2024

DIOGO FERREIRA DE MONTALVÃO E SILVA

MOTION TRACKING METHODOLOGIES

Dissertação apresentada ao IADE- Faculdade de Design, Tecnologia e Comunicação da Universidade Europeia, para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Computação Criativa e Inteligência Artificial realizada sob a orientação científica do Doutor João Alfredo Fazendeiro Fernandes Dias, Professor Associado do IADE- Faculdade de Design, Tecnologia e Comunicação da Universidade Europeia

Acknowledgments

Firstly, I want to acknowledge my advisor, João Dias, for his encouragement, guidance, and follow-up during the development process of this thesis, for which I am grateful.

I am also profoundly grateful to the faculty professors for their valuable input and constructive criticism. Your diverse perspectives and suggestions have greatly enriched the quality of this work.

Lastly, a special thanks to my family and friends, your encouragement has been consistent throughout this whole process.

Resumo

É possível melhorar a precisão e a confiabilidade do rastreamento de movimento alterando o método dependendo do ambiente? No rastreamento de movimento, as técnicas para encontrar um alvo em cada ambiente têm aumentando em paralelo à tecnologia usada para realizar essa tarefa. Com cada uma delas melhorando, ainda há algumas que são melhores que outras, quando se leva em consideração o objetivo e o ambiente. Este artigo aponta para criar uma estrutura que ajude no processo de seleção para descobrir qual método de rastreamento de movimento se adapta melhor, dependendo do ambiente, suas condições e o alvo. Este estudo usa dados de outras pesquisas, artigos e documentos, considerando o ambiente em que foi testado e suas condições, bem como a capacidade geral do método em questão quando introduzido com diferentes variáveis e em diferentes cenários. Categorizar os métodos de rastreamento de movimento e entender suas capacidades em vários cenários proporciona muitos benefícios na otimização da seleção e aplicação dos mesmos. A estrutura não só simplifica o processo de seleção, mas também melhora a precisão e a adaptabilidade. Em última análise, essa abordagem leva a soluções de rastreamento de movimento mais eficientes, confiáveis e específicas para cada necessidade.

Palavras-chave

Rastreamento de movimento, Métodos de rastreamento, Avaliação de métodos, Processo de decisão.

Abstract

Is it possible to improve motion tracking accuracy and reliability by changing the tracking method depending on the environment? In motion tracking, the techniques to find a target in each environment have been increasing in parallel to the technology used to accomplish this task. With each of them improving, there are still some that are better than others, when taking into consideration the goal and the environment. This paper aims to create a framework that can help with the selection process to find out which motion tracking method fits better depending on the environment, its conditions and the target. This study uses data from other research, papers and articles, considering the environment in which it was tested and its conditions, as well as the overall capability of the method in question when introduced with different variables and in different scenarios. Categorizing motion tracking methods and understanding their capabilities in various scenarios provides many benefits in optimizing the selection and application of the same. The framework not only simplifies the selection process but also improves accuracy and adaptability. Ultimately, this approach leads to more efficient, reliable, and specific motion tracking solutions for each need.

Keywords

Motion tracking, Tracking Methods, Computer Vision, Machine Learning, Framework, Performance

INDEX

Acronyms

AI - Artificial Intelligence

AR - Augmented Reality

CGI - Computer-Generated Imagery

CNN - Convolutional Neural Networks

CSS - Cascading Style Sheets

GAN - Generative Adversarial Networks

HTML - HyperText Markup Language

IMU - Inertial Measurement Units

LSTM - Long Short-Term Memory

RNN- Recurrent Neural Networks

SfM - Structure from Motion

VR - Virtual Reality

Figures List

Table List

1. INTRODUCTION

In this first chapter, we present an overview of the components that form the foundation of this thesis. The chapter is structured to provide an exploration of the research theme, delineate the specific objectives that guide the study, articulate the central hypothesis, and outline the document's overall structure.

1.1. Theme Contextualization

In the topic of technological advancements, motion tracking methodologies have emerged as tools as a means to accomplish countless projects, revolutionizing various fields across diverse environments. This dissertation resolves in the comprehensive exploration of motion tracking methods and their applications, categorizing their performances in various scenarios. From controlled situations environments to dynamic and crowded outdoor spaces, where unknown variables may always appear, this study navigates the field of motion tracking, with particular focus on the methodologies that have most potential to innovate in various fields (Azuma, 1997).

The environmental conditions in which each motion tracking method operates significantly shape its efficiency, presenting numerous challenges that necessitate continuous technological evolution. In controlled environments, where lighting conditions are stable, optical tracking systems operate seamlessly, yielding near-perfect results. However, these systems face various issues outdoors, where unpredictable natural lighting, shadows, and environmental interferences can compromise accuracy, leading to unwanted malfunctions. For instance, Inertial Measurement Units (IMUs), while portable, face challenges such as cumulative errors and drift, especially in dynamic outdoor settings. Additionally, depth-sensing technologies can struggle with the complexities of outdoor backgrounds, hindering accurate depth perception. To surmount these challenges, technology in motion tracking is advancing rapidly. Innovations like sensor fusion, which combines data from multiple sensors, aim to enhance robustness and accuracy. Machine learning algorithms are employed to adaptively compensate for environmental disturbances, fostering a more resilient tracking performance. As technology evolves, the integration of environmental awareness and adaptive algorithms signifies a promising evolution in these technologies (Zhang, He, & Duan, 2020).

Artificial Intelligence (AI) seamlessly integrates with the motion tracking methods within this project, enhancing their capabilities and expanding their potential applications. AI-driven approaches, particularly Deep Learning-based methods, have revolutionized motion tracking by providing more accuracy, adaptability, and real-time processing capabilities. These methods leverage large datasets and neural networks to detect and track objects with remarkable precision, even in complex and dynamic environments. AI is also involved in traditional methods by improving feature extraction, matching, and prediction accuracy. This integration of AI not only boosts the performance of individual tracking methods but also ensures that the framework adapts to new challenges and evolving requirements (Leal-Taixé, Canton-Ferrer, & Schindler, 2015).

Although in this paper we shall look at the possibility that when facing these unknown outdoors variables, switching of the methodology in use increases the accuracy of the results depending on the system and final goal. With all of these method possibilities it also comes with a difficulty to understand which would perform the best depending on the conditions of one’s scenario. For this reason, the framework is created to ease this process, and facilitate the decision making with accuracy, structure and an adaptable approach, with selectable inputs that will point to the method that will align with requirements. Overall, this framework enhances the efficiency, reliability, and versatility of motion tracking solutions, driving innovation and improving outcomes across various industries. In this paper it is approached the methodologies revolving in three big groups, Tradicional Computer Vision, Machine learning and Hybrid methods. Within these groups there will also be a subtopic to categorize the methods which are the marker-based, markerless, and inertial, whether that can be improved by adding markers to the environment (Billinghurst, Clark, & Lee, 2015).

1.2. Objectives

The primary objective of this thesis is to develop a framework that recommends tracking methodologies based on specific inputs. This framework aims to improve the selection process by matching given parameters (inputs) with the most suitable tracking methods (outputs), ensuring optimal performance in the specific scenario. The various tracking methodologies will be researched to achieve this, evaluating their performance effectiveness and adaptability across different conditions. Through comparison of these methods, the thesis seeks to establish which methodologies perform best in specific scenarios, ultimately providing a robust, user-friendly tool.

The overall objectives of this thesis are the following:

- Research & achieve Results on motion tracking methodologies in different scenarios
    - Research their backgrounds & related work
    - Agglomerate them in groups and subtopics
    - Extract data & achieve results on their performance in different scenarios
- Create a framework
    - Construct a system of decision-making
    - Implement in a web app (HTML, CSS, JavaScript)
- Evaluate motion tracking technologies in the fields of ethics and future

1.3. Thesis Hypothesis

The research questions of this thesis are: "Can a framework be developed that accurately matches specific input parameters to the most effective tracking methodologies for various scenarios?", "Is it possible to speed the selection process of choosing a tracking methodology?" and "Is it possible to improve motion tracking accuracy and reliability by changing the tracking method depending on the environment?". This thesis hypothesis is that such a framework can be created and will significantly improve both the accuracy and speed of selecting appropriate tracking methods for a specific environment and purpose. It is anticipated that the framework will be able to analyze input parameters and accurately identify the tracking methodologies that have the best performance for the given scenario. This will be created through extensive research, based on the research there will be an evaluation of all tracking methods across a range of conditions. The success of the hypothesis will be measured by the framework's precision to identify the, indeed, the best method based on their performance. This way we can address the research questions and contribute valuable insights to the field.

1.4. Document Structure

This thesis is structured following the traditional document format, ensuring a comprehensive and systematic presentation of the research. The document begins with a Title Page, followed by an Abstract that summarizes the research question, purpose, data used, keywords, and conclusions. The Index outlines the organization of the thesis, leading into the Introduction, which introduces the research topic, questions, and objectives (Chapter 1 - Introduction). A thorough Literature Review follows, analyzing existing research and identifying gaps addressed by this study (Chapter 2 - Background & Related Work). The Methodology chapter consists of the data research and reason for the choosing methodologies, as well as potential limitations (Chapter 3 - Development Methodology). The Results chapter presents the findings, which are then interpreted and discussed in the Discussion chapter (Chapter 4 - Performance Evaluation). The Conclusion summarizes the key results, suggests directions for future research, and evaluates the ethics of motion tracking and its impact on society (Chapter 5 - Conclusion & Future Perspectives). References/Bibliography lists all sources cited, and Appendices provide supplementary materials, such as raw data and detailed calculations, ensuring a clear and detailed documentation of the research process and outcomes (Chapter 6 - Bibliography).

2. BACKGROUND & RELATED WORK

The following chapter discusses the background and related work that informs and situates this study of motion tracking methodologies. The research is structured to cover the essential theories, methodologies, and findings that have shaped this field's current state of the art. The chapter is organized to investigate each method, their base groups and motion tracking, while noting their areas of expertise and known products in the market. Starting in this last one, motion tracking, taking note of the groups and subtopics, followed by each group within, traditional computer vision, machine learning and, finally, hybrid methods.

2.1 Motion Tracking

In the most basic sense, it is the technology of capturing the actions of a continuously moving target using either sensors or cameras to record its movement (Billinghurst, Clark, & Lee, 2015). These recordings can be used in various fields with multiple applications, such as film and video games, where they are used to create character animations by mapping actors' movements into their digital characters (Guo & Zhong, 2022). Motion tracking creates immersive environments in VR by converting users' movements into digital experiences.

Motion tracking encompasses various methodologies constructed to perform specific contexts and applications. IMUs leverage accelerometers and gyroscopes to capture movement data, which are suitable for portable devices and sports applications (Madgwick et al., 2011). Optical tracking relies on cameras and markers to monitor spatial changes, proving valuable in virtual reality setups and biomechanics research (Azuma, 1997; Billinghurst et al., 2015). Magnetic tracking employs sensors sensitive to magnetic fields, finding utility in indoor navigation and specific medical applications (Han et al., 2014). Depth sensing technologies, epitomized by systems like Microsoft's Kinect, utilize infrared sensors to measure distances and track movements in real time, contributing to interactive gaming and 3D modeling (Redmon et al., 2016). Fusing these diverse methodologies creates a rich tapestry of motion tracking tools, each contributing unique strengths to various domains (Zhang, He, & Duan, 2020; Hasanujjaman, Chowdhury, & Jang, 2023).

When identifying these methodologies, arranging them into groups (Figure 1) and evaluating each of them on their advantages and disadvantages is necessary.

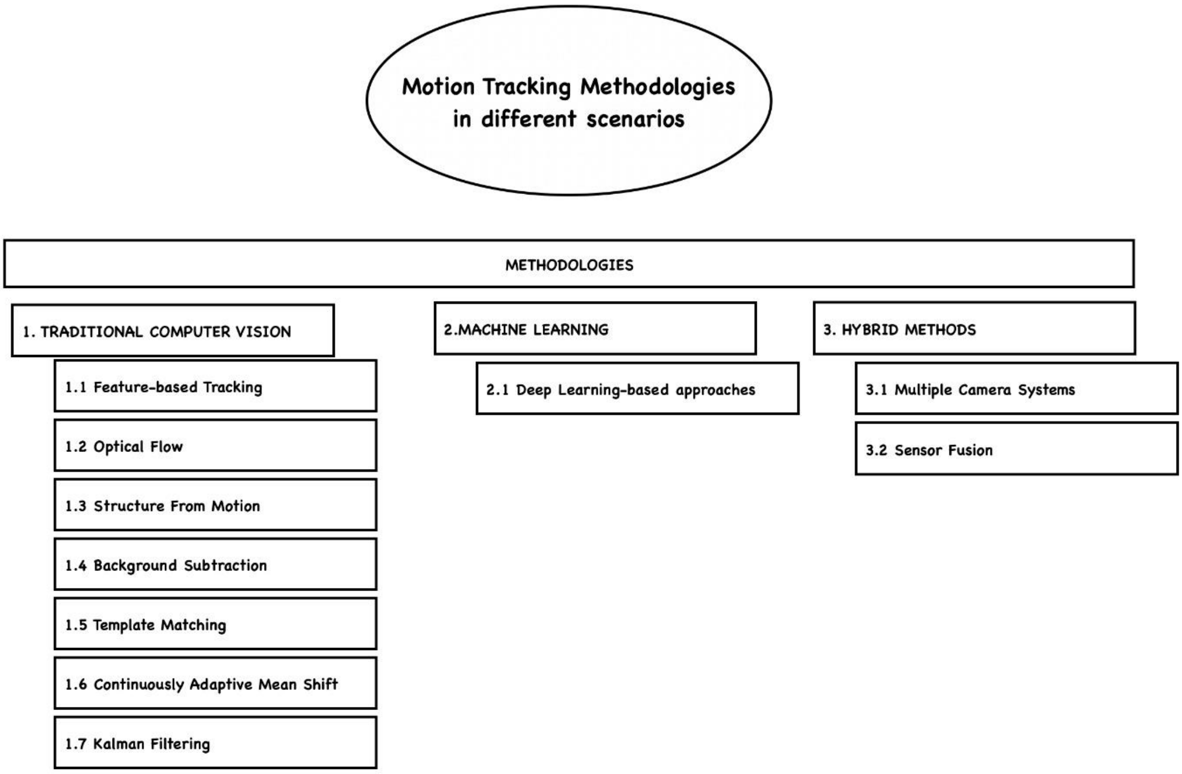

Figure 1 - Methodologies in consideration of this paper in their base groups.

2.1.1. Sub-categories

When subcategorizing methodologies, one must note three fundamental differences between them. Methods can either be marker-based, markerless, or inertial. However, markerless and marker-based are almost on the same page, whereas inertial is entirely apart from these two. Some methods can perform in a markerless environment but perform better in a marker-based application.

2.1.1.1. Marker-Based

Marker-based motion tracking relies on using identifiable markers placed on objects to track their movement. These markers can be visual patterns or physical objects with distinct features. Cameras capture the motion of these markers in a scene, and computer vision algorithms analyze their positions to determine object trajectories. This approach is often employed in animation, virtual reality, and biomechanics motion capture (Azuma, 1997). The advantage of marker-based tracking lies in its precision and ability to capture intricate movements, making it a preferred choice in applications where fine-grained motion analysis is crucial. However, it may require controlled environments and specific setups for optimal performance (Billinghurst, Clark, & Lee, 2015). When field testing outdoors, there will be many outcomes where if the markers are not read in the system, the accuracy falls apart and loses all purpose. In conclusion, this method is not recommended when dealing with uncontrolled environments but will present good results in the opposite scenario (Foxlin, 2002). However, for this study's purpose, this marker-based subtopic will be the method that either inherently relies on markers or can be significantly improved by using them. For example, while feature-based tracking and template matching are fundamentally markerless, they can also be adapted to use markers for enhanced precision, especially in controlled environments or where high accuracy is required (Zhang, He, & Duan, 2020).

2.1.1.2. Markerless

Markerless motion tracking eliminates the need for physical markers by relying on algorithms to analyze and interpret a scene's natural features and patterns. When placing markers on objects or subjects is impractical, such as in real-world environments or spontaneous interactions, these practices become extra handy (Lowe, 2004). These algorithms often involve feature extraction, pattern recognition, and depth sensing. This flexibility makes it suitable for applications like gesture recognition, 3D reconstruction, and augmented reality, where capturing natural motion is essential (Redmon et al., 2016). While markerless tracking offers versatility, it can pose challenges in accurately capturing complex movements and may require sophisticated algorithms to handle various environmental conditions. In the end, markers can always increase the accuracy of these methods (Zhang, He, & Duan, 2020).

2.1.1.3. Inertial

Inertial motion tracking relies on sensors like accelerometers and gyroscopes to capture and calculate the user's position by measuring an object's acceleration and angular velocity. By integrating these sensor readings over time, the system can estimate the object's position and orientation accurately, supposing there are no occlusions of the sensors (Madgwick et al., 2011). This approach is beneficial in scenarios with limited visual cues, such as in GPS-denied environments, enclosed spaces or extremely crowded scenarios. Inertial motion tracking is commonly employed in wearable devices, virtual reality headsets, and sports equipment for activities like fitness tracking (Foxlin, 2002). While inertial tracking provides real-time data and is less affected by environmental factors, it may be prone to drift over time, requiring occasional corrections from external references for improved accuracy (Han et al., 2014).

2.1.2. Environments

The deployment of motion tracking methodologies across diverse environments introduces a complex list of challenges researchers and developers must navigate. Optical tracking systems may thrive in controlled laboratory settings, offering high precision and detailed spatial information. However, these systems face significant challenges when transitioning to uncontrolled outdoor environments where factors such as varying lighting conditions, natural elements, and unpredictable obstacles pose formidable challenges to their efficacy. The difficulty lies not only in selecting a methodology that meets the precision requirements of the specific tracking system but also in ensuring adaptability to the dynamic and unpredictable variables inherent in diverse environments.

2.1.2.1. Controlled

In motion tracking, the significance of controlled environments cannot be overstated. These environments are meticulously designed to regulate external variables such as lighting, background, and object movement, thereby significantly enhancing the accuracy and reliability of motion tracking methodologies. This level of environmental control reduces the unpredictability commonly associated with uncontrolled, real-world scenarios, facilitating more precise and reliable tracking outcomes (Azuma, 1997).

Consider, for example, film, animation and television studios that employ green screens and consistent lighting setups to track actors or objects meticulously. This controlled setup enables the seamless integration of computer-generated imagery (CGI) in post-production, achieving visual effects nearly indistinguishable from real life (Guo & Zhong, 2022). Similarly, scientific research often relies on laboratory settings with uniform lighting and backgrounds to conduct motion tracking experiments. Such an environment ensures that variable environmental factors do not adulterate the movement data, leading to more valid and reliable research outcomes. Precise assembly and handling processes are paramount in the manufacturing and robotics industry. Controlled factory environments allow for the accurate tracking of machinery and robotic arms, ensuring efficiency and precision in production lines (Cognex, 2019).

The efficacy of motion tracking in these controlled settings can be attributed to several key factors. Consistent lighting is paramount; it enables the tracking systems to distinguish the object of interest from its background more accurately. Variations in lighting can introduce unwanted shadows and highlights, potentially misleading tracking algorithms. Additionally, minimizing external interferences such as vibrations, unexpected sounds, or reflective surfaces is crucial in maintaining the integrity of the tracking data (Stauffer & Grimson, 1999).

For more precise outcomes in controlled environments, the use of high-resolution cameras and sensitive sensors is highly recommended. These tools can capture more detailed movement data, thereby refining the motion tracking process. Regular calibration of the tracking system is also essential; it ensures sustained sensitivity and accuracy to the movements of interest, compensating for any minor environmental changes over time. Moreover, the integration of multiple tracking systems—combining optical with inertial technologies, for instance—can offer a more robust solution by compensating for the limitations inherent in individual tracking methods (Han et al., 2014).

Controlled environments are not just necessary in motion tracking, they are a catalyst for pushing the boundaries of what can be achieved. By managing external variables, these settings not only enhance the performance of motion tracking systems but also pave the way for advancements in visual effects, scientific research, and manufacturing processes. As technology continues to evolve, the integration of advanced hardware, sophisticated software solutions, and precise calibration techniques will undoubtedly unlock new possibilities, further enhancing motion tracking accuracy and reliability in these meticulously curated environments (Zhang, He, & Duan, 2020).

2.1.2.2. Uncontrolled

Navigating motion tracking challenges in uncontrolled environments presents a distinct set of complexities. Unlike their controlled counterparts, these settings are characterized by their unpredictability, with factors such as changing lighting conditions, dynamic backgrounds, and the unscripted movement of objects. This variability introduces significant challenges to motion tracking systems, demanding high adaptability and robustness to ensure accurate tracking outcomes (Billinghurst, Clark, & Lee, 2015).

Uncontrolled environments are commonly encountered in real-world scenarios, ranging from outdoor landscapes to bustling urban settings. For instance, tracking wildlife in their natural habitats requires coping with varying weather conditions, diverse terrains, and the animals' spontaneous behaviors. Similarly, urban surveillance systems must contend with the dynamic nature of city life, including fluctuating lighting from day to night, seasonal changes, and the constant movement of people and vehicles. In sports analytics, capturing athletes' performance outdoors involves dealing with shadows, crowd movements, and the fast pace of the game, all of which can impact the quality of motion tracking (Fani et al., 2017).

The inherent unpredictability of these environments necessitates motion tracking methods that can dynamically adjust to changing conditions. In this context, deep learning-based approaches have emerged as a powerful tool. These methods leverage extensive datasets to train models, enabling them to recognize and track objects accurately despite variations in appearance, lighting, or background. The adaptability of deep learning models allows them to discern the object of interest from a cluttered or changing background, a critical capability in unpredictable settings (Goodfellow et al., 2014).

Moreover, integrating sensor fusion techniques enhances motion tracking in uncontrolled environments. Combining data from multiple sources, such as optical cameras, infrared sensors, and GPS, achieves a more comprehensive understanding of the scene. This multi-modal approach mitigates the limitations of individual sensors, providing a resilient solution capable of delivering accurate tracking data under less-than-ideal conditions ([Hasanujjaman, Chowdhury, & Jang, 2023).

Achieving reliable motion tracking in uncontrolled environments also relies on algorithms capable of filtering noise and managing inconsistencies in the data collected. These algorithms play a crucial role in distinguishing between relevant movement patterns and background noise, ensuring that the tracking process remains focused on the object of interest despite external disturbances (Stauffer & Grimson, 1999).

Uncontrolled environments pose significant challenges to motion tracking due to their inherent unpredictability and complexity. Successfully navigating these challenges requires leveraging advanced technologies such as deep learning and sensor fusion, coupled with robust algorithms designed to filter noise and handle variability. These strategies enable motion tracking systems to adapt dynamically to changing conditions, ensuring accurate and reliable tracking outcomes even in the most challenging real-world scenarios. As technology advances, the development of more sophisticated models and algorithms will continue to improve the efficacy of motion tracking in uncontrolled environments, expanding its applicability across diverse fields and applications (Zhang, He, & Duan, 2020).

2.1.3. Areas of Expertise

Motion tracking technology encompasses several areas of expertise, making it a versatile tool across various industries. In animation and filmmaking, motion tracking—often referred to as motion capture (mocap)—is crucial for creating lifelike digital characters by capturing actors' movements and translating them into CGI performances. Notable applications include films like "Avatar" and "The Lord of the Rings," where mocap has brought complex characters to life with remarkable realism (What Is Motion Tracking? - a Guide to Understanding Motion Tracking Technology - Sanguine Production, 2023; Crawford, 2024; DeGuzman, 2023).

In AR and VR, motion tracking enhances user experience by accurately mapping physical movements to virtual environments, allowing for more immersive interactions. This technology is integral to virtual reality (VR) headsets, which track head and body movements to create a seamless virtual experience (What Is Motion Tracking? - a Guide to Understanding Motion Tracking Technology - Sanguine Production, 2023).

Sports science also benefits from motion tracking by analyzing athletes' movements to optimize performance and prevent injuries. This data-driven approach helps refine techniques and develop training programs tailored to individual athletes' needs (Crawford, 2024).

In healthcare, motion tracking is used in rehabilitation to monitor patients' progress and ensure proper exercise execution. By providing real-time feedback, it helps therapists adjust treatments and improve recovery outcomes (Guo, Yong & Zhong, Chan., 2022).

Furthermore, motion tracking is essential in robotics and human-computer interaction, enhancing machine responsiveness and accuracy. This technology enables robots to interact more naturally with their environment and improves the precision of control systems in various applications (What Is Motion Tracking? - a Guide to Understanding Motion Tracking Technology - Sanguine Production, 2023; Crawford, 2024).

Overall, motion tracking's ability to capture and analyze movement patterns is transforming multiple fields, making it an invaluable tool for innovation and efficiency across diverse applications.

2.1.4. What it does

Motion tracking is a sophisticated technology used to follow and record the movement of objects or people within a video frame. By employing a combination of sensors, algorithms, and computer vision techniques, motion tracking systems capture and analyze motion data, which can be applied to various applications. These sensors often include infrared cameras, accelerometers, gyroscopes, and magnetometers, working together to measure position, velocity, and acceleration in real time (What Is Motion Tracking? - a Guide to Understanding Motion Tracking Technology - Sanguine Production, 2023; Motion Tracking in after Effects | Adobe, n.d.) .

In practical terms, motion tracking is widely used in fields such as film production, video games, and VR. For instance, in film production, motion tracking enables the seamless integration of CGI with live-action footage, allowing for the creation of realistic visual effects (What Is Motion Tracking? - a Guide to Understanding Motion Tracking Technology - Sanguine Production, 2023; The 6 Types of Motion Tracking in after Effects – 2022 Guide |

Evercast Blog, n.d.). This technology enhances player experience in video games by accurately capturing and reflecting their movements, leading to more immersive gameplay (Motion Tracking in after Effects | Adobe, n.d.). Similarly, VR applications utilize motion tracking to monitor and respond to users' head movements, thus providing a more interactive experience (Introduction to Motion Tracking, n.d.).

2.1.5. Products References

Motion tracking technology has become increasingly prevalent across various industries, with several products being the most known and widely used today. Among these, Microsoft HoloLens and Apple ARKit (Apple, 2020; (Microsoft HoloLens | Mixed Reality Technology for Business, n.d.) are particularly prominent. Microsoft HoloLens, a mixed-reality headset, leverages advanced motion tracking to provide immersive augmented and virtual reality experiences, making it a leading tool in fields like education, healthcare, and design. Similarly, Apple ARKit has revolutionized the development of augmented reality applications on iOS devices, enabling precise motion tracking through sophisticated computer vision algorithms.

2.2 Traditional Computer Vision Techniques

Motion tracking in computer vision involves the analysis of successive frames of video to understand and interpret the movement of objects within a scene. Computer vision algorithms utilize techniques such as optical flow, feature tracking, and background subtraction to identify and follow the trajectory of objects over time. This method is mostly applied in surveillance, object recognition, and augmented reality, where the accurate monitoring of motion provides crucial information for understanding and interacting with the environment. Computer vision-based motion tracking plays a pivotal role in various industries, contributing to robotics, autonomous vehicles, and human-computer interaction advancements.

2.2.1. Methodologies in Consideration

As a part of the traditional computer vision techniques, the methods belonging to this group are the following:

Feature-based Tracking:

- Point Feature Tracking
- Object Recognition and Tracking

Optical Flow:

- Optical flow algorithms

Structure from Motion (SfM):

- 3D Reconstruction

Background Subtraction:

- Foreground Detection

Template Matching:

- Template-based Tracking

CamShift (Continuously Adaptive Mean Shift):

- Object Tracking

Kalman Filtering:

- Kalman Filters

2.2.1.1. Feature-based Tracking

Feature-based Tracking relies on identifying and tracking distinct features (such as corners or textured areas) across successive frames of video. It requires the initial detection of these features in the first frame and then follows their movement by comparing their appearance in subsequent frames. This method is particularly useful for tracking objects in scenes where the object's motion is relatively small and the features remain visible and unchanged (Lowe, 2004).

2.2.1.2. Optical Flow

Optical Flow calculates the motion tracking of the target by comparing two consecutive frames, which can be caused by either the movement of the target itself or by the camera. It requires a sequence of images where the motion is small enough that the objects' appearance does not change significantly between frames. The method estimates the motion at each pixel, creating a flow field that represents the movement direction and speed of every part of the image (Horn, 1981).

2.2.1.3. Structure from Motion

SfM involves reconstructing a three-dimensional structure of a scenario or an object. This method utilizes a series of images from different viewpoints and works by finding matching features across these images. By analyzing the apparent motion of these features relative to the camera, SfM algorithms can deduce the depth and structure of the scene and reconstruct it digitally (Snavely, 2008).

2.2.1.4. Background Subtraction

Background Subtraction is a technique used to segment moving objects from the static background in video sequences. It requires a clear view of the background or a model to compare against the current frame. When comparing the image from the background and again when an object appears in the frame, the object gets detected, effectively separating moving objects from the background (Stauffer, 1999).

2.2.1.5. Template Matching

Template Matching searches for parts of an image that match a template image. It requires a template or a small image of the object to be tracked and compared to every location in the larger image to find the best match. This method is straightforward but can struggle with orientation, scale, or lighting changes (Lewis, 1995).

2.2.1.6. CamShift

CamShift (Continuously Adaptive Meanshift) extends the Mean Shift algorithm to allow for the tracking of objects with changing size and orientation. It requires an initial location and the object's color distribution. The algorithm then adjusts the search window size and orientation to continuously track the object moving and changing shape (Bradski, 1998).

2.2.1.7. Kalman Filtering

Kalman Filtering is a predictive model used in motion tracking. It requires an initial estimate of the object’s state (such as its position and velocity) and a model of the system’s dynamics and noise characteristics. The filter predicts the object’s future state and updates its estimates with each new measurement to improve accuracy over time (Kalman, 1960).

2.2.2. Areas of Expertise

Traditional computer vision techniques for motion tracking encompass several areas of expertise that collectively enhance the ability to track and analyze motion accurately. Key areas include image processing, where algorithms are developed to filter and enhance visual data. Digital image processing encompasses various techniques to enhance visual data for subsequent analysis (Gonzalez & Woods, 2018). Object detection implicates identifying and locating objects between frames. This approach achieves high frame rates and detection accuracy by combining simple features using a cascade of classifiers (Viola & Jones, 2001). Feature extraction focuses on identifying relevant features like edges, textures, and shapes. The approach identifies and matches local features invariant to scale and orientation, which are crucial for robust object recognition (Lowe, 1999).

Additionally, motion estimation techniques are crucial for predicting and understanding movement patterns. This iterative technique aligns images by minimizing the error between a template and a target image, facilitating accurate motion estimation (Lucas & Kanade, 1981). Pattern recognition enables the classification and interpretation of visual data. Pattern recognition systems categorize input data into predefined classes based on learned patterns (Duda et al., 2001). These areas of expertise are supported by advancements in machine learning, which enhance the adaptability and accuracy of motion tracking systems. Together, these domains contribute to a robust framework for applications in surveillance, augmented reality, robotics, and human-computer interaction, demonstrating the multifaceted nature of traditional computer vision methodologies in motion tracking.

2.2.3. Products References

Apple Inc., with its ARKit (Apple, 2020), has democratized augmented reality, transforming how people interact with digital content in their physical environments. By enabling developers to create augmented reality (AR) applications easily, Apple has facilitated educational tools that make learning more interactive and gaming experiences that blend the virtual with the real. The potential for further societal impact lies in expanding AR’s accessibility and utility, making it a tool for more than just entertainment, but also for practical applications in healthcare, education, and beyond.

Adobe’s mastery of Optical Flow in After Effects (Motion Tracking in after Effects | Adobe, n.d.) has revolutionized the creative industry, enabling filmmakers and visual artists to produce content that was once deemed impossible without significant resources. This technology empowers content creators to tell stories in more engaging ways, potentially enhancing educational content and social campaigns with more impactful visuals. Adobe's continued innovation can further support independent creators and small studios, leveling the playing field in the creative industry.

Autodesk's utilization of SfM in tools like ReCap has redefined heritage conservation, urban planning, and construction. By enabling accurate 3D reconstructions of historical sites and landscapes, Autodesk helps preserve cultural heritage and aids in disaster recovery by rapidly assessing and planning reconstructions. Expanding the application of SfM to more collaborative, cloud-based platforms could further democratize access to this technology, allowing communities worldwide to engage in preserving their heritage and planning their futures. (ReCap Software | Get Prices & Buy ReCap pro 2023 | Autodesk, 2020)

Axis Communications, through its advanced surveillance technologies, has enhanced public safety and security. By improving the reliability and efficiency of surveillance systems, Axis helps protect public spaces, reducing crime rates and increasing community safety. The next step involves integrating AI and ethical considerations more deeply into surveillance practices, ensuring privacy protection and fostering a safer, more secure society without intrusive monitoring. (Axis Communications - Leader in Network Cameras and Other IP Networking Solutions | Axis Communications, n.d.)

Cognex's template matching technology in industrial settings ensures sustainable manufacturing practices by reducing waste. By further refining the accuracy and speed of their machine vision systems, Cognex can aid industries in achieving higher efficiency and lower environmental impact, promoting a move towards more sustainable industrial practices. (Cognex | Machine Vision and Barcode Readers, 2019)

Intel's OpenVINO toolkit and the CamShift algorithm (Intel® Distribution of OpenVINOTM Toolkit, n.d.) have pushed the boundaries of real-time video analytics, opening new avenues for personalized customer experiences and enhanced security measures. Intel's ongoing development in AI and machine learning could further personalize user interactions without compromising privacy, creating more secure and engaging environments in retail, healthcare, and public services.

Garmin Ltd.'s precision in location tracking through Kalman Filtering enhances outdoor activities and personal fitness, encouraging healthier lifestyles. By continuing to innovate in wearable technology, Garmin can provide more insightful health data, promoting wellness and preventive healthcare. (GPS Accuracy | Garmin Centro de Assistência, n.d.)

2.3. Machine Learning

Motion tracking in machine learning involves the application of algorithms that can learn and predict the movement patterns of objects over time. By analyzing labeled datasets, machines recognized as can be trained to recognize and track analyzing motion behaviors. These models can then generalize their understanding of new, unseen data, enabling robust and adaptive motion tracking systems. Machine learning enhances motion tracking accuracy by continuously refining predictions based on real-time information. This method is particularly useful in dynamic environments where the motion patterns may be complex or unpredictable. The ability of machine learning to adapt and improve over time makes it a powerful tool for motion tracking in various applications, including video analysis, sports analytics, and healthcare monitoring.

2.3.1. Methodologies in Consideration

As a part of the machine learning techniques, the method belonging to this group is as follows:

Deep Learning-based Approaches:

- Convolutional Neural Networks (CNNs)
- Object detection and tracking using deep learning

2.3.1.1. Deep Learning-based Approaches

Deep Learning-based Approaches utilize neural networks to learn how to recognize and track objects from large amounts of data. These methods require extensive datasets for training but are highly effective in handling variations in object appearance, lighting, and occlusions, making them robust across different tracking scenarios (Leal-Taixé, 2015).

2.3.2. Areas of expertise

Deep learning-based approaches represent a cutting-edge area of expertise within machine learning techniques for motion tracking, leveraging neural networks to achieve unprecedented accuracy and robustness. Central to these approaches are CNNs, which excel at processing and analyzing visual data to detect and track motion. CNNs automatically learn hierarchical representations of input images, improving recognition accuracy (LeCun et al., 1998). By learning hierarchical feature representations directly from raw data, CNNs can identify complex patterns and relationships that traditional methods might miss. Recurrent neural networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, are also pivotal in modeling temporal dependencies and predicting future motion based on past observations. LSTM networks address the vanishing gradient problem, making them practical for modeling long-term dependencies in sequential data (Hochreiter & Schmidhuber, 1997).

Additionally, techniques such as Generative Adversarial Networks (GANs) enhance the realism and precision of motion tracking by generating high-quality synthetic data for training and validation. (Goodfellow et al., 2014). These deep learning methodologies are particularly effective in handling large-scale datasets, adapting to diverse and dynamic environments, and achieving real-time performance, making them indispensable for autonomous vehicles, surveillance, sports analytics, and virtual reality applications. The integration of deep learning in motion tracking not only improves accuracy but also enables the development of more intelligent and adaptive systems.

2.3.3. Products References

With its deep learning platforms, NVIDIA Corporation (World Leader in AI Computing, n.d.) drives advancements in autonomous vehicles and innovative city initiatives, aiming for safer, more efficient urban environments. NVIDIA's focus on ethical AI and collaboration with urban planners can accelerate the adoption of intelligent technologies in cities, improving traffic management, reducing emissions, and enhancing public services.

2.4. Hybrid Methods

Hybrid techniques combine the strengths of both traditional computer vision methods and machine learning algorithms. By integrating feature-based tracking from computer vision with machine learning capabilities, hybrid techniques offer a more comprehensive approach to motion analysis. These systems can adapt to various scenarios, leveraging predefined rules for specific tasks while learning from data to enhance performance. This synergy allows for more accurate and flexible motion tracking across diverse environments. Hybrid approaches are efficient in applications where a combination of rule-based and learned behaviors is essential, such as in robotics, where precise and adaptable motion tracking is crucial for navigation and interaction with the surroundings.

2.4.1. Methodologies in Consideration

Finally, consisting in the hybrid techniques, the methods a part of this group are the following:

Multiple Camera Systems:

- Stereo Vision and Multi-view Geometry

Sensor Fusion:

- Combining camera data with other sensor data for enhanced tracking accuracy.

2.4.1.1. Multiple Camera Systems

Multiple Camera Systems use data from several cameras to track objects in three dimensions. This method requires multiple synchronized cameras positioned around the scene. By analyzing the differences in images taken from these various viewpoints, it is possible to triangulate the position of objects in the scene, allowing for accurate 3D tracking (D'Angelo, 2012).

2.4.1.2. Sensor Fusion

Sensor Fusion combines information from multiple sources, such as cameras, radar, and lidar, to track objects more accurately than any single sensor. It requires the integration of data that may have different formats, scales, or perspectives. This method improves tracking accuracy and reliability by leveraging the strengths of each sensor type to compensate for their weaknesses (Hasanujjaman, 2023).

2.4.2. Areas of expertise

Hybrid motion tracking methods represent an innovative convergence of multiple tracking techniques, leveraging the strengths of each to address the limitations inherent in individual approaches. These methods often combine optical tracking with inertial sensors, integrating the precise positional data from cameras with the orientation and movement data from IMUs to enhance accuracy and robustness (Bleser & Stricker, 2008). By fusing data from various sources, hybrid approaches can effectively manage occlusions, reduce drift, and improve tracking reliability in challenging environments (Zhang et al., 2020). Another common hybrid technique involves the integration of radio frequency (RF) tracking with vision-based systems, where RF signals provide coarse localization, and visual data refines the tracking accuracy (Foxlin, 2002). Machine learning algorithms, intense learning-based ones, are increasingly employed in hybrid systems to intelligently combine and interpret diverse data streams, leading to more adaptive and resilient tracking solutions (Han et al., 2014). These methodologies are particularly valuable in AR, VR, autonomous navigation, and advanced human-computer interaction applications, ensuring seamless and continuous tracking performance in complex, real-world scenarios (Ojeda & Borenstein, 2007).

2.4.3. Products References

Sony Corporation’s multi-camera systems enrich sports analytics and entertainment, offering audiences unparalleled perspectives and insights. By advancing the integration of VR and AR with their camera systems, Sony can transform spectator experiences, making remote viewing as immersive as being at the event and connecting people across distances in new and exciting ways. (Xu & Huang, 2022)

Tesla, Inc.‘s advancements in sensor fusion technology are not just about creating autonomous vehicles; they are about envisioning a future with fewer traffic accidents, reduced congestion, and lower emissions. Tesla’s continuous improvement in autonomous technology and its commitment to sustainability could transform urban mobility, fostering environments that prioritize people over cars and promoting a more sustainable, efficient, and safe urban landscape.(Hasanujjaman et al., 2023)

2.5. Open issues & challenges

Incorporating motion-tracking technologies into various facets of daily life, leading companies' contributions have driven technological advancements and wielded a profound impact on society. Through their innovative applications, these companies are reshaping industries, enhancing user experiences, and addressing complex societal challenges.

These companies, through their pioneering work in motion tracking technologies, are not merely advancing their respective fields but are also making significant contributions to societal well-being, environmental sustainability, and cultural preservation. By focusing on ethical considerations, accessibility, and sustainability, they can amplify their positive impact on society, ensuring that technology serves as a force for good, improving lives and safeguarding the planet for future generations.

2.6. Recap

Table 1 summarizes chapter 2.1, with each motion tracking organized with their major group and where they fit with each subtopic.

Motion Tracking Method	**Major Group**	**Marker-based**	**Markerless**	**Inertial**
Feature-based Tracking	Traditional Computer Vision Techniques	✓	✓	x
Optical Flow	Traditional Computer Vision Techniques	x	✓	x
SfM	Traditional Computer Vision Techniques	✓	✓	x
Background Subtraction	Traditional Computer Vision Techniques	✓	✓	x

Template Matching	Traditional Computer Vision Techniques	✓	✓	x
CamShift	Traditional Computer Vision Techniques	✓	✓	x
Kalman Filtering	Traditional Computer Vision Techniques	✓	✓	x
Deep Learning-based Approaches	Machine Learning Techniques	✓	✓	x
Multiple Camera Systems	Hybrid Approaches	✓	✓	x
Sensor Fusion	Hybrid Approaches	✓	✓	✓

Table 1 - Categorization of each motion tracking method in their base groups and subtopics.

3. DEVELOPMENT METHODOLOGY

This chapter details the research design, methods, materials, and procedures employed to develop a framework for identifying and recommending the most suitable tracking methodologies based on specific input parameters. It alsoIt also outlines the approach taken to achieve the research objectives, which primarily involves synthesizing data from existing research papers, articles, and books rather than conducting original empirical testing.

This study adopts a qualitative research approach, focusing on a comprehensive literature review to gather and synthesize existing knowledge on tracking methodologies. The research design is structured to analyze and evaluate various tracking methods based on previously published data, providing a robust framework for identifying the most effective methodologies under different conditions.

The data collection for this study is meticulous, involving an extensive review of existing research papers, articles, and books on motion tracking methodologies. This process includes identifying key characteristics, performance metrics, and contextual applications of different tracking methods.

The data analysis methods used in this research include qualitative synthesis, which involves the systematic analysis and synthesis of qualitative data from the reviewed literature to identify common themes and findings. Comparative analysis is conducted to compare different tracking methodologies based on their reported performance metrics and contextual suitability. Additionally, content analysis is used to extract relevant data points and performance indicators from the reviewed literature to inform the framework development.

The specific motion tracking methods chosen for this thesis were carefully selected based on their proven effectiveness and versatility across a wide range of applications. Feature-based Tracking and Optical Flow are fundamental techniques that offer reliable performance in various conditions, making them suitable for foundational tracking needs. Structure from Motion (SfM) and Multiple Camera Systems excel in 3D reconstruction and spatial analysis, crucial for applications requiring depth and spatial understanding. Background Subtraction and Template Matching provide robust solutions for controlled environments, particularly in static and semi-static scenarios. CamShift and Kalman Filtering are well-regarded for their real-time adaptability and predictive capabilities, essential for dynamic and rapidly changing environments. Deep Learning-based Approaches represent the cutting-edge in motion tracking, offering unparalleled accuracy and adaptability through neural networks. Sensor Fusion, on the other hand, integrates data from multiple sources to enhance tracking accuracy and robustness, particularly in complex environments. These methods were chosen over others due to their comprehensive coverage of tracking needs, proven reliability in diverse scenarios, and their ability to integrate with advanced analytical techniques, ensuring the final result remains robust and versatile.

The research procedures are as follows: first, a comprehensive search of academic databases is conducted to identify relevant literature on motion tracking methodologies. Essential information from the collected literature is then extracted, including descriptions of tracking methods, performance metrics, and application scenarios. This extracted data is synthesized and analyzed to identify patterns and correlations between input parameters and tracking performance. Based on the synthesized data, the framework is developed, incorporating the most effective methodologies for different input parameters. Finally, the framework is validated through critical analysis and comparison with established methods and guidelines presented in the literature.

Ethical considerations are minimal towards the framework as this research does not involve human participants or original data collection. One thing to keep in mind is the ethical practices of these methodologies and the basic sense of motion tracking. The primary ethical consideration is ensuring proper citation and acknowledgment of all sources used in the literature review to maintain academic integrity.

The potential limitations of this research include the dependence on existing papers, which means the framework's accuracy is bound upon the quality and scope of the research.

Additionally, the absence of original empirical testing may limit the ability to validate the framework in real-world scenarios.

To ensure reliability and validity, a comprehensive search strategy is employed to gather a wide range of relevant literature. Triangulation is used to cross-reference data from multiple sources to ensure the robustness of findings, and critical appraisal is conducted to assess the quality and relevance of the reviewed literature, ensuring the validity of the synthesized data.

This chapter has detailed the research design, methods, materials, and procedures used to develop a framework for identifying and recommending tracking methodologies based on specific input parameters. By relying on an extensive review and synthesis of existing research, this study aims to provide a robust and reliable framework capable of enhancing the selection process for motion tracking technologies.

In Figure 2, the timeline shows the essential points of the development of this thesis, as shown before, expanded from november from 2023 to june of 2024.

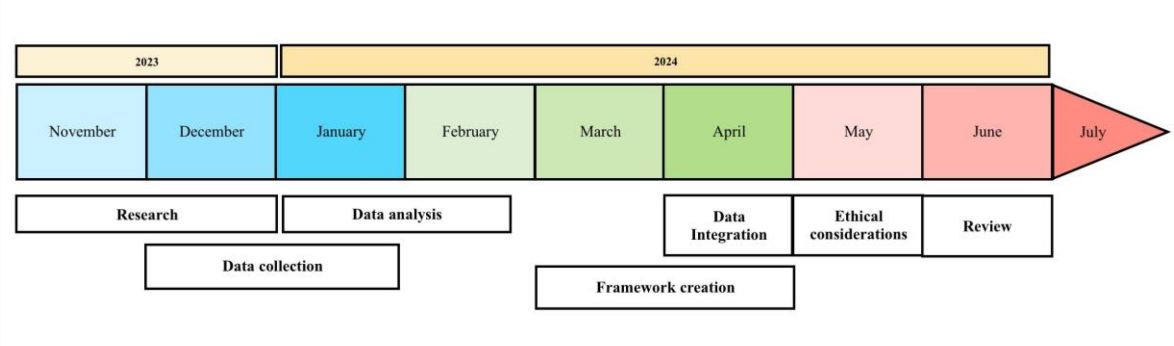

Figure 2 - Development timeline

4. MTDM Framework

The developed framework for selecting motion tracking methodologies is designed to facilitate decision-making by allowing users to input specific parameters related to their project needs. This system employs a 3-point scale for each input parameter, such as distance, to evaluate and recommend the most suitable tracking methods. For example, when considering the distance parameter, users can select from short, medium, or long distance based on their requirements. Each tracking method is assigned points based on its performance in these scenarios. For instance, background subtraction demonstrates excellent performance in short-distance scenarios, particularly in controlled environments with stationary camera setups. As a result, selecting the short-distance option deposits 3 points into the background subtraction method, reflecting its suitability. Conversely, it performs poorly in scenarios where the distance to the target increases, thus receiving no points for medium or long distances. This point-based system ensures that each method is evaluated according to its strengths and weaknesses as revealed by performance results. Additionally, certain project requirements may inherently make specific methods more compatible. For instance, if a project necessitates the use of sensors for enhanced tracking accuracy, sensor fusion becomes the most compatible option as it uniquely incorporates data from multiple sensors to improve performance. This approach allows users to make decisions by aligning their project needs with the most effective tracking methodologies.

The framework is implemented on a web page using HTML and Java, providing an accessible and user-friendly interface. At the beginning of the user experience, an option will allow users to specify a particular scenario or usability requirement. This feature was created to streamline the decision-making process by directly navigating users to the most appropriate tracking method based on their needs. Users can bypass the detailed input process and quickly identify the most suitable motion tracking methodology by selecting their scenario or intended use case, ensuring an efficient and tailored solution. This web-based framework aims to make the selection process intuitive and efficient, leveraging the versatility of HTML and Java to deliver a robust, interactive tool for users in various fields.

4.1 Inputs & Outputs

At the beginning of the framework, users will have the option to choose from predefined scenarios, such as news, robotics, 3D reconstruction, autonomous vehicles, multi-view geometry, multiple live targets, sports analytics, VR/AR, healthcare and rehabilitation, and security and surveillance, instead of customizing their inputs.

This decision is based on extensive research and analysis of the capabilities and applications of various motion tracking methodologies. For instance, deep learning-based approaches have shown significant effectiveness in real-time object detection for news broadcasting (Redmon et al., 2016). At the same time, sensor fusion and Kalman filtering are critical for accurate tracking in robotics (Kalman, 1960). SfM and multiple camera systems are highly suitable for 3D reconstruction due to their robust 3D modeling capabilities (Snavely et al., 2008, Furukawa & Ponce, 2010). Autonomous vehicles benefit significantly from sensor fusion and deep learning for navigation and object detection (Chen, 2017, Kendall & Cipolla, 2016). Multi-view geometry is best served by SfM and multiple camera systems, enabling precise geometric calculations (Hartley & Zisserman, 2004).

For tracking multiple live targets, deep learning and multiple camera systems provide robust solutions (Wojke et al., 2017; Leal-Taixé et al, 2015). Sports analytics utilize multiple camera systems and deep learning for detailed player tracking and performance analysis, which helps with the depth of the environment and near real-time results (Fani, 2017; McCarthy & Collins, 2018). VR and AR applications require accurate and low-latency tracking, achievable through sensor fusion and deep learning (Azuma, 1997; Billinghurst, 2015). In healthcare and rehabilitation, sensor fusion and Kalman filtering are essential for patient monitoring and therapy (Giggins et al., 2013; Pérez, 2010). Finally, security and surveillance applications are enhanced by deep learning and multiple camera systems, providing comprehensive monitoring capabilities (Hampapur et al., 2005; Zhan et al., 2008). These conclusions are drawn from a detailed review of relevant literature and practical implementations, ensuring that the predefined scenarios align with the most effective methodologies for each application.

The framework for selecting motion tracking methodologies is designed with a range of inputs designed to match the user's specific project requirements. Whereas the outputs are the methodologies taken into consideration, which. depending on the inputs chosen, there will be given the five most compatible methodologies (Figure 3)

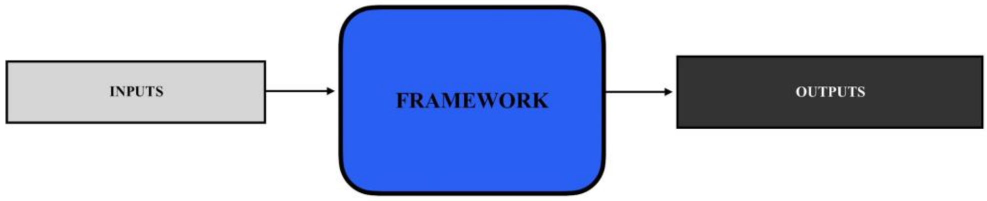

Figure 3 - Framework input and output connection

These inputs include Environment, Environmental Conditions, Distance, Processing Speed, Target, Integration Needs, Cost Constraints, and Time Constraints. Users can specify the Environment as Indoor or Outdoor, affecting the choice of methods suited to these different settings. Environmental Conditions allow for selection between Controlled, Dynamic, or Crowded environments, which influence the tracking method's effectiveness. Distance inputs let users choose between short, medium, and long distances, directing the framework to recommend methods that perform optimally within the specified range and avoiding overdoing implementation for better results. Processing Speed options of Low, Medium, and High ensure that the selected methodology aligns with the computational resources and real-time requirements of the project.

The Target input differentiates between Live Target, Object, Multiple Targets, and Environment, guiding the choice of methods based on the complexity and nature of the tracking subject.

Integration Needs are crucial, with options including None, Markers, and Sensors. Suppose the project specifies the need for Sensors, in that case, the framework will ultimately identify sensor fusion as the most compatible method due to its unique capability of combining data from multiple sensors to enhance tracking accuracy and robustness. In the situation where a user can utilize markers in the environment, a more compatible method will be recommended than if they choose not to. Cost Constraints, set to None or Low-cost, help in selecting methods that fit the budgetary limits.

Finally, Time Constraints of None or Quick Implementation prioritize methods that can be deployed within the project's timeline. This comprehensive set of inputs allows the framework to deliver tailored recommendations, ensuring the most effective and efficient tracking solutions for various scenarios. (Figure 4)

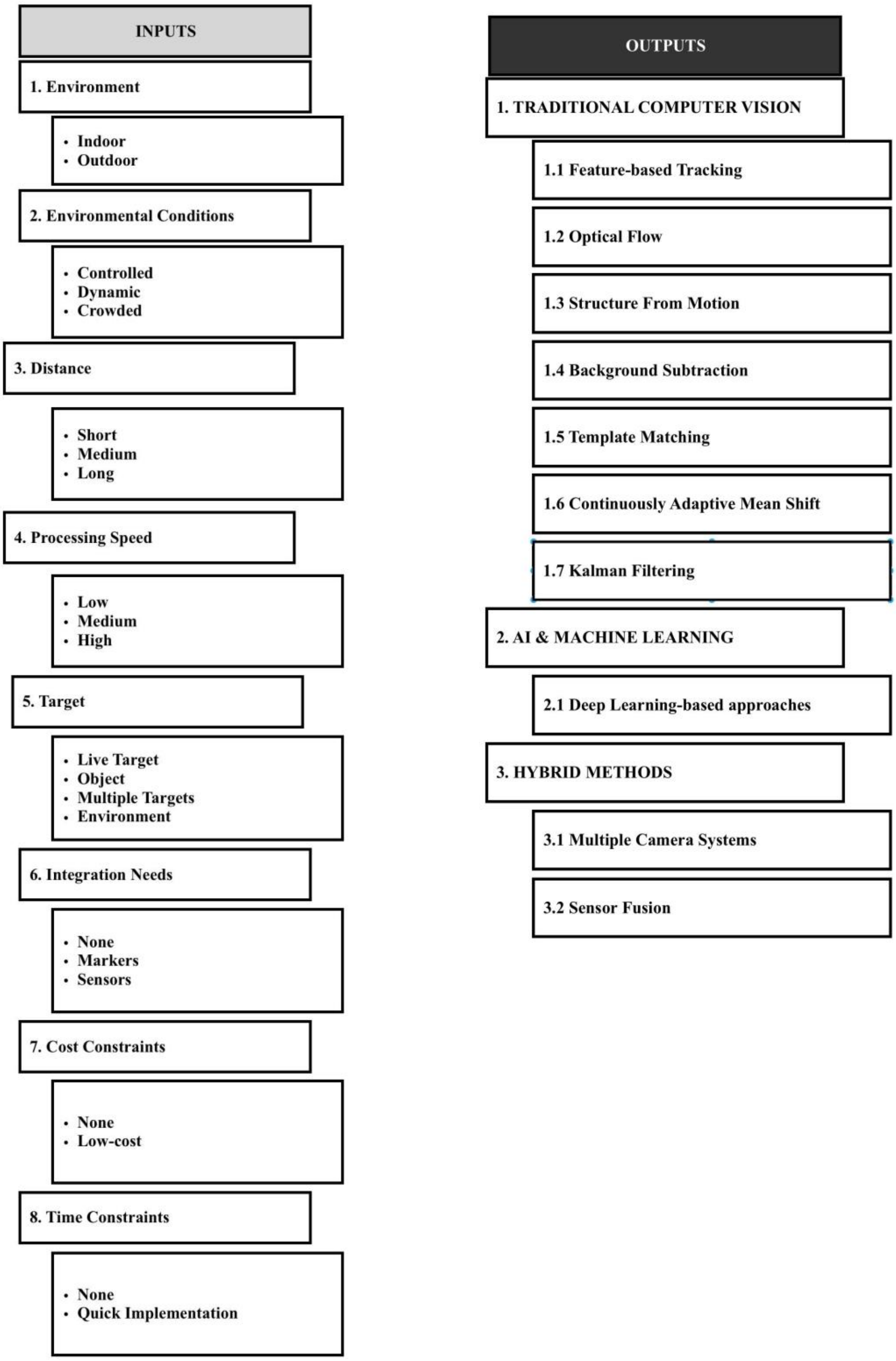

Figure 4 - Enumeration and showcase of all input and outputs of the framework

4.2 Decision Making

This chapter provides a detailed evaluation of various motion tracking methodologies, assigning points based on a 3-point scale across different scenarios. These scenarios encompass a wide range of factors that influence the effectiveness and suitability of tracking methods, including environmental conditions, distance, tracking speed, target type, integration needs, cost constraints, and time constraints, visible in Figure 5.

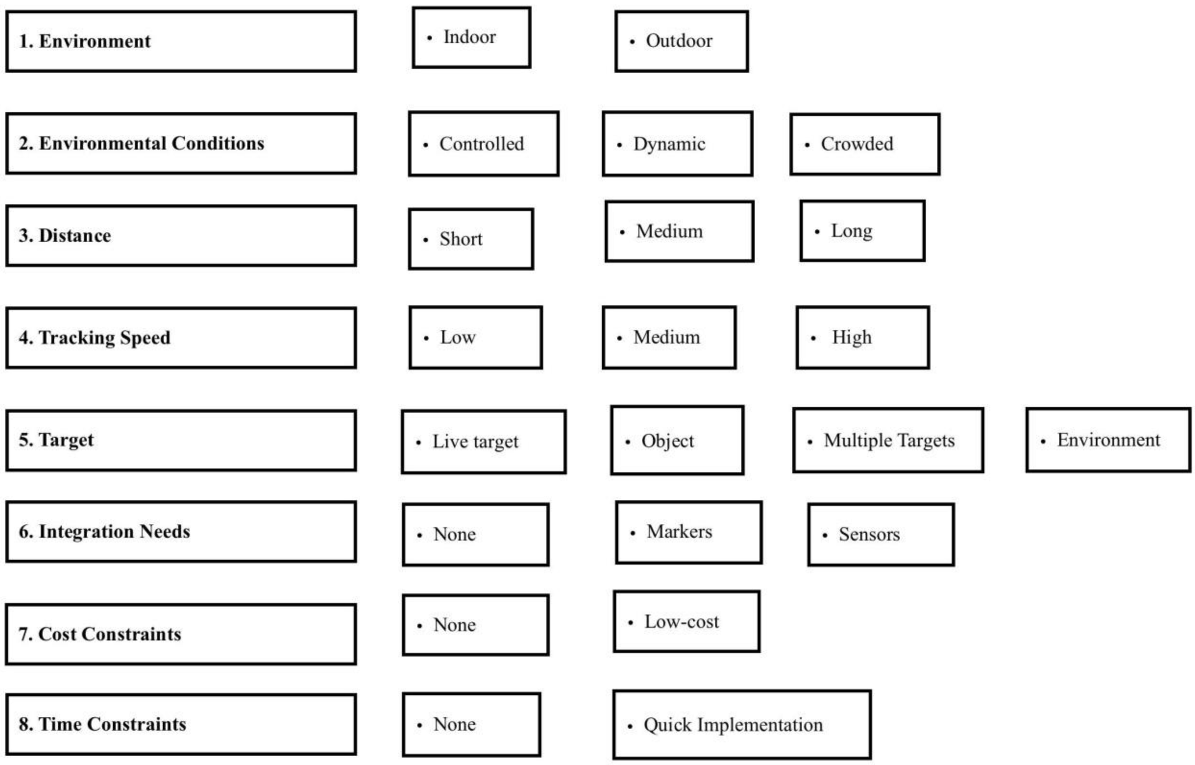

Figure 5 - Inputs of the framework.

The point system is designed to reflect the strengths and limitations of each method as derived from extensive literature review and analysis of academic articles and papers. This evaluation aims to guide users in selecting the most appropriate motion tracking technique tailored to their specific project requirements. Each method is scored to highlight its relative performance in different contexts, facilitating informed decision-making in the selection process.

The Tables 2,3,4 and 5, as well as Figure 6, 7 and 8, showcase the framework use of a point system to evaluate various motion tracking methodologies across different scenarios. Each method receives a score ranging from 0 to 3 from each input chosen, with each score representing the method's effectiveness in each scenario. Here's what each point value signifies:

- **+0**: The method is ineffective or unsuitable for the scenario. It performs poorly and is not recommended for use in this context.
- **+1**: The method is minimally effective. It may work under certain conditions but generally lacks reliability and robustness in this scenario.
- **+2**: The method is moderately effective. It performs well in most cases but may have some limitations or conditions where its performance is suboptimal.
- **+3**: The method is highly effective and well-suited for the scenario. It provides reliable, robust performance and is strongly recommended for use in this context.

	Environment		**Environmental Conditions**		
Motion Tracking Method	**Indoor**	**Outdoor**	**Controlled**	**Dynamic**	**Crowded**
Feature-based Tracking	+3	+1	+3	0	+2
Optical Flow	+3	+2	+3	+3	0
Structure from Motion	+3	+3	+3	+3	0
Background Subtraction	+3	+0	+3	0	0
Template Matching	+3	+2	+3	0	0
CamShift	+3	+2	+3	0	0

Kalman Filtering	+3	+3	+3	+2	0
Deep Learning-based	+3	+3	+3	+3	+3
Multiple Camera Systems	+3	+3	+3	+3	0
Sensor Fusion	+3	+3	+3	+3	+3

Table 2 - Framework 3-point system distribution on Environment and Environmental Conditions inputs.

	Distance			**Tracking Speed**		
Motion Tracking Method	**Short**	**Medium**	**Long**	**Low**	**Medium**	**High**
Feature-based Tracking	+2	+2	+1	+3	+3	0
Optical Flow	+3	+1	+1	+3	0	0
Structure from Motion	+3	+3	+2	+3	0	0
Background Subtraction	+3	0	0	+3	+3	+3
Template Matching	+2	+1	0	+3	+3	0
CamShift	+2	+2	+1	+3	+3	0
Kalman Filtering	+2	+2	+1	+3	+3	+3
Deep Learning-based	+3	+3	+3	+2	+1	0

Multiple Camera Systems	+3	+3	+2	+3	0	0
Sensor Fusion	+3	+3	+3	+3	+2	+1

Table 3 - Framework 3-point system distribution on Distance and Tracking Speed inputs.

	Target				**Integration Needs**		
Motion Tracking Method	**Live target**	**Object**	**Multiple targets**	**Environ ment**	**None**	**Marker s**	**Sensors**
Feature-based Tracking	+1	+2	0	0	+2	+3	0
Optical Flow	+3	0	0	0	+3	0	0
Structure from Motion	0	+3	0	+3	+2	+3	0
Background Subtraction	+3	0	0	0	+2	+3	0
Template Matching	0	+3	+1	0	+2	+3	0
CamShift	+2	+3	0	0	+2	+3	0
Kalman Filtering	+3	+3	0	0	+2	+3	0
Deep Learning-based	+3	+3	+3	+3	+2	+3	0
Multiple Camera Systems	+3	+3	+2	+2	+2	+3	0
Sensor Fusion	+3	+3	+3	+3	0	0	+3

Table 4 - Framework 3-point system distribution on Target and Integration Needs inputs.

	Cost Constraints		Time Constraints	
Motion Tracking Method	**None**	**Low-Cost**	**None**	**Quick Implementation**
Feature-based Tracking	+3	+1	+3	+1
Optical Flow	+3	+1	+3	+1
SfM	+3	0	+3	0
Background Subtraction	+3	+3	+3	+3
Template Matching	+3	+3	+3	+3
CamShift	+3	+2	+3	+1
Kalman Filtering	+3	+2	+3	+1
Deep Learning-based Approaches	+3	0	+3	0
Multiple Camera Systems	+3	0	+3	0
Sensor Fusion	+3	0	+3	0

Table 5 - Framework 3-point system distribution on Time and Cost Constrictions inputs.

4.3 WebApp

The web app is hosted on an Apache server, a web server software that handles client requests and delivers the web app's content to user's browsers. The HTML provides the structure of the web app, CSS ensures it is visually appealing and user-friendly, and JavaScript is where the framework point system is implemented. The web app can be accessed from anywhere, if it has an internet connection, and provides a seamless and efficient user experience.

Figure 9 shows the first thing user's step upon when accessing the web app, the Home page, where they must choose between two choices. They can either choose a predetermined scenario or application for the motion tracking or choose custom design.

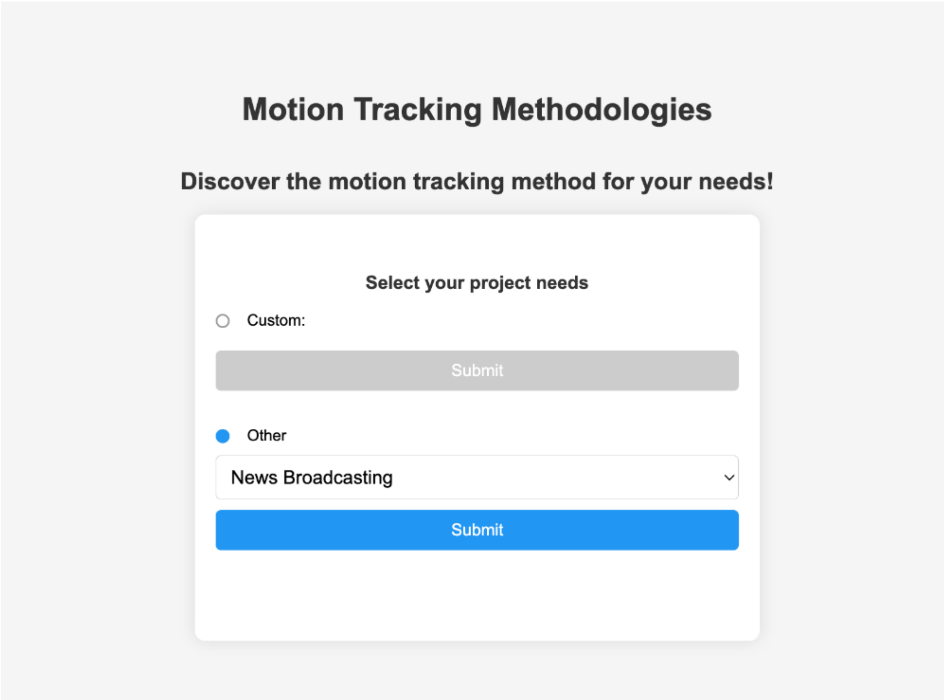

Figure 6 - Web app home page.

If the user decides to check the "other button", an option previously disabled becomes available, they will have the chance to choose the predetermined scenario (Figure 10).

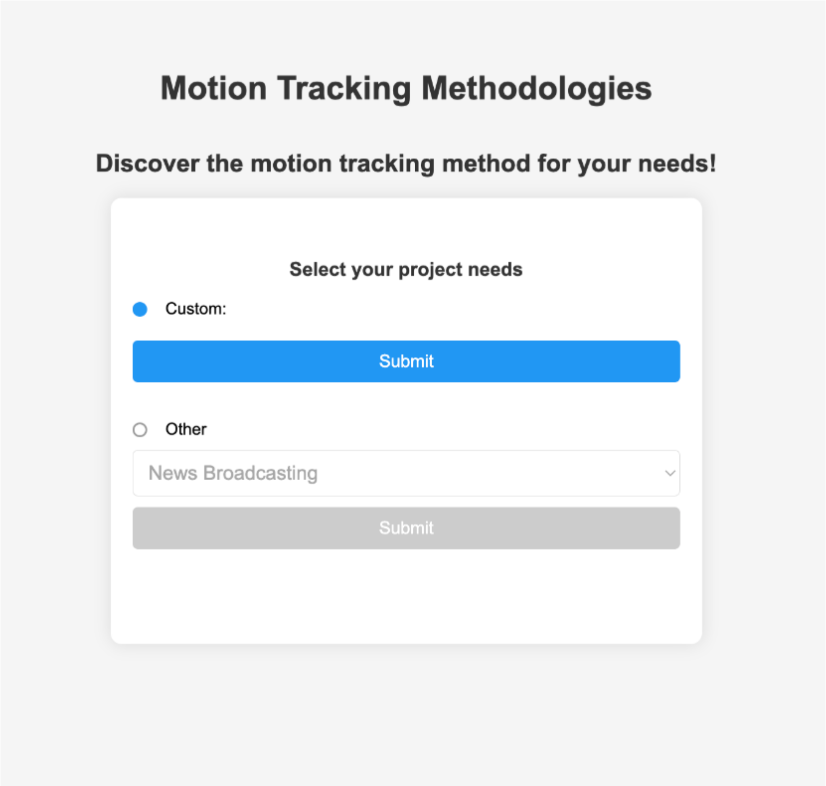

Figure 7 - Home Page when “Other” button is selected.

These scenarios were created based on the strengths and weaknesses table, and are the following: News Broadcasting, Robotics, 3D Reconstruction, Autonomous Vehicles, Multi-view Geometry, Multiple Live Targets, Sports Analytics, VR & AR, Healthcare and Rehabilitation, Security and Surveillance.

In Figure 11 we can see when one of these options is chosen, where the 3 most compatible motion tracking methods are in display. In this case we go directly to the Results page.

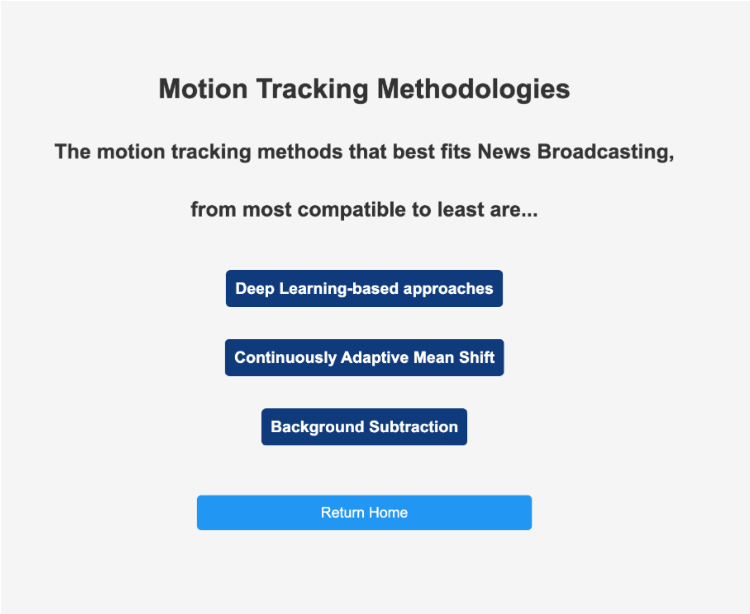

Figure 8 - Predetermined scenario in results page.

We then go back to the Home page, using the “Return Home” button, and if we select the Custom option, we will navigate to the questionnaire page, where we can see all of the inputs and choose between them to create our own scenario, illustrated in Figure 12.

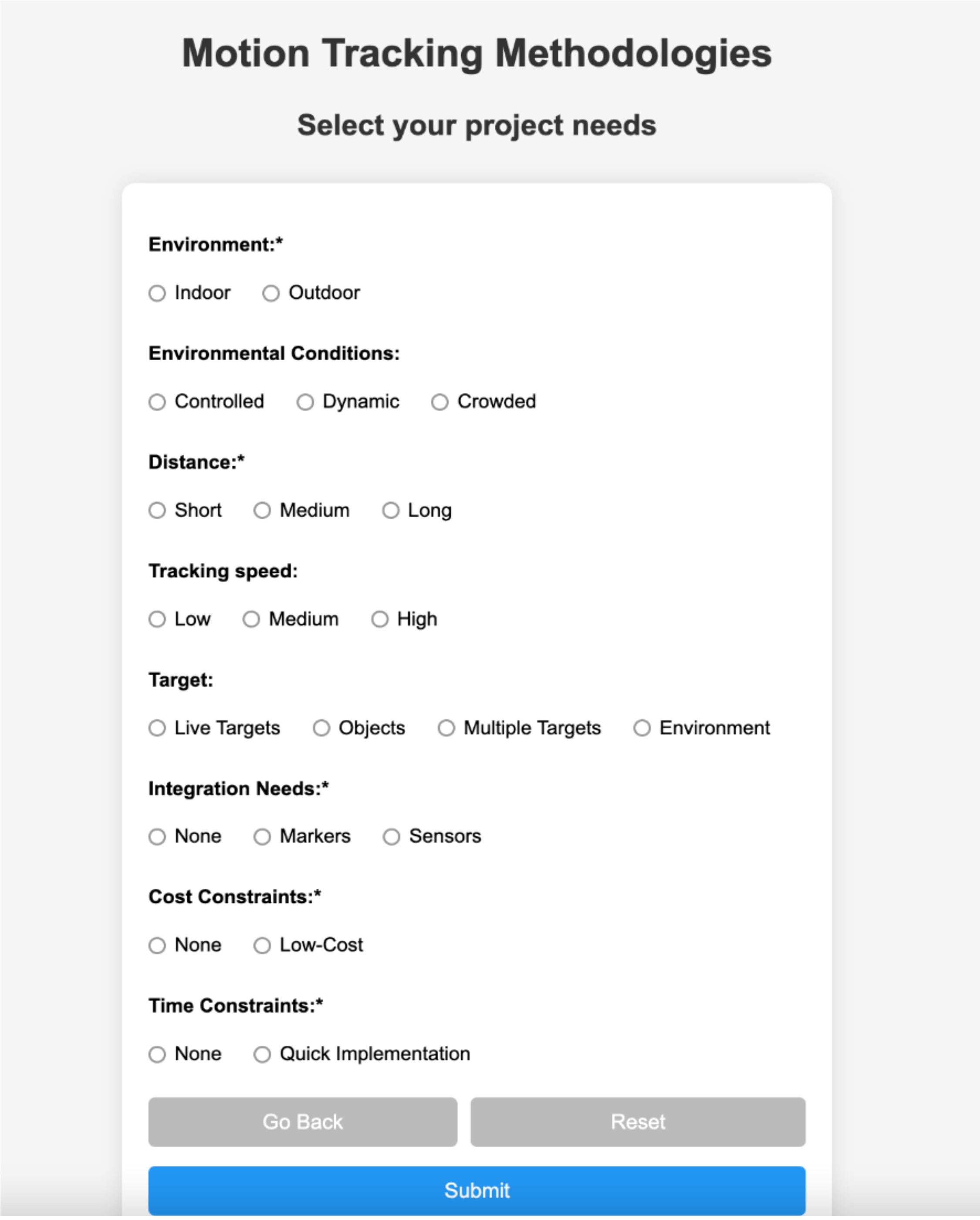

Figure 9 - Questionnaire page with all input options.

When choosing the custom inputs, instead of having three most compatible methods, like in the predetermined scenarios, there are five methods on display, Figure 13. This difference is due to not knowing the exact scenario, so giving more options for the user to investigate may help the process if one is not fully adaptable to the scenario in cause.

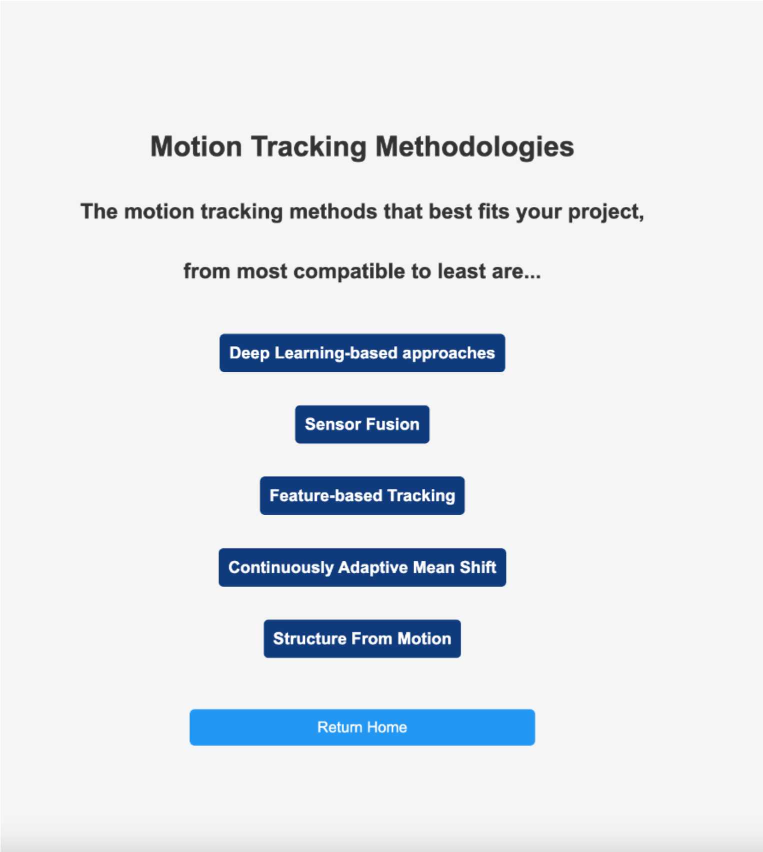

Figure 10 - Results page after custom input selection.

Finally, in the results page it is possible to click on the method to know more about it, see an image of the method being used in an application and there are three links for the user to discover some products that utilize the correspondent method, Figure 14.

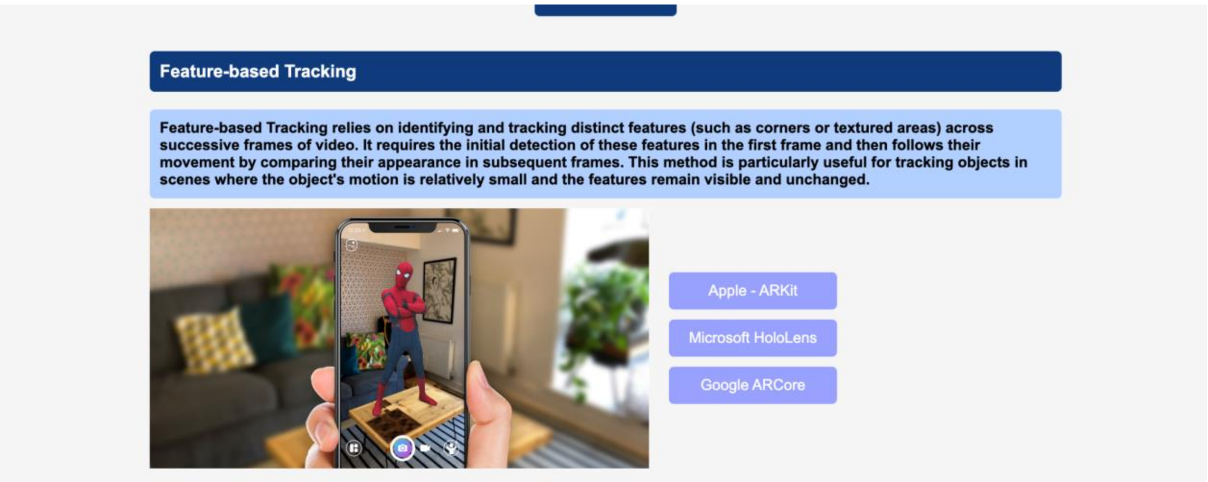

Figure 11 - Feature-based Tracking in Results page.

To understand the decision-making process visually, Figure 15 illustrates the functionality of the framework through a diagram that highlights the process using only two motion tracking methods, feature-based tracking and deep learning approaches, and for the first two criteria, Environment and Environmental Conditions, for simplified exemplification. Different points will be assigned to each method depending on which input is selected.

For example, upon selecting "Indoor" for Environment and "Controlled" for Environmental Conditions, the framework evaluates the suitability of each method based on these inputs. In this specific case, both methods are equal in result. The diagram visually represents how these inputs influence the framework's recommendation process, leading to the identification of the most suitable motion tracking method for the given scenario. This visual depiction underscores the framework's capacity to tailor its recommendations based on specific user inputs, ensuring optimal method selection.

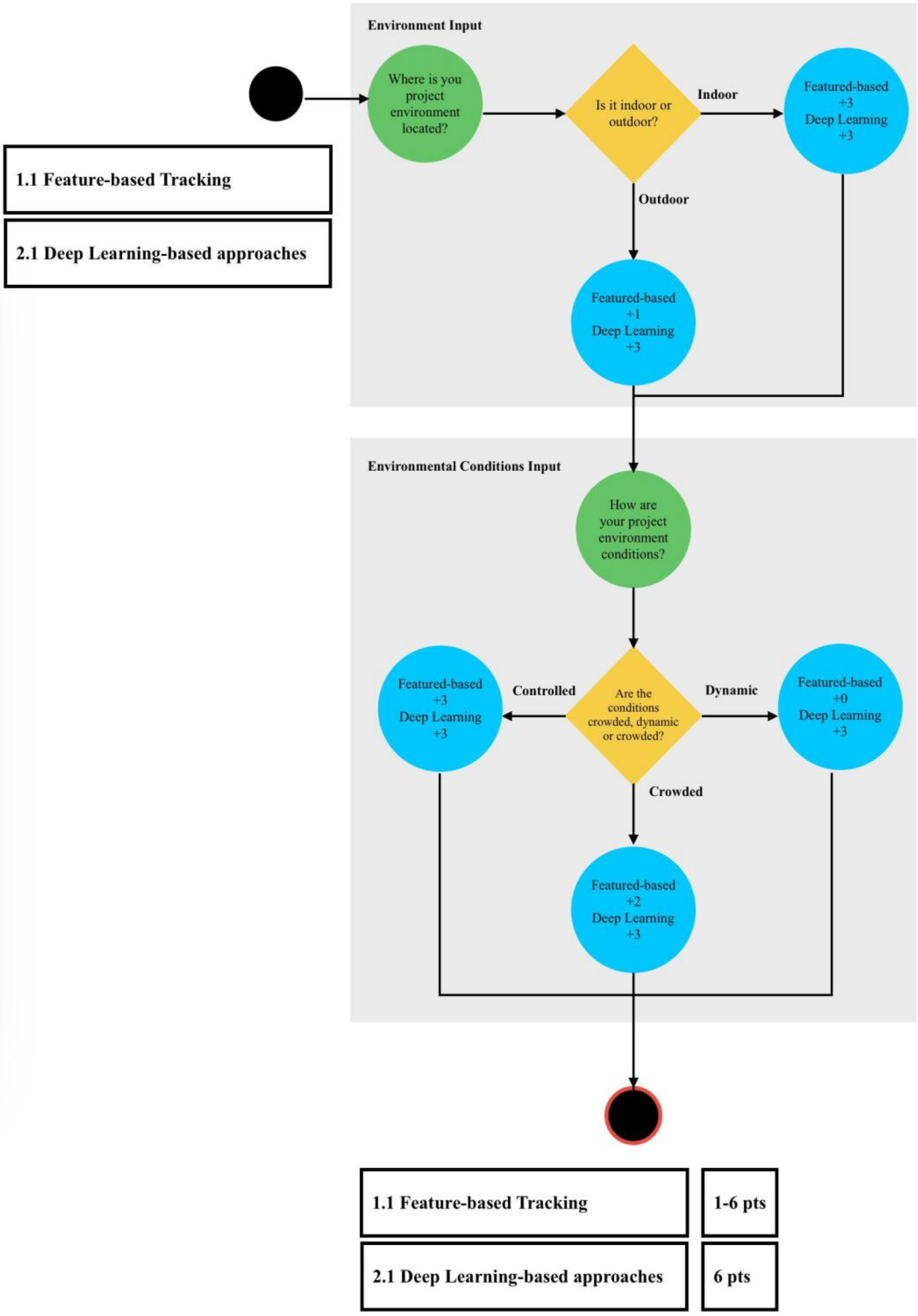

Figure 12 - Diagram showcasing functionality of the framework.

5. PERFORMANCE EVALUATION

This chapter presents a comprehensive evaluation of the performance of the motion tracking methodologies investigated in this thesis. The primary objective is to provide a detailed analysis of the results obtained from extensive research, assessments and discussion of their implications within the context of the study. The chapter is structured into three main sections: Performance Results, Assessment, and Discussion.

5.1. Performance Results

This chapter presents the performance results of various motion tracking methodologies based on extensive literature review and synthesis of existing research. The evaluation covers feature-based tracking, optical flow, SfM, background subtraction, template matching, CamShift, Kalman filtering, deep learning-based approaches, multiple camera systems, and sensor fusion. Each method is assessed in terms of strengths and weaknesses, distance capabilities, tracking speed, and time and cost of implementation.

Feature-based tracking involves identifying and tracking distinctive features within an image, such as corners, edges, or textures. The strengths of this method include its robustness to changes in lighting and partial occlusions, making it effective in dynamic environments. However, it requires high computational power and can struggle with featureless or repetitive patterns. In terms of distance capabilities, feature-based tracking is effective for short to medium distances, though performance degrades with increased distance. Tracking speed is moderate to high, depending on the algorithm and computational resources available. The time and cost of implementation are moderate, as it requires significant computational resources and sophisticated algorithms (Lowe, 2004).

Optical flow estimates motion by analyzing the apparent motion of objects between consecutive frames. This method works well for dense motion fields and is effective in smooth motion scenarios. However, it is sensitive to noise and sudden changes, and it is computationally intensive. Optical flow is generally effective for short to medium distances. The tracking speed is moderate, dependent on the complexity of the scene and the computational power. The time and cost of implementation are high, requiring significant processing power and advanced algorithms (Horn & Schunck, 1981).

SfM reconstructs 3D structures from 2D image sequences taken from different viewpoints. Its strengths include providing detailed 3D reconstruction and being effective for large-scale environments. However, it is computationally intensive and sensitive to feature matching errors. SfM is effective for medium to long distances. The tracking speed is low due to high computational requirements. The time and cost of implementation are high, requiring extensive computational resources and sophisticated algorithms (Snavely, 2008).

Background subtraction involves segmenting moving objects from a static background. Its strengths include simplicity and efficiency, making it effective in controlled environments with static backgrounds. However, it struggles with dynamic backgrounds and varying lighting conditions and is sensitive to noise. Background subtraction is effective for short to medium distances. The tracking speed is high due to its simplicity. The time and cost of implementation are low, with straightforward implementation and minimal computational requirements (Stauffer & Grimson, 1999).

Template matching involves finding a sub-image or template within a larger image. This method is mostly simple and easy to implement, making it effective for known objects in controlled environments. However, it is sensitive to changes in scale, rotation, and lighting, and it can be computationally expensive for large templates. Template matching is effective for short distances. The tracking speed is moderate, decreasing with template size and search area. The time and cost of implementation are low to moderate, with straightforward implementation but potential computational demands (Lewis, 1995).

CamShift is an algorithm that adapts the search window size based on the color distribution of the target. Its strengths include robustness to scale and orientation changes, making it efficient for real-time tracking. However, it is sensitive to background clutter and color changes, and it can lose track if the object changes color. CamShift is effective for short to medium distances.

The tracking speed is high, designed for real-time applications. The time and cost of implementation are moderate, requiring specific algorithmic knowledge but efficient computationally (Bradski, 1998).

Kalman filtering is a recursive algorithm used for estimating the state of a dynamic system from noisy measurements. Its strengths include robustness to noise and measurement errors, and it is effective for linear dynamic systems. However, it is limited to linear models, and performance degrades with non-linear dynamics. Kalman filtering is effective for medium to long distances. The tracking speed is high, with efficient computation suitable for real-time applications. The time and cost of implementation are moderate, requiring an understanding of linear algebra and system dynamics (Kalman, 1960).

Deep learning-based approaches leverage neural networks to model and predict motion. These methods are highly accurate and adaptable to complex and dynamic environments, capable of learning from large datasets. However, they require extensive computational resources and large labeled datasets, and can be prone to overfitting. Deep learning-based approaches are effective for a wide range of distances. The tracking speed is variable, potentially high with optimized models and hardware. The time and cost of implementation are high, necessitating significant investment in data collection, model training, and computational infrastructure (Wang & Gupta, 2018).

Multiple camera systems use data from several cameras to enhance tracking accuracy and coverage. Their strengths include providing comprehensive coverage and redundancy, making them effective for large and complex environments. However, they have high setup and calibration complexity, and require significant computational and storage requirements. Multiple camera systems are effective for a wide range of distances. The tracking speed is moderate to high, depending on system configuration and computational resources. The time and cost of implementation are high, requiring substantial investment in hardware and system integration (D’Angelo & Dugelay, 2012).

Sensor fusion achieves its result by combining data from multiple sensors in order to improve tracking performance. Its strengths include robustness to noise and failures, providing complementary information for enhanced accuracy. However, it involves complex integration and calibration, and significant computational requirements. Sensor fusion is effective for a wide range of distances. The tracking speed is high, with efficient data fusion algorithms enabling real-time tracking. The time and cost of implementation are high, requiring sophisticated algorithms and diverse sensor setups (Madgwick et al., 2011).

This chapter has presented the performance results of various motion tracking methodologies based on existing research. Each method's strengths, weaknesses, distance capabilities, tracking speed, and time and cost of implementation were assessed. The findings provide a comprehensive understanding of the suitability and performance of different tracking techniques, guiding the development of a robust framework for selecting the most appropriate method based on specific input parameters.

5.2. Assessment and Discussion

This chapter provides a thorough assessment and discussion of the performance and applicability of the motion tracking methodologies explored in this thesis through data analysis. The objective is to showcase the analysis of the strengths and weaknesses of each method, analyze their distance capabilities, examine their tracking speed, and consider the time and cost associated with their implementation.

5.2.1. Strengths and Weaknesses

In the world of motion tracking technology, understanding how different methods perform across various settings is crucial. Specifically, we examine how these technologies fare in controlled versus uncontrolled environments. Controlled environments, such as indoor settings where lighting and background elements can be precisely managed, allow certain tracking methods to excel. For example, background subtraction and template matching, with their precision and reliability, are particularly effective under these conditions because they rely on consistency in the scene to detect and track objects (Stauffer & Grimson, 1999; Lewis, 1995).

Conversely, in uncontrolled environments, which are characterized by natural lighting variations, dynamic backgrounds, and unpredictable object movements, other tracking techniques prove to be more robust. Deep learning-based approaches, with their remarkable adaptability and resilience, stand out in their ability to identify objects despite the environmental changes (Leal-Taixé, Canton-Ferrer, & Schindler, 2015). These methods benefit from extensive training on diverse datasets, enabling them to recognize and follow objects accurately even in complex outdoor scenes (Redmon et al., 2016).

Moreover, by sensor fusion combining information from various sensors, this approach compensates for the limitations of individual methods, offering a more comprehensive and reliable tracking solution suitable for a wider range of environments (Hasanujjaman, Chowdhury, & Jang, 2023).

This analysis highlights the importance of selecting appropriate motion tracking methods based on the specific requirements of the environment. Whether the setting is controlled with predictable parameters or uncontrolled with numerous variables, understanding the strengths and limitations of each tracking technology is essential for achieving optimal performance (Azuma, 1997; Billinghurst, Clark, & Lee, 2015).

Table 6 provides a high-level overview of the strengths and weaknesses associated with each motion tracking method. It's important to consider that the effectiveness of a method can greatly depend on the specific context and requirements of the application it is being used for.

Motion Tracking Method	Strengths	Weaknesses
Feature-based Tracking	High accuracy in stable environments Efficient for tracking small or well-defined features	Susceptible to occlusion Can struggle with rapid motion or drastic appearance changes
Optical Flow	Good for tracking movement patterns over time Works well with textured objects,	Sensitive to lighting changes High computational cost
SfM	Enables 3D reconstruction from camera movements Useful for mapping environments	Requires multiple images from different viewpoints Computationally intensive
Background Subtraction	Effective for stationary camera setups Simple to implement for foreground detection	Sensitive to dynamic backgrounds and lighting changes Not suitable for moving cameras
Template Matching	Easy to implement Effective for matching and tracking unoccluded objects	Not robust to scale or rotation changes Fails with occlusion or background clutter
CamShift	Adapts to the object size and rotation Effective for color-based tracking	Depends heavily on the color distribution Not effective in complex scenes
Kalman Filtering	Predictive tracking in linear systems Works well with noisy measurements	Assumes a linear model which may not always be accurate Requires tuning of noise parameters
Deep Learning-based	High accuracy in complex scenarios	Requires large datasets for training

	Robust to occlusions and variable conditions	High computational resources
Multiple Camera Systems	Enhances depth perception through stereo vision Provides multiple perspectives for tracking	Complex setup and calibration Increased cost and computational requirements
Sensor Fusion	Combines multiple data sources for improved accuracy Robust to individual sensor limitations	Complexity in data integration Requires synchronization between different sensors

Table 6 - Main strengths and weaknesses of each motion tracking.

5.2.2 Distance Capabilities

Distance input is a critical factor in the performance of motion tracking methodologies, as methods vary in performance depending on the range at which they are used. For instance, Feature-based Tracking and Optical Flow are generally more effective at short to medium distances where the resolution of features and the consistency of motion patterns are more noticeable (Lowe, 2004; Horn & Schunck, 1981). In contrast, methods like SfM and Multiple Camera Systems perform better at long distances, as they can leverage spatial data from multiple viewpoints to reconstruct and track objects in a three-dimensional space (Snavely, Seitz, & Szeliski, 2008; D’Angelo & Dugelay, 2012). Background Subtraction and Template Matching tend to perform better in short-range, controlled environments where consistency and accuracy are maintained. On the other hand, Deep Learning-based Approaches and Sensor Fusion techniques can adapt to various distances by utilizing advanced algorithms and integrating data from multiple sensors to maintain high accuracy across different ranges (Leal-Taixé, Canton-Ferrer, & Schindler, 2015; Hasanujjaman, Chowdhury, & Jang, 2023). The importance of the distance input lies in its ability to guide the selection of the most appropriate tracking method, ensuring optimal performance and reliability. By accurately considering distance, the framework can tailor the tracking solution to the application's specific needs, whether it involves close-up interactions or long-range observations, thereby enhancing overall efficiency and effectiveness (Azuma, 1997; Billinghurst, Clark, & Lee, 2015).

Table 7 provides the overview of the performances of each motion tracking method in each distance setting. It's important to consider that the effectiveness of a method can greatly depend on the specific context and requirements of the application it is being used for.

Motion Tracking Method	**Controlled Environments**	**Medium Distance Outside**	**Long Distance Outside**	**Additional Environments**
Feature-based Tracking	Good	Good	Moderate	Crowded environments (with careful implementation)
Optical Flow	Excellent	Moderate	Moderate	Dynamic backgrounds
Structure from Motion	Excellent	Excellent	Good	3D Reconstruction, Aerial mapping
Background Subtraction	Excellent	Poor	Poor	Stationary camera setups
Template Matching	Good	Moderate	Poor	Controlled object tracking
CamShift	Good	Good	Moderate	Color-dominant scenes
Kalman Filtering	Good	Good	Moderate	Predictive tracking, Robotics
Deep Learning-based	Excellent	Excellent	Excellent	Complex scenes, Crowded environments
Multiple Camera Systems	Excellent	Excellent	Good	Stereo vision, Multi-view geometry

Sensor Fusion	Excellent	Excellent	Excellent	Enhanced tracking accuracy, Autonomous vehicles

Table 7 - Assessment of environment conditions of each motion tracking.

5.2.3 Tracking Speed

Transitioning from the distance in the environment to the aspect of tracking speed further refines our understanding and selection of motion tracking methods. While the adaptability to different environments sets the stage for potential candidates, the speed of tracking often becomes the deciding factor, especially in real-time applications where quick response times are crucial (Stauffer & Grimson, 1999; Kalman, 1960). The connection between environmental complexity and operational speed highlights a critical balance that must be achieved for effective motion tracking. As we move towards the table of tracking speed, this consideration takes an important role, providing a detailed breakdown of how each method behaves not just in the context of where it can be used, but also in terms of how quickly it can adapt and respond to the project needs (Azuma, 1997; Billinghurst, Clark, & Lee, 2015).

Table 8 provides the overview of the average tracking speed associated with each motion tracking method in comparison with each other (Horn & Schunck, 1981; Snavely, Seitz, & Szeliski, 2008). It’s important to consider that the effectiveness of a method can greatly depend on the specific context and requirements of the application it is being used for. ([Leal-Taixé, Canton-Ferrer, & Schindler, 2015; Hasanujjaman, Chowdhury, & Jang, 2023)

Motion Tracking Method	Speed of Tracking
Feature-based Tracking	Medium
Optical Flow	Low
Structure from Motion	Low
Background Subtraction	High
Template Matching	Medium
CamShift	Medium
Kalman Filtering	High
Deep Learning-based	Medium to Low
Multiple Camera Systems	Low
Sensor Fusion	Medium to High

Table 8 - Generalized idea average speed tracking of each motion tracking method.

Low-speed tracking methods exceed the average processing times and take a few seconds to achieve a result. These methods are typically used in applications where precision and depth of data are more critical than immediate response, such as complex environmental interactions or analytical assessments like post-production tracking.

Medium-speed tracking methods usually process data within a moderate timeframe. These are suitable for applications where some delay is tolerable and can be balanced against other requirements, such as higher accuracy.

High-speed tracking methods are characterized by their ability to process and update tracking information in near real-time. These methods are optimal for applications where immediate response is critical, such as interactive systems, real-time surveillance, or news broadcasts where tracking accuracy and immediate results are essential.

5.2.4. Time & Cost of Implementation

The time and cost of implementation are essential in the decision-making process when choosing motion tracking methodologies. These elements directly influence the accessibility and usability of motion tracking technologies across various sectors. For industries like healthcare and entertainment, where cost efficiency and rapid deployment are vital, selecting a motion tracking system that balances performance with affordability and ease of integration can significantly impact overall success (Giggins, Persson, & Caulfield, 2013; Guo & Zhong, 2022). A quicker and less expensive setup enables businesses and institutions to implement motion tracking solutions without extensive disruption to existing operations, fostering innovation and technological advancement. Furthermore, smaller enterprises and researchers can leverage these powerful tools by reducing time and financial barriers and democratizing access to cutting-edge technology (Hasanujjaman, Chowdhury, & Jang, 2023). Thus, evaluating the time and cost of implementation is essential for ensuring economic viability, promoting broader adoption and maximizing the potential of motion tracking systems in diverse applications (Stauffer & Grimson, 1999; Lewis, 1995). The framework can select the most appropriate and reliable tracking method by incorporating speed as an essential input (Azuma, 1997).

Table 9 provides the overview of the average time and cost associated with each motion tracking method implementation. It's important to consider that the effectiveness of a method can greatly depend on the specific context and requirements of the application it is being used for.

Motion Tracking Method	Time to Implement	Cost to Implement
Feature-based Tracking	Medium	Low to Medium
Optical Flow	Medium	Medium
SfM	High	High

Background Subtraction	Low	Low
Template Matching	Low	Low
CamShift	Medium	Low to Medium
Kalman Filtering	Medium	Low to Medium
Deep Learning-based	High	High
Multiple Camera Systems	High	High
Sensor Fusion	High	High

Table 9 - Generalized idea of average time and cost of implementation of each motion tracking method.

Time to Implement:

- **Low**: Can be implemented relatively quickly, often with existing libraries and without extensive customization.
- **Medium**: Requires moderate development time, including some customization and optimization for specific applications.
- **High**: Involves extensive development and testing time, potentially including the creation of custom models or algorithms, and extensive dataset collection and training for deep learning approaches.

Cost to Implement:

- **Low**: Generally inexpensive, using off-the-shelf software and minimal specialized hardware.
- **Medium**: May require commercial software licenses, moderate hardware resources, or some level of custom development.

- **High**: Likely requires significant investment in specialized hardware, software licenses, custom development, and potentially large datasets for training machine learning models.

It is important to note that these assessments are approximations. The actual time and cost can vary widely depending on the scope of the project, the complexity of the environment in which the tracking is being implemented, the precision required, and whether the development is custom-made or already developed and simply requires a small integration. Additionally, ongoing advancements in technology and software libraries can also influence both the time and cost needed to implement these methods, potentially lowering barriers to entry over time.

6. CONCLUSION & FUTURE PERSPECTIVES

This chapter provides a comprehensive conclusion to the research conducted in this thesis, summarizing the key findings and their implications. Additionally, it explores the ethical considerations and societal impact of motion tracking and presents potential future directions for further investigation. The chapter is structured into three main sections: Conclusion, Ethics View & Impact on Society, and Future Perspectives.

6.1. Conclusion

In conclusion, exploring motion tracking methods through various insights, like their inherent strengths and weaknesses, their suitability for controlled versus uncontrolled environments, or their tracking speed, reveals a field where the optimal choice of technology is contingent upon many input factors. Methods like background subtraction and template matching have demonstrated considerable strength in controlled environments due to their reliance on consistency and predictability. Conversely, deep learning-based approaches have emerged as frontrunners in uncontrolled environments, adept at handling complex scenarios such as long distances and crowded settings, thanks to their robust adaptability and learning capabilities. When the conversation shifts towards tracking speed, methods like Kalman Filtering and background subtraction stand out for their rapid response times, making them particularly suitable for applications requiring real-time performance. However, it is essential to note that the fastest tracking capabilities often have reciprocations regarding setup complexity and resource requirements.

Traditional computer vision techniques, while sometimes less complex than deep learning counterparts, offer a cost-effective solution for various applications, balancing performance with financial feasibility.

Taking all factors into consideration—accuracy, adaptability to distance, speed, and cost—the choice of a motion tracking method becomes a balancing act. Deep learning-based approaches, despite their higher initial cost and computational demands, offer unparalleled versatility and performance across a high range of environments, making them one of the best options for scenarios where accuracy cannot be compromised. Meanwhile, for projects where cost and efficiency are vital, traditional methods like background subtraction and Kalman Filtering provide a viable and effective solution, demonstrating the practical implications of our research in real-world scenarios.

Ultimately, the decision when selecting a motion tracking method requires careful consideration of the application's specific needs, the operational environment, and the available resources. Considering these factors against the comprehensive overview provided by research that results in the tables of strengths and weaknesses, environmental suitability, and tracking speed, one can navigate the complexities of motion tracking technology to select the most appropriate method for their needs, ensuring optimal performance in any given scenario.

The motion tracking framework developed in this thesis represents a significant advancement in the selection and application of motion tracking methodologies. The comprehensive evaluation of methods such as Feature-based Tracking, Optical Flow, SfM, Background Subtraction, Template Matching, CamShift, Kalman Filtering, Deep Learning-based Approaches, Multiple Camera Systems, and Sensor Fusion ensures that the most suitable methodology is selected for each specific scenario. Overall, this framework not only simplifies the complex process of choosing the right motion tracking method but also sets a foundation for future enhancements and integrations, ultimately contributing to more efficient, reliable, and versatile motion tracking solutions.

6.2. Ethics view & Impact on Society

The increasing ubiquity of motion tracking technologies presents numerous ethical challenges and profoundly impacts various sectors of society. This chapter explores the current and future ethical landscape surrounding motion tracking, highlighting privacy, consent, data security, and bias concerns. Additionally, it examines the societal impact of the proposed motion tracking framework, detailing its influence on industries, healthcare, education, public safety, and everyday life.

Privacy is one of the most ethical issues associated with motion tracking that needs extra attention. These technologies can collect detailed data about individuals' movements and behaviors, leading to significant privacy invasions if not managed properly. Concerns revolve around how this data is collected, stored, and used, especially in public spaces. Without suitable safeguards, there is a risk that individuals' movements could be monitored without their knowledge or consent, leading to potential misuse of data.

As motion tracking continues to evolve, the potential for privacy breaches could also increase. Emerging technologies such as wearable devices, smart home systems, and augmented reality applications could collect continuous data streams, raising concerns about who can access this information. Ensuring strong privacy protections will be crucial to mitigating these risks.

Another critical ethical consideration is the issue of informed consent. Individuals must be fully aware and agree to using motion tracking technologies that collect data about their movements. This involves clear communication about the data usage and how it is managed. There is often a need for more transparency, and users may not fully understand the extent of the data collection or the implications of its use, especially when using extensive accord agreements.

In the future, obtaining informed consent may become even more challenging as motion tracking technologies become more integrated into everyday life. For example, ensuring that everyone has informed consent can be difficult in environments where multiple people are being tracked simultaneously, such as public transportation systems. Developing standardized protocols for obtaining and managing consent will be essential to address this issue.

Data security collected by motion tracking technologies is another primary ethical concern. Motion tracking data can be susceptible, revealing not only an individual's location but also behavior patterns and personal habits. Properly storing and protecting this data is critical to preventing data breaches and misuse.

As the quantity and quality of collected data increases, the need for robust data security protocols will become even more important for future society.

Motion tracking technologies can also be subject to biases that affect their accuracy and fairness. For instance, tracking algorithms may perform differently based on an individual's physical characteristics, such as body shape, skin color, or mobility patterns. This can lead to disparities in the effectiveness of motion tracking systems across different demographic groups.

Addressing these issues is crucial to ensuring that motion tracking technologies are fair and equitable for all, which involves ongoing research and development to identify and mitigate sources of bias in tracking algorithms. In the future, as these technologies become more prevalent in areas such as law enforcement and employment, ensuring that they do not perpetuate existing inequalities will be a significant ethical challenge.

As motion tracking technologies continue to evolve, new ethical concerns are likely to emerge. The use of AI to analyze motion tracking data, for instance, could raise questions about accountability and transparency. Ensuring that these systems are explainable and that their decision-making processes are transparent will be critical to maintaining trust.

Another potential future concern is the use of motion tracking for surveillance purposes, especially if combined with face recognition. While these technologies can enhance public safety, they also have the potential to be used for invasive surveillance, leading to a loss of individual freedom. Balancing the benefits of motion tracking for security with the need to protect civil liberties will be an ongoing ethical dilemma.

The framework can also have a profound impact on public safety and security. By helping agencies choose the most suitable tracking methods, the framework can enhance situational awareness, improve response times, and increase the overall effectiveness of public safety operations. This can lead to safer communities, better crime prevention, and more efficient disaster management.

Motion tracking technologies are becoming more prevalent in everyday life, like consumer electronics, fitness tracking, and smart home systems. The framework can guide developers in selecting the best tracking methods to improve user experiences and product functionality. This can lead to more intuitive and reliable consumer products, enhancing convenience, health monitoring, and entertainment.

The ethical implications and societal impact of motion tracking technologies are vast and multifaceted. Addressing issues such as privacy, informed consent, data security, and bias is essential to ensure responsible use of these technologies. The proposed framework can drive innovation and efficiency across various sectors.

6.3. Future Perspectives

As motion tracking technologies evolve and become increasingly sophisticated, these systems' potential applications and benefits will expand significantly. This chapter explores future perspectives for the framework developed in this thesis, focusing on creating standards for more advanced frameworks that can integrate multiple motion tracking methods within the same project.

The primary future perspective for this thesis is establishing a standardized framework for motion tracking methodologies. The current framework provides a systematic approach to selecting the most suitable motion tracking method based on specific project requirements.

However, as technologies advance, there will be a growing need for frameworks that integrate multiple methods within a single project.

This approach of integrating multiple motion tracking methods would allow for the combination of different techniques, each contributing its strengths to improve the system's overall performance. For instance, feature-based tracking could be used with deep learning-based approaches to enhance accuracy in complex environments. At the same time, sensor fusion could provide additional robustness by integrating data from various sensors.

Integrating multiple methods would enable the creation of hybrid systems more adaptable to a wide range of scenarios. Such systems could dynamically switch between methods or use them simultaneously, depending on the specific requirements of the task at hand. This would result in more reliable and accurate motion tracking, especially in challenging environments where a single method might fall short.

Developing standardized protocols and interfaces will ensure that different tracking technologies can work together seamlessly. This will involve defining standard data formats, communication protocols, and integration guidelines, allowing easy operation between various motion tracking systems.

Standardization efforts will also need to address data synchronization, calibration, and real-time processing issues. Ensuring that data from different sources can be accurately combined and processed in real-time will be essential for successfully implementing multi-method frameworks. Collaborative efforts between industry, academia, and standardization bodies will be necessary to develop and promote these standards.

Integrating multiple motion tracking methods within a standardized framework can significantly enhance performance and accuracy. Hybrid systems, by combining the strengths of different techniques, a. Hybrid systems can achieve higher levels of precision and reliability while minimizing the weaknesses. For example, integrating optical flow with structure from motion could improve tracking accuracy in dynamic environments, while incorporating sensor fusion could enhance robustness against occlusions and environmental changes.

Within the framework, combining an implementation of AI and machine learning with the optimization of multiple methods, as well as market products and technologies, in the field of motion tracking, can enhance much further the overall process to the point that only the whole implementation and decision process can be designed within the framework. The algorithms can analyze the performance of different tracking techniques in real-time and dynamically adjust their usage based on the specific conditions of the environment.

The future perspectives for the motion tracking framework developed in this thesis are both exciting and promising. By creating a standardized framework that supports the whole integration process, it can be achieved significant advancements in accuracy, robustness, and versatility. New possibilities will open up for applications in healthcare, industry, AR/VR, and beyond, ultimately contributing to the development of more advanced and effective motion tracking solutions. Through collaborative efforts in the evolution of tracking technologies, we can look forward to a future where motion tracking systems are more powerful and adaptable than ever before.

7. BIBLIOGRAPHY

Apple. (2020). ARKit 3 - Augmented Reality - Apple Developer. Apple Developer. https://developer.apple.com/augmented-reality/arkit/

Axis Communications - Leader in network cameras and other IP networking solutions | Axis Communications. (n.d.). Www.axis.com. https://www.axis.com/

Azuma, R. T. (1997). A survey of augmented reality. Presence: Teleoperators & Virtual Environments, 6(4), 355-385. https://doi.org/10.1162/pres.1997.6.4.355

Billinghurst, M., Clark, A., & Lee, G. (2015). A survey of augmented reality. Foundations and Trends® in Human–Computer Interaction, 8(2-3), 73-272. https://doi.org/10.1561/1100000049

Bleser, G., & Stricker, D. (2008). Advanced tracking through efficient image processing and visual-inertial sensor fusion. Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, 13-22. https://doi.org/10.1109/ISMAR.2008.4637328

Bradski, G. R. (1998). Computer vision face tracking for use in a perceptual user interface. Intel Technology Journal, 2(2), 12-21.

Chen, L., Tian, Y., He, Z., Yi, D., & Xu, Y. (2017). Monocular vision-based obstacle detection for a micro air vehicle in indoor environments. Sensors, 17(8), 1876. https://doi.org/10.3390/s17081876

Cognex | Machine Vision and Barcode Readers. (2019). Cognex.com. https://www.cognex.com/

Crawford, M. (2024, February 8). What is motion capture (Mo-Cap): the definitive guide. Filmmaking Lifestyle. https://filmlifestyle.com/what-is-motion-capture/

DeGuzman, K. (2023, September 7). What is Mocap — The Science and Art Behind Motion Capture. StudioBinder. https://www.studiobinder.com/blog/what-is-mocap-definition/

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). Pattern classification (2nd ed.). Wiley-Interscience.

D'Angelo, G., & Dugelay, J. L. (2012). Multi-camera people tracking by collaborative particle filters and data fusion. Signal Processing: Image Communication, 27(11), 1277-1291. https://doi.org/10.1016/j.image.2012.08.002

Fani, S., Dabiri, S., Heaslip, K., & Stevanovic, A. (2017). Computer vision in sports: A critical review. Multimedia Tools and Applications, 76(21), 23049-23069. https://doi.org/10.1007/s11042-016-4267-0

Foxlin, E. (2002). Pedestrian tracking with shoe-mounted inertial sensors. IEEE Computer Graphics and Applications, 25(6), 38-46. https://doi.org/10.1109/MCG.2005.140

Furukawa, Y., & Ponce, J. (2010). Accurate, dense, and robust multiview stereopsis. LSTMIEEE Transactions on Pattern Analysis and Machine Intelligence, 32(8), 1362-1376. https://doi.org/10.1109/TPAMI.2009.161

Giggins, O. M., Persson, U. M., & Caulfield, B. (2013). Biofeedback in rehabilitation. Journal of NeuroEngineering and Rehabilitation, 10(1), 60. https://doi.org/10.1186/1743-0003-10-60

Gonzalez, R. C., & Woods, R. E. (2018). Digital Image Processing (4th ed.). Pearson.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. In Advances in Neural Information Processing Systems (Vol. 27, pp. 2672-2680).

GPS Accuracy | Garmin Centro de Assistência. (n.d.). Support.garmin.com. Retrieved June 20, 2024, from https://support.garmin.com/pt-PT/?faq=aZc8RezeAb9LjCDpJplTY7

Guo, Yong & Zhong, Chan. (2022). Motion Capture Technology and Its Applications in Film and Television Animation. Advances in Multimedia. 2022. 1-10. 10.1155/2022/6392168.

Han, G., Xiao, Y., & Peng, Y. (2014). An indoor localization system based on hybrid WiFi/BLE and inertial sensors. IEEE Access, 6, 7020-7030. https://doi.org/10.1109/ACCESS.2018.2876975

Hartley, R., & Zisserman, A. (2004). Multiple view geometry in computer vision. Cambridge University Press.

Hasanujjaman, M., Chowdhury, M. Z., & Jang, Y. M. (2023). Sensor Fusion in Autonomous Vehicle with Traffic Surveillance Camera System: Detection, Localization, and AI Networking. Sensors, 23(6), 3335. https://doi.org/10.3390/s23063335

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735-1780.

Horn, B. K., & Schunck, B. G. (1981). Determining optical flow. Artificial Intelligence, 17(1-3), 185-203. https://doi.org/10.1016/0004-3702(81)90024-2

Intel® Distribution of OpenVINOTM Toolkit. (n.d.). Intel. https://www.intel.com/content/www/us/en/developer/tools/openvino-toolkit/overview.html

Introduction to Motion Tracking. (n.d.). Corel Discovery Center. Retrieved June 20, 2024, from https://learn.corel.com/tutorials/introduction-to-motion-tracking/

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. Journal of Basic Engineering, 82(1), 35-45. https://doi.org/10.1115/1.3662552

Kendall, A., & Cipolla, R. (2016). Modelling uncertainty in deep learning for camera relocalization. Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), 4762-4769. https://doi.org/10.1109/ICRA.2016.7487679

Leal-Taixé, L., Canton-Ferrer, C., & Schindler, K. (2015). Learning by tracking: Siamese CNN for robust target association. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 33-40. https://doi.org/10.1109/CVPRW.2016.41

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.

Lewis, J. P. (1995). Fast template matching. Vision Interface, 95(120123), 15-19.

Lowe, D. G. (1999). Object recognition from local scale-invariant features. In Proceedings of the Seventh IEEE International Conference on Computer Vision (Vol. 2, pp. 1150-1157).

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2), 91-110. https://doi.org/10.1023/B:VISI.0000029664.99615.94

Lucas, B. D., & Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. In Proceedings of the 7th International Joint Conference on Artificial Intelligence (Vol. 2, pp. 674-679).

Madgwick, S. O. H., Harrison, A. J. L., & Vaidyanathan, R. (2011). Estimation of IMU and MARG orientation using a gradient descent algorithm. 2011 IEEE International Conference on Rehabilitation Robotics, 1-7. https://doi.org/10.1109/ICORR.2011.5975346

McCarthy, C., & Collins, K. (2018). Initial results from using deep learning and computer vision techniques to analyze player tracking data in the Australian Football League. Proceedings of the MIT Sloan Sports Analytics Conference.

Microsoft HoloLens | Mixed Reality Technology for Business. (n.d.). Www.microsoft.com. https://www.microsoft.com/en/hololens

Motion tracking in After Effects | Adobe. (n.d.). Www.adobe.com. Retrieved June 20, 2024, from https://www.adobe.com/products/aftereffects/motion-tracking.html

Ojeda, L., & Borenstein, J. (2007). Non-GPS navigation for security personnel and first responders. Journal of Navigation, 60(3), 391-407. https://doi.org/10.1017/S0373463307004290

Pérez, R., Costa, Ú., Torrent, M., Solana, J., Opisso, E., Cáceres, C., ... & Tormos, J. M. (2010). Virtual reality for sensorimotor rehabilitation post stroke: Design principles and evidence. Studies in Health Technology and Informatics, 154,

ReCap Software | Get Prices & Buy ReCap Pro 2023 | Autodesk. (2020, November 30). Autodesk.com. https://www.autodesk.com/products/recap/overview?term=1-YEAR&tab=subscription&plc=RECAP

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 779-788. https://doi.org/10.1109/CVPR.2016.91

Snavely, N., Seitz, S. M., & Szeliski, R. (2008). Modeling the world from internet photo collections. International Journal of Computer Vision, 80(2), 189-210. https://doi.org/10.1007/s11263-007-0107-3

Stauffer, C., & Grimson, W. E. L. (1999). Adaptive background mixture models for real-time tracking. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 246-252. https://doi.org/10.1109/CVPR.1999.784637

The 6 types of motion tracking in After Effects – 2022 Guide | Evercast Blog. (n.d.). Www.evercast.us. https://www.evercast.us/blog/tracking-in-after-effects

Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) (Vol. 1, pp. I-511-I-518).

Wang, X., & Gupta, A. (2018). Videos as space-time region graphs. Proceedings of the European Conference on Computer Vision (ECCV), 399-417. https://doi.org/10.1007/978-3-030-01216-8_24

What is Motion Tracking? - A Guide to Understanding Motion Tracking Technology - Sanguine Production. (2023, July 23). Sanguine Production. https://sanguineproduction.com/blog/post-production/what-is-motion-tracking-a-guide-to-understanding-motion-tracking-techn

Wojke, N., Bewley, A., & Paulus, D. (2017). Simple online and realtime tracking with a deep association metric. 2017 IEEE International Conference on Image Processing (ICIP), 3645-3649. https://doi.org/10.1109/ICIP.2017.8296962

World Leader in AI Computing. (n.d.). NVIDIA. https://www.nvidia.com/en-eu/

Xu, W., & Huang, A. (2022). Multi-camera operating room activity analysis for workflow analysis. In SPIE Medical Imaging.

https://www.sony.com/en/SonyInfo/research/publications/Multi-camera-operating-room-activity-analysis-for-workflow-analysis/

Zhang, Z., He, T., & Duan, L. (2020). A hybrid motion tracking system combining visual and inertial sensors for precise indoor navigation. IEEE Sensors Journal, 20(6), 3301-3312. https://doi.org/10.1109/JSEN.2019.2957630