

**GROUND TRUTH:  
ON THE CONTINGENCY OF MACHINE  
VISION AS SUBSTITUTE FOR VISUAL  
COGNITION INSIDE THREE-DIMENSIONAL  
INTERMEDIATE SPACES**

Aral Cincim 2019/2020

*The entire text is available for download at <https://www.aralincim.com/Ground-Truth>*

*This article was presented as part of my master graduation work, the text in itself (as of WS2020) is an ongoing component of my research on spatial cognition inside computational environments*



## An Unapt Introduction to Depth Sensing and Spatial Cognition

This work in its studious compendium attempts to substantiate a virtual-spatial environment, which relies on the human visual system and its cognitive components, in particular, the presumably indigent human sensory cognition and/or the mechanisms of the human eye as an optical interface through the biological augmentation of sensory data elucidation. The primary concern of this project has been prefabricated into multifarious volumes, contingent on the functions of the apprehension of visual information and the faulty thresholds of the human eye as an optical system which relies on the psycho-visually redundant properties of quantizers<sup>1</sup>, such as the subspace approach in spectral color science<sup>2</sup>, through the digitization of transmitted sensory data. This integral process of spatial cognition and *it* being devoured constantly into segments by the peripheral nervous system has certain similarities, when one has the impetus to correlate this structure with reference to our occipital lobes, where visual stimuli would have been enumerated later to be unraveled by the retinofugal projection along the optic tract.

The procedure itself reconnoiters the transpiring habitation of head mounted displays (HMD) being attached to depth sensors, as a utility of three-dimensional structural mapping<sup>3</sup> in respect to machine vision and its instrument-environment correspondences in the field of stereoscopic projection where the extent of our interpretation of reality could perchance be transformed into the perspective of an autonomous and possibly deceptively latent video feedback. This appliance, as a deduction, formulizes the triangulation of stereoscopic image displacement in flat co-planar moving images. At this early development stage, as of mid-April 2019, I have been experimenting with a primitive HMD-smartphone setup, inside which an APK in Unity real-time development platform running Google's ARCore<sup>4</sup> implements the back-facing camera for plane detection as a monoscopic utensil which envisages two-dimensional perspective as moving images in real-time.

The comparison between human spatial perception regarding vision and the systematic of machine-computer aided visual environments provides the fundamental basis of this project, which intentionally does not emanate from a scientific standpoint but nevertheless I glean the necessity to amalgamate disciplines of formal sciences *apropos* of geospatial informatics as a

---

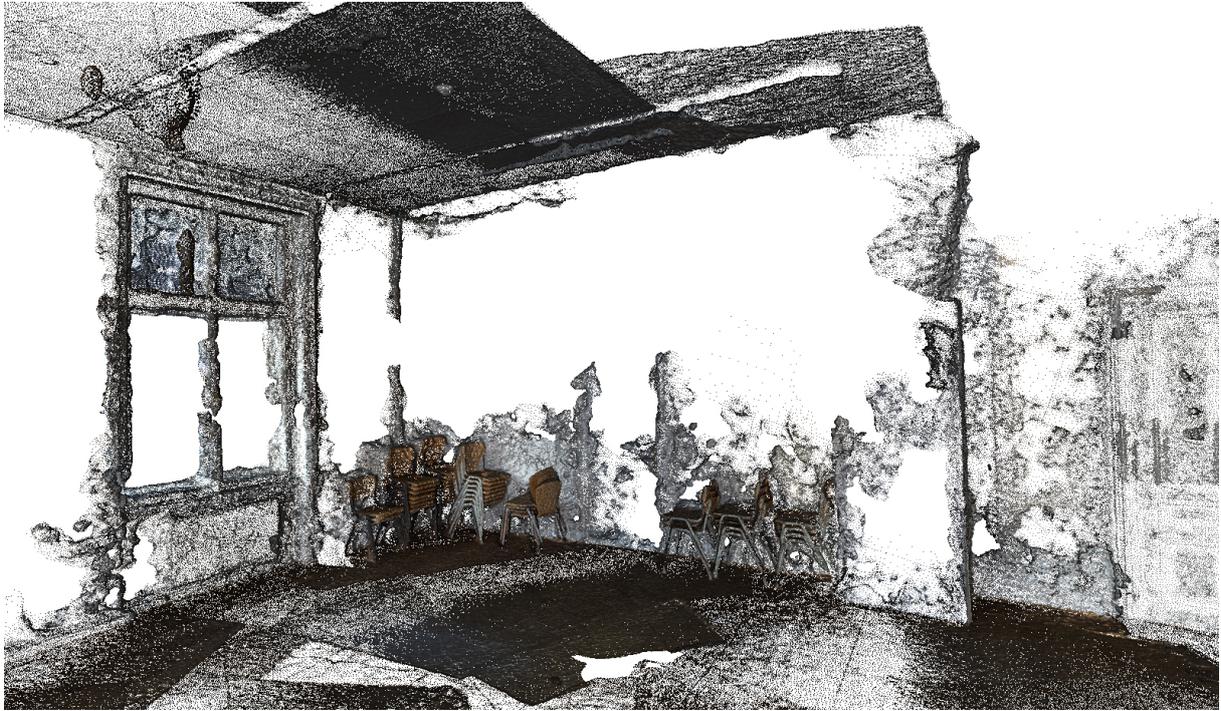
<sup>1</sup> Quantizers map an input amplitude to an output amplitude, as a device or as an algorithmic function

<sup>2</sup> Advances in Independent Component Analysis and Learning Machines, Chapter 10, Academic Press, (2015)

<sup>3</sup> Structural Mapping is often used in structural geology for terrain identification purposes in field analysis

<sup>4</sup> Google's platform for constructing augmented environments, enabling the smartphone of the user to sense its environment and interlace with the stipulated visual information, <https://developers.google.com/ar/discover/>

methodological tool, strictly speaking, this auxiliary tool will entirely be subsumed within the retrieval span of the 3-D rearrangement of an epipolar geometry (points, polygons, vertices and so on) inside the current exhibition space.



*Figure 1: Dense Point Cloud Agglomeration Module (DPCAM) consisting of 3.648.449 vertices after numerous iterations with the Kinect v.1 Sensor, Real-time Appearance-based Mapping of Room 201, Vordere Zollamtstraße 7, Aral Cimcim, December 2019*

*“The human visual system constantly adapts to different luminance levels when viewing natural scenes. We present a model of the visual adaptation, which supports displaying the high dynamic range content on the low dynamic range displays. In this solution, an eye tracker captures the location of the observer’s gaze. Temporary adaptation luminance is then determined as the impact of the light area surrounding the gaze point. Finally, the high dynamic range video frame is tone mapped and displayed on the screen in real-time. We use a model of local adaptation, which predicts how the adaptation signal is integrated in the retina, including both time-course and spatial extent of the visual adaptation. The applied tone mapping technique uses a global compression curve, the shape of which is adapted to the local luminance value. This approach mimics a natural phenomenon of the visual adaptation occurring in human eyes.”<sup>5</sup>*

There have already been experiments in the academic territory of video formats related to computer graphics and image processing, which have been establishing themselves exponentially,

---

<sup>5</sup> High Dynamic Range Video, From Acquisition to Display and Applications, Editors - Dufaux F., Le Callet P., Mantiuk R., Mrak M., Academic Press, ISBN: 9780081004128, (2016), pp. 347-369

such as HDR - high dynamic range video, where tone mapping is an essential condition for being viewed on a display device that circulates on a limited color gamut and dynamic range. The increase in the advancements in the new adaptations of visual sensory information has already proven the imperative exigency for an enhanced state of human visual buffer<sup>6</sup>, beyond the dynamic scope of what the evolution of the human eye has inaugurated so far, though not far enough according to Dan-E. Nilsson and Susanne Pelger, who in their research on the future biological development of the human eye propose that “it would take less than 364,000 years for a camera eye to evolve from a light-sensitive patch”<sup>7</sup>

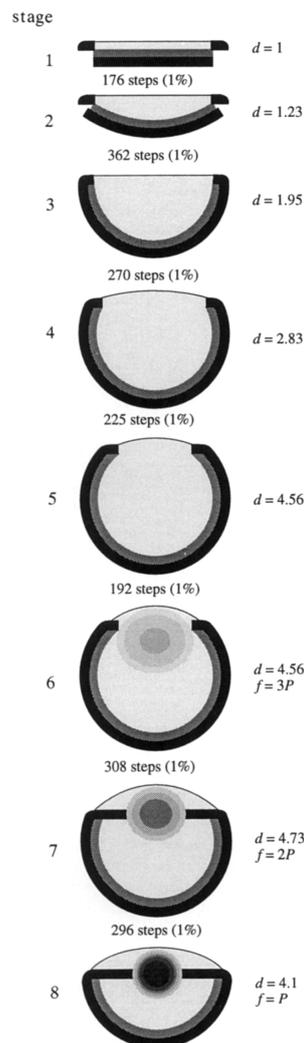


Figure 2: Representative Stages of an “erroneous” Model Sequence of Eye Evolution according to D. Nilsson and S. Pelger

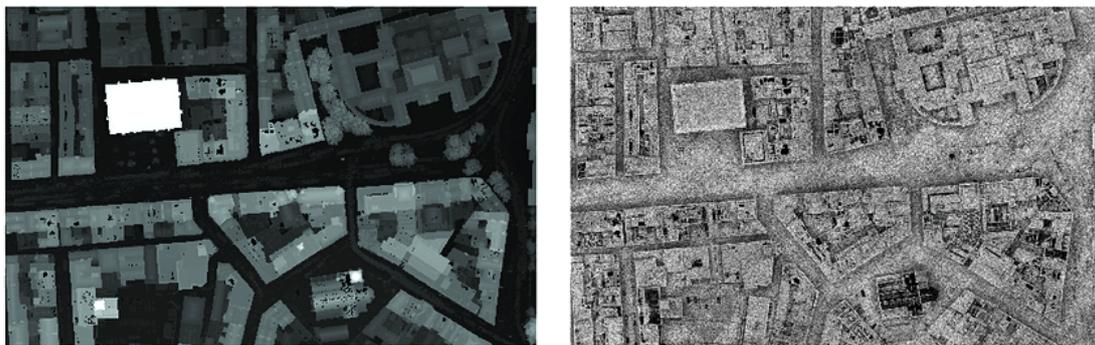
<sup>6</sup> Visual Images Preserve Metric Spatial Information: Evidence from Studies of Image Scanning, Stephen M. Kosslyn, Thomas M. Ball, Journal of Experimental Psychology: Human Perception and Performance (1978), Vol. 4, No. 1, pp. 47-60, *The quasi-pictorial theory of imagery* describes the methodology of a visual buffer with surface and depth representations being two of the fundamental components according to Kosslyn’s work on the matter, Image and Mind, Harvard University Press, (1980)

<sup>7</sup> A Pessimistic Estimate of the Required Time for an Eye to Evolve, Dan-E. Nilsson, Susanne Pelger, (1994)

Imperative schemes of stereo imaging, pointedly stereoscopy<sup>8</sup> and the utensils of depth recovery have been employed henceforth the analytical transformation of the human mind, as early as, according to cognitive archeologists, the initiations of the pictorial occlusion images in Paleolithic cave paintings signifying the artistic capabilities of the, then prehistoric beings, regardless of their construction of the huddled, primitive knowledge and the paradigmatic piquancy that might be evaluated from this primeval example of a galactic zoo, until this age of computational informatics, much has changed in the peripheral field of view of the human eye, in regard to inspecting the duality of still images side by side and being under the enchantment of this somewhat optical phantasmagoria of the three-dimensional behavioral modernity that encompasses us.

*“Though we have considerable information about brain sizes in fossil species, and a little information about brain shapes, the relevant anatomical information, the internal micro-architecture of these brains, has left no fossil trail. With respect to fossil brains, we will never find the “smoking gun”-the first brain capable of language. We will only have access to circumstantial information.”*<sup>9</sup>

Casting some artificial light through the entrance of the *Grotte de Lascaux*<sup>10</sup> on the denouement of human cognitive skills, the question that intrigues me the most in this relation to human biology, possibly, is the bodily progression of the projection of a mental image, *videlicet*, the prehistoric development of language and the paradigms attached to it. Though somewhat misleading might be the nineteenth century theories of the phylogenic rankings of species and the



*Figures 3-4: Visualization of Urban Airborne Laser Scanning Data with Occlusion Images*

*Courtesy of ISPRS Journal of Photogrammetry and Remote Sensing*

*104:77-87, June 2015*

---

<sup>8</sup> Stereoscopy can be reckoned as the annexation of the sense of depth onto two-dimensional placidly still imagery, the term dates back to the eighteenth century achievements in the field of optical physics by Charles Wheatstone and David Brewster

<sup>9</sup> *The Symbolic Species, The Co-Evolution of Language and The Brain*, Terence W. Deacon, (1997)

<sup>10</sup> *Grotte de Lascaux*, Paleolithic cave situated in southwestern Dordogne region of France, where a significant collection of nearly 600 prehistoric paintings were discovered in 1940, GPS Coordinates: 45°03'13"N 1°10'12"E

addendum of the theories on humans being the more advanced, articulate and capable of *all*.

According to neuroanthropologist Terrence W. Deacon, neural sciences while on the one hand, researching the ways in which the human brain works, solved some facets of the mysteries on how the brain functions and have already mapped some of the cognitive tasks to associated regions of the brain, following the trail of this data, computational simulations of neural networks<sup>11</sup> have already been constructed, on the other hand there has already been a lacking prospect of theorizing the pattern of the nature of the human mind in the origins of language which would then be applied on a global scale, not all the intersections of how the human mind works are clearly defined. He also suggests that over a period of several million years the brain size of our ancestors has expanded three times as of before due to the growing use of symbols as in a set of relationships between interpretative conditions of antecedently acquired stimuli in a precognitive fashion. What Deacon formulizes is that the responses that are developed as part of daily tasks in accordance to associative learning are results of a hierarchical interpretative process.

Contextualizing a visual system of symbols as in the circumstances of large data sets, namely Big Data<sup>12</sup> MapReduce<sup>13</sup> models consisting of the map, shuffle and reduce functions seem not so far-fetched than theorizing a descriptive model of human cognitive skills, in the sense that the cross-section of anthropology, artificial intelligence and neuroscience have been adjacent to the representational structures of the mind. Inexorably the fundamental topic of this research does not focus on the formation of a language system as part of analytic philosophy, rather on the sampling of, for the purpose of this introduction, the analytical evaluation of the algebraic perspective of our encirclements and our indispensable positions therein. The expansion of depth sensor<sup>14</sup> technologies and structural mapping devices using remote sensing methods have proven to be, in regard to the surface patterns and the source of emitted-refracted light that clouds the

---

<sup>11</sup> The neural system consists of three components: receptors, effectors and neural network. The receptors reap the stimuli either within or from the periphery, afterwards they deliver the data into the neurons in the dispersal of electrical impulses

<sup>12</sup> Big Data refers to the extraction of structured or unstructured dispatches from voluminous data sets that are too complex to be summated computationally by conventional data-processing software

<sup>13</sup> MapReduce is the software framework for compiling applications which refine ample amounts of multi-terabyte data on large clusters of hardware in a tolerant and dependable demeanor

<sup>14</sup> Depth Sensors have been used extensively for facial detection, recognition and verification by Smartphone manufacturers and telecommunications equipment companies, i.e., the camera module from Qualcomm Incorporated can sense three-dimensional depth in real-time and simultaneously generates a 3-D point-cloud of data in commuting industrial and domiciliary situations. Microsoft's Kinect devices, also the one that I have been using in this project can generate real-time point clouds with the combination of a depth sensor and a high definition camera module with the SDK developed by Microsoft, though the Kinect versions 1&2 are not being manufactured anymore, instead the new iteration currently in development is named as Azure, also being used in the HoloLens mixed reality glasses as an allotment of Project Baraboo and HMDs

peripheries that are being rendered in terms of spectral, temporal and spatial resolution in global scales, powerful tools of visualization. Following the recent scientific developments in the field of machine vision and relying on similar computational variables where cameras that are part of a machine vision system<sup>15</sup> (MVS) are constantly recognizing and estimating the values of the digitalized input data, in other words, these systems are substitutes of the human visual sense and judgment capabilities with recording devices, a camera instead of an eye.

It has not only been about the devices themselves but also about what these substitute machines are capable of, in a more factually logical paragon, what are not humanly discernable but “experienceable”, such as electromagnetic waves, gamma rays, infrared light, signals with different wavelengths inside the electromagnetic spectrum that are fluctuating spatially between consecutively corresponding points. Jim Bell, the lead scientist for the Pancam color imaging system<sup>16</sup> which was used on the twin NASA Mars exploration rovers (MER) – Spirit and Opportunity, which were robotic field geologists equipped with stereoscopic cameras, points out that they actually try to avoid the term “true color”<sup>17</sup> since nobody precisely knows what the “truth” is on Mars, in virtue of the dynamic properties of color. Incipiently true color imagery was designed to display our environments in color similar to what our eyes naturally see. It is primarily dependent on the RGB color model with an improved green channel that is facilitating the rapid delineation of surface types, providing the ability to distinguish differences in surface geometry and environment variables in distance-dependent spatial imaging. From a skeptical exemplary standpoint, with reference to appearance, reality and the specification of true color, how visual information is perceived is not exactly how material-paraphernalia really are or simply what we optically perceive does not exactly mean that what we see is reality withal color realism<sup>18</sup> in regard to the reciprocal true color plights being ambiguously subjective.

---

<sup>15</sup> “Where human vision is best for qualitative interpretation of a complex, unstructured scene, machine vision excels at quantitative measurement of a structured scene because of its speed, accuracy, and repeatability. For example, on a production line, a machine vision system can inspect hundreds, or even thousands, of parts per minute. A machine vision system built around the right camera resolution and optics can easily inspect object details too small to be seen by the human eye” Introduction to Machine Vision, Cognex Corporation, Retrieved from the Internet Article, <https://www.cognex.com/what-is/machine-vision/benefits>

<sup>16</sup> Pancam imaging system is the panoramic camera system mounted on the rover maneuver point at average human-eyesight level that is capable of 1024 x 1024 pixels or 20/20 human vision

<sup>17</sup> A case for natural color imagery from geostationary satellites, and an approximation for the GOES-R ABI, Steven D. Miller, Christopher C. Schmidt, Timothy J. Schmit, Donald W. Hillger, International Journal of Remote Sensing (2011) / True color is also sometimes regard as the specification of the color of a pixel on a display screen using a 24-bit value, which allows the possibility of up to 16,777,216 possible colors

<sup>18</sup> “If someone with normal color vision looks at a tomato in good light, the tomato will appear to have a distinctive property - a property that strawberries and cherries also appear to have, and which we call “red” in English. The problem of color realism is posed by the following two questions. First, do objects like tomatoes, strawberries and radishes really have the distinctive property that they do appear to have? Second, what is this property?” Byrne & Hilbert (2003), p. 4

Alex Byrne and David Hilbert<sup>19</sup> in their joint article on the realistic properties of color, attempt to define how color science grapples with the precarious problematic of physicalism<sup>20</sup>, the entailed supervenience of everything. Physical attributes of objects that are responsible for the appearance of color have already been extensively scrutinized, the processing of color in the visual system, the evolution of color vision and its deficiencies, atypical conditions of viewing true color to name a few. According to scientists engaged intensively in the field of color and cognitive neurology, one of the mostly conferred answers for the problem of color possibly is that it is the product of a process taking place in the brain in a still-not-exactly-pinpointed location, therefore surmising the fact that nothing is in reality colored. Neurobiological overtures on the mechanisms of color suggest that the human nervous system instead of analyzing colors, exerts itself to obtain visual data in outer, extraneous milieux and reckons confiding on the appropriated data of the surface structure and the wavelengths of the refracted light in those outlying frameworks to resemble an appraisal of color.

Scholars surveying the sundry sub-branches of vision science might have other configurations of completion, the scholarly approach of a physicist when inquired to measure the color of an appreciable object would highly likely be attempting to grasp the optical properties of the material whereas a psychologist would investigate the processes supervening inside the consciousness of an observer/perceiver, where a non-analytic philosopher<sup>21</sup> would fixate on the visual awareness of the object and the sense datum<sup>22</sup> adhered to it rather than affirming that the phenomenon of a circular afterimage<sup>23</sup> constituting inside an optical system might be responsible for an elusive neural event without remarking the cross-platform intrusions of scientific reasoning or as Bram Stoker put it, “...it is the fault of our science that it wants to explain all; and if it explain not, then it says there is nothing to explain.”<sup>24</sup> Deeming the necessity of the possibilities of an anti-dispositionalist manifesto<sup>25</sup> indicated by Byrne & Hilbert on the scientific fact that colors might be secondary qualities in respect to the research conducted during the seventeenth and eighteenth century,

---

<sup>19</sup> Color realism and color science, Behavioral and Brain Sciences (2003) 26, 3–64, Byrne & Hilbert

<sup>20</sup> “Physicalism is not a doctrine about universals or other abstract objects, but about the empirical world, and specifically about minds. It says that mental entities, properties, relations and facts are all really physical. The mental is physicalism’s chief target; but one we think it does not hit” There is No Question of Physicalism, Tim Crane & D. H. Mellor, Mind, New Series, Vol. 99, No. 394 (1990), pp. 185-206

<sup>21</sup> Non-analytic philosophy also described as continental philosophy, focuses mainly on language and meaning rather than embracing the construed prospect of natural sciences

<sup>22</sup> Sense data are sensations that are detached from any particular information that they might be conveying from their originating source in the external, material world

<sup>23</sup> Afterimages are basically what the brain sees after the eyes stop looking, appearing in complementary shapes such as hexagons, derived from the German word “Nachbild”, (1875-80)

<sup>24</sup> Dracula, Bram Stoker, (1897), p. 272

<sup>25</sup> With a strong propensity against the virtue of situationism and the person-situation debate associated to it, an imperative remark on this topic might be the Stanford prison experiment supervised in 1971 by Philip Zimbardo

where it has been rephrased and henceforward controverted that it is based on an erroneous promise of perception, broaches one into the Kantian question of primary quality distinction, the credo of transcendental idealism.<sup>26</sup>

In Critique of Pure Reason, Kant bases his argument on the premise that human experience of materiality can be reduced to an appellation which he defines as “transcendental unity of apperception”<sup>27</sup>, the philosophical theory associated with this understanding of the material world that surrounds us as human beings is postulated on the fact that this material world exists inside a certain space and we, our “selves” delineate paths through this world, as he contemplates in the chapter of Refutation of Idealism<sup>28</sup> declaring the existence of objects in an external space retaining the qualities of either being indemonstrable or false. Reexamining the notion of supervenience physicalism<sup>29</sup>, pursuant to counterfactual philosopher David Lewis, the global property of a dot-matrix image does have certain components, such as symmetry and clutter while being composed of dots and non-dots at each point of the matrix, corresponding of nothing but a dot pattern on a global scale. The spatiotemporal arrangement of points<sup>30</sup> as a three-dimensional dominion, inside which each single point has its own properties instantiated either by variables such as mass and color or by the properties of the point itself, has an underlying similarity with the generation of point cloud compare software through which a polygonal mesh is created by photogrammetric methods and matching algorithms, a subcategory in respect to airborne spatial data is OPALS<sup>31</sup>, the processing chain for acquired data using laser scanning methods such as LIDAR<sup>32</sup>, the package-oriented modules of the software is being used for trajectory georeferencing. The retina layer of the human eye perpetuates unswervingly from the brain cells, enclosing the endowment to convey the convened light rays into neural signals along the optic nerve path, functioning moderately similarly to the solid-state image sensors of the twentieth century that revolve around the optical components which are embedded in image recording devices.

---

<sup>26</sup> Transcendental idealism also regarded as formalistic idealism is the doctrine founded by German philosopher Immanuel Kant in the eighteenth century basing time and space on human concepts of perception

<sup>27</sup> Critique of Pure Reason, Immanuel Kant, (1781-1787)

<sup>28</sup> Kant unveils this term in the preface of Critique of Pure Reason highlighting that the things that exist outside us must be beheld only as faith and if someone doubts this assumption then “we” would skirmish with the counteract of satisfactory proof

<sup>29</sup> Metaphysical approach attributed to the ancient Greek philosopher Thales also to Berkeley in the eighteenth century, nature is a single substance and thus water or basically everything is mental according to Berkeley’s idealism

<sup>30</sup> On the Plurality of Worlds, David K. Lewis, (1986), pp. 14-17

<sup>31</sup> OPALS stands for Orientation and Processing of Airborne Laser Scanning, it is bifurcated into diminutive modules which can be combined freely for large data-sets, i.e., the grid system *opalsGrid* can be used with simple moving-plane methods of *n* number of adjacent neighbors on a surface

<sup>32</sup> LIDAR stands for Light Detection and Ranging, used as a pulsed laser utility for remote sensing applications

Electrophysiologists use the term “single-cell recording”<sup>33</sup> to denominate the functions of the brain and its receptive properties in alternating environments, capturing the changes during neural activity and locomotion of the visual cortex. How the brain represents information along the visual pathway by the cells encoding the transmitted visual data through neural ensembles has been used as a fundamental point for analyzing human cognition and creating a cortical mesh that can be used in brain computer interfaces (BCI). Gesture based computer interaction has proven that it can be a naturally intuitive method for the communication between machines and humans thanks to the scientific progress in the field of ubiquitous computing and vision-based methods such as eye tracking, rendering human body parts as receptacles that contain the interface information. Appliances akin to Kinect work as three dimensional cameras that can capture a stream of pixels containing color data about the depth of each pixel which contains the values of the distance between the sensor and the user operating in the same direction. In the view of the verity that BCIs and machine-generated visual experiences create an alternative space which can then be regarded as a multi-sensory lattice of Xs, Ys and Zs, a reciprocal territory of transmogrification and therefore according to an applied mathematical approach<sup>34</sup>, might be considered as an appending temporal space, a space that is in between two extremes, *neither here nor there*. How physical space is represented in the brain has been investigated by many in the past five decades, contributing to the discovery of grid cells, place-modulated neurons with multiple firing locations that convey a triangular array which covers the entire surface of a two-dimensional space, being pondered as an important ingredient for the metric navigation skills of the brain.

The principle assumptions of Aristotelian philosophical subdivisions of terrestrial and celestial<sup>35</sup> were based on acumens, as it were, if something displaces itself, it is because something else disarranges it. In consonance with these assumptions of corporeal movement inside a three-dimensional coordinate geometry, it is axiomatically possible that the forces that impel movement cannot be executed inside the boundaries of any void. The existential characteristics of the admeasurements of semantically generated point cloud environments prove that “self-supervised, machine-guided reconstruction methods”<sup>36</sup> have been a plausibility where the physical subdivisions of the description of intermediate spaces fail to administer us with adequate spatial unravelments.

---

<sup>33</sup> The method that is used to observe the current or voltage inside a single neuron by inserting depth electrodes

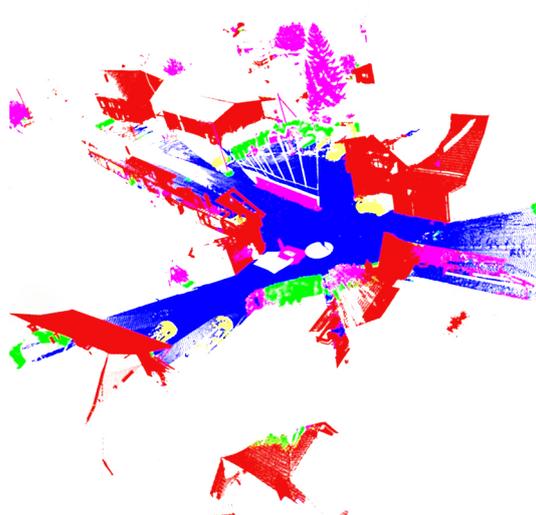
<sup>34</sup> Short Lecture Notes: Interpolation Theory and Function Spaces, Helmut Abels, University of Regensburg, (2011)

<sup>35</sup> Aristotle Physics, Books III and IV, Oxford University Press, (1983) p. 42

<sup>36</sup> Self-Supervised Deep Learning on Point Clouds by Reconstructing Space, Jonathan Sauder & Bjarne Sievers, Hasso Plattner Institute Potsdam, (2019), retrieved digitally from <https://arxiv.org/pdf/1901.08396.pdf>



Figures 5-6: Rapid Prototyping with a Raspberry Pi 3 – Smartphone HMD Configuration running on Latent Wireless Optical Feedback, Aral Cimcim, May 2019, Vordere Zolamtsstraße 7, Vienna



Figures 7-8: Semantic Segmentation of a Point Cloud, Dense Cloud with RGB (left) vs. Prediction Map with PointNet implementation (right), Images courtesy of the open-source library of the Open3d Frontend, <http://www.open3d.org/>

(The circulation of data acquired by the workstation-setup which is being transmitted (as of January 2020) to the HMD has been performing with a slight delay due to the length of the USB-3 cable (15m USB-A to B) since the depth sensor I've been using was manufactured with the intention of *it* operating as a static object rather than working inside a highly kinetic environment.)

*“Animals with eyes like ours evolved from ancestors with no eyes at all. An extreme saltationist<sup>37</sup> might postulate that the evolution took place in a single mutational step. A parent had no eye at all, just bare skin where the eye might be. He had a freak offspring with a fully developed eye, complete with variable focus lens, iris diaphragm for 'stopping down', retina with millions of three-color photoreceptors, all with nerves correctly connected up in the brain to provide him with correct, binocular, stereoscopic color vision.”<sup>38</sup>*

The operative mechanism of the ocular system of any specie functions with the actuality of light and its importance on the perception of space is covered massively by eminence<sup>39</sup>. Light as a form of electromagnetic radiation which occurs over an immense range of wavelengths, propagating at certain amplitudes and thresholds transmits spatial and temporal information which coincides with the biological evolution of the optical weldment of the human eye, therefore permeating its layers and attaching a substantial meaning to anything surpassed by it. However the wave theory of light has not been proven to be sufficient enough for light travelling in low intensities, since light being part of the electromagnetic spectrum, which is the collection of radio waves, x-rays, visible light and gamma rays, has long been debated whether to be described as a particle or as a wave composition.

*“Many animal tissues/cells are photosensitive, yet only two types of photoreceptors (i.e., opsins and cryptochromes) have been discovered in metazoans. The question arises as to whether unknown types of photoreceptors exist in the animal kingdom. LITE-1, a seven-transmembrane gustatory receptor (GR) homolog, mediates UV-light-induced avoidance behavior in *C. elegans*. However, it is not known whether LITE-1 functions as a chemoreceptor or photoreceptor.”<sup>40</sup>*

The whole case of the investigations with physiological optics inside the spheres of research on the mechanisms of living organisms has been a briskly expanding exemplary field, since the early implementations on the topic of depth perception in stereo vision and the photosensitive

---

<sup>37</sup> Reciting on the theories that the evolution of species emanates in major stages by the sudden transformation of ancestral species into a descendant of a different strain, rather than by the gradual accumulation of small alterations

<sup>38</sup> The Blind Watchmaker, Chapter 9, Richard Dawkins, scandalous collaboration story of the computational model of Nilssen & Pelger is mentioned here in detail, <https://evolutionnews.org/2018/05/eyes-in-a-twinkling/>, David Berlinski's comments on the issue can be found in the Commentary magazine article available under the webpage: [https://www.researchgate.net/profile/Nicholas\\_Matzke/publication/298133093\\_A\\_scientific\\_scandal/links/57d114d708ae5f03b489357a.pdf](https://www.researchgate.net/profile/Nicholas_Matzke/publication/298133093_A_scientific_scandal/links/57d114d708ae5f03b489357a.pdf), Commentary July-August (2003)

<sup>39</sup> Eminence as in an elevated surface topography, a piece of rising ground

<sup>40</sup> The *C. elegans* Taste Receptor Homolog LITE-1 Is a Photoreceptor, Cell Press, an Imprint of Elsevier, (2016)

layers of the mundanely curved retinas of the mammals with neocortices<sup>41</sup>, in addition to, more to be discovered photoreceptors of metazoans and the non-biologically ramifying effect of stereoscopic photographs<sup>42</sup> of the mid-nineteenth century have all contributed to the ontogeny of new disciplines on the visual cognition of three-dimensional, data-based spaces. Approximating the terminologies while being cognizant of mediated spaces and the media that are infused to these multifarious habitats are the results of the consolidation of the public and private spaces, as it were, the organic epitome of synecdoche and the auxiliary, topological manifolds of acuteness.

Indulging deeper into the granulation of human stereoscopic vision where, theoretically speaking, the visceral segment that renders the whole process of viewing has been based on the deductive fact that the technical media of reproduction of the twentieth century is based upon light as a recording source of a prodigious quantity, throughout history after the invention of light-sensitive materials and photomechanical abraders such as the dynamics of the camera obscura<sup>43</sup> and its real-time, upside-down delineation of the three-dimensional macrocosm of *je ne sais quoi* that inundate us.

This fact, on its own, when merged with the quotation related to the scandalous collaboration between Dawkins and the pessimistic, computer-generated model of the research on the future development of the human eye from a single light-sensitive patch, at the beginning of this text (p. 5) makes me cogitate, in a captivating demeanor, how would all these devices of reproduction evolve in time, when we consider that we, as humans with the probability of serious genetic mutations such as neonatal retinoblastoma<sup>44</sup> or color vision deficiency<sup>45</sup>, if our optical perception had been transubstantiated anamorphically similar to that of the eyesight of the cyclopes in Greek mythology, the giant one-eyed beings e.g. *Polyphemus* with, what I presume, a greatly embellished sense of orthographic projection or what if any algorithmic adoxography was read with a camera-like eye in a time period, in the distant future, following 364,000 years thenceforth, how would it have been *there and then* elicited ?

---

<sup>41</sup> The Neocortex is a deeply coordinated structure present in all mammals that processes sensory, motor, language, emotional and associative data, being the largest part of the mammalian brain this structure commands functions such as cognition, spatial reasoning and sensory perception

<sup>42</sup> A planar image with a picture plane does not emulate exact distance or size, it does not distinctly separate every surface of the stratum which the image is composed of and it surely does not exhibit the depth outline of an object

<sup>43</sup> Nineteenth century optical device often used by artists to make quick sketches in the field, originally used with a screen by scientists, mathematicians and astronomers in the eleventh century alongside with the pinhole camera

<sup>44</sup> Neonatal cancer is defined as a malignant and abnormal growth that is diagnosed during the first four weeks of life, developing from the immature cells of prenatal retinas

<sup>45</sup> Red-green color deformities are the most common form of color vision deficiency related to the differentiation between shades of red, yellow, and green. Blue-yellow color vision defects, which are more uncommon, cause issues with differentiating shades of blue and green also causing difficulty demarcating dark blue from black

Optimal encoding of visual orientation data-memorandums is conditional to the realistic representations of natural environments to reduce the disparities between the palpable aspects of spatial tectonics and common materiality of everyday urbanity. Due to this dependency of spatial knowledge within the observer's attention span as in, a mathematical coordinate system, humans pursue the compulsion to experience their environment through physical presence, on the one hand transmuting space into a more complex entity, on the other hand moving outside the boundaries of a geographically presumed and digitalized arrangement. Steering away slightly from the standpoint of the interaction between biological observers and the re-mediated spaces that they linger in, in an impermanent idiosyncrasy, following a somewhat oscillating rationale, hereby reconciling into the characteristics of the main variable in this dually cyclopean equation, the electromagnetic spectrum of light.

Light expediting as a wave while wending our monocular field of view is considered as an esoteric but fairly measurable, brandishing and perplexing electromagnetic field which keeps our microwave ovens running, computer hard drives use it to store data on rotating disks, x-ray and MRI scanning devices are substrated on it while electric motors use it along coils to produce displacement, moreover magnetotactic bacteria residing in seas have threads of magnetite crystals that bolster them to navigate depending on the orientation of the northern or the southern hemisphere, not only the antennas in our mobile phones but also the relays of base stations emit electromagnetic fields keeping us denuded and imbibed to these disparate lengths and breadths of waves throughout our entire lives. Radio waves have wavelengths from a few millimeters to kilometers where the wavelength of visible light is relatively short, ranging from 400 – 700 nanometers. The wave model of light is used to define waves as moving energy from one physical point to another at a speed that is independent of its intensity or wavelength whereas not moving mass, interfering and diffracting as it passes through different smaller scale environments, bending and spreading through or around obstacles.

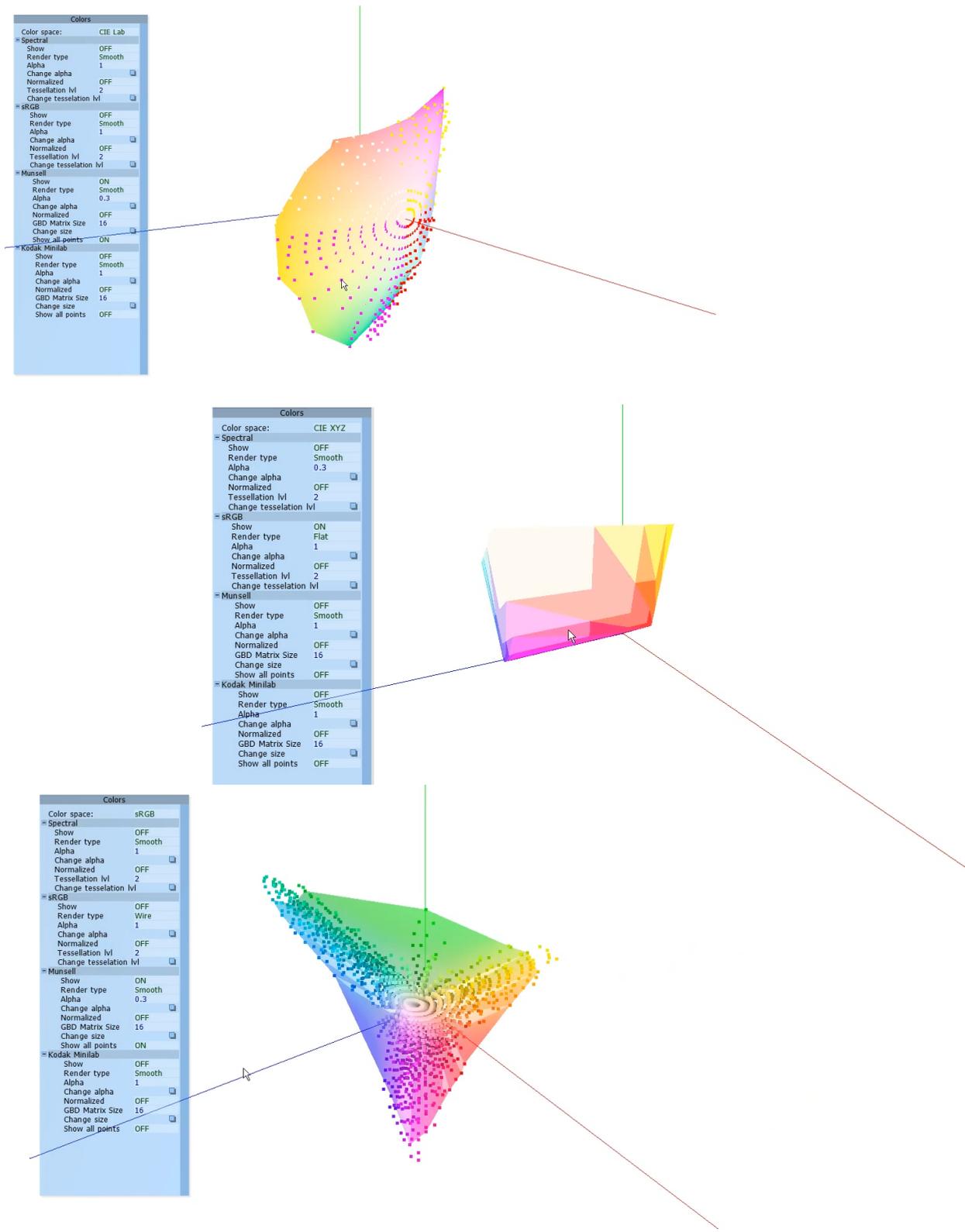
These wave like characteristics of light regarding motion arises when a periodic disturbance of some kind is propagated through a medium, such as pressure variations through air or the variations of the magnetic fields in space, to each medium a certain velocity is attributed at which the disturbance travels, whether sinusoidal or planar depending on dimensionality. If the void mentioned in the tenth page is considered as a component of a truly empty space as the vacuum according to modern physics inside which not existing as an entity but as a photonic structure exceedingly covering everything in its range, light itself travels at a straight line at a fixed and radiometrically measurable speed, disarranging from a supposedly imaginary point A to B. During

this journey two fundamental events may take place, either light gets deflected through different media, being bent depending on the indices of refraction without altering its frequency or it bounces off a certain surface and gets reflected. It is the occurrence of refraction that makes all the optical mechanisms around us possible, including our own eyes when we consider light on the basis of a waveform that sojourns through and throughout space.

Aristotle put forward the postulate adroitly formulating that, *in nature abhors a vacuum*<sup>46</sup>, unfilled spaces are unnatural and therefore defy the laws of nature and physics. The impossibility of motion inside a vacuum encompassment orates that the empty space should be entirely symmetrical spawning the obstacle that there should not be a specific direction in order for any given substantiating body to dislocate itself in preference, culminating in a state of stasis. Surely the philosophical suppositions, as far as the birth of modern physics is concerned, had long been empirically examined by Laplace, Euler and Fourier to name a few in the seventeenth century, in the fields of linear algebra and vector calculus as in the mathematical approach to the definitions of constants and derivatives and posteriorly undertaken by the nineteenth century atomists and the abundant study of geometrostatics. Circling back to Aristotle's approximation of empty space, where the equations that are describing a substantial realm of material objects are a contingency of mathematical symmetry, if the void does not confront with defiance it can continue towards any direction and a body being enclosed inside has the possibility to aggrandize in any direction following its container concurrently. Here the inference that if the dimension of the void is where the body exists then the void might perhaps be comprising surfaces and from the surface thereafter being segmented into the cloud of points. Therefore from the subset of *ifs* foments the question if the body and void can be divided into smaller components, if then when a body is delineated into prorated surfaces, then each individual surface would mitigate inside a void of its own, condoning to the antecedent integral argument as to empty spaces being filled whilst being bestowed upon the laws of nature.

---

<sup>46</sup> "The void is thought to be place with nothing in it. The reason for this is that people take what exists to be body, and hold that while every body is in place, void is place in which there is no body, so that where there is no body, there must be void." Complete Works of Aristotle, Volume 1: The Revised Oxford Translation, Princeton University Press, (1984) p. 363, propounded in Aristotle's Physics Book IV



Figures 9-11: Real-time representation of disparate color assemblages (all visible colors, sRGB, Munsell) on CIE XYZ<sup>47</sup> - CIE Lab and sRGB digital color spaces, Images courtesy of the repository available at <https://github.com/fabio-miranda/ColorGamut>, F. Miranda, (2010)

<sup>47</sup> CIE stands for *Commission internationale de l'éclairage*, the International commission on illumination, CIE XYZ was mathematically ascertained in 1931, CIE L\*a\*b\* is a combination of lightness L and chrominance values a-b

Optometrists designate interpupillary distance as the measurement of the centers of the pupils of two eyeballs, bracketing in males with the mean of 65.32 millimeters with a standard deviation of 1.50 and in females 61.53 millimeters with a standard deviation of 2.66 in consonance with the academic normative assessments using an infrared eye tracker and center point stimuli<sup>48</sup> on 416 participants (68% white, 17% black, 8% Hispanic, 1% Native American, %6 non-disclosed), calculating the relevance of accuracy on the account of IPD being closely associated with stereoscopic functions of the brain which is dependent on the combination of two stereoscopically separated images for human stereo vision. The values procured by indispensable pupillary measurements contribute prominent data for the fields of ophthalmology and enhanced optical magnification lenses, moreover one of the single predominant emphasis of all derives from our visual acuity for circadian activities and our day-to-day stomping grounds of recognition.

In the present commercial head-mounted virtual reality displays, there are two types of pupillary distance choices, either the user accesses the interface with the adjustment of the real IPD by palpably moving the ocular lenses inside the unit which ranges commonly from 54 to 72 millimeters or with the implemented virtual IPD running in the software, which effects the sense of scale in respect to the yielded digital environment. Tangible interfaces of the current HMD – Fresnel lens settings, e.g. Oculus Rift has been working on correcting the distortion variables on the subject of different lens types and individual eye relief, Oculus SDK<sup>49</sup>, the version that I had used in February 2019 for the prototyping of the GearVR<sup>50</sup> headset (p.11) refined by Samsung was conditional to the same software development kit established also by Oculus to function correspondently with Unity, employs distortion correction automatically within the compositor contrivance, maintaining the latency-reducing timewarp<sup>51</sup> and presenting corrected frames to the headset. To put it briefly, the edifice of the wearable interface must have indubitable conditions for the observer, in particular, image latency, pixel density, positional tracking and motion parallax correction to warrant a commutual and second nature *liaison* between machine vision and the human optical system.

---

<sup>48</sup> The Reliability, Validity, and Normative Data of Interpupillary Distance and Pupil Diameter Using Eye-Tracking Technology, Murray P. N., Hunfalvay M., Bolte T., *Translational Vision Science & Technology*, Vol. 6 No. 4, (2017), retrieved digitally from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5497600/>

<sup>49</sup> SDK stands for the acronym Software Development Kit, a set of computational tools or pre-defined pieces of code furnished by software vendors by means of which developers can use to build applications for specific platforms such as mobile mixed reality apps

<sup>50</sup> Samsung's standalone mobile HMD platform which was initiated in cooperation with Oculus in 2014

<sup>51</sup> Sundries of time-warping have been deployed in VR headsets to lower the perceived latency values by re-projecting an already rendered frame just before sending it to the display to compensate for the change in the head rotation of the observer, the attribute inside the Oculus SDK was adjoined to the software in 2014 with the public release of v0.3.1

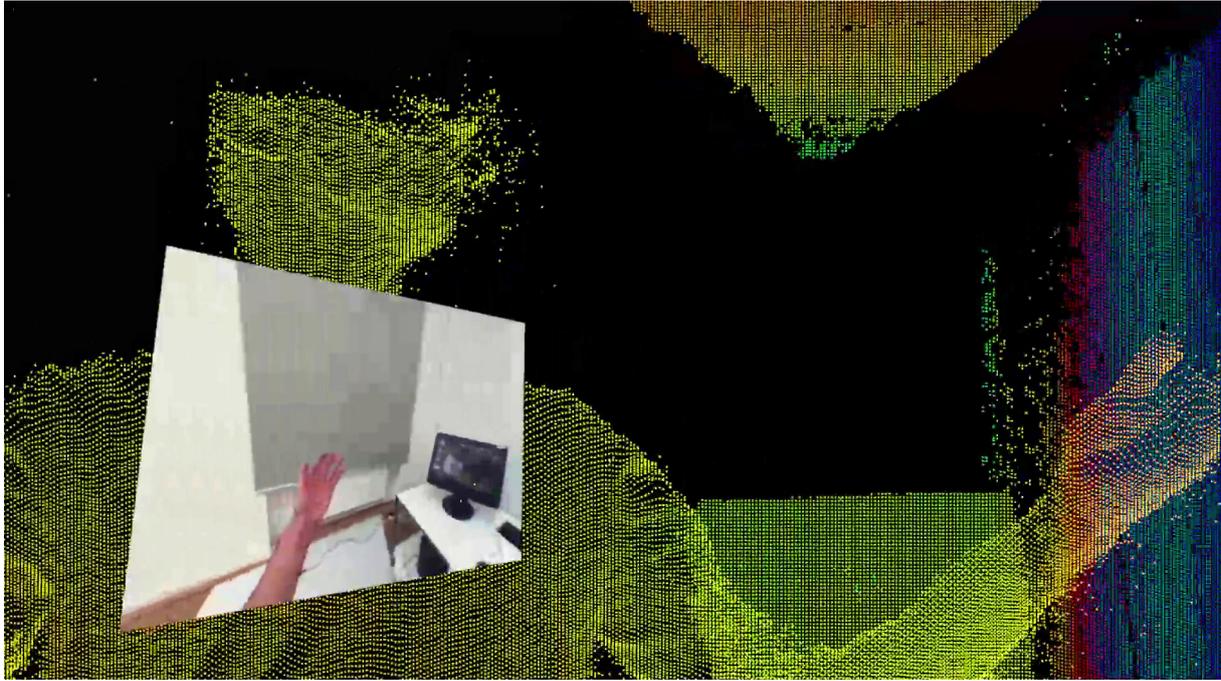


Figure 12: Early testing stages of the reconstructed digital space with dual video streams of reference from a single camera, Vordere Zollamtsstraße 7, (2019)

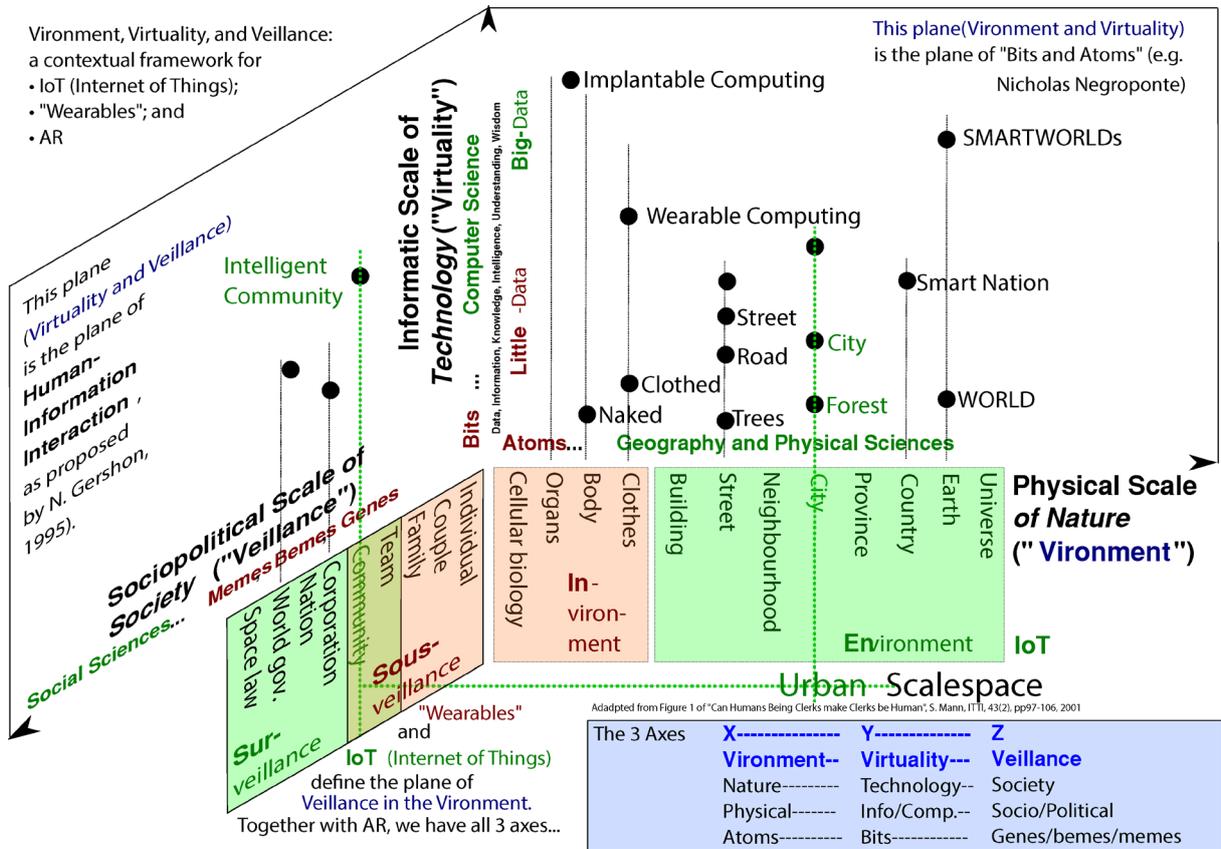


Figure 13: Fundamental Principles and Concepts of AR; Vironment, Virtuality and Veillance, Image Courtesy of Steve Mann, retrieved from <http://wearcam.org/ar/> Prof. Steve Mann of the Department of Electrical and Computer Engineering of the University of Toronto is considered as the father of wearable computing, peruses predominantly in fields such as periodicity in image processing, mediated reality and optical perception devices, <https://www.ece.utoronto.ca/people/mann-s/>, (2019)



## Bibliography

*Aristotle Physics, Books III and IV*, Oxford University Press, (1983)

*A case for natural color imagery from geostationary satellites, and an approximation for the GOES-R ABI*, Steven D. Miller, Christopher C. Schmidt, Timothy J. Schmit, Donald W. Hillger, *International Journal of Remote Sensing*, (2011)

*A Pessimistic Estimate of the Required Time for an Eye to Evolve*, Dan-E. Nilsson, Susanne Pelger, (1994)

*Advances in Independent Component Analysis and Learning Machines, Chapter 10*, E. Bingham, S. Kaski, J. Laaksonen, J. Lampinen, Academic Press, (2015)

*An Introductory Course in Computational Neuroscience*, Paul Miller, (2018)

*Beyond Art: A Third Culture: A Comparative Study in Cultures, Art and Science in 20th Century Austria and Hungary*, Peter Weibel, (2005)

*Big Data Is Not a Monolith*, Sugimoto R. C., R. H. Ekebia, Mattioli M., (2016)

*Color realism and color science*, *Behavioral and Brain Sciences* 26, 3–64, Byrne & Hilbert, (2003)

*Complete Works of Aristotle, Volume 1: The Revised Oxford Translation*, Princeton University Press, (1984)

*Computational Vision - Information Processing in Perception and Visual Behavior*, Hanspeter A. Mallot, (2000)

*Critique of Pure Reason*, Immanuel Kant, (1781-1787)

*Disappearing Architecture: From Real to Virtual to Quantum*, Georg Flachbart, Peter Weibel, (2005)

*Dracula*, Bram Stoker, (1897)

*High Dynamic Range Video, From Acquisition to Display and Applications*, Editors - Dufaux F., Le Callet P., Mantiuk R., Mrak M., Academic Press, (2016)

*Into the Universe of Technical Images*, Vilém Flusser, (1985)

*Introduction to Machine Vision*, Cognex Corporation, Retrieved from the Internet Article, <https://www.cognex.com/what-is/machine-vision/benefits>

*Leibniz's Metaphysics of Time and Space*, Michael Futch, (2008)

*On the Plurality of Worlds*, David K. Lewis, (1986)

*Perceptrons: An Introduction to Computational Geometry*, Expanded Edition, Marvin Minsky, Seymour A. Papert, (1987)

*Self-Supervised Deep Learning on Point Clouds by Reconstructing Space*, Jonathan Sauder & Bjarne Sievers, Hasso Plattner Institute Potsdam, (2019)

*Short Lecture Notes: Interpolation Theory and Function Spaces*, Helmut Abels, University of Regensburg, (2011)

*The Symbolic Species, The Co-Evolution of Language and The Brain*, Terence W. Deacon, (1997)

*There is No Question of Physicalism*, Tim Crane & D. H. Mellor, *Mind*, New Series, Vol. 99, No. 394, (1990)

*Visual Images Preserve Metric Spatial Information: Evidence from Studies of Image Scanning*, Stephen M. Kosslyn, Thomas M. Ball, *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 4, (1978)

*Visual Space Perception*, Maurice Hershenson, MIT Press, (1998)