ABSTRACT

Big Data is increasingly deployed for the capturing, mapping, and analysis of the built environment. These layers of information are used to produce highly convincing representations of the built environment. This paper explores how these processes, specifically those deployed by Google Earth, are used to translate data into simulated environments with the goal of better understanding how designers might begin to produce physical constructs that provide resistance to accurate image capture and simulation. We first provide some examples of ways in which artists and designers have been approaching similar processes of image aggregation to produce new kinds of images that flatten time and hide formal specificity, then describe the processes by which one particular image aggregation engine (Google Earth) produces simulated three-dimensional environments. Finally, we present a case study in which we developed a strategy for interfering with these processes through the production reverse engineered physical prototypes.
INTRODUCTION

The role of algorithms for the collection, aggregation and translation of data into actionable outputs is increasingly embedded within the design space of politics, arts, sports and, for the purposes of this paper, architecture. In 2012, Nathan Silver’s 538 blog correctly predicted the exact Electoral College counts for the presidential election using algorithmic aggregation techniques that compiled and “translated” disparate polls to produce highly accurate projections. A decade earlier, he was one of the pioneers of the “stathead” revolution in sports that produced valuations and projections on players, based not on historically significant categories of statistics, but rather on those statistics which best predicted successful outcomes over an entire season or career. From weather to financial markets, these expansive processes of accumulation and interpretation are leading to increasingly detailed and sometimes accurate simulations of complex systems. Big Data is enabling surveillance regimes and resistance movements simultaneously as we discovered in Edward Snowden’s release of National Security Administration surveillance data and in the Gezi Park protests in Istanbul (Tufekci 2014).

At the same time, these techniques of aggregation, dissection and blending are increasingly visible in the design disciplines as processes for the production of communicative surfaces. Facebreeder, a 2004 project by Theodore Spyropoulus and Vasilis Stroumpakos, used a physical image capturing device (digital camera mounted to a head rest) to capture a series of similar if not identical photos of participants. These images were then divided using a nine square grid, separated, shuffled (while maintaining their coordinate system), and reconfigured into new facial constructions (Figure 2). These faces were projected onto a surface which essentially erased their association with the physical forms of the original faces or the devices used to capture the images. The artist Jason Salavon’s series of digital overlays of self-similar photographs take another approach. In his works “Every Playboy Centerfold” (1999) and “The Decades” (2001) and “Homes for Sale” (2002), he takes advantage of the fact that centerfolds and real estate are consistently photographed from the same distance and at the same angle to produce a series of images that provide a fuzzy accumulation of visual data that can be used to interpret trends and tendencies within subjects being investigated (Figures 3 & 4). These composite images take what are normally highly articulated images of familiar formal constructs and through the process of superimposition and blending create smoothed out blended images that allow only the shared characteristics of these subjects to become clear.
Shinseungback Kimyonghun uses customized software to capture faces from every twenty-four frames of films to produce a composite or averaged face that, in the words of the artist, “reflects the visual mood of the movie”. Perhaps the more accurate description would be that they reflect the process of their production as one that eliminates all surrounding spatial and atmospheric data in favor of human-centered data. The processes used to create this images hint at the differences between computer or Internet vision that can be very specifically programmed to capture or ignore certain types of material and foreground others. In these instances, these processes are used to produce representational outputs that are still focused on human vision. That is to say that the rapidly deployed computer vision, which queries very specific aspects of large numbers of images is being used to simulate a condition of what we might call simultaneous representation. We see all important or selected data occurring through time at once and in a simulated condition that produces a new piece of composite data.

The projects presented here produce a type of resistance to data-driven compositional algorithms, which focus on producing hyperaccurate representational spaces. The research takes a two-step approach towards this end. The first is a series of tests that establish the rules present in the algorithms deployed by Google to produce their various representational environments. The second stage of the work looks specifically at ways in which objects (and eventually buildings and even cities) could be designed in such a way as to produce “glitches” in these mechanisms. The intent is to produce anomalous outputs that could be used either as camouflage or more interestingly as emergent design systems.

This strategy of corrupting or as we would call it, introducing resistance, into computationally (digitally parametric) driven fabrication of material structures has been widely explored, notably by Andrew Kudless in his P-Wall installations, and described by Fure as Digital Materiallurgy (Fure, 2011). In this paper, we present research that focuses less on material resistance to highly precise computational protocols and tooling techniques and more on understanding how to introduce visual resistance into the large-scale data collection and representational tools.

THE WORLD ACCORDING TO GOOGLE—SMOOTHING DIFFERENCE

As the physical environment is increasingly mediated through algorithms and computational protocols, multiple levels and scales of data are superimposed to create a digital identity mapped on a single body. The digital identity of an individual is a composite dataset that defines an entity’s relationship across the numerous systems that define this identity. The built environment possesses a particular identity(s) within its physical context that is accurate to the intentions of the architect and its users and additional identities within its digital manifestations. The digital identity of the built environment is created through the use of nonhuman machines (satellites, cameras, sensors, lasers) that sense aspects of the physical environment, converting it into data to be consumed through the protocols of digital representation. Digital identity is no longer authored by intention but through the governing logics of the systems tasked with creating a virtual reality. Identity becomes subject to mediation and augmentation of the digital and as such is reconstructed in the image of the systems that create its digital identity. The conversion of the built environment into the digital is no more prevalent or ubiquitous than the one in Google Earth, a platform that explicitly molds composite datasets towards its own intentions of creating a Google-authored environment.

Google Earth is not a series of snapshots stitched together to create a singular globe. Rather, the “images” of Google Earth are composition datasets that have been algorithmically assembled and processed to create a ‘photographic’ representation of a 3D navigable environment. Generated from a multitude of privately sourced satellite images, mapping agency data, government images and documents, Google Earth creates a digitally curated photo collage so visually “precise” that it has become an influential platform for architectural site analysis. What is at stake in regards to architecture’s relationship with Google Earth is that
the highly articulated outputs of this process create such “real” representations of these datasets, that it is difficult for the consumers of these images (in our case, designers) are unable to accurately interpret or question what they are looking at. In other words, the opaque nature of these algorithms and the “realness” of the representational interface reduce the user’s ability to critically evaluate the environment in ways that a visit to the site itself might.

In order for Google Earth to create a “realistic” representation of the built environment, its governing algorithms must operate behind a veil of anonymity. However, there are moments when the algorithm exposes itself, when a breakdown in data mapping leads to a visual “glitch”, a moment when the built environment fails to be accurately represented. These anomalies create unexpected deformations within the digital identity of the physical environment. Although these deformations challenge the cognitive understanding of the built environment, the anomalies are accurate to the protocols of the system and its processing of information. Architecture is thus faced with the challenge of engaging the protocols of digital representation in order to control the amount of noise generated as computational entities attempt to recreate extant environments.

The governing logics of Google Earth create an interiority that validates the digital identity of the built environment, allowing the digital to appropriate data towards its own ends and creating a virtual representation that may differ from a physical reality. The differential virtual and physical realities present architecture with an opportunity to embrace techniques that communicate specific information towards intentional ends. As such, this research focuses on generating information that results in data overlaps and inconsistencies within the algorithms of digital representation in an attempt to expose what the algorithm actually sees. The research is an attempt at reclaiming representational autonomy over architecture digital identity, purposefully creating “glitches” to force contemplation within processes of digital production and consumption.

SIFT: THE MATERIAL OF COMPUTER VISION

This research acknowledges that there are many specific protocols within the realm of computer vision and the many machines that sense and convert the physical environment into the digital. Within this broad expanse of digital protocols, this research has identified SIFT descriptors and their implications in digital image processing and application within Google Earth as the primary focus of intervention.
Scale Invariant Feature Transformation (SIFT) descriptors convert 2D digital images into data allowing for image stitching, stereo matching, feature identification, dense image correlation, facial recognition and movement tracking (Liu, et al. 2011). Developed by David Lowe (1999, 2004), SIFT descriptors are extracted at each pixel of an input image and encoded with contextual information converting images into data for algorithmic processing. SIFT’s ability to uniquely identify geometric features throughout a series of images allows for multiple datasets to be accurately superimposed towards ends of data interpolation and alignment. It is within these protocols that SIFT descriptors present architecture with an opportunity to interfere in the formation of its digital identity, augmenting the relationships within composite datasets to increase or decrease the legibility of its digital identity.

Developed off the SIFT Flow software package run in Matlab, this research utilizes optical flow methods of matching SIFT descriptors to increase the accuracy of scene alignment between dynamic images (Liu, et al 2008). Computer images are understood as the function \( f(x,y) \), where \( x \) and \( y \) are the coordinates of image pixels. SIFT functions impose a Gaussian Scale Space to the image, adding a third coordinate \( \sigma \), which is the Gaussian kernel for an isotropic smoothing of the image. SIFT algorithms sample all pixels of an input image across their spatial coordinates \( (x,y) \) and relative to its smoothing scale coordinates \( \sigma \) generating a pyramid of progressively blurred images (Vedaldi 2013).

Adjacent images within the pyramid are subtracted from each other to produce a Difference-of-Gaussian (DoG) used for determining the maxima and minima Gaussian difference between image pixels and the resulting identification of recognizable features referred to as Keypoints. Identified Keypoints are assigned a consistent orientation based on the local properties of an image, allowing a SIFT features invariance relative to image rotation (Lowe 2004). Orientation is assigned relative to a Keypoints position within a 360 degree histogram.

In order to account for inconsistencies between multiple images resulting from different changes in 3D viewpoint or illumination levels SIFT, descriptors allow for procedures that facilitate the interpolation of non-correlative and missing data within an image set. To accomplish this task, SIFT descriptors assigns a 3D vector to an image’s pixels that can be mapped to adjacent images within a set, creating a warped image that is the result of the computer vision logics accounting for the missing data between two images. In this way, the warped image can be understood as a visualization of the algorithms at work. When sufficient data exists to accurately correlate two images, minimal distortion occurs; however, as data inconsistencies between images increase, the amount of associated noise grows. The warped image is the algorithm at work, trying to account for non-contiguous data.

---

9 SIFT Descriptor Process. (Parker 2014)

10 Peak averaging. (Parker 2014)
The algorithmic protocols of computer vision interpolation are demonstrated in Figure 12. The top two images of Casa del Fascio, taken from different perspectives, have been digitally processed removing background data to eliminate extraneous noise. Once the noise has been removed the images are converted to SIFT descriptors that facilitate accurate image filtration and alignment. The identification and correlate SIFT descriptors suggests areas of data consistency that are superimposed on top of each other. The process of superimposition creates 3D vectors that warp the base image relative to the coinciding SIFT descriptors of the input image. The final image shows the resulting image distortion, where areas with low correlating data are more drastically deformed then areas exhibiting high degrees of data correlation. The warped image is what the algorithm “sees” as it tries to interpolate missing information between images.

FROM 2D TO 3D

The use of SIFT descriptors within computer vision allows for interpolation of data; it becomes an indeterminate materiality within the governing logics of the digital. In order for architecture, a thing concerned with the creation of 3D form, to deploy this data, it must find ways to expand the capabilities of SIFT extrapolating its material effects from the 2D realm of image processing to the 3D domain of form-making. To this end, the research extracted the previously assigned 3D spatial vectors and uses them as deformers within 3D space. 3D vectors are codified relative to a HSB color wheel situated over an x,y graph with white, or the center of the wheel, located at the 0,0,0 coordinates (Figure 13). Vector direction is assigned relative to color and magnitude of the vector with saturation, creating a full range of 3D vectors which can be mapped to digital surfaces.

Mapping vectors relative to color (Figure 14) allows us to apply deformations onto surfaces, controlling the magnitude and direction of deformation relative to the contrast between hue and saturation. This process allows for controlled application of positive and negative vectors, multiple color palettes and the application of white, or a zero-vector to codify a 3D surface. It allows the design to a certain capacity that controls the anomaly within the algorithm.
ENCODING DATA TO FACILITATE ANOMALY

The authors speculate that the application of 3D vector translational data through the use of color mapping to a surface or primitive volume will allow us to influence the amount of image deformation as computer vision algorithms try to interpolate missing data within a dynamic image set. We propose that mapping 2D color to a surface is representative of the texture mapping processes of Google’s Universal Texture (the algorithms responsible for texture mapping within Google Earth’s 3D environment); however, instead of mapping photo realistic textures to laser scanned 3D meshes, we are using gradient information to transform three-dimensional surfaces and painting the resulting geometry in an attempt to create inconsistencies within the algorithm obscuring computer vision.

To test this hypothesis, we applied deformation color maps to geometric primitives. The cube with its inherent properties is predictably easy to map for computer vision. Its consistent topological relationships make the interpolation of missing data relatively accurate for the algorithm. As such, it presents an ideal geometry to test the effectiveness of our methodology. The modeled cube was rotated six degrees at a time, producing sixty snapshots throughout a 360-degree rotation. These snapshots were compared and evaluated relative to the developed methodology at different frequencies of rotation. The following results show a small sample within the larger set of more than 100 tests (Figure 15).

The resulting warped images on the right hand side of (Figure 15) show the produced anomaly. There is a massive compression of data through the warping process. The algorithm in many cases unexpectedly flattens data; however, the visual representation is inconsistent with anticipated results given previous studies using existing architectural form. The bottom two examples in (Figure 11) exhibit large dissimilarities between the two input images to be mapped yet yield the least visual deformation in the interpolated image. The previous three studies exhibit an understandable relationship relative to rotation; however, the algorithm has trouble processing and interpolating this information into accurate geometry.

(Figure 16) shows the application of the entire algorithm, the extraction and matching of SIFT descriptors, the application of 3D spatial vectors and the resulting warped image. The algorithm is functioning exactly as expected at all stages; however, the resulting warped image is inconsistent...
with the anticipated results relative to the 3D spatial vectors or the image. In this case, the resulting image can be considered a “glitch”. The geometry of the cube is successful in creating inconsistencies within the algorithm’s ability to interpolate missing data. These images represent an exposure of the algorithm’s vision. Inconsistent and alien to our own, mediated through computation, the unexpected representation is not a glitch, it is the difference in data processing between the algorithm and the anticipated result congruent with the physical world. It is the compression of data into a unique anomaly.

CONCLUSIONS AND IMPLICATIONS

The research presented here has focused on the production of techniques for exploiting the processing capacities of big data machines like Google in order to produce formal variation that may be resistant to assimilation within this same machine. The research and data produced here focus on an image-based paradigm for achieving this. While the digital models were produced in three-dimensional model space, they were evaluated by the algorithms in their two-dimensional state as renderings of those volumetric cubes. Additional research that involved material fabrication of these models (Figures 17) is ongoing but reinforces the
findings of the two-dimensional simulations. This research has developed a set of techniques that could form the basis for a more rigorous experimentation in the production of form finding techniques that introduce variability into the architectural project with the performative function of data camouflage. In addition to the building code, energy and programmatic criteria that increasingly make up the parametric generative model of architecture, we could begin addressing the issue of formal languages that resist capture by Google cars, Flickr accounts and other data collection sources based on visual data. These stealth buildings could begin to re-privilege proximate human experience of architecture over the mediated street view. The anonymous building might no longer be that, which blends seamlessly into its built context, but rather one that produces formal anomalies meant to allow for disappearance or obfuscation within the context of massive data clouds.

Of course the more likely use of these techniques is towards the production of first two-dimensional representations that combine large pools of images of self-similar (in terms of program, massing, material etc.) architectural projects into what we might call a “BuildingBreeder” Google interface. These images could then be reversely engineered to produce interior spatial conditions based on similar combination strategies using images of hundreds of floor plans. This could lead to a self-reinforcing Parametricism of reduced anomalies and decreased innovation. This research is advocating for the individual sporting event, in which the anomaly of an unexpected outcome is possible as opposed to the view of a large body of self-reinforcing and increasingly predictable outcomes of large-scale data simulations. The argument for resistance through anomalies is one that embraces the technology that makes both the homogenous and the heterogeneous futures possible as a means for hacking the future.
REFERENCES


IMAGE CREDITS

Figure 1. Parker, Matthew and Johnson, Jason (2014) Digitally Fabricated Prototypes.

Figure 2. Spyropoulus, Theodore, Stroumpakos, Vasilis (2004) Facebreeder, London.

Figure 3. Salavon, Jason (2002), Every Playboy Centerfold, The Decades (normalized).

Figure 4. Salavon, Jason (1997-2002), Homes For Sale.

Figure 5. Kimyonghun, Shinseungback (2013), The Matrix, Portrait.

Figure 6. (Image retrieved from Google Earth, 04/04/2014).

Figure 7. (Image retrieved from http://rhizome.org/editorial/2012/jul/31/universal-texture/, 14/01/2014).

Figure 8. (Image retrieved from http://rhizome.org/editorial/2012/jul/31/universal-texture/, 14/01/2014).

Figure 9. Parker, Matthew (2014) SIFT Descriptor Process Diagram.

Figure 10. Parker, Matthew (2014) Peak Averaging Diagram.

Figure 11. Parker, Matthew (2014) Keypoint Vector Extraction and Matching.

Figure 12. Parker, Matthew (2014) Casa Del Fascio: Case Study.

Figure 13. Parker, Matthew (2014) HSB Color Wheel Vector Map.

Figure 14. Parker, Matthew (2014) Color Displacement Map.

Figure 15. Parker, Matthew (2014) SIFT Glitch.

Figure 16. Parker, Matthew (2014) Compressions and Anomalies.

JASON SCOTT JOHNSON has practiced and taught architecture in North and South America and Europe. He is a founding partner of the Laboratory for Integrative Design at the University of Calgary, where he is an Assistant Professor in the Faculty of Environmental Design. Jason is the director of the research-based Minus Architecture Studio.

MATTHEW PARKER has worked as a research assistant for the University of Calgary’s Laboratory for Integrative Design (LID), contributing to competitions and publications that have been internationally recognized. His current research focuses on computer vision, SIFT materiality and the protocols employed by digital systems that sense and convert physical space into data for computational processing. Parker is developing strategies that allow for codified architectural surfaces to exchange information between the multiple systems that govern architecture’s physical and digital identities. Matthew Parker holds a Masters of Architecture degree from the University of Calgary and is currently engaged in a Post-Professional Thesis undertaken through the University of Calgary’s Faculty of Environmental Design.