CASE STUDY

Dynamic Pose Detection and Motion Application to 3D Avatars

An innovative tech start-up in the gaming and virtual meetings industry approached SpringCT with a challenge to develop a solution for capturing human motion from video streams and applying it to 3D avatars in real-time. The client envisioned an immersive, interactive platform where user movements could be mirrored seamlessly by digital avatars.

This project presented complexities in achieving smooth, real-time animations at 60 frames per second. Leveraging its technical expertise, SpringCT designed and implemented a library module, **MoCap**, using Google's MediaPipe and custom algorithms. This lightweight solution brought advanced motion capture capabilities to standard webcams.

Product Features

The **MoCap** library provides an intuitive and efficient framework for motion mapping of the upper body, including the spine, hands, head, and face. The key features include

- **Hands Motion Capture**
  Real-time mapping of left- and right-hand movements with high accuracy.
- **Spine and Head Tracking**
  Captures spine and head rotations (left, right, forward, backward) and applies them to the avatar.
- **Facial Expression Mapping**
  Reflects real-time face gestures, such as smiles, eyebrow movements, and eye blinking.
- **Ease of Integration**
  A modular design allows seamless embedding into any web application.

Technologies Used

- **MediaPipe :** For accurate pose detection from video streams.
- **TypeScript :** To manage pose landmark data flow and animation logic.
- **Blender :** For studying 3D avatar anatomy and ensuring realistic motion rendering.
- **React.js :** To create an interactive and responsive user interface for testing and deployment.

Key Technical Achievements

- **Depth Estimation**
  Google's MediaPipe provides 2D coordinates for body landmarks, making it challenging to accurately infer 3D movements. Custom depth estimation techniques were devised to address this limitation.
- **Body Occlusion**
  Instances of partial body occlusion due to clothing or other objects posed challenges for precise motion detection. Fine-tuning algorithms and leveraging MediaPipe's resilience to occlusions mitigated this issue.
- **Synchronization and Real-Time Processing**
  Ensuring smooth and accurate movement synchronization at 60 FPS required optimizing animation logic and managing computational resources effectively.
- **Avatar Anatomy Compatibility**
  Translating human poses into anatomically realistic movements for a 3D avatar involved careful study and adaptation of avatar joint structures using Blender.

Results

- **Smooth Real-Time Mapping:** The system achieved low-latency, real-time detection and mapping of upper body movements, including simple motions like waving and smiling.
- **High Accuracy:** The library provided reliable motion tracking despite body occlusions, reflecting movements accurately in the avatar.
- User-Friendly Integration: The modular MoCap library enabled straightforward implementation in diverse applications, including virtual meetings and gaming.
- **Challenges Addressed:** While complex motions such as rapid rotations occasionally introduced minor lag, the system proved highly efficient for most use cases.

Conclusion

- SpringCT successfully demonstrated its expertise by delivering a cutting-edge dynamic pose detection solution that animates avatars in real-time.
- By leveraging Google's MediaPipe and innovative techniques for depth estimation and animation, the MoCap library offers an accessible and powerful tool for immersive user interaction.
- This solution is poised to revolutionize applications in gaming, fitness tracking, and virtual meetings, showcasing SpringCT's ability to address complex technological challenges with precision.