⁵⁷

On April 1, 2016, Google announced its plan for the first actual reality headset named Google Cardboard Plastic. Of course, if you notice the date or have heard of Google's annual April Fools' prank video, you soon realize that the headset is nothing but a viewer-shaped plastic box. Still, the pranking nature does not make its underlying argument any less true. Indeed, from panorama, diorama, and stereoscope of the nineteenth century to virtual reality of the twenty-first century, "everything along the way is a step towards trying to get more immersed in the experience." Just as photography and film, so too is Street View yet another "something" that is put between us and the world. This crude claim may not be the best definition of Street View, but it suggests to us an intermedial lineage that we can trace to better understand it as a new medium of visual representation. Based on this intermedial approach, what we have traced in this research, in response to the main research questions, can be reiterated as follows:

First, the intra-frame temporalities of individual digital photographs combine to form a continuous timeline when the images are stitched together. This, along with physical movement involved in the processes of image production and perception, becomes the basis for an inevitable presentation of time and a non-cinematic narrative unfolding. Then, the virtual space presented in Street View is one of navigation, whose relationship with the viewer is defined by the mobilized virtual gaze and the panoramic perception. This virtual navigation is gaining special popularity among museum institutions, whose exhibitions are increasingly being remediated as indexical cultural interfaces. Finally, the algorithmic turn, in superseding the photographic paradigm as visual economies of the future, signifies a

¹⁵⁷ Google, 2016.

paradigmatic change in the representation of physical reality. As a reaction to the ubiquity of algorithmic intervention in geomedia, many artists resort to performative mapping and the screenshot aesthetic, the latter of which shifts our focus back to single-frame images.

Although the arguments made in the three chapters all revolve around Street View, I hope nonetheless to emphasize the possibility of extending some of the key observations beyond the program itself. For instance, the distinction made between simultaneous and non-simultaneous duration complicates the temporal scheme of time-based media works and our experience of time when interacting with these works. Similarly, instead of a rigid dichotomy between motion and stillness, the notions of inter-frame and intra-frame temporalities provide an alternative theoretical framework in which we can understand the assemblage of single-frame images. For yet another example, the terms of indexical and unindexical interfaces tap into the realm of computer space, establishing one possible connection between the physical space that is governed by geometric rules and the virtual space that adopts an entirely different logic. These terms, I hope, not only clarify some of the idiosyncrasies of Street View, but would also be pertinent to future media research in a more general manner.

In addition to limitations of time and scope, one major challenge facing this research is the constant changes in Street View, some aspects of which, as well as related discussions, did not surface until late in the process of completing the research. Similarly, and yet in sad contrast, Photosynth was completely shut down earlier this year. Again, we need to remind ourselves that the new media nature and its supposed separation of database and interface suggests a different ephemerality, in which the research objects (the interface *as it is*) may someday become even less accessible than an ancient artifact. If this research has managed to establish a provisional connection between Street View and its immediate predecessors, it is my genuine hope that possible future researches may continue keeping track of the development of the program and expand on or revise this tentative connection as well as the arguments based on it.

List of Illustrations

Figures

All URLs were accessed on July 15, 2017.

- 0.1 A rosette camera set with fifteen cameras, 2010, see Anguelov et al., 2010, 34.
- 0.2 A Street View car, <https://goo.gl/oaPs9X>.
- 0.3 Screenshot from Google Street View, Witte Singel, Leiden, <https://goo.gl/9wrfZy>.

- 1.1 Sam Taylor-Johnson, *Five Revolutionary Seconds IX*, 1997, <https://goo.gl/bcnWRN>.
- 1.2 Maarten Vanvolsem, *Specious-present 2 (Seas)*, 2006, <https://goo.gl/CLfN2L>.
- 1.3a Demonstration of Snapshot and the “concentration” of time in Sugimoto’s works.
- 1.3b Demonstration of the “multiplication” of time in Vanvolsem’s works.
- 1.4 Edward Ruscha, *Every Building on the Sunset Strip*, 1966, <https://goo.gl/YushLC>.
- 1.5 Demonstration of the exposure process with a rolling shutter.
- 1.6 Jacques-Henri Lartigue, *Le Grand Prix A.C.F.*, 1913, <https://goo.gl/zSP9kx>.
- 1.7 Demonstration of image distortion in photographing fast moving objects with a rolling shutter.
- 1.8 Demonstration of intra-frame temporal axes weaved into one single timeline.
- 1.9a Screenshot from Google Maps, Reuvenplaats, Leiden, <https://goo.gl/e1aLj2>.
- 1.9b Screenshot from Google Earth, Witte Singel, Leiden, <https://goo.gl/LyUhyg>.
- 1.10 Robin Hewlett and Ben Kinsley, *Street with a View*, 2008, installation shot at Centro di Cultura Contemporanea Strozzi, Firenze, <https://goo.gl/L4aH1w>.
- 1.11a Screenshot from Google Street View, 607 Sampsonia Way, Pittsburgh, <https://goo.gl/DJdEPc>.
- 1.11b Screenshot from Google Street View, 243 Sampsonia Way, Pittsburgh, <https://goo.gl/QFtFuC>.
- 1.11c Screenshot from Google Street View, 237 Sampsonia Way, Pittsburgh, <https://goo.gl/nFRhDB>.

- 2.1 *Aspen Interactive Movie Map*, 1978-1981, screenshot from video, <https://goo.gl/f3HeBV>.
- 2.2 Four navigational modes in Photosynth, 2013, screenshot from video, <https://goo.gl/W79sBv>.
- 2.3 Google Cultural Institute and its major components.
- 2.4 The Art Camera, 2012, screenshot from video, <https://goo.gl/S3uBoE>.
- 2.5 A Street View trolley, <https://goo.gl/oaPs9X>.
- 2.6a Museum View in Van Gogh Museum, screenshot from Arts & Culture, <https://goo.gl/e65fY4>.
- 2.6b Detailed view of Vincent van Gogh’s *The Bedroom*, 1888, screenshot from Arts & Culture, <https://goo.gl/uHfPZu>.
- 2.7a Cover panel of an online exhibition, screenshot from Arts & Culture, <https://goo.gl/spwKcy>.
- 2.7b Exhibition interface from a previous version, screenshot from video, <https://goo.gl/8UEgUz>.
- 2.8a Curator Table, screenshot from Arts & Culture, <https://goo.gl/4d81VB>.
- 2.8b Curator Table, screenshot from Arts & Culture, <https://goo.gl/FWYMMU>.
- 2.9a X Degrees of Separation, selection view, screenshot from Arts & Culture, <https://goo.gl/xtqscS>.
- 2.9b X Degrees of Separation, result view, screenshot from Arts & Culture, <https://goo.gl/Np8TDG>.
- 2.10 Panel 32 from Warburg’s *Mnemosyne Atlas*, <https://goo.gl/s8Qmz2>.
- 2.11 Screenshot from Google Street View, Nagiso, Nagano Prefecture, <https://goo.gl/AuizdZ>.

- 3.1 Perspectives in Photosynth, screenshot from video, <https://goo.gl/NXMIUF>.
- 3.2 Clement Valla, *Postcards from Google Earth*, 2010-ongoing, <https://goo.gl/7nogD5>.
- 3.3a Clement Valla, *Postcards from Google Earth*, 2010-ongoing, installation shot, <https://goo.gl/N4xDyo>.

- 3.3b Clement Valla, *The Universal Texture*, 2012-ongoing, installation shot, <https://goo.gl/xYGd14>.
- 3.3c Clement Valla, *Postcards from Google Earth*, 2010-ongoing, installation shot, <https://goo.gl/N4xDyo>.
- 3.4a Google's publishing process: collecting imagery, <https://goo.gl/oaPs9X>.
- 3.4b Google's publishing process: aligning imagery, <https://goo.gl/oaPs9X>.
- 3.4c Google's publishing process: turning photos into 360 photos, <https://goo.gl/oaPs9X>.
- 3.4d Google's publishing process: showing the right image, <https://goo.gl/oaPs9X>.

Tables

- 1.1 Temporal schemes of photography based on (non-)instantaneity and (non-)simultaneity.
- 2.1 Principles of cinematic vs. televisual spectatorship, see Friedberg, 1994, 133-137.
- 3.1 A comparative analysis between photographic images and screenshot images.

Bibliography

- Albanesius, Chloe. 2011. *Google Blocks Street View Shooter Game*. December 14. Accessed May 4, 2017. <http://www.pcmag.com/article2/0,2817,2397622,00.asp>.
- Allan, Derek. 2010. "André Malraux, the Art Museum, and the Digital Musée Imaginaire." Presentation at the symposium Imaging Identity at National Portrait Gallery, Canberra, July 15-17. Accessed May 15, 2017. <http://www.home.netspeed.com.au/derek.allan/musee%20imaginaire.htm>.
- Allen, Matthew. 2016. "Screenshot Aesthetic." In *MOS: Selected Works*, edited by Michael Meredith and Hilary Sample, 271-276. New York: Princeton Architectural Press.
- Andrews, Lew. 1995. *Story and Space in Renaissance Art: The Rebirth of Continuous Narrative*. Cambridge: Cambridge University Press.
- Anguelov, Dragomir, Carole Dulong, Daniel Filip, Christian Frueh, Stéphane Lafon, Richard Lyon, Abhijit Ogale, Luc Vincent, and Josh Weaver. 2010. "Google Street View: Capturing the World at Street Level." *Computer* 43: 32-38.
- Baker, George. 2005. "Photography's Expanded Field." *October* 114: 120-140. <http://www.jstor.org/stable/3397627>.
- Balsom, Erika. 2013. *Exhibiting Cinema in Contemporary Art*. Amsterdam: Amsterdam University Press.
- Barthes, Roland. 1982. *Camera Lucida: Reflections on Photography*. Translated by Richard Howard. New York: Hill and Wang.
- Bazin, André. 2005. *What is Cinema?* Vol. 1. Berkeley: University of California Press.
- Bobo, Ervin. 1986. "Focus on Screen Shots." *Run*, September: 40-45.
- Bolter, Jay David, and Richard Grusin. 2000. *Remediation: Understanding New Media*. Cambridge: MIT Press.
- Burgin, Victor. 2009. "The time of the panorama." In *Situational Aesthetics: Selected Writings by Victor Burgin*, edited by Alexander Streitberger, 293-312. Leuven: Leuven University Press.
- Campany, David. 2007. "Safety in Numbness: Some Remarks on Problems of 'Late Photography'." In *The Cinematic*, edited by David Campany, 185-194. London: Whitechapel.
- Carrozzino, Marcello, and Massimo Bergamasco. 2010. "Beyond virtual museums: Experiencing immersive virtual reality in real museums." *Journal of Cultural Heritage* 11 (4): 452-458. doi:10.1016/j.culher.2010.04.001.
- Cartier-Bresson, Henri. 1952. *The Decisive Moment: Photography by Henri Cartier-Bresson*. New York: Simon and Schuster.
- Cavell, Stanley. 1979. *The World Viewed: Reflections on the Ontology of Film*. Cambridge: Harvard University Press.
- Cerf, Vint. 2017. *Preserving digital art: How will it survive?* May 30. Accessed June 15, 2017. <https://www.blog.google/topics/google-europe/preserving-digital-art/>.
- Cohen, Brianne, and Alexander Streitberger. 2016. *The Photofilmic: Entangled Images in Contemporary Art and Visual Culture*. Leuven: Leuven University Press.
- Colomina, Beatriz. 1996. *Privacy and Publicity: Modern Architecture as Mass Media*. Cambridge: The MIT Press.
- Comrie, Bernard. 1976. *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press.
- Deleuze, Gilles. 1992. "Postscript on the Societies of Control." *October* 59: 3-7.
- Duve, Thierry de. 1978. "Time Exposure and Snapshot: The Photograph as Paradox." *October* 5: 113-125.
- Elsaesser, Thomas. 2011. "Stop/Motion." In *Between Stillness and Motion: Film, Photography, Algorithms*, edited by Eivind Røssaak, 109-122. Amsterdam: Amsterdam University Press.
- Ernst, Wolfgang. 2000. "Archi(ve)textures of Museology." In *Museums and Memory*, edited by Susan A. Crane, 17-34. Stanford: Stanford University Press.
- Fang, Hui. 2017. *Using Deep Learning to Create Professional-Level Photographs*. July 13. Accessed July 15, 2017. <https://research.googleblog.com/2017/07/using-deep-learning-to-create.html>.

- Foucault, Michel. 1995. *Discipline and Punish: The Birth of the Prison*. Translated by Alan Sheridan. New York: Vintage Books.
- Frampton, Hollis. 1983. *Circles of Confusion: Film Photography Video: Texts 1968-1980*. Rochester: Visual Studies Workshop Press.
- Friedberg, Anne. 1994. *Window Shopping: Cinema and the Postmodern*. Berkeley: University of California Press.
- Giorgetta, Davide, and Valerio Nicoletti. 2016. "Suddenly in the Digital." In *Photodump*, edited by Domenico Quaranta, 11-82. lulu.com.
- Google Cultural Institute. 2015. *About the Google Cultural Institute*. Accessed May 2, 2017. <https://www.google.com/culturalinstitute/about/partners/>.
- Google. 2016. *Google Cardboard Plastic*. April 1. Accessed June 24, 2017. <https://www.google.com/get/cardboard/plastic/>.
- Green, David. 2006. "Marking Time: Photography, Film and Temporalities of the Image." In *Stillness and Time: Photography and the Moving Image*, edited by David Green and Joanna Lowry, 9-21. Brighton: Photoworks.
- Gunning, Tom. 2008. "What's the Point of an Index? or, Faking Photographs." In *Still Moving: Between Cinema and Photography*, edited by Karen Beckman and Jean Ma, 23-40. Durham: Duke University Press Books.
- Henning, Michelle. 2007. "Legibility and Affect: Museums as New Media." In *Exhibition Experiments*, edited by Sharon Macdonald and Paul Basu, 25-46. Malden: Blackwell Publishing.
- Hesselberth, Pepita. 2012. "From Eubject-effect to Presence-effect: A Deictic Approach to the Cinematic." *NECSUS: European Journal to Media Studies* 1 (2): 241-267. doi:10.5117/NECSUS2012.2.HESS.
- Hoelzl, Ingrid, and Remi Marie. 2015. *Softimage: Towards a New Theory of the Digital Image*. Bristol: Intellect.
- Huhtamo, Erkki. 2010. "On the Origins of the Virtual Museum." In *Museums in a Digital Age*, edited by Ross Parry, 121-135. London: Routledge.
- Johnson, Christopher D. 2013. *About the Mnemosyne Atlas*. Accessed May 17, 2017. <https://warburg.library.cornell.edu/about>.
- Kinsley, Ben. 2008. *Street with a View (2008)*. Accessed March 28, 2017. <http://benkinsley.com/street-with-a-view/>.
- Krauss, Rosalind. 1982. "Photography's Discursive Spaces: Landscape/View." *Art Journal* 42 (4): 311-319. doi:10.2307/776691.
- Krauss, Rosalind. 1990. "The Cultural Logic of the Late Capitalist Museum." *October* 54: 3-17. doi:10.2307/778666.
- Lapenta, Francesco. 2011. "Geomedia: On Location-Based Media, the Changing Status of Collective Image Production and the Emergence of Social Navigation Systems." *Visual Studies* 26 (1): 14-24. doi:10.1080/1472586X.2011.548485.
- Larsen, Larry. 2013. *A Photosynth 2 Overview with David Gedy*. December 10. Accessed May 12, 2017. <https://channel9.msdn.com/posts/A-Photosynth-2-Overview-with-David-Gedy>.
- Lessing, Gotthold Ephraim. 1836. *Laocoon; Or The Limits of Poetry and Painting*. Translated by William Ross. London: J. Ridgway & Sons.
- Malraux, André. 1974. *The Voices of Silence*. Translated by Stuart Gilbert. St Albans: Paladin.
- Manovich, Lev. 2001. *The Language of New Media*. Cambridge: The MIT Press.
- Metz, Christian. 1991. *Film Language: A Semiotics of the Cinema*. Translated by Michael Taylor. Chicago: University of Chicago Press.
- Moholy-Nagy, Laszlo. 1969. *Painting, Photography, Film*. Translated by Janet Seligman. London: Lund Humphries.
- Morris, Robert. 1968. "Notes on Sculpture." In *Minimal Art: A Critical Anthology*, edited by Gregory Battcock, 222-235. New York: E.P. Dutton.
- Naimark, Michael. 2006. "Aspen the Verb: Musings on Heritage and Virtuality." *Presence* 15 (3): 330-335.

- Ng-Chan, Taien. 2016. "Detouring the Map." In *Engaging Affects, Thinking Feelings: Social, Political and Artistic Practices*, edited by Susan Driver, Kara Stone and Melanie Patenaude, 62-78. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Niccolai, Marzia. 2017. *Searching for art just got better. Where will you start?* May 31. Accessed June 24, 2017. <https://www.blog.google/products/search/searching-art-just-got-better-where-will-you-start/>.
- Official Google Blog. 2014. *Go back in time with Street View*. April 23. Accessed March 19, 2017. <https://googleblog.blogspot.nl/2014/04/go-back-in-time-with-street-view.html>.
- Petersen, Phred, and Andrew Davidhazy. 2013. "Streak and Strip Photography." In *The Focal Encyclopedia of Photography: Digital Imaging, Theory and Applications, History, and Science*, edited by Michael R. Peres, 616-620. Burlington: Focal Press.
- Rhizome. n.d. *Excellences & Perfections: Preserving social media with Webrecorder*. Accessed June 25, 2017. <https://www.google.com/culturalinstitute/beta/exhibit/WQJSTPJVh-JgJw>.
- Schivelbusch, Wolfgang. 2014. *The Railway Journey: The Industrialization of Time and Space in the Nineteenth Century*. Oakland: University of California Press.
- Shaviro, Steven. 2010. *Post-Cinematic Affect*. Winchester: Zero Books.
- Sreenivasan, Sree, interview by Ravi Mattu. 2015. *Museums in the digital world - 3D Printing & Virtual Reality* Financial Times, (August 20). Accessed May 31, 2017. <https://www.youtube.com/watch?v=gJuouf7jKz8>.
- Sternberger, Dolf, and Joachim Neugroschel. 1977. "Panorama of the 19th Century." *October* 4: 3-20.
- Szarkowski, John. 2009. *The Photographer's Eye*. New York: The Museum of Modern Art.
- Thielman, Tristan. 2010. "Locative Media and Mediated Localities: An Introduction to Media Geography." *Aether: The Journal of Media Geography* 5A: 1-17.
- Timeto, Federica. 2015. "Locating Media, Performing Spatiality: A Nonrepresentational Approach to Locative Media." In *Locative Media*, edited by Rowan Wilken and Gerard Goggin, 94-106. New York: Routledge.
- Uricchio, William. 2011. "The algorithmic turn: photosynth, augmented reality and the changing implications of the image." *Visual Studies* 26 (1): 25-35. doi:10.1080/1472586X.2011.548486.
- Valla, Clement. 2013. "The Universal Texture." In *Best of Rhizome 2012*, edited by Joanne Mcneil, 143-150. lulu.com.
- Van Gelder, Hilde, and Helen Westgeest. 2011. *Photography Theory in Historical Perspective*. Chichester: Wiley-Blackwell.
- Vanvolsem, Maarten. 2005. "Hinting at an experience of time in still photography." *Journal of Visual Art Practice* 4 (1): 49-56. doi:10.1386/jvap.4.1.49/1.
- . 2011. *The Art of Strip Photography: Making Still Images with a Moving Camera*. Leuven: Leuven University Press.
- Weber, Marc. 2012. *Going Places: A History of Surrogate Travel and Google Maps with Street View*. June 22. Accessed May 11, 2017. <http://www.computerhistory.org/atcm/going-places-a-history-of-google-maps-with-street-view/>.
- Williams, Raymond. 2003. *Television: Technology and Cultural Form*. London: Routledge.
- Wollen, Peter. 2007. "Fire and Ice." In *The Cinematic*, edited by David Campamy, 108-113. London: Whitechapel.