Beyond Human Motion: Exploring Advanced Techniques and Challenges in Animal Motion Capture for Cinematography

Maximilian Rueth

Abstract

Animal motion capture in cinematography presents significant challenges, particularly in achieving non-invasive tracking while maintaining natural behaviors. Although human motion capture has seen considerable advancements, applying similar techniques to animals requires innovative solutions. This paper explores a variety of methods, including markerless capture systems for underwater creatures, long-range data collection for wildlife, and advanced tracking approaches for large four-legged animals. Through a comparative analysis, we evaluate these methods' effectiveness and potential integration into cinematic works while revealing their limitations. Further advancements are essential, particularly in improving accessibility, enhancing automation, and addressing ethical considerations, to ensure these technologies can be more broadly and effectively applied across the film industry and enhance the realism and immersion of animal portrayals in cinema.

1. Introduction

Motion Capture technology has long been a cornerstone in the video game and film industries, enabling the creation of lifelike human characters and complex virtual environments. Over the past two decades, advancements have primarily focused on human or articulated characters, leading to significant improvements in detailed character animation [MPQG18]. With the rise of Artificial Intelligence, creating 3D animations for characters from simple video uploads has become possible in minutes through several capture tools [Rok24]. As Motion Capture technology continues to advance, it's applications are expanding beyond human subjects to encompass the dynamic and complex movements of animals.

Animal motion capture holds immense potential across multiple disciplines, from conservation efforts and enhancing animal welfare to the analysis of movement and neural activities [MSD*24, fIS24]. In cinematography, the integration of animal motion capture is essential for creating lifelike and compelling visual narratives. By replicating animal movements with precision, animators and filmmakers can bring fantastical creatures to life with unprecedented realism, enriching the viewer's immersive experience. Even in abstract animation styles, capturing animal motion provides animators with valuable references, improving animation quality and saving significant time. Despite its promise, applying motion capture to animals presents a unique set of challenges and remains comparatively underdeveloped [LMK*24]. Unlike human subjects, animals pose specific difficulties, such as the complexity of data collection in natural habitats and the logistics of hardware setup—like placing markers on animals such as horses and dogs [LMK*24, BOX24]. These challenges are further complicated by the lack of comprehensive datasets, making it difficult to work with motion data-based AI algorithms and develop accurate, detailed motion studies.

Existing methods for tracking animal movement, such as GPS-IMU collars, camera traps, and implanted neural recording devices, offer some insights but often fail to capture fine-grained details like precise body poses [MSD*24, fIS24]. Moreover, these methods frequently require animals to be sedated or confined indoors, which can induce stress and alter natural behaviors,defeating the purpose of preserving authentic animal movement. These challenges are particularly pronounced in film production, where time and resources are often limited, making rehearsal schedules demanding. This drives the need for reliable capture techniques that are both time-efficient and beneficial to the entire production process.

In this paper, we explore the current state of animal motion capture, highlighting its potential and challenges. We examine various modern methods and approaches, their limitations, and emerging solutions that could revolutionize the implementation of animal motion capture in film. By doing so, we aim to illuminate the future of animal motion capture and its application in the creative industries.

2. Related Work

While animal motion capture technology has not yet reached the level of sophistication seen in human motion capture, it has nonetheless experienced notable advancements and applications across several fields. Many of the emerging techniques originate from medical, neurological, and behavioral studies, where extensive research is conducted on animal movement analytics. However, as is the case with human motion capture, there is no single, standardized method for recording animal movement. Instead, various approaches exist, each presenting specific advantages and limitations depending on the context in which they are applied. The inherent diversity in animal physiologies and habitats introduces additional complexities that must be accounted for in the design and execution of these systems.

Numerous studies have proposed different motion capture methods or software, often focusing on their individual systems while drawing comparisons with existing techniques to highlight their respective benefits [MPQG18, MSD*24]. These studies predominantly approach animal motion capture from an observational or analytical perspective, aiming to study behavior or to address medical questions, rather than examining how these techniques could be integrated into cinematic workflows. Existing comparative papers include discussions of motion capture and animation systems, including lower-end technologies like the Microsoft Kinect and Optical Motion Capture, as well as their respective limitations [Sch12]. Other studies examine the implementation of motion capture within film and television production, addressing factors such as cost, efficiency, quality, and usability [GZ22, Sch12]. However, the majority of these discussions center around human motion capture, an area that is already well-documented and established within film production. In contrast, research publications on animal motion capture rarely consider multiple methods in comparison, and discussions specific to its use in a cinematographic context are particularly hard to find. This gap in the literature leaves much to be explored regarding the practical application of these techniques in filmmaking.

In response to these gaps, this paper seeks to explore a range of animal motion capture methods, highlighting recent advancements and novel solutions to the unique challenges presented by capturing animal movement. Through a detailed analysis of these methods, we aim to demonstrate how they might be integrated into cinematic production pipelines, identify their limitations, and suggest directions for future development. By doing so, this paper aims to provide a forward-looking perspective on the future of animal motion capture technology within the film industry.

3. Common Methods and Difficulties

When discussing motion capture, understanding the fundamental methodologies is crucial before exploring their application to animals: marker-based and markerless systems. These approaches are the foundation of motion capture technology and serve as the basis for capturing movement in various contexts, including medical research, sports analysis, and cinematography. While both methods have been extensively developed for human subjects, they reveal a unique set of challenges that must be addressed to achieve realistic and high quality animations, when adapted for animals. By first examining these methods and their limitations, we can better grasp the specific difficulties in animal motion capture and the innovations needed to overcome them in the film industry.

3.1. Marker-based Motion Capture

Marker-based motion capture is a traditional and widely used technique that involves placing physical markers on the subject's body to track its movement. This method provides high precision in capturing motion, as the markers are tracked by multiple cameras arranged in a controlled environment [LMK*24]. The data captured from these markers allow for the accurate reconstruction of the subject's movements in three-dimensional space. This method is typically divided into two categories: autarkic and non-autarkic systems. Autarkic systems involve wearable suits embedded with sensors that send motion data directly to the capture device. They are highly flexible and suitable for capturing rough, broad movements but may lack fine detail. Non-autarkic systems use external devices, such as synchronized cameras, to capture the motion of markers placed on the subject. The Vicon system, for example, is a prominent non-autarkic system that uses a network of cameras to record and reconstruct movement with high accuracy [Jä24].

Marker-based systems present several challenges when applied to animals. The process of attaching markers can be intrusive, causing animals to resist or exhibit altered behavior, which undermines the goal of capturing natural motion. Additionally, markers can shift or fall off during dynamic movements, especially in high-energy animals, leading to data inaccuracies. Larger animals, such as horses, require many markers, increasing setup complexity and the risk of errors [LMK*24]. Furthermore, marker-based systems are typically designed for controlled indoor environments, making it difficult to capture accurate data in outdoor or natural settings where lighting, obstructions, and environmental variability can interfere with the process.

3.2. Markerless Motion Capture

Markerless motion capture has gained significant traction in recent years, driven by advances in computer vision, machine learning, and depth-sensing technologies [LMK*24, MSD*24]. Unlike marker-based systems, markerless motion capture does not require physical markers on the subject. Instead, it relies on a combination of cameras and sophisticated algorithms to track and reconstruct the subject's movement by analyzing visual data [MPQG18]. This technology can extract motion from video sequences, which can then be processed to generate a 3D model of the movement.

While offering a non-invasive alternative to traditional methods, markerless systems present several significant challenges when applied to animals. Animal movements are often complex and unpredictable, making it difficult for algorithms to track motion accurately without physical markers. Occlusions, where parts of the body block others from view, frequently lead to incomplete or inaccurate data [MSD*24]. Developing species-specific skeletal models is also time-consuming and resource-intensive, given the wide range of animal shapes and sizes. Additionally, capturing precise depth information in outdoor or poorly lit environments remains difficult, often resulting in less accurate 3D reconstructions [MSD*24]. Markerless systems are also sensitive to environmental factors like lighting and background complexity, further complicating the capture process in natural settings. These challenges make it difficult to achieve consistent, high-quality motion data in uncontrolled environments, especially for use in cinematic productions.

3.3. Specific Difficulties in Animal Motion Capture

Capturing animal motion presents unique challenges that go beyond those faced in human motion capture. Animals have a much broader range of body structures, sizes, and movement patterns, making it difficult to apply standard motion capture techniques developed for humans. This diversity requires highly specialized approaches for each species, which can be both time-consuming and resource-intensive. In addition to physical differences, animals exhibit behaviors that are often unpredictable, making it challenging to capture consistent, repeatable movements. The need to preserve natural behavior during capture is another critical challenge. For marker-based systems, the attachment of markers can cause discomfort or stress, leading to altered or unnatural motion, undermining the authenticity of the captured data. Markerless systems, while non-invasive, struggle with occlusions, environmental noise, and the complexity of interpreting movement in uncontrolled outdoor settings, where animals are often found. Lighting, terrain, and background interference further complicate the process, particularly in wildlife cinematography, where capturing animals in their natural habitats is essential [MSD*24]. Additionally, ethical considerations must be taken into account, as ensuring the welfare of animals during motion capture is a priority. Techniques that involve sedation, confinement, or invasive measures to facilitate capture may harm the animals or compromise the integrity of the data by influencing their natural behaviors [MSD*24]. These specific difficulties highlight the complexity of animal motion capture and the need for innovative, non-invasive solutions that can adapt to diverse environments and preserve the authenticity of animal movements.

4. Unique Challenges and Solutions

This section presents a selection of potential motion capture methods that demonstrate promise for integration into a cinematographic pipeline. These methods differ significantly in their approaches, each addressing unique challenges inherent in animal motion capture. By showcasing these examples, the aim is to highlight their innovative aspects and explore their potential for future development.

4.1. Fish Motion Capture

Capturing the motion of underwater animals like fish is especially challenging due to the difficulty of attaching markers and the optical distortions introduced by water. The fish motion capture system, described in detail by Meng et al. [MPQG18], provides a markerless solution, utilizing a monocular camera to track fish movements in water, avoiding the need for complex setups or physical markers. The method starts by positioning a camera above a tank to capture the fish's movement. An adaptive background subtraction algorithm isolates the fish from its environment, removing visual noise [MPQG18]. The system then extracts the fish's contour and uses elliptical Fourier coefficients to smooth the contour and accurately model the fish's medial axis (spine) [MPQG18]. This step is critical for understanding the motion dynamics, as the fish's body flexibility requires precise tracking of its curved movements. Once the data is captured, the system employs a two-stage retargeting process [MPQG18]. First, the body motion is transferred to a digital fish model, ensuring the overall movement is replicated. Second, fin movements are mapped separately using junction points, maintaining accurate positioning and motion consistency. This method allows animators to fine-tune the motion if necessary, making it adaptable for various production needs [MPQG18]. The fish motion capture system's markerless approach and ability to capture realistic underwater motion with a single camera make it an accessible tool for both low-budget projects and larger productions. While it lacks the depth information of multi-camera systems, it provides a practical solution for capturing fluid aquatic motion without the complexity of traditional setups, especially in scenes where cost and simplicity are key factors [MPQG18].

4.2. WildPose 3D

Capturing the natural movements of wildlife poses significant challenges, especially when animals need to be observed without human interference. Traditional motion capture methods, such as attaching markers or filming at close range, often disturb the animals and alter their behavior. WildPose, developed by Muramatsu et al. [MSD*24], solves this by using a non-invasive, long-range system that integrates a zoom-lens camera with LiDAR (Light Detection and Ranging) technology, enabling high-quality capture from a distance of up to 200 meters. WildPose's technical foundation lies in its ability to capture synchronized 2D and 3D data simultaneously. The zoom-lens camera records high-resolution 2D footage, while the LiDAR system generates a 3D point cloud that provides depth information [MSD*24]. This combination allows for accurate reconstruction of the animal's movements in 3D space without the need for physical markers. The captured data is then processed using a segmentation model (e.g., the Segment Anything model) to extract the precise silhouettes of the animals from the 2D footage, which are back-projected into the 3D space created by the LiDAR data. This ensures that both the fine details and broader movement patterns are captured with high accuracy. This setup is particularly beneficial for filming wildlife in their natural habitats, where close proximity would interfere with their behavior. By allowing data collection at long distances, WildPose minimizes human interaction, preserving the authenticity of the animals' movements [MSD*24]. However, this system's reliance on advanced technology such as LiDAR and the complexity of synchronizing visual and depth data may limit its accessibility for lower-budget productions. Despite these challenges, WildPose represents a significant advancement in non-invasive motion capture, providing filmmakers with the ability to capture detailed, naturalistic animal behavior in remote or wild environments. This method opens new possibilities for wildlife documentaries and other film genres that rely on authentic animal movements [MSD*24].

4.3. PFERD Dataset

Capturing the motion of large quadrupeds, such as horses, presents unique challenges due to their size, dynamic movement, and complex biomechanics. Traditional motion capture systems often struggle with occlusion and the need to track both skeletal and soft tissue movements. The PFERD (Poses for Equine Research Dataset) system, developed by Li et al. [LMK*24], addresses these challenges by using an advanced optical motion capture system designed specifically for large animals. The PFERD dataset employs

56 synchronized cameras, capturing motion at a high frame rate of 240 Hz, ensuring that even fast, dynamic movements are recorded with minimal occlusion. Over 100 reflective markers are strategically placed on the horse's body to track not only skeletal motion but also soft tissue deformation and hoof dynamics [LMK*24]. These markers are categorized into groups—skeletal model markers, soft tissue markers, and hoof markers—to provide detailed, multi-level tracking of the horse's anatomy. This setup captures a wide range of equine behaviors, from basic gaits like walking and trotting to more complex movements like rearing and pirouetting [LMK*24]. Once the data is captured, the PFERD system ensures high-resolution motion capture by minimizing occlusion through its multi-camera array. The combination of high frame rate and multiple angles allows for precise 3D reconstruction of both rigid body movement and the subtle dynamics of muscle and skin deformation [LMK*24]. This level of detail is essential for productions requiring highly realistic depictions of equine motion, particularly in scenes where precise biomechanics are critical. While the PFERD system offers unparalleled accuracy, it does come with some logistical challenges. The process of attaching over 100 markers can be time-consuming and sensitive to marker slippage, especially during rapid, high-energy movements [LMK*24]. Additionally, the high volume of data generated by such a system demands significant computational resources for processing, which can pose limitations for productions with tight schedules or limited post-production infrastructure. Despite these challenges, the PFERD dataset remains an invaluable resource for films requiring detailed equine motion, and its techniques can be adapted for other large quadrupeds. The system's ability to capture both skeletal and soft tissue dynamics makes it a versatile tool for creating highly realistic animations [LMK*24].

5. Comparing Results and Discussion

In this section, we critically evaluate and compare the three motion capture methods introduced in the previous section to assess their effectiveness, applicability, and potential for future development within a cinematographic context. By exploring the strengths and limitations of each method, we aim to determine their suitability for integration into the film industry's production pipeline and identify key factors necessary for their refinement and wider implementation.

5.1. Technological Capabilities and Limitations

The fish motion capture system offers a unique markerless solution to capturing the complex movements of fish using a single monocular camera. This approach is notably cost-effective, making it accessible for smaller productions or independent animators who require a low-budget method for realistic fish animations. However, while this technique effectively captures basic body movements, it falls short in delivering the high level of detail and precision that multi-camera systems or marker-based methods can provide. The reliance on a single camera limits the ability to capture depth and complex movements, which might reduce its effectiveness in high-end cinematic productions requiring highly detailed underwater sequences.

WildPose represents a significant advancement with its integration of LiDAR and zoom-lens cameras, enabling the capture of detailed 3D motion data from a distance. This method is particularly beneficial for wildlife documentaries or any cinematic work that requires the natural, undisturbed behavior of animals. By allowing long-range, non-invasive data collection, WildPose minimizes human interference, preserving the authenticity of the animals' movements [MSD*24]. However, the system's complexity and resource demands may be excessive for simpler projects or controlled studio environments, where less sophisticated methods could suffice. The need for specialized equipment and expertise to operate the system might also limit its adoption in mainstream film production.

PFERD-Dataset stands out for its comprehensive capture of equine motion, using an advanced optical motion capture system with a large number of cameras and markers. This method is particularly valuable for productions that require highly accurate depictions of horse movements, such as historical dramas or epic fantasy films. The high frame rate and multiple angles minimize occlusions and ensure detailed 3D reconstruction of even the most subtle movements [LMK*24]. However, the process of attaching over 100 markers to a horse is both time-consuming and prone to error, especially during rapid or complex movements. Additionally, the extensive data generated requires significant computational resources for processing, which could pose challenges in fast-paced production environments.

5.2. Application in Cinematographic Contexts

In a cinematographic context, the fish motion capture system is best suited for projects requiring realistic yet straightforward fish animations, particularly in underwater scenes where traditional motion capture methods are impractical. Its affordability and simplicity make it an attractive option for independent films or smaller studios. However, in large-scale productions where high fidelity and detailed motion are important, the limitations of this method might become apparent, potentially requiring supplementary techniques to achieve the desired level of realism.

WildPose has the potential to revolutionize how wildlife is depicted in film, offering a powerful tool for capturing natural animal behavior from a distance. This makes it particularly valuable for nature documentaries or any film requiring authentic animal interactions in their natural habitats. However, its utility may be limited in controlled studio environments or for scenes where animals are performing specific, directed actions that can be more reliable and accurately captured using traditional methods.

The PFERD-Dataset is indispensable for films that feature horses or comparable four-legged animals prominently, providing unmatched detail in capturing the complex biomechanics of equine movement. Its comprehensive data makes it particularly useful for high-budget productions that can afford the time and resources required for such a detailed capture process. However, its complexity and the potential for marker-related errors may limit its use in projects with tighter budgets or where less precision is required.

5.3. Key Factors for Further Development

Future developments in animal motion capture should focus on several key areas:

- Accessibility and Cost: To broaden the application of these technologies, efforts should be made to reduce costs and improve accessibility. Enhancing the precision of markerless approaches, like those used in the fish motion capture system, or simplifying the equipment requirements for systems like WildPose, could make these methods more viable for a wider range of productions.
- Automation and Efficiency: The automation of processes, particularly in systems like the PFERD-Dataset, where manual marker placement is time-intensive, could significantly streamline production workflows. Advances in AI and machine learning could play a crucial role in automating motion capture processes, reducing the need for manual intervention and speeding up data processing.
- Integration with Existing Pipelines: Ensuring that these motion capture methods can seamlessly integrate into existing cinematic pipelines is essential. This could involve standardizing data formats or improving compatibility with popular animation and editing software, making it easier for filmmakers to incorporate these technologies into their workflows.

5.4. Future Directions

Looking to the future, there is considerable potential for the development of hybrid systems that combine the strengths of different motion capture methods. For instance, integrating the non-invasive, long-range capabilities of WildPose with the detailed, marker-based data from the PFERD-Dataset could create a versatile and comprehensive motion capture solution, capable of handling a wide range of animal movements in various cinematic contexts. Advances in AI and deep learning are also likely to play a significant role in the future of animal motion capture. These technologies could enhance markerless capture methods, improving accuracy and reducing the need for complex setups. Additionally, AI could aid in refining motion retargeting across different species, making it easier to apply captured movements to various digital models. Finally, ethical considerations will become increasingly important as motion capture technologies evolve. The development of non-invasive techniques that do not disturb or harm animals should be a priority, ensuring that the advancement of this technology aligns with principles of animal welfare.

6. Conclusion

Animal motion capture presents unique technological challenges compared to well-established human motion capture systems, primarily due to the diversity of animal physiologies, behaviors, and habitats. This paper explores several advanced techniques designed to address these challenges, focusing on their applicability within cinematography. Marker-based systems, such as the PFERD Dataset, offer unparalleled precision for capturing detailed biomechanical movements of large quadrupeds, though they are labor-intensive and sensitive to errors like marker slippage. On the other hand, markerless approaches like the fish motion capture system and WildPose provide more flexible and less intrusive alternatives, though they still face limitations in accuracy, particularly with complex movements and occlusions.

The primary challenges identified—ranging from maintaining natural animal behavior during capture to the complexity of setting up and operating these systems—highlight the need for ongoing advancements. Emerging solutions, such as long-range systems that minimize human interference (e.g., WildPose), represent significant progress toward non-invasive and scalable motion capture technologies.

Looking ahead, the future of animal motion capture in cinematography lies in the development of hybrid systems that combine the precision of marker-based approaches with the non-invasiveness of markerless systems. Further research and innovation in automation, accessibility, and integration with existing film pipelines are essential to overcoming the current technological barriers. By addressing these needs, filmmakers will be able to create more realistic and immersive portrayals of animals, ultimately enhancing visual narratives and the audience's cinematic experience.

References

[BOX24] BOXX TECHNOLOGIES: Animal motion capture data developments & the impact on animation, 2024. Accessed: 2024-09-05. URL: https://boxx.com/blog/emerging-technologies/animal-motion-capture-data-developments--amp-the-impact- 1

[fIS24] FOR INTELLIGENT SYSTEMS M. P. I.: Markerless animal motion capture, 2024. Accessed: 2024-09-05. URL: https://ps.is.mpg.de/research_projects/markerless-animal-mocap. 1

[GZ22] GUO Y., ZHONG C.: Motion capture technology and its applications in film and television animation. *Advances in Multimedia 2022* (2022), 1–10. URL: https://doi.org/10.1155/2022/6392168, doi:10.1155/2022/6392168. 2

[Jä24] JÄGER M.: Vicon motion capture system. https://forschung.rwu.de/node/479, 2024. Accessed: 2024-09-05. 2

[LMK*24] LI C., MELLBIN Y., KROGAGER J., POLIKOVSKY S., HOLMBERG M., GHORBANI N., BLACK M. J., KJELLSTRÖM H., ZUFFI S., HERNLUND E.: The poses for equine research dataset (pferd). *Scientific Data 11* (2024), 497. URL: https://doi.org/10.1038/s41597-024-03312-1, doi:10.1038/s41597-024-03312-1. 1, 2, 3, 4

[MPQG18] MENG X., PAN J., QIN H., GE P.: Real-time fish animation generation by monocular camera. *Computers & Graphics 71* (2018), 55–65. URL: https://doi.org/10.1016/j.cag.2017.12.004, doi:10.1016/j.cag.2017.12.004. 1, 2, 3

[MSD*24] MURAMATSU N., SHIN S., DENG Q., MARKHAM A., PATEL A.: Wildpose: A long-range 3d wildlife motion capture system. *bioRxiv* (2024). Preprint under CC-BY-ND 4.0 International license. URL: https://doi.org/10.1101/2024.02.05.578861, doi:10.1101/2024.02.05.578861. 1, 2, 3, 4

[Rok24] ROKOKO: Rokoko vision, 2024. Accessed: 2024-09-05. URL: https://www.rokoko.com/products/vision. 1

[Sch12] SCHMIDT R.: *Self-made Motion Capture: Comparing Current Motion Capture Methods Using Kinect and Optical Motion Capture Systems*. Bachelor's thesis, Ludwig-Maximilians-Universität München, 2012. Accessed: 2024-09-05. URL: https://www.researchgate.net/publication/301231445_Self-made_Motion_Capture_Comparing_current_motion_capture_methods_using_Kinect_and_Optical_Motion_Capture_Systems. 2