Prof. Albert Lang

Johannes Pointner

Christian Schmidts

Abgabedatum: 04.04.2025

Operational Analysis of Photogrammetry

Tracing the Transductive

Master Thesis

MA Design & Computation

Universität der Künste Berlin

Technische Universität Berlin

Joel Tenenberg

Matrikel Nr.: 4044938

joel@tenenberg.net

Contents

Introduction

“Mapping activates territory.” (Kitchin & Dodge, 2007, p. 18)

The ability to transform objects, spaces and entire territories into analyzable, reproducible data has become a foundational process in computation. From logistics and navigation to the automatization of decision-making and real-time visualization, countless applications rely on the continuous translation of the physical world into a structured, digital model. Systems that rely on these models are as common as Google Maps, whose available 3D terrain data is generated through a complex entanglement of image processes (*Google Maps 101: How Imagery Powers Our Map*, n.d.; *How We Built Immersive View for Routes on Maps*, 2023), or as niche as salmon livestock control through autonomous underwater pods, which use cameras to monitor and treat parasitic infestations in salmon populations (*The Stingray System*, 2017). The core of these systems is photogrammetry, a chain of algorithmic operations that utilize overlapping images to reconstruct spatial information.

Photogrammetry, along with other computer vision processes, comprises on the transformation of visual input into machine readable data, a shift in which the visible becomes computational and thereby increasingly invisible to human perception and intervention (Parikka, 2023). In this sense, these processes can be linked to a black box, which conceals its inner workings, excludes interactions from the human sphere and completely transfers them into the computational sphere. This alienation of technical processes was already diagnosed by Gilbert Simondon in 1958 as symptomatic of the growing rift between cultural and technical perceptions (Simondon, 2017, p. 256). There are also epistemological and political implications to the black-box nature of algorithmic processes. Namely, their internal logic of decision-making and quantification processes, which determine what is contained in a scan, are obscured from view. This opacity reinforces a hegemonial structure of inclusion and exclusion (Downey, 2021). After the process of extracting spatial information out of image data, the transformation into the computational realm concludes with a point cloud, a defined collection of points with coordinate and color data. This process resembles a mapping of spatial features from the real world into the digital domain. As critical cartographic theory has long argued about geographic maps (Harley, 1989; Kitchin & Dodge, 2007), mapping and photogrammetry construct space and information, they do not merely reflect it. Based on the parallels between both operations, photogrammetry actively structures the relation between humans and their spatial environment within its own emergent operational regimes and frameworks.

This thesis analyzes photogrammetry as a technical object, tracing its underlying mechanism and identifying its key operative moments. By mapping these operations and making the internal decision structure visible again, this work aims for a deeper understanding of image operations and the work performed by them. The urgency for this work arises out of the subliminal reality-shaping aspects of image operations and their impact on surveillance, data extraction and automated decision-making. The implications of these operations extend to questions of agency in technology, the human relation to technology and inherent power dynamics, which are discussed in this thesis. The methodological framework for this analysis of photogrammetry draws on Friedrich and Hoel's ideas, put forth in their article "Operational analysis: A method for observing and analyzing digital media operations"(2023). Friedrich and Hoel's approach facilitates the characterization of media operations, which are entangled in various agencies and diffuse operational workings. They conceptualize a tool set to situate these operational agencies in different structural domains. This method guides the theoretical lens and practical execution of this project, allowing for a layered investigation of each sub-operation of the process of photogrammetry. These sub-operations span over the input of images through to the reconstruction of the sparse point cloud. Covering the fundamentals used in almost every camera-based scanning process.

At the core of the thesis the research process finds its object-like manifestation in a practical reconstruction of a photogrammetry setup: an installation that continuously scans the space, going through all analyzed sub-steps in computing the sparse point cloud, while simultaneously visualizing the algorithmic operations on screens. Because the operations run in real-time, a direct impact through the change of values in the code enables the variable and transparent character of the process. The installation is an instrument to resemble the transformation from black-box to glass-box and foster the interaction with seemingly monolithic algorithms. Its focus lies not on the scanned object but rather on the procedure of scanning itself.

The objective is to analyze the extent of operational regimes within the photogrammetry process and to identify the agencies embedded in the algorithmic structure through their direct impact on processed data. The transformation from a black-box to a glass box is facilitated by rebuilding the code, allowing previously invariable values to become variable, thereby enabling an examination of their impact on the pipeline's output. This thesis seeks to answer the question:

> How do operative moments shape photogrammetry, and can they be theoretically conceptualized within a media-theoretical framework?

Structure

The effort to analyze photogrammetry spans over various theoretical and methodological approaches, which cannot always be distinctly separated. The structure of this work is as follows: first, a theoretical and conceptual framework is established in which photogrammetry can be situated as a technical object. Secondly, while acknowledging the entangled consequences of scanning, the counteractive approach of mapping is introduced to regain or maintain knowledge of the process. This approach involves the main methodology of this work, the operational analysis (Friedrich & Hoel, 2023). The third part forms the fusion of the ideas previously outlined in the application of the operational analysis on photogrammetry.

The theoretical framework includes correlations to photogrammetry prior to the actual process of scanning. Therefore, Chapter 1 is dedicated to an introduction into image theory, or more precisely the theory of image-induced actions. The agency which photogrammetry incorporates can be traced back to the input frames. Images, which structure actions of operations around themselves, consequently contributing to the structuring nature of photogrammetry. The idea of a structuring technical entity can also be found in the philosophy of Gilbert Simondon, who introduces the concept of transduction as part of the individuation of technical objects. Friedrich and Hoel based their analysis on this notion, calling it adaptive mediation and making it a fundamental argument for an intertwined connection between technical objects and their surroundings. The need for the identification of such mediation in photogrammetry arises out of algorithmic opacity and its effects on human-machine interaction. Supposedly objective computational-aided mapping processes are deeply connected with neocolonial structures of power. Giving reason to characterize these operations and form potential counter measures.

In Chapter 2, the subject of diagrammatic mapping is introduced as a method of opposing exploitation evoked through black-box algorithms. The chapter gives an insight into the theoretical approaches to diagrams as knowledge creation and transfer mechanisms. Then, mapping is employed as an analytical methodology to render opaque algorithmic structures transparent. To this end, the methodology of operational analysis as developed by Friedrich and Hoel (2023) is presented in detail.

The bespoke analysis is then applied in Chapter 3, forming the central part of the thesis. Here, each operative moment of the photogrammetry pipeline is analyzed in the order it contributes to the creation of the sparse point cloud. Each step is situated within the broader process, examined through the operational analysis and visualized through a complementary

map (Fig. 6) of the workflow. This diagrammatic mapping accompanies the written work, allowing for a multi-layered representation of the system under investigation.
Chapter 4 provides a conclusion on the research process and sheds light on the efficiency of the operational analysis, considering the advantages and disadvantages of the method. It will also provide an insight into current research in the field of spatial computer sensing, suggesting a possible transfer of knowledge from this work to future research.

1. Operational Opacity

To analyze the process of photogrammetry, this work connects a range of theoretical approaches, with the goal of breaching the ephemeral character of technological operations. Based on the temporal and causal order of photogrammetry, the starting point of this endeavor is the operational image. A term summarizing those images, which constitute the beginning of an operation, like the camera input of drones to avoid obstacles or facial recognition through phone cameras. In the same operational realm lies also the concept of adaptive mediation, the idea of exchanging and structuring information inside a given system. In this case mediating information between an input and output or between camera and executed action. A notion originally derived from Simondon's thoughts about the individuation of technical objects in his book "On the Mode of Existence of Technical Objects" (Du mode d'existence des objets techniques) from 1958 (Simondon, 2017)[1], and further refined through the work of A. S. Aurora Hoel, especially in her article "Technicity and the Virtual" (2022).
Based on the theories, in this work, photogrammetry is approached as a technical object undergoing individuation, acknowledging it not solely as a passive instrument but as an emergent structure. This involves a reorientation of the perspective on algorithmic structures, such as photogrammetry, focusing less on reducing them to their uses, and more on their active role in the production of labor. The consequences of such technical mediation are shown on the implications of algorithmic quantization of the sense-able world into discrete, processable data. Applying numeric thresholds and values, which highlight the agency in algorithmic processing to include or exclude information. Directly connecting photogrammetry and image processing to the exploiting workings of post-colonization.

[1] This work cites the 2017 English translation by Cecile Malaspina and John Rogove, published by Univocal/University of Minnesota Press.

1.1. Operational Images

Upon initiation of the process of photogrammetry, the first operation is applied to the images captured on camera. The efficiency of the algorithms used is highly dependent on the quality of the images, defining these images as operational guiding and being utilized instead of aesthetically depicting (Farocki, 2004, p. 17). Farocki originally coined the term "operational images" in relation to military imaging systems, where images no longer required human viewing to execute an operational force such as the guidance of drones or missile systems. The change of the image purpose interrupts the traditional human to image relation. To comprehend this transformation, it is useful to compare the operational image with the non-operational image.

In traditional human-made images, real-world objects are represented as symbols through the lens of the creator. Deciphering these symbols relies on both the creator's original perception and the viewers interpretation (Flusser, 2018, p. 14). Consequently, the meaning of an image emerges from the interplay between the reaction to the actual object and its depicted form. The essence of a traditional image is inherently relational and dependent on its surrounding context (Alloa, 2021, p. 4). The ontology of non-operational images is determined by the human-centered power of interpreting what the image represents.

Operational images now invert this power structure. Rather than relying on human-centered interpretation, they are embedded in multilayered algorithmic processes. From the moment the camera sensor transfers data up to the final execution of an action, the operational image is never intended to be seen by the human eye. It is fragmented into a grid of pixels in which each is converted into a value of brightness, losing its color and becoming quantified in the process. In this quantifying process, data like color is regarded as merely another level of information. If this information is not required for the subsequent operations, it is excluded. Image analysis only syphons from the image what is computationally necessary.

The remaining information of brightness gradients serves as a starting point from which various algorithmic interpretations emerge. In this chain reaction of data being encoded and decoded choices are made. The impact of these choices extends beyond the algorithmic pipeline and, if an action follows, it manifests as an intervention in the tangible world.

The operational image, created by machines for machines, undermines the human centered sovereignty of interpretation about the image. The way data is read from these images is no longer dependent on aesthetic or anthropomorphic values. Instead, the decoding is entirely embedded in the laws of algorithmic processes. Rendering the role of human perception irrelevant and the operational image invisible (Downey, 2021). An operational image, if visible, is always specifically created for human vision, and therefore a symbolic representation of the operational image (Paglen, 2014, p. 73). This circumstance is applicable to all figures in this work. The image of the operational image acts as a mediator, a semantic bridge between human and technology.

Furthermore, the ontology of the operative image itself is directly correlated to the relation from image to operation. Christina Varvia describes the fine difference in which images can relate to an operation in her article "Image-sections: The evidentiary capacity of images to sample the lifeworld and have an operative life" (2021) as either operative or operational. Citing the dictionary definitions (2021, p. 211), she argues that "operational" corresponds to a hierarchically dependent level of impact and value in the chain of actions of the operation. Emphasizing the integral character of the operational as opposed to the initiating character of the operative, which expands its relationship not only to the operation in question, but also to the factors that make the operation possible in the first place. Varvia demonstrates that the impact of images is not limited to the operation built upon them but is woven deeper into the inter-related dependencies on which operations are carried out, ranging from the socio-political sphere to the laws of physics. The operative image organizes a multitude of different interactions around itself. The concept of reshaping and organizing different relational functions of images can be directly connected to Simondon's theory about the individuation of technical objects through the creation of their own "associated milieu" (Simondon, 2017, p. 59).

Operational images carry visual information of the object they depict and mediate that information to a subsequent operation. Both Flusser (2018, p. 9) and Simondon (2017, p. 9) align on the idea of the mediating character of images, even though they approach the topic of the mediator from different perspectives. Where Flusser focuses on the actual image as a product of technical process and inducer of mediation, Simondon takes a step back and situates the mediative, which he calls the transductive, aspect inside the ontology of technology itself. Enabling a media theoretical approach that includes all technical operations building upon the operative image to be categorized as mediators, including photogrammetry.

1.2. Technical Objects

The need for a closer look at the technical object arises from the tension between culture and technology. The French philosopher Gilbert Simondon thought about this tension and the following alienation of work caused by technology as early as the middle of the 20th century, although his "Theory of the Mode of Existence of Technical Objects" has been in

English translation only since 2017. The subsequent analysis explores Simondon's theory of the development of technical objects and applies it to the operation of photogrammetry, aiming to categorize it and contribute to a more comprehensive understanding of algorithmic entities and their relationship to humans.

Simondon's argument starts with the statement that culture functions as a bulwark against technology, because of its limited understanding of technological processes (2017, p. 10). The concept of "culture" encompasses a broad spectrum of meanings. These range from the idolization of technology and salvation fantasies in which the technological cyborg represents redemption from a patriarchally dominated world, as proposed by Donna Haraway (2016), to complete technological abstinence, as exemplified by groups such as the Luddites (O'Rourke et al., 2013). Nonetheless, in both cases, the technical element is assumed to be in a position contrary to the human. Consequently, the actions and work performed by a technical process are regarded as non-human and thereby alienated. This notion can be traced back to the beginning of the Industrial Revolution, at which point machines began to execute operations themselves, rather than only amplifying them (Simondon, 2017, p. 132). This marked the first instance of machine agency competing with human laborers, a theme that already appeared in the analysis of operational images. The human worker is no longer directly exercising the work on the matter, instead they instruct the technical object to execute the task of interest. The person operating is close to the operation but never at its core because the technical operation stays hidden. The person tasked with scanning prepares for the act of scanning, but they never actually analyze images or construct a point cloud themselves (2017, p. 249). To overcome this alienation, it is possible to metaphorically work oneself into the algorithm. This incorporation requires a comprehensive understanding of the moments in which operations change, delete or transfer data. Afterwards, the relationship with the operation may still be distant, but the operation itself will no longer be hidden.

To understand the inner mechanics of the photogrammetry pipeline, one must consider two of Simondon's ideas about technological beings, which are the foundation of his theory of individuation: modulation and transduction. Technical objects, after Simondon, have a symbiotic relationship with their surroundings. They not only receive transformative moments but actively participate in forming them. His thoughts take a challenging stance against the Aristotelian hylomorphic scheme[2], because of the passive role of matter in it, being depicted as only receiving.

[2] after which every object consists of its matter (hýlē) and form (morphḗ).

In order to address these limitations, Simondon introduces the concept of modulation, which involves continuous evolution and entanglement of matter over time. Distributing the agency in the human–machine–material interaction anew.

Modulation refers to the continuous, adaptive interactions between technical objects and their environments, leading to ongoing individuation. This idea highlights the dynamic nature of technical objects, which evolve through persistent engagement with their external milieu (Simondon, 2017, p. 140). Technical objects do not materialize fully formed but rather emerge gradually from the ongoing integration and refinement of prior technologies, slowly transforming into their current final form. Photogrammetry, as a technical system, does not function as a monolithic entity but as an assemblage of distinct algorithmic operations, such as feature detection, image alignment and depth estimation, each of which contributes to its overall modulation. The adaptation of the whole process depends on the modulation of its parts. That also means that the description of the ontology of photogrammetry is rather a description of the ontology of the sub-algorithms used in photogrammetry. Each sub-algorithm contributes to the system's individuation by affecting structural transformations that collectively define the technical object's ever-changing transformative identity.

This continuous evolution is accompanied by transduction, the process by which activity emerges within a domain. Transduction structures itself and subsequent operations progressively (1995). The transduction can be seen as a regulating process which enables the transformation of potential energy into actualized energy, without being energy itself. It is tightly coupled to the transmission of information, because the regulation is achieved through the input of external information into the system (2017, p. 169). What Simondon describes here is an altering power of higher magnitude that has an impact on a realm which itself is not a part of. This theory opens a space of hierarchy in technical operations, enabling the definition of key moments in which this regulating aspect takes place. In technical systems like photogrammetry, transduction enables the integration and coordination of various sub-algorithms, each influencing and refining the system's overall functionality. The alignment of images in the Structure-from-Motion (SfM) pipeline follows the transductive schema, where the spatial coherence of one set of points determines the positioning of the next. In this example, spatial coherence is the information that enables transduction of the potential of image data.

In conclusion, modulation describes the adaptive mode of existence of algorithmic structures, while transduction describes their inner structuring and self-organization principles. Together modulation and transduction facilitate the individuation of technical objects. This synergy enables adaptive mediation, in which technical objects like photogrammetry not only process data but actively shape and are shaped by their operational contexts.

Friedrich and Hoel further refine the concept of transduction and incorporate it in their notion of adaptive mediation of technical objects (2023, p. 56). Here the enabling and structuring character of transduction is set in a "middle-order environment"(2023, p. 65), because of the unaffiliated essence of transduction towards the system it regulates. Additionally, the structuring quality shows itself as a mode of compatibility and communication between the sub-parts (2023, p. 65). In the same manner as transduction enables the identification of operative moments through the definition of a structuring power, the adaptive mediation enables the identification of the dynamic compatibilities of the sub-processes through the definition of their operational entanglement. This focus on structure and inter-compatibility will govern the applied analysis later in this work.

1.3. Black-Box and Alienation

Photogrammetry is a series of algorithms that quantify and qualify based on their input operational images. When this operation is solely seen as a tool, through the lens of its utilization, its inner workings remain hidden, and the work done by the operation is alienated. This correlation was already mentioned by Gilbert Simondon as a motivation to look closer to the individuation of the technological object.

> "It is the essential part that is missing, the active center of the technical operations that remains veil." (Simondon, 2017, p. 249)

This notion of the center remaining veil is the reason that algorithms are labeled as black-boxes, reinforcing their initial alienation. The analogy to the black-box refers to the concealment of the technical operation, where only the inputs and outputs are available for analysis while the inner processes remain opaque (Gabriel et al., 2024, p. 319). The term "black-box" originated from early radar systems developed in World War II ("Radar for Airlines," 1945). A time in which technological advances were under extreme security measures to prevent them from being compromised by the enemy. Even though the term is used differently in these scenarios, both usages point towards a concealed operation in which development, function and internal logic remain obscured or inaccessible.

If actors are actively labeling operations as black-boxes it immensely influences our perception of algorithmic operations. It provokes the picture of a box that is impenetrable, can resist multiple amplitudes of g-forces and stays untouched by extreme heat or blazing cold. A box as a hermetic concealment which can only be opened and accessed by powerful entities in possession of the right key. The problem with black-boxed algorithmic systems is that the concealment makes it difficult for external agents to contest and challenge the internal logic of algorithms' operative decision-making. This opacity directly feeds into power asymmetries, in which control over algorithmic processes becomes monopolized by those with access to their inner workings, while others are excluded and disempowered. If these systems are intentionally or structurally designed to reinforce hegemonic configurations, then the inability to challenge them results in the maintenance and naturalization of those configurations.

Surveillance systems such as the *Afghan Automated Biometric Identification System* (AABIS), developed by the US Department of Homeland Security and NATO and deployed in Afghanistan in 2004 (Jacobsen, 2021), are an example of a data-driven hegemonic system. Such structures rely on biometric data in the form of iris scans, fingerprints and facial photographs obtained through image processing technology. The use of AABIS was intended to monitor the Afghan army and keep Taliban militants out of the state forces. However, the transfer of power after the withdrawal of Western forces in August 2021 left this information in the hands of the Taliban, which now resembles the data of traitors who collaborated with the US army (*New Evidence That Biometric Data Systems Imperil Afghans | Human Rights Watch*, 2022). This example shows the direct consequences of black-box scanning systems and the handling of personal data without the necessary safeguards. Enabling a system of hegemonic power without the possibility of regaining ownership of the data collected.

The logic of these systems is not neutral, but rather algorithmic, quantizing, detecting, logging and making visible only what is computationally processable. This selective visibility is a subtle bias that does not completely change the contents of the operation but does decide what is valuable to keep in the data stream. An example is the numerous thresholds in the photogrammetry pipeline that determine which types of data are permissible. Such thresholds are based on statistical success across average-use cases, resulting in the marginalization of edge cases and niche phenomena. As a result, the scanned or mapped object becomes not only a representation but a reduction, an algorithmic effigy of what can be extracted and commodified. Therefore, the hazard lies not in the practice itself but in the relation to the scanned, the mapped, the other. As Downey (2021, p. 2) articulates:

> "The will to calculate, measure, and qualify the other–the ambition to 'fix' the other as an objectified, calculable and thereafter commodifiable entity–is the link between the deterministic rationale of colonial discourse in the 18th and 19th

> centuries and the biopolitical re-inscription of the algorithmically quantifiable other through the technologies of surveillance employed by neocolonial powers today."

This observation highlights how rationalist mapping practices and algorithmic objectivity are deeply entangled with historical patterns of domination and control. Importantly, this extractive logic extends beyond overtly militarized applications such as drone targeting or aerial surveillance. Even simple space-depicting photogrammetry, without executive power or political intent, shares the same underlying operational logic. The representational mode is not shaped by the intention but by the infrastructural nature of the algorithm itself, which encodes specific assumptions about space, value and visibility. The final product of scanning applications is then used to train machine-learning algorithms or execute actions which then inherit the same attitude towards their target objective.

The function of image operations is not only to represent but to control and prepare data for further manipulation. Photogrammetry enables actors to collect data of the planet with imagery and spatially query this information. It is to siphon and to reproduce the siphoned. To generate computable space within an otherwise indeterminate world. This mode of operation is comparable with the AABIS system, which rendered the Afghan state forces a calculable domain based on scanned data.

To work critically with the scanning process is to open the black-box, to render visible the operative logics that have remained hidden. Each operational step can be analyzed, contextualized and therefore traced to its origins. No technical aspect is beyond comprehension. All technical aspects exist within a broader environment of design, constraints, and possibilities. Algorithms in themselves are nothing more than a chain of commands or actions that lead to a certain output (Chabert, 1999, p. 1). This attitude is not exclusive to algorithmic systems but can be applied to everyday tools and technical artifacts. An example is a common household drill. While most users understand that it is electrically powered, equipped with a trigger and gearing mechanisms, the precise interaction of its internal components, such as how energy is transferred and regulated, is rarely understood without direct investigation. The internal logic of the device remains concealed until it is physically opened and examined. In this sense, the process of opening and mapping the drill mirrors the work required to understand algorithmic systems.

In the most consumer-friendly applications, it is indeed a company-owned secret how the actual algorithms work and how they transfer and compute data. The usage is beginner-friendly and easy to understand because users trade freedom of choice for streamlined, predecided workflows, sacrificing control over the scanned data and, ultimately, personal user data (Stallman, 2010). This problem underlines the need for self-determined usage of technology and data. Especially in the realm of coding applications, the open-source community has established itself as an alternative to corporate solutions and many different applications and code snippets are available online under open distribution licenses (pluja, 2020/2025).

The open-source philosophy, which emerged in the 1970s, plays an important role in the Simondian understanding of the becoming of technical objects. Without free distribution, the exchange of sub-algorithmic parts would not be possible, thus complicating the active modulation of technical objects. Inferentially the forking and copying of already-established applications proves the individuation of technical objects which can be seen as a genesis of their sub-parts. In this context, the technical object does not merely emerge but undergoes distinct stages of development, each contributing to its gradual formation (Simondon, 2017, p. 40). This is also demonstrated in the context of photogrammetry, where the code used in this work is built from a series of pre-built components. The transition from *Parallel Tracking and Mapping* (PTAM) (Klein & Murray, 2007) to ORB-SLAM3 (Campos et al., 2020) resembles the concretization of *Simultaneous Localization and Mapping* (SLAM) systems. Rather than being coded from one initial idea, ORB-SLAM3 is built upon the operational architecture of its predecessor. With each new iteration including a richer number of complementary algorithms, like second-order organizational structures and error corrections. In this way, the technical object evolves not through replacement, but through recursive integration, layering new functions atop stabilized structures. This understanding demonstrates that the entire black-box is made of distinct parts that collectively execute the operation of the hierarchical higher object, making the findings of a differentiated and precise analysis of sub-parts fundamental to understanding the technical object.

2. Counter Action

Fusing the concept of technical operation as an assemblage of its sub-parts and the need to transform the operation into a glass-box to make it accountable, this thesis will employ an analysis and a map of the technical media operation of photogrammetry. This mapping approach forms a counter action to detect the dependencies and biases of technical objects through unraveling them. Even though the main focus of this thesis is not the diagram, its use in the niche connection between maps and algorithmic operations is valuable, as many examples of mapped operations have shown (*Anatomy of an AI System*, n.d.; *Critical Atlas of Internet*, n.d.). In this work, the process of mapping became a method of knowledge acquisition. Arranging the algorithmic components spatially revealed dependencies, flows and redundancies that remained hidden in textual analysis. Decisions about the placement, proximity and layering of the sub-parts in the map depend on the logic of the operation. The map was not created after understanding the system, but was an essential part of how the understanding was formed. In the context of this work, mapping resists classification as either theory or method. It is both a way of thinking about structure and a way of making structure visible. This hybrid character is what makes it useful for investigating the operational workings of photogrammetry. Mapping helped to make sense of photogrammetry as well as communicate these findings to an audience.

2.1. Diagrammatic Mapping

Network analysis can be challenging due to the complexity and number of connections between components and the data these components manipulate. In such cases, the use of diagrams or diagrammatic representations, like abstracted maps, can be helpful in providing a visual and spatial representation of the process. This approach requires a deliberate examination of each component of the process at multiple levels. The creation of a map involves considering not only the component itself, but also its spatial relationship to other components and the nature of its connections to its neighbors. Mapping is not only a methodology for didactics but also actively creates knowledge (Krämer & Ljungberg, 2016, p. 14). Therefore, the purpose of a map is dual. One outcome is the final product that is shown to an audience. Mediating the contents between analysis and recipient. The other is the process of creation, which enables a deeper understanding of each component through spatial arrangement – an understanding that can't be achieved through text.

Interdisciplinary concepts such as data extractivism which reach deep into social, political and environmental spheres are difficult to grasp without the establishment of a shared conceptual ground. As Christina Ljungberg argues, maps can serve as "prostheses for cognitive explorations", tools that support the visualization and externalization of abstract knowledge structures (2016, p. 155) creating such a shared ground. An example of this principle is Vladan Joler's "New Extractivism" (2020), a visual collection that utilizes mapping and diagrammatic components. Joler describes his work as "an assemblage of concepts and allegories" (2020). Similarly, in the context of algorithmic data predictions, diagrams can be used as a design practice and tool for showcasing inner logics, making them available for research and investigation (Benque, 2020).

These examples show that a mapping approach, combined with a diagrammatic representation of concepts, can enable externalization of internal logics and render opaque processes available to investigation and critique. Mapping facilitates a mode of knowledge, where abstract systems such as algorithmic pipelines become accessible through spatialized and relational visualizations. This theoretical foundation underpins the mapping of photogrammetry pursued in this thesis: not only as a written analysis, but as a visual assemblage that seeks to spatialize the internal logic, decision layers and ontological structuring of the photogrammetric process.

However, despite all the epistemic affordances of maps and diagrams, their role as instruments of power must also be acknowledged. In the same way in which critique is applied to the inclusion and exclusion of content in images and image operations (Varvia (Images By Amel Alzakout), 2021, p. 209), mapping must also be understood as an act of selection, abstraction and value encoding. As J.B. Harley (1989) emphasizes, the power of maps lies not only in what they depict, but in the value systems embedded in their construction.

> "Maps and territories are co-constructed. Space is constituted through mapping practices, among many others, so that maps are not a reflection of the world, but a re-creation of it." (Kitchin & Dodge, 2007, p. 18)

The epistemic potential of mapping as both visualization and exchange of thoughts forms not only a theoretical lens through which operations can be understood, but also a methodological framework for this thesis. The following chapter will further explore the topic of mapping as an analytical strategy. This method will be actively applied to photogrammetry to reveal relational structures and interdependencies. The subjective core of mapping thereby remains and is also true for the mapping of photogrammetry. This work cannot avoid actively framing what it presents, which necessarily includes and excludes, simplifies and encodes.

2.2. Operational Analysis

To deepen the understanding of the epistemic, technical and cultural dimensions of scanning processes based on image operations, it is necessary to move beyond representational frameworks that interpret scans as simply mimetic representations of reality. Instead, this thesis draws on the methodological approach of the operational analysis as developed by Kathrin Friedrich and A. S. Aurora Hoel (2023)[3]. Their approach offers a way to observe and analyze digital media operations as dynamic, situated entanglements of human and non-human agency. The operational analysis emphasizes that digital media technologies do not simply represent or capture reality. Instead, they actively intervene in it, presenting forms of mediation that co-shape both technical processes and the cultural meanings derived from them. Friedrich and Hoel's approach is based on the technical–philosophical theory of Gilbert Simondon. Conceiving technical objects as products of ongoing individuation, through which they co-construct technical environments. Their mapping approach views algorithmic operations not merely as mathematical functions, but as socio-cultural beings embedded in a large network of actors.

> "We call this method operational analysis. By proposing this method, we seek to integrate theoretical considerations about the operational and interventional aspects of digital media technologies, including considerations of their inner workings, while simultaneously factoring in their contexts of use." (Friedrich & Hoel, 2023, p. 3)

Operational analysis forms a cornerstone in the argumentation for photogrammetry as an active mediator between human and non-human agency. It can be seen as a step towards breaching the black-box-like appearance of algorithms through a mapping of their layered agencies, hybrid environments and recurrent operational regimes (2023, p. 16).

It might be appealing to describe the operational functioning of algorithms through a linear sequence of the steps of each sub-process. However, this approach would ignore the intertwined reality of the algorithm. Intertwined not only with itself in a complex order of action and reaction but also intertwined within a socio-cultural fabric. This is evident by the immediate effect image operations have on the ways spaces are perceived, bodies are

[3] The operational analysis embodies a theoretical, methodological and practical approach in one. Where in this chapter the theoretical background and methodological framework is discussed, the actual applied practice is described in chapter (3).

categorized, and decisions are delegated to computational processes. This multi-layered understanding of media operations highlights the need for an environment of analysis in which the sub-parts can have simultaneous and non-linear connections to each other while remaining in a causal relation. This is where an analytical and mapping approach can help to grasp the workings of media operations and demystify them.

The structure of the operation of interest is not predetermined, but modular and dependent on the use case. Sub-processes can be added or exchanged to give the pipeline a different direction. Understanding which software module or sub-process prepares which functionality is key to understanding the pipeline, and thus the processing of the data it receives and establishes. To enable this understanding in front of a seemingly unstructured background, Friedrich and Hoel concretized different parts of their method.

The operational analysis begins by defining the task of interest mediated by the digital media operation (2023, p. 59). This task serves as an analytical focus point and will guide the overall embedment of the media operation and the selection of relevant sub-operations. The closer the examination of the sub-processes of media operations, the more evident the complex entanglement becomes on a broader scale. At this high level of resolution, the interdependence among operations is so pronounced that micro-level distinctions blur, making it difficult to isolate individual elements. On this scale it becomes unclear why the analysis should start with the camera frames rather than considering the camera sensor, the history of photography, or why it should end with the point cloud instead of the dense point cloud or other spatial data from automated computer sensing. To refine the focus of analysis, the operational analyst selects a specific task that represents the core objective of the operations and follows every step until its completion.

> "Having decided on the task of interest, the operational analyst proceeds to ask whether the media operation under scrutiny is carried out in one go, or whether, as is often the case, it is carried out through a series of subtasks, executed at different times and places." (Friedrich & Hoel, 2023, p. 11)

The operative moment is the precise instant in which the media operation is actively involved in a specific action or processing step. It is here that the transformative nature of the operation can be pinpointed, showing how technical and material elements are intertwined in a concrete situation. These moments are guiding the diagrammatic mapping as a structuring system. Semantically organizing operations of finer detail under their domain.

> "To get a grasp of a certain media operation, the operational analyst needs to consider it along two axes: temporally along the task trajectory (horizontal

> analytical resolution), and in depth within each operative moment (vertical analytical resolution)." (Friedrich & Hoel, 2023, p. 13)

The operative moments themselves are situated inside a two-dimensional coordinate system defined by the task of interest and the research question guiding the analysis. The different resolutions enable a heterogeneous categorization of the sub-processes and are the key factor for a multi-dimensional analysis. The analytical resolution serves as a background on which the different parts of the operation can be arranged and brought into context.

To situate the Simondonian idea of transduction within the analysis, it is necessary to introduce a component resembling the structuring qualia. Friedrich and Hoel propose the alignment grid, an element that establishes an overarching framework to which the different operative moments can be positioned (2023, p. 59). Because it transfers and organizes information inside and between the operations, the grid itself acts as an adaptive mediator. In that way it adds a recursive and self-referential layer to the analysis. As previously mentioned, a part of the functionality of the individual operative moments can only be explained through their multilayered dependency which constitutes their agency. To map these moments does not necessarily mean to disentangle the dependencies but rather to put them in relation to the alignment grid. It serves as background against which the different agencies can be classified.

Through the introduction of the analytical zoom lens, a meta-level retracing of the scope of analysis within the analysis is possible.

> "An analytical zoom lens, that is, a dynamic and inquisitive magnifying lens that guides the operational analyst and allows her to move back and forth and to zoom in and out of operational layers in a systematic and shareable way." (Friedrich & Hoel, 2023, p. 15)

Through framing the analysis as a tool, they also acknowledge the situated knowledge (D. Haraway, 1988) produced through positioning the operational analyst above the operation, rather than as a part of it. The analytical zoom lens alludes to the awareness of the creation of an analysis from a certain viewpoint and respects this inventible relation between executing power and object of interest.

3. Applied Operational Analysis on Photogrammetry

The operational analysis as proposed by Friedrich and Hoel (2023) does not hold space for the interchangeability of the photogrammetry pipeline. In which different parts can be exchanged and augmented, for example by deep learning algorithm or left out completely when different hardware is used. Other variations of photogrammetry use stereo cameras or bypass the visual aspect completely, by using time-of-flight laser measurements like LiDAR sensors. Instead of approaching the analysis on a coarse level of detail, the operation of interest is defined precisely. The actual analysis is applied to the front-end algorithms of monocular visual ORB-SLAM3 (2020), a Structure-from-Motion implementation built in TouchDesigner using the OpenCV library in Python. This description gives an insight into the intertwined workings of the many different environments used in this project. The chosen research scope limits the possible multitude of algorithms in use. The focus is directed towards elemental computer vision techniques. Ensuring that the analysis remains relevant for understanding the fundamentals of image processing while not becoming overly specialized.

ORB-SLAM is a sub-category of photogrammetry, sharing the same basic algorithmic structure, while being able to be performed in real time. The continuous data stream enables a direct response to parameter adjustments in the code. This renders these value changes more perceptible and facilitates the glass-box transformation in the practical application of the analysis. The accompanying diagrammatic map of the analysis (Fig. 6) should complement and act as a visual aid to the written component.

The analysis will trace the key components of the photogrammetry pipeline, with particular emphasis on the operational transitions between algorithmic segments to comprehend the critical steps involved in data transformation. The objective is to identify and scrutinize biases inherent in decisions based on non-algorithmic factors. Examples are the number of feature points detected, or the threshold applied in the RANSAC filter. These pre- and human-made decisions are also situated inside a bigger and multilayered assemblage of algorithmic and computational actors, which make an isolated description of the operative moment less informative. This work therefore adopts the mapping approach to identify and trace the agencies at work, with the aim of revealing their interdependencies and enabling an overarching understanding of the impact of their encoded biases. This classification of sub-elements enables a dismantling of the whole into its smaller, intertwined parts with defined features. Which in turn allows a precise understanding of each of those parts, and the mechanics of their assemblage. The method of media operation analysis must be undertaken through various distinct lenses. The following sections briefly describe the task of interest, operative moment, alignment grid, analytical resolution and analytical zoom lens in the context of ORB-SLAM3.

The task of interest of the analyzed operation is SLAM (Durrant-Whyte & Bailey, 2006), which stands for Simultaneous Localization and Mapping and is a series of algorithmic operations frequently used in autonomous robotics. SLAM enables the sensing of spatial data through combining imagery and inertial movement of a camera setup, successfully transforming 2D data into 3D data. This makes SLAM especially useful to establish knowledge of the surroundings where global positioning data or pre-made maps are unavailable.

The distinct operative moments are described in the following sub-chapters and are chosen based on the main epistemological, ontological and data transforming moments. The first stage of separating the pipeline into sub-tasks and defining the operative moments is the division into first-order and second-order operations. This division refers to the character of the manipulation of data. Through first-order operations, SLAM can process feature points from incoming frames, match the feature points and estimate the camera pose concluding in a first loose mapping of the surrounding. The addition of new data through computations of different inputs characterizes the first-order algorithms. Second-order operations finetune the pre-computed data through comparison and refinement in the means of lop closure or optimized camera alignment. This analysis focuses on the first-order algorithms, emphasizing how 2D image data is converted and arranged into 3D space.

Throughout the analysis multiple alignment grids are identified, which enable the communication between various spaces of computation. Examples are the different space matrices enabling the transformation between image, camera and world space as well as the cartesian coordinate system, in which the triangulated points form the sparse point cloud.

The horizontal analytical resolution corresponds to the causal and temporal sequence of operations. That does not mean that each step is linear. Recursions are a substantial part of the photogrammetry pipeline that queue themselves up in the overarching causality of the pipeline. Whereas the vertical analytical resolution refers to the inner workings of each sub-step. Dividing these two operative moments enables a more in-depth and specific analysis, because the actual linear causality of the algorithmic steps can be seen next to the mesh-like behavior of its sub-steps. Connecting these two levels is a core part of addressing the ephemeral appearance of complex algorithms.

The vertical analytical resolution corresponds to the depth of analysis for each sub-step. Divided into micro, the technical algorithmic layer; meso, the communication layer between the human and non-human agency; and macro, the overarching ontological and epistemological layer resulting from the previous layers.

The analytical zoom lens changes with each operational moment and with each examined relation of the sub-parts. The process of photogrammetry includes communication between algorithms that act on different scales. Especially when looking at the image processing and the overall construction of the point cloud out of many of these image processing steps, the relations between these different scaled parts become diffuse. Under the analytical zoom lens, one can "zoom in" the process, temporarily hiding connections inside the operations, which appear again as the analysis "zooms out".

3.1. Image Frames

The start of the algorithmic Structure-from-Motion (SfM) process is situated on the meso level of operations, caught between the digital and physical worlds. While the actual light-to-pixel conversion is a technical step, physical dependencies also have a substantial impact. The quality and quantity of detected and matched features are dependent on the lighting conditions, exposure parameters and camera motion. Because the applied field of photogrammetry and odometrical applications is wide, the ways in which the input camera is operated can vary widely. While in odometry, the camera is mounted on an autonomous vehicle which is the center of the operation, like a drone or automatic driving car, photogrammetry can also be processed on handheld devices. In the practical installation of this work, the camera is mounted on a stepper motor, moving it 360 degrees in one direction before switching direction and rotating it the opposite way. Automating the motion of the camera enables a more precise variation of other, more computational-orientated parameters.

3.2. Feature Detection

The feature detection algorithm computes each frame as a unit, with its pixels forming sub-units. This means that the semantic visual structure of an image is converted into an array of numbers that correlate to each pixel and its color values. To narrow this array down, computer vision processes at this level only use the grayscale version of an image limiting the values saved to the brightness level of each pixel.

ORB-SLAM computes a continuous flow of images. To create a common ground on which these images can be compared, a robust and invariant comparison structure is established. The logic behind feature detection is to find points in an image that have enough detail to be found in another image. In this case the ORB[4] (Rublee et al., 2011) algorithm for extracting key points is used. The main functionality of the ORB algorithm is to detect feature points in an image and to describe these features with a binary string making them comparable between two images. ORB establishes data upon which higher order algorithms can build. ORB is an opensource feature extraction algorithm developed by OpenCV Labs, making it a suitable alternative to other algorithms like SIFT, which has a patent issued by the University of British Columbia. The algorithm itself is a combination of two other more specialized algorithms plus added modifications, showcasing the modulating character. The base is laid out by the Feature from Accelerated Segments Test (FAST) (Rosten & Drummond, 2006, pp. 430–443) key point detector and the Binary Robust Independent Elementary Features (BRIEF) (Hutchison et al., 2010) key point descriptor. The first finds the key points in an image and the latter encodes them in a unique descriptor to be matched later down the pipeline. The combination of those two parts is not flawless and requires additional algorithms in between. For instance, BRIEF is not able to respect the orientation in the description of the key point, making the descriptor not rotation invariant and in continuity the key points will not be recognized if the comparable frame is tilted.

FAST is a real-time corner detection algorithm optimized for speed. It detects corners by analyzing the intensity of 16 pixels arranged in a circle around a candidate pixel. A pixel is classified as a corner if a defined number of contiguous pixels in this circle is either significantly darker or significantly brighter than the candidate pixel, exceeding a user-defined intensity threshold. The algorithm optimizes this process by first checking four key pixels before evaluating the full set of 16 pixels. The number of continuous pixels and the threshold are both adjustable in the code, opening the possibility to fine-tune these values for a controlled working of the corner detection. Since FAST detects only key point locations without orientation or scale information, it is often combined with descriptor methods like BRIEF for feature matching.

The unique description of each key point through BRIEF relies on a comparison of two pixels surrounding the key point. To be precise, the algorithm compares the pixel intensity of two predefined points. If the intensity is higher in point A, this pixel pair adds a 1 to the binary string. Otherwise, a 0 is added. The number of compared pixel pairs determines the length of the binary string. This binary description then resembles a unique fingerprint of the key point. The position of the compared pixels is predetermined through a random but reproducible pattern. There are many different sampling methods, ORB is using a gaussian

[4] *Oriented FAST and Rotated BRIEF*

distribution with a value dependent on the circumference of the defined area around the key point. Using a gaussian distribution centered at the key point increases the likelihood of selecting pixels near the key point, ensuring a more stable description. To solve the previously mentioned missing rotation invariance, ORB precomputes the sampling pattern at different rotational angles and saves these patterns. For each detected key point, ORB estimates the rotation using the intensity centroid method (Rosin, 1999), which determines the dominant intensity direction of a patch surrounding the key point. ORB now uses the sampling pattern which aligns the best with the angle of the calculated orientation, successfully adding a descriptor with an embedded angle.

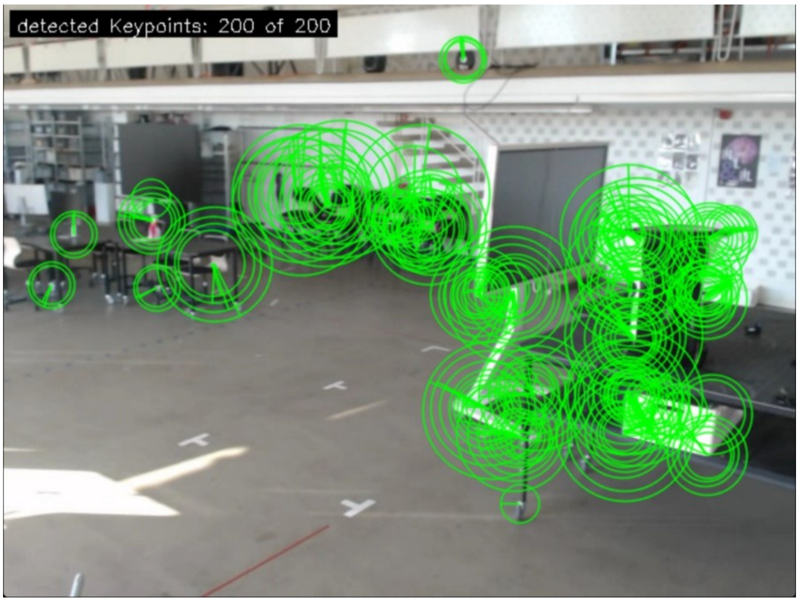

Fig. 1 ORB feature detection with key point visualization – Screenshot, TouchDesigner

Guided by the operational analysis, most of the feature detection process is situated at the micro level, the technical algorithmic layer. At the meso layer, they remain influenced by environmental factors that impact image quality, including scene lighting and camera parameters. The importance of understanding these fundamental and mathematically based operations arises from their significance later in the computation. The key point detection and description constitutes a system of comparison, enabling geometric consistency and alignment for subsequent multi-view reconstruction. Making these steps a fundamental basis on which an intertwined and communicating computation is even possible. The duality of algorithmic quantization already exists on this level of computation, one factor being the rigid dependency on brightness values, the other being the still modifiable thresholds and variables. The interplay of these two organizing layers can be tested through actively intervening in the code and observing the change in the final product. An interesting observation during the adjustments of these values was the sensibility of these operative moments. Even minimal adjustments to the parameters led to cascading failures, disrupting the feature detection and destabilizing the entire photogrammetry pipeline, as later operative moments like feature matching and pose estimation rely on a stable set of key points. The fragility of this multilayered operational structure underlines the interdependence of algorithmic parameters and their direct impact on large-scale 3D reconstruction. The feature detection and description algorithm of ORB bridges the gap between pixel and binary. Changing the character of the object of interest and making it available for a faster computational comparison.

The image is rendered invisible to the human eye by being transformed into an abstract, numerical form that can be further processed through mathematical transformations. Here the essence of Farockis operational image is evident; the original image is functionalized, no longer serving as a depicted image, but as a dataset for an operative system. The transformation into binary values (0 and 1) is not merely a conversion, but an epistemological redefinition of visibility: the world is made legible as a set of calculations. This creates an operational regime in which the code and its logic become the primary organization principles. The operational moment of feature matching proves that it is no longer the pixel values that matter, but the relational structures of the data.

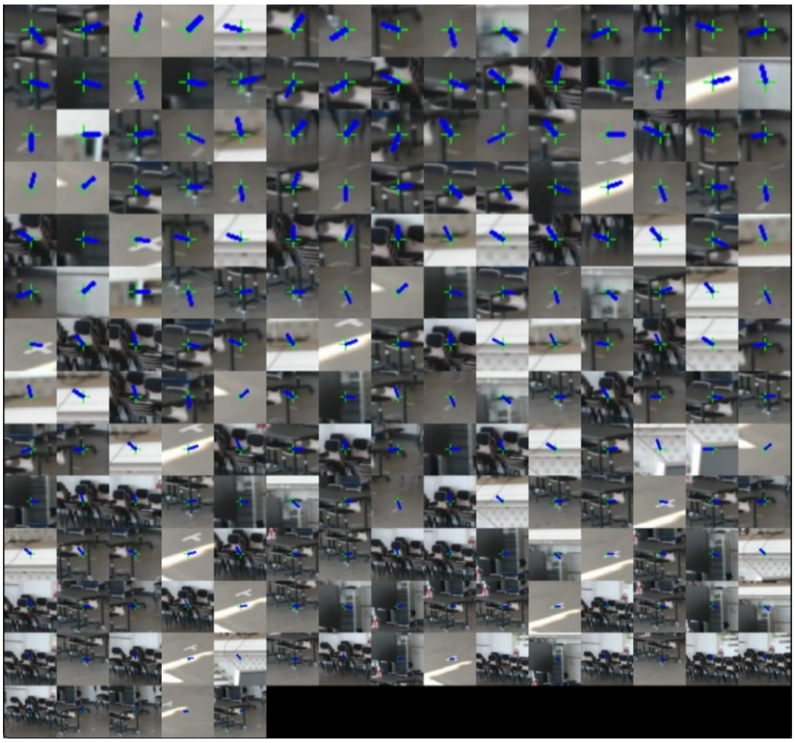

Fig. 2 Visualization of feature points with gradient direction – Screenshot, TouchDesigner

3.3. Feature Matching

Shifting to binary strings as the object on which operations are applied introduces new ways to measure qualia that could not previously be translated into the numerical – algorithmic environment. Feature matching describes the operation in which these qualia are matched between different images, finding pairs of similar pixel areas between two images. This step can be described as the first zoom out of single image processing, integrating multiple images into the operation. This is achieved not only by using more flat data but also by connecting points of interest between these images, adding data in the form of connections.

Feature matching uses the Hamming distance (Hamming, 1950) to measure the similarity of detected key points. It was introduced to measure bit-wise differences between binary sequences. Hamming's white paper is the foundation of modern error detection and binary-based similarity measures. The algorithm compares two binary strings and adds a value every time the binaries do not match. A Hamming distance of four therefore indicates four different pairs of binaries, with the rest of the digits in the two compared strings being similar. Because the binary string is derived from the distribution of brightness level of the pixels around the detected key point, a low Hamming distance indicates a similar distribution of brightness around two key points between two frames. A low Hamming distance indicates a successful match, thereby establishing a correlation between the images.

Throughout the pipeline of ORB-SLAM, the relation of the different algorithms to the computed data changes. One moment of changing relation is introduced after measuring the Hamming distance. The following algorithmic operations are no longer characterized through a transformation of image data into computational available data, but through their operational base on the data from previous steps. This means that from this point on, the image is completely invisible, transformed into the computational space where the following operations never refer to the original pixel data.

After the feature points are matched, they resemble a set of point correspondences between two images. Some of these matches are wrong due to noise, repeated patterns, or lack of distinctive features. To ensure a robust matching result, ORB-SLAM applies RANSAC (random sample consensus) (Fischler & Bolles, 1981) to the set of matched features. RANSAC is a resampling algorithm to estimate a model within a series of measured values with errors and outliers. Outliers refer to points outside of the estimated model.

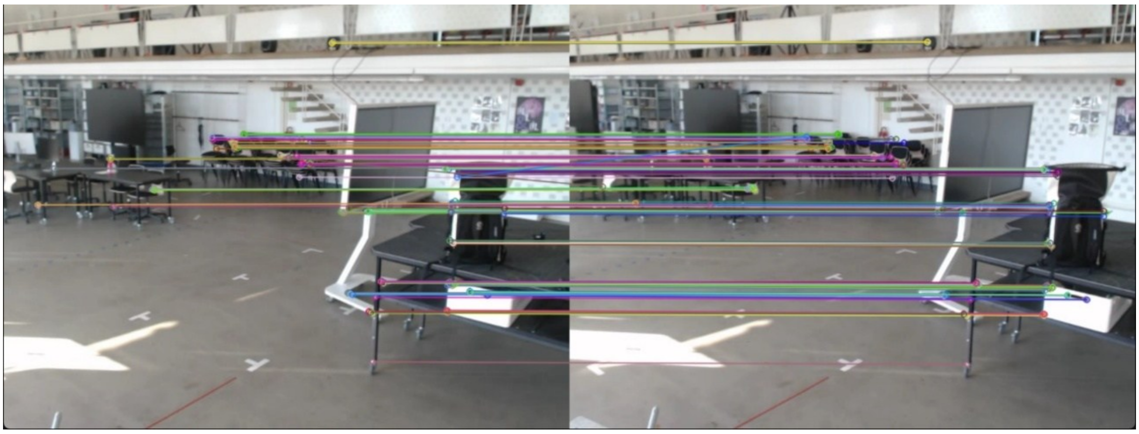

Fig. 3 Feature matching between two frames – Screenshot, TouchDesigner

To define the model and how RANSAC operates, it makes sense to take a step back and recapitulate on the task of interest as introduced by Friedrich and Hoel (2023), the initial goal of the overarching operation: the reconstruction of 2D points in a 3D space. The goal of finding pairs of features in different images is to retrace the transformation of the features across multiple images, which is directly linked to the transformation of the camera in space. The model that comprises the base for calculating the correspondence of feature points between images is the epipolar geometry, which is based on perspective dependencies in Euclidean space. Epipolar geometry emerges as a relational structure within the image alignment space, defining the correlation between two camera perspectives[5]. It is based on the simple assumption that the projections of the key point in two compared camera image planes only differ because of the relative motion between the two cameras. The optical centers of two cameras and the feature point of interest form the epipolar plane in space. The line on which the image frame and the epipolar plane intersect is the epipolar line. It resembles a geometrical approximation of where the projection of the feature point in both images must lay.

[5] In the following, the term "camera" does not refer to a physical imaging device but to the calculated camera pose associated with each image frame. A simplified camera model is used, which consists of the focal point, focal length and image dimensions.

Fig. 4 RANSAC applied on feature matching results; Outliers are indicated with a red circle. – Screenshot, TouchDesigner

This calculation incorporates intrinsic camera parameters which correspond to the actual physical properties of the camera in use like the principal point, focal length and distortion coefficients. This information is accurately calibrated and then integrated into the computational space as the fundamental matrix (Zhang, 2000). After the properties of the input images, this moment defines a point in which external parameters have an impact on the photogrammetry pipeline. Interestingly, the calibration of cameras can be done with the OpenCV Library and is based on the same image processing algorithms as photogrammetry. Epipolar dependencies act as a space of mediation, structuring through enforcing geometric consistency and a coherent spatial relationship onto a metastable set of feature correspondences. Through its structuring character, RANSAC functions as a transductive component in the pipeline of photogrammetry.

3.4. Pose Estimation

Pose estimation describes the computation of position and orientation of the camera through the previously established correspondences between the points in space and their projection on the image plane (Marchand et al., 2016). This operational moment defines the core concept of stereo vision and Structure-from-Motion operations and marks the first instance of true spatial computation, where previously 2D inter-frame relations are restructured into a 3D spatial configuration.

In the process of deducing camera poses, the operation navigates through multiple spatial reference systems. These do not represent different spaces, but different coordinate systems within the same geometric space, each with their own independent reference point and relational logic. Each transformation between these coordinate spaces changes the spatial relation of the positional data, situating it in different relational domains. This multifaceted structure contains the world coordinate frame, the global reference frame that defines the scene. Its origin can be arbitrarily chosen, out of clarity the first camera position or the start of the mapping operation is commonly selected.

Each image frame has an associated camera pose, which defines a local coordinate system, the camera coordinate frame. In this space, the optical center of the camera is the origin, and the axes are oriented according to the intrinsic orientation of the camera at the time of capture. Objects described through this coordinate frame are always relative to the camera, fundamental for establishing a connection between the 3D space and the 2D space of the image plane.

The image plane represents the projection of the scene and is thereby reduced to two dimensions. It is fixed within the camera coordinate frame, but establishes its own referential space, the image coordinate frame. All of the previously explained key point detection and description algorithms use this frame as reference. These different spaces are not isolated but connected and interlaced through a series of transformations, steered by different matrices. The extrinsic matrix traverses between the world frame and the camera frame by including the rotation and translation values between camera poses. To navigate between the camera frame and image frame, the intrinsic matrix transforms the data through the intrinsic parameters of the camera. In the terminology of Friedrich and Hoel, these coordinate systems can be understood as alignment grids representing structured regimes that organize how data is situated, interpreted and transformed within the photogrammetry pipeline. Shifting between these grids is not a simple translation; it is a reconfiguration of spatial relations, which allows image data to be mapped into a computationally coherent spatial structure.

Building upon the geometrical coherence of the feature matching operation, pose estimation is a transductive moment in which the system individuates itself as spatially coherent. Not only frame to frame but also from frame to the overall scene. This process establishes another structured domain from previously computed data. The fundamental constraints achieved by epipolar geometry do not disappear but are taken over into an organizational regime of greater magnitude.

Based on this established spatial domain, the computation of points in space resumes. From the feature matching process, the known variables include two matched frames. The translation between the corresponding camera positions is computed using the essential matrix[6] which estimates the relative camera poses. By taking a coherent feature pair and the

---

[6] The essential matrix is a 3 x 3 matrix that relates corresponding points in image pairs, containing the translation and rotation vector.

camera position in the world frame, the corresponding point in the 3D space can be triangulated. The result is a first sparse point cloud to which more points are added. Now the known variables include points in the world frame plus the 2D matches in the image frame. Taking these added values into account, the Perspective-n-Point (PnP) method is used to solve for a new camera position. A fundamental aspect of the pose estimation moment is that the pool of available information grows with each computation, making each new computation better informed than the last. In the same way that the Hamming distance products were refined by the RANSAC algorithm, the PnP products are also refined. Here, the reference model is the essential matrix, indicating that only a certain camera position is physically possible. Points added to the point cloud using the PnP method must therefore align with this camera position.

While the mathematical operations involved in triangulation and pose recovery, like essential matrix decomposition and the PnP method are well established in computer vision, the focus here is not on their formal derivation, but on the operational logic that governs how they structure and propagate spatial information within the operation. This stage of the photogrammetry pipeline reveals how algorithmic sub-operations, while individually structured as linear processes, form an interdependent meshwork of operations. Each operative moment feeds into and reconfigures the conditions for the next, demonstrating a dynamic logic of accumulation, modulation, and spatial transduction.

3.5. Sparse Point Cloud Mapping

The culmination of the mesh-like structuring of data is the combining moment, at which the accumulation of data generated in previous sub-steps is visualized in a durable and overarching representation: the sparse point cloud. The sparse point cloud represents the spatial memory as a global point map. It primarily displays all triangulated points, recreating the scanned scene in a digital environment. It can also contain information on previous frames and the camera path in space. In the continuous incoming sequence of frames, not every image is used, instead the photogrammetry pipeline saves operational checkpoints in the form of keyframes. These frames are characterized through a high parallax relative to the last frame, a certain number of new feature matches and a stable pose estimation. They establish a finer data structure in which camera poses can be traced back, a utility used by second-order algorithms that recursively adjust the point cloud. These recursive operating algorithms are highly dependent on a greater amount of captured data.

Fig. 5 Sparse point cloud – Screenshot, TouchDesigner

The mapping process repeatedly performs the same operational task of feature detection, feature matching, pose estimation and triangulation, while every new camera and correlating points are added to the existing map. The adjustable parameters of the code in the installation that physically accompanies this research include the maximum of saved points to research the implications the size of the point cloud exerts on the pre-steps.

4. Discussion

This analysis successfully mapped the entangled operational steps of photogrammetry without ignoring the interconnected dependencies of the system. The mapping approach is conceived as an investigative activity, in which the algorithmic relations are spatialized, and a layer of comprehension is added to the written work. Acknowledging the downsides and implications of creating a map, this diagrammatic map of photogrammetry must be seen as a situated artifact, shaped by the analytical zoom lens through which the work was approached.

The work focuses on the transduction of external data into the system's emerging order. The focus of the research can be extended to include the transduction of internally generated second-order data but requires more time and resources to do so. This also formulates a critical point of the operational analysis. The exponential increase in complexity of technical objects through their continuous modulation necessitates a corresponding length of time to unravel them. Theoretically, any media operation can be analyzed; in practice, it must be evaluated in how far a complete analysis of a complex assemblage is useful to its understanding. The solution to this problem lies in the method itself. In the same way that the operation is dissected and divided into sub-parts, the corresponding analysis can be built upon each other, creating a mirroring structure of analyzing content that grows with each addition. Research on second-order algorithms can therefore build on this analysis.

Another point of discussion is the impact of the integration of machine learning and artificial intelligence on operational analysis as a method. As deep learning becomes a fundamental component in the most recent and high-performance real-time 3D reconstruction systems, not only the amount but also the character of individuation of these operations changes. Approaches such as MASt3R (Murai et al., 2024) demonstrate how transformer-based modules can significantly enhance the robustness, speed and generalization of pose estimation and 3D structure inference. These systems modulate and optimize sub-processes within the photogrammetry pipeline, leading to more efficient results even under challenging input conditions (Leroy et al., 2024). However, deep learning-based systems have intentionally been excluded from the focus of this work. While these algorithms offer performance advantages, they also reinforce the black-box character of technical operations. The tuning of parameters, previously visible as threshold values or configurable algorithmic steps, is now handled by learned internal weights, which are opaque to the user and resistant to interpretation. More broadly, the evolution of technical objects is increasingly entangled with the evolution of machine learning itself.

From a Simondonian perspective, deep learning introduces a new order of transduction, as it not only adjusts internal parameters, but takes over the organization of entire workflows, including the flow of photogrammetry. These models do not merely automate tasks; they restructure the logic of the pipeline, often without exposing the underlying operational mechanisms. Their own process of individuation, how they come to function and organize themselves, still must be fully mapped. This also becomes visible in generative approaches to 3D reconstruction, where deep learning models use normal and depth maps, often from single input views, to synthesize fully novel 3D structures (Xiang et al., 2024). In these cases, the model no longer enhances individual steps in a geometric process but replaces the process entirely with a synthetic approximation.

Just as knowledge about sub-algorithms contributes to a deeper understanding of their higher order technical object and counters the alienation of its performed work, so too does the ignorance of these sub-processes intensify this alienation. The ongoing implementation of machine learning as a sub-part in technical objects introduces an expansion of the contents needed to be mapped to fully understand the object. It complicates the functioning of the operational analysis and may even demand a new framework for interpretation and critique.

5. Conclusion

The motivation for this research originated from the elusive nature of photogrammetry applications. From autonomous sensing systems to consumer-grade mobile apps, they produce fast and resilient spatial data. This perceived accessibility and ease of usage paired with an opaque functioning conceal computational abstractions and normative assumptions that are inherent in the process. Although based on physical principles such as perspective and Euclidean space, the photogrammetry process evokes an illusion of objective reasoning. This illusion collapses under critical analysis as proposed in this work, showing how the black-boxing of operations obscures the choices made as data is transferred from the physical to the digital domain. Images are not analyzed semantically but operationally, thereby reduced to features valuable for the following operations. This quantification is not neutral, instead it enforces constraints on what is decoded and reconstructed.

Drawing on Simondon's theory of individuation (Simondon, 2017) and Friedrich and Hoel's concept of the adaptive mediation (Friedrich & Hoel, 2023), photogrammetry emerges as a technical object in a continuous state of becoming. Modulating itself and transducing and mediating information in and around itself. This research dissected the photogrammetry pipeline into sub-operations and it emerged that an investigation based on a linear written text is incapable of containing the full scope of the cross connections thereby detected. For this reason, the diagrammatic map of the photogrammetric landscape has been developed alongside the thesis. The mapping process has been fundamental to both the research and the contemplation of this analysis, through the identification of operative moments. The diagrammatic map and the operational analysis rendered these moments visible, in terms of both their function and interdependence. From feature detection up to the sparse point cloud construction, each moment does not merely operate on data but restructures the computational conditions of the next one. The sub-operations are not just executing algorithmic processes, but they carry a structuring agency, which is initially invisible.

This agency is exercised through a regime of quantification. Algorithmic logic decides what counts as valid data to construct a spatial coherence. The decisions are based on physical constraints but are also decided through embedded values and inaccessible parameters. Creating an opaque operation. The black-box nature of these processes contributes to an alienation between their products and the users operating them. A phenomenon that Simondon described as a rupture in the technical object's social integration (Simondon, 2017, p. 256). One aim of this research has been to counteract that rupture by successfully transforming the pipeline into a glass box: a transparent system with an accessible control of parameters.

The reconstruction of the photogrammetry process supports this transformation, by visually displaying each transductive stage. Through enabling real-time manipulation of embedded parameters, the effects of interventions across the algorithmic chain can be traced. The reconstruction does not present a more efficient pipeline but a theoretical zoom lens to investigate operative moments and their implications. The resulting understanding is not limited to photogrammetry and can be applied to similar structures in other algorithmic applications. The fundamental principles underlying computer vision operations stay the same, as well as mode of transduction.

In conclusion, this thesis advocates for a shift in attitude toward algorithmic systems and the work they perform. Rather than accepting them as given or static, black-box operations must be investigated. This requires granular engagement with the operation, to challenge biases obscured by opacity and render their reasoning visible. The research process has revealed technical objects as continuously transforming and self-structuring entities in constant exchange with their surroundings, including humans. Perceiving them as evolving out of already-known contexts and processes enables a more nuanced perspective on technical objects. It allows for the reclamation of authority over extracted data and its products and resists the alienation of algorithmic work.

References

Alloa, E. (2021). A Lesser Being. From Louis Marin to Simondon and Back. *The Nordic Journal of Aesthetics*, *30*(61–62), 8–13. https://doi.org/10.7146/nja.v30i61-62.127841

Anatomy of an AI System. (n.d.). Retrieved March 5, 2025, from https://anatomyof.ai/

Benque, D. (2020). *Case board, traces, & chicanes: Diagrams for an archaeology of algorithmic prediction through critical design practice* [Thesis, Royal College of Art]. https://researchonline.rca.ac.uk/id/eprint/4420

Campos, C., Elvira, R., Rodríguez, J. J. G., Montiel, J. M. M., & Tardós, J. D. (2020). *ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial and Multi-Map SLAM*. https://doi.org/10.48550/ARXIV.2007.11898

Chabert, J.-L. (Ed.). (1999). *A History of Algorithms*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-18192-4

Critical Atlas of Internet. (n.d.). Retrieved March 5, 2025, from https://louisedrulhe.fr/internet-atlas/

Downey, A. (2021). The Algorithmic Apparatus of Neo-Colonialism: Or, Can We Hold "Operational Images" to Account? *The Nordic Journal of Aesthetics*, *30*(61–62), 78–82. https://doi.org/10.7146/nja.v30i61-62.127862

Durrant-Whyte, H., & Bailey, T. (2006). Simultaneous localization and mapping: Part I. *IEEE Robotics & Automation Magazine*, *13*(2), 99–110. https://doi.org/10.1109/MRA.2006.1638022

Farocki, H. (2004). *Phantom Images*. *29*. https://public.journals.yorku.ca/index.php/public/article/view/30354

Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, *24*(6), 381–395. https://doi.org/10.1145/358669.358692

Flusser, V. (2018). *Für eine Philosophie der Fotografie* (12. Auflage). European Photography.

Friedrich, K., & Hoel, A. S. A. (2023). Operational analysis: A method for observing and analyzing digital media operations. *New Media & Society*, *25*(1), 50–71. https://doi.org/10.1177/1461444821998645

Gabriel, G., Mittelstraß, J., & Carrier, M. (Eds.). (2024). *Enzyklopädie Philosophie und Wissenschaftstheorie*. J.B. Metzler.

Google Maps 101: How imagery powers our map. (n.d.). Retrieved March 26, 2025, from https://blog.google/products/maps/google-maps-101-how-imagery-powers-our-map/

Hamming, R. W. (1950). Error Detecting and Error Correcting Codes. *Bell System Technical Journal*, *29*(2), 147–160. https://doi.org/10.1002/j.1538-7305.1950.tb00463.x

Haraway, D. (1988). Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies*, *14*(3), 575. https://doi.org/10.2307/3178066

Haraway, D. J. (2016). A Cyborg Manifesto. In *Manifestly Haraway* (pp. 3–90). University of Minnesota Press. https://doi.org/10.5749/minnesota/9780816650477.003.0001

Harley, J. B. (1989). Deconstructing the Map. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *26*(2), 1–20. https://doi.org/10.3138/E635-7827-1757-9T53

Hoel, A. S. A. (2022). Technicity and the Virtual. *Humanities*, *11*(6), 135. https://doi.org/10.3390/h11060135

How we built Immersive View for routes on Maps. (2023, November 2). Google. https://blog.google/products/maps/google-maps-immersive-view-routes/

Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J. M., Mattern, F., Mitchell, J. C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B., Sudan, M., Terzopoulos, D., Tygar, D., Vardi, M. Y., Weikum, G., Calonder, M., Lepetit, V., Strecha, C., & Fua,

P. (2010). BRIEF: Binary Robust Independent Elementary Features. In K. Daniilidis, P. Maragos, & N. Paragios (Eds.), *Computer Vision – ECCV 2010* (Vol. 6314, pp. 778–792). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-15561-1_56

Jacobsen, K. L. (2021). Biometric data flows and unintended consequences of counterterrorism. *International Review of the Red Cross*, *103*(916–917), 619–652. https://doi.org/10.1017/S1816383121000928

Kitchin, R., & Dodge, M. (2007). Rethinking maps. *Progress in Human Geography*, *31*(3), 331–344. https://doi.org/10.1177/0309132507077082

Klein, G., & Murray, D. (2007). Parallel Tracking and Mapping for Small AR Workspaces. *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, 1–10. https://doi.org/10.1109/ISMAR.2007.4538852

Krämer, S., & Ljungberg, C. (Eds.). (2016). *Thinking with Diagrams: The Semiotic Basis of Human Cognition*. De Gruyter. https://doi.org/10.1515/9781501503757

Leroy, V., Cabon, Y., & Revaud, J. (2024). *Grounding Image Matching in 3D with MASt3R* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2406.09756

Marchand, E., Uchiyama, H., & Spindler, F. (2016). Pose Estimation for Augmented Reality: A Hands-On Survey. *IEEE Transactions on Visualization and Computer Graphics*, *22*(12), 2633–2651. https://doi.org/10.1109/TVCG.2015.2513408

Murai, R., Dexheimer, E., & Davison, A. J. (2024). *MASt3R-SLAM: Real-Time Dense SLAM with 3D Reconstruction Priors* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2412.12392

New Evidence that Biometric Data Systems Imperil Afghans | Human Rights Watch. (2022, March 30). https://www.hrw.org/news/2022/03/30/new-evidence-biometric-data-systems-imperil-afghans

New Extractivism. (2020). New Extractivism. https://extractivism.online/

O'Rourke, K. H., Rahman, A. S., & Taylor, A. M. (2013). Luddites, the industrial revolution, and the demographic transition. *Journal of Economic Growth*, *18*(4), 373–409. https://doi.org/10.1007/s10887-013-9096-y

Paglen, T. (2014). *Operational Images*. *59*, 72–73.

Parikka, J. (2023). *Operational images: From the visual to the invisual*. University of Minnesota Press.

pluja. (2025). *Pluja/awesome-privacy* [Computer software]. https://github.com/pluja/awesome-privacy (Original work published 2020)

Radar for Airlines. (1945). *Flight*. https://web.archive.org/web/20190214002854/https://www.flightglobal.com/pdfarchive/view/1946/1946%20-%200844.html

Rosin, P. L. (1999). Measuring Corner Properties. *Computer Vision and Image Understanding*, *73*(2), 291–307. https://doi.org/10.1006/cviu.1998.0719

Rosten, E., & Drummond, T. (2006). Machine Learning for High-Speed Corner Detection. In A. Leonardis, H. Bischof, & A. Pinz (Eds.), *Computer Vision – ECCV 2006* (Vol. 3951, pp. 430–443). Springer Berlin Heidelberg. https://doi.org/10.1007/11744023_34

Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. (2011). ORB: An efficient alternative to SIFT or SURF. *2011 International Conference on Computer Vision*, 2564–2571. https://doi.org/10.1109/ICCV.2011.6126544

Simondon, G. (1995). The Genesis of the Individual. In *Incorporations* (2. print, pp. 297–319). ZONE.

Simondon, G. (with Univocal Publishing). (2017). *On the mode of existence of technical objects* (C. Malaspina & J. Rogove, Trans.; First University of Minnesota Press Edition 2017).

Stallman, R. M. (2010, March 18). What Does That Server Really Serve? *Boston Review*. https://www.bostonreview.net/articles/what-does-that-server-really-serve/

The Stingray system. (2017, October 23). Stingray Marine Solutions AS - Bedre Dyrehelse. Økt Kunnskap. Uten Håndtering Av Fisken. https://www.stingray.no/delousing-with-laser/?lang=en

Varvia (Images By Amel Alzakout), C. (2021). Image-Sections: The Evidentiary Capacity of Images to Sample the Lifeworld and Have an Operative Life. *The Nordic Journal of Aesthetics*, *30*(61–62), 206–229. https://doi.org/10.7146/nja.v30i61-62.127900

Xiang, J., Lv, Z., Xu, S., Deng, Y., Wang, R., Zhang, B., Chen, D., Tong, X., & Yang, J. (2024). *Structured 3D Latents for Scalable and Versatile 3D Generation* (arXiv:2412.01506). arXiv. https://doi.org/10.48550/arXiv.2412.01506

Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(11), 1330–1334. https://doi.org/10.1109/34.888718

Appendix

Figures

All screenshots displayed in this thesis are captured from the personal project that implements segments of ORB-SLAM3 code (Campos et al., 2020) within the TouchDesigner environment.

Tenenberg, J. (2025). Fig. 1 *ORB feature detection with key point visualization* (screenshot),

Fig. 2 *Visualization of feature points with gradient direction* (screenshot),

Fig. 3 *Feature matching between two frames* (screenshot),

Fig. 4 *RANSAC applied on feature matching results* (screenshot),

Fig. 5 *Sparse point cloud* (screenshot).

Fig. 6 *Operational analysis of photogrammetry,* (diagrammatic map)

AI Tools

- ChatGPT Modell 4o, OpenAi: https://chatgpt.com/g/g-p-676d414a54b88191b83ed5b244e2388b-masterarbeit/project
    - Structuring assistance
    - Research assistance
- DeepL Translate, Deepl SE: https://www.deepl.com/de/translator
    - Translation of terminology
- Copilot o3-mini (Preview), in Visual Studio Code: https://github.com/features/copilot
    - Coding assistance

Declaration of Authorship

I hereby declare that I have written this thesis by myself, without the help of third parties and exclusively using the sources and aids listed.
I have marked as such all passages that are taken from the sources and aids used, either unchanged or in terms of content. Where generative AI tools were used, I have stated it. After using AI tools, I have reviewed and edited the content, and I take full responsibility for the content of the submitted work.
I further declare that I have not yet submitted the work in the same or a similar form to any other examination authority.

Map of Operational Analysis of Photogrammetry (Fig. 6)

Fig. 6 *Operational analysis of photogrammetry,* (diagrammatic map)

Operational Analysis of Photogrammetry

Tracing the Transductive