

Outline:

- **Motion Capture**
 - ❑ Definition and Importance
 - ❑ Challenges
 - ❑ Application
- **Motion Tracking and Body Tracking**
 - ❑ Definition and Importance
- **Human Pose Estimation (HPS)**
 - ❑ Definition and Importance
 - ❑ Challenges
 - ❑ Application
 - ❑ Algorithms and Methods
 - ❑ Optimization, Metaheuristics, and **Differential Evolution (DE)**
 - Definition and Importance
 - Challenges
 - Application
 - HPS by DE



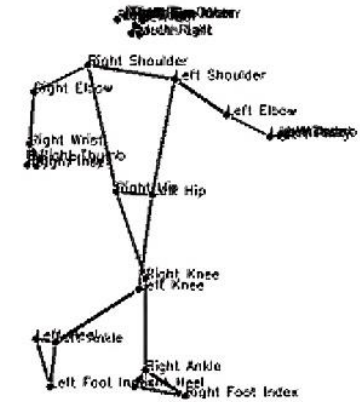
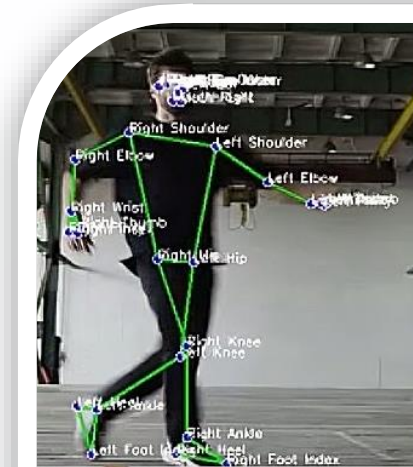
Initialize Population

Mutation

Crossover

Selection

Check Convergence



Result Sample

By: Seyed Muhammad Hossein Mousavi

mosavi.a.i.buali@gmail.com

Lugano, Switzerland

September 2024



• Motion Capture

❑ Definition and Importance

- Motion capture (mocap) is the **process of recording the movement** of objects or people, often using markers, sensors, or cameras.
- It captures precise body movements, which are then mapped onto digital models, making it crucial for realistic **animation in films, video games, and virtual reality**.
- **In healthcare and sports, mocap aids in analyzing physical performance, rehabilitation, and injury prevention.**
- In robotics and AI, it **provides data** for **human-robot interaction** and **motion prediction**.
- Ultimately, it enhances authenticity and accuracy in applications that require realistic motion representation.

❑ Challenges

- Managing **data accuracy**
- Dealing with occlusion when markers or **body parts are hidden from cameras**
- Handling high data **complexity for real-time processing**
- **Motion capture setups can be expensive** and **require controlled environments**, which **limits accessibility and portability**.
- **Data noise and alignment issues** can complicate post-processing, while achieving natural, adaptable animations often requires skilled technical expertise to interpret and refine captured movements.

❑ Applications

- **Entertainment** – Creating **realistic animations** in movies, games, and virtual reality.
- **Sports** – **Analyzing athlete performance**, enhancing training, and preventing injuries.
- **Healthcare** – **Aiding rehabilitation**, posture correction, and gait analysis.
- **Robotics** – **Teaching robots human-like motion** and enhancing human-robot interaction.
- **Research** – Studying biomechanics, cognitive science, and **human behavior modeling**.



• Motion Tracking and Body Tracking

□ Definition and Importance

- **Motion tracking** refers to the process of **capturing and following the movement of objects or individuals within a defined space**.
- This technology **uses sensors, cameras, or software algorithms to record the position, velocity, and trajectory of moving targets**, which can then be analyzed or mapped onto digital avatars.
- It is critical in fields like film, augmented reality, and sports, as it provides real-time or recorded motion data that enhances visual realism, enables interactive experiences, and allows detailed analysis of dynamic movement.
- **Body tracking** specifically focuses on capturing and analyzing the movement of a **person's entire body**, often by monitoring multiple **joints or body segments**.
- **Unlike general motion tracking, which could involve any moving object, body tracking is tailored for human anatomy and biomechanics**, capturing details like joint rotations, posture, and gait.
- This is particularly important in healthcare for gait analysis and rehabilitation, in fitness for posture correction, and in entertainment for creating realistic character animations by transferring human body dynamics to digital models.
- Body tracking allows precise human-centric analysis and representation, which is vital for fields that rely on authentic and adaptive human motion.
- Challenges and applications are same as motion capture.



- **Human Pose Estimation (HPS)**

- **Definition and Importance**

- Human pose estimation is a computer vision technique that identifies and tracks the positions and orientations of human joints in images or videos.
- It involves detecting key body parts such as elbows, knees, and hips, and often represents them as coordinates in 2D or 3D space.
- The goal is to understand human body configurations and movements, even in complex environments with multiple individuals.

- **Challenges**

- **Occlusion:** Body parts can be obscured by other objects or people in the image, making it difficult for algorithms to accurately locate key points. Overlapping figures and interactions, such as people hugging or carrying objects, further complicate pose estimation.
- **Variability in Human Anatomy:** There is a wide range of body shapes, sizes, and proportions among different individuals. This variability can lead to inaccuracies in models that are not trained on sufficiently diverse datasets.
- **Complex and Rapid Movements:** Fast-moving subjects or those engaging in complex movements, such as athletes or dancers, pose a challenge because the pose can change significantly between frames of video.
- **Viewpoint and Perspective:** Changes in camera angle, distance, or perspective can alter the appearance of body parts, making consistent pose estimation difficult across different viewpoints.
- **Lighting and Environmental Conditions:** Poor lighting or extreme lighting conditions can obscure body parts, making it hard for algorithms to detect them. Shadows and reflective surfaces can also mislead pose estimation algorithms.



- **Human Pose Estimation (HPS)**

- ❑ **Challenges**

- **Background Clutter:** Busy or dynamic backgrounds can confuse algorithms, which may mistakenly identify background features as parts of the human body.
- **Real-Time Processing Needs:** For applications like interactive gaming, augmented reality, or surveillance, human pose estimation must be performed in real-time. Achieving high accuracy while minimizing latency is technically challenging.

- ❑ **Application**

- **Augmented Reality (AR) and Virtual Reality (VR):** Pose estimation is crucial in AR and VR environments for tracking user movements and integrating them with digital content. This is essential for immersive experiences in gaming, virtual try-ons in shopping apps, or interactive educational tools.
- **Healthcare:** In physical therapy and rehabilitation, pose estimation can help assess and monitor patients' progress by tracking their movements to ensure exercises are performed correctly. It's also used in surgical training simulations and can aid in early diagnosis of movement disorder diseases like Parkinson's.
- **Sports Analysis:** Coaches and athletes use pose estimation to analyze body mechanics during training. It provides insights into performance improvement and injury prevention by analyzing postures and movements in sports activities.
- **Entertainment and Animation:** In film and gaming, human pose estimation facilitates animation and special effects by capturing human actors' movements and translating them into digital character models, streamlining the production of animated films and video games.
- **Surveillance and Security:** Pose estimation can enhance surveillance systems by detecting unusual or threatening behaviors, like someone falling or acting aggressively, improving security monitoring in public spaces.



- **Human Pose Estimation (HPS)**

- Application

- **Robotics:** Robots that interact with humans, like those used in assistive technologies or manufacturing, can use pose estimation to better understand and predict human actions, making interactions safer and more efficient.
- **Automotive:** In advanced driver-assistance systems (ADAS), pose estimation can help determine the state of the driver, like drowsiness or distraction, and trigger alerts or corrective actions to prevent accidents.
- **Fitness and Wellness:** Apps that track workouts and guide users through exercises often use pose estimation to provide feedback on form and technique, enhancing the effectiveness of home fitness programs.
- **Human-Computer Interaction (HCI):** Pose estimation enables more intuitive HCI by allowing computers and devices to respond to user gestures and movements, facilitating hands-free control and more natural user interfaces.
- **Fashion and Retail:** Virtual fitting rooms use pose estimation to allow customers to try on clothes virtually, providing a convenient and personalized shopping experience that can increase customer satisfaction and reduce returns.



- **Human Pose Estimation (HPS)**

- Algorithms and Methods

- **MediaPipe** is a versatile framework developed by Google for building multimodal machine learning pipelines, such as human pose estimation.
 - ✓ It is great in real-time analysis, **tracking 33 landmarks across the face, hands, and body with high accuracy**, ideal for applications in fitness and sports analytics.
 - ✓ The framework supports deployment across multiple platforms including Android, iOS, and web browsers, making it accessible for a wide range of applications.
 - ✓ MediaPipe integrates seamlessly with other vision tasks like face and hand tracking, enhancing applications in AR and interactive systems.
 - ✓ It's optimized for performance with efficient, **lightweight** models suitable for mobile devices, bolstered by a supportive community and extensive documentation for developers.
- **Convolutional Neural Networks (CNNs)**: CNNs are the backbone of many modern pose estimation systems. They are good in capturing spatial hierarchies in images and are commonly used to predict the likelihood of joints at each pixel.
- **OpenPose**: Developed at Carnegie Mellon University, **OpenPose is one of the first real-time systems to jointly detect human body, hand, facial, and foot keypoints (in total, 135 keypoints) on single images.**
 - ✓ It uses a multi-stage CNN called part affinity fields to predict a set of 2D vectors that encode the position and orientation of limbs across the image.
- **PoseNet and MoveNet**: These are **lightweight** models developed by Google, designed to work in real-time on both the web and mobile devices. They provide good trade-offs between speed and accuracy, and are easily accessible for developers through TensorFlow.js.



- **Human Pose Estimation (HPS)**

- Algorithms and Methods

- **DeepCut and DeeperCut:** These methods focus on part detection and spatial modeling of limbs using CNNs. They were among the first to use deep learning for pose estimation, optimizing joint detection using integer linear programming.
- **AlphaPose:** This algorithm achieves high accuracy in pose estimation by proposing a symmetric spatial transformer network and a parametric pose non-maximum suppression (NMS) that refines the pose predictions by suppressing less likely predicted keypoints.
- **Stacked Hourglass Networks:** Named for their architecture shape, these networks process data at multiple scales and integrate features across all scales. They repeatedly downsample and then upsample the image features, refining the predictions through each 'hourglass' module.
- **Mask R-CNN:** Originally designed for object detection and instance segmentation, Mask R-CNN has been extended to pose detection. It generates bounding boxes for each detected person and then predicts keypoints within each box.
- **EfficientPose:** An adaptation of the EfficientNet architecture for pose estimation, this method scales the baseline network to achieve a good balance between speed and accuracy, making it suitable for real-time applications.
- **3D Pose Estimation Methods:** Techniques like Volumetric CNNs, which predict 3D poses by generating probability volumes for each joint, and methods utilizing depth sensors or multiple cameras to estimate the 3D coordinates of body joints directly.
- **Graph Convolutional Networks (GCNs):** These are used for structured prediction in pose estimation, capturing relationships between joints by treating the body as a graph. This is particularly useful for complex articulations and interactions between body parts.



- **Human Pose Estimation (HPS)**

- ❑ Optimization, Metaheuristics, and Differential Evolution (DE)

- Definition and Importance

- ✓ **Optimization** involves finding the best solution from a set of possible choices that maximizes or minimizes a particular objective. It's crucial in various fields like engineering, finance, and machine learning, where decisions need to be made to enhance performance or efficiency. Techniques range from classical methods like gradient descent to advanced algorithms such as genetic algorithms and simulated annealing. Optimization is also key in training neural networks, where the goal is to minimize a loss function to improve model accuracy. The process can be constrained or unconstrained, depending on whether certain conditions must be met.
 - ✓ **Metaheuristics** are high-level problem-solving frameworks that employ strategies to explore and exploit search spaces efficiently, **often used when exact methods are impractical**. These methods include genetic algorithms, simulated annealing, differential evolution, ant colony optimization, and more, which are inspired by natural processes and behaviors. **Metaheuristics do not guarantee an optimal solution but are effective for finding good-enough solutions quickly for complex optimization problems.** They are widely applicable across disciplines, from logistics and scheduling to engineering design and artificial intelligence. The flexibility and robustness of metaheuristics make them valuable tools in tackling large-scale and multi-modal optimization challenges.



- **Human Pose Estimation (HPS)**

- **Optimization, Metaheuristics, and Differential Evolution (DE)**

- **Definition and Importance**

- ✓ **Differential Evolution (DE)** is a type of evolutionary algorithm and metaheuristic used for global optimization. It optimizes a problem by iteratively improving a candidate solution concerning a given measure of quality.
 - ✓ DE uses mechanisms **similar to genetic algorithms**, involving mutation, crossover, and selection processes. It works with populations of candidate solutions, and these are subjected to perturbation by the differential mutation of vector differences within the population. This is followed by crossover to generate trial candidates, which are then evaluated, and the best candidates are selected for the next generation.
 - ✓ DE is particularly noted for its **simplicity, speed, and robustness**, making it effective in diverse optimization scenarios.

- **Challenges - Optimization**

- ✓ **Scalability:** Handling large-scale problems efficiently without excessive computation time or memory usage.
 - ✓ **Convergence:** Ensuring the algorithm converges to the optimal solution, especially in non-convex problems where local minima can trap algorithms.
 - ✓ **Multi-objectivity:** Balancing multiple competing objectives which may require trade-offs, as in many real-world applications.
 - ✓ **Dynamic environments:** Adapting to changes in real-time where the optimization landscape may change, necessitating continuous recalibration.
 - ✓ **Noisy and incomplete data:** Dealing with uncertainties and errors in the data that can lead to suboptimal or infeasible solutions.



- **Human Pose Estimation (HPS)**

- Optimization, Metaheuristics, and Differential Evolution (DE)

- Challenges - DE

- ✓ **Parameter Selection:** Choosing appropriate values for parameters such as crossover probability and differential weight is crucial, as these significantly influence the convergence behavior and effectiveness of the algorithm.
 - ✓ **Premature Convergence:** DE can suffer from premature convergence, where the population may converge too quickly to a suboptimal solution, especially in complex, multimodal landscapes.
 - ✓ **Scaling to High Dimensions:** As the problem dimensionality increases, DE may require adjustments in strategy and parameter tuning to maintain efficiency and avoid excessive computational costs.
 - ✓ **Balance Between Exploration and Exploitation:** Achieving an effective balance is critical in DE. Too much exploration can lead to slow convergence, while too much exploitation can cause premature convergence to local optima.
 - ✓ **Handling Constraints:** Integrating constraints directly into the DE process can be challenging, requiring specialized techniques to handle boundary conditions and other constraint types effectively.



- **Human Pose Estimation (HPS)**

- Optimization, Metaheuristics, and Differential Evolution (DE)

- Applications

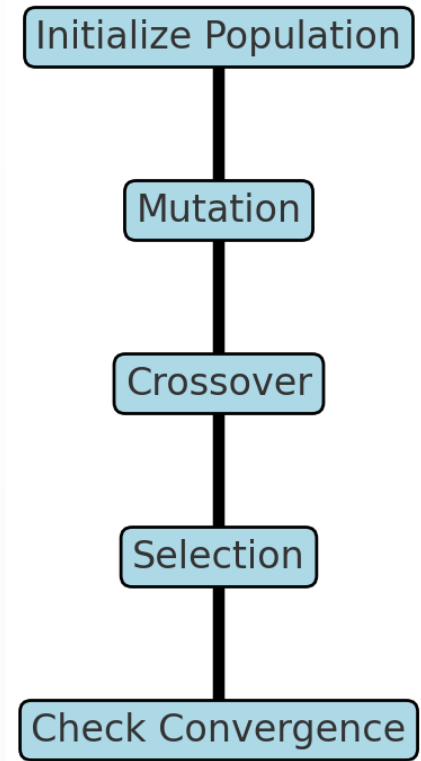
- ✓ **Engineering Design:** DE is widely used for optimizing design parameters in mechanical, electrical, and structural engineering, such as optimizing the shapes of aerospace components for stress resistance or fuel efficiency.
 - ✓ **Image Processing:** In digital image processing, DE can be used for tasks like image segmentation, feature extraction, and image enhancement to improve the quality of visual data.
 - ✓ **Control Systems:** DE is applied in the tuning of control system parameters to optimize the performance of controllers in industrial automation, robotics, and vehicle systems.
 - ✓ **Power Systems:** In the energy sector, DE helps in optimizing the operation of power systems, including load dispatch, renewable energy integration, and power system stabilization.
 - ✓ **Bioinformatics:** DE is used in bioinformatics for sequence alignment, protein structure prediction, and gene expression data analysis, where it helps in understanding biological processes and disease mechanisms.
 - ✓ **Financial Modeling:** DE assists in portfolio optimization, risk management, and option pricing by finding optimal solutions under complex financial scenarios.
 - ✓ **Machine Learning:** It is used for feature selection and neural network training, where DE optimizes the selection of features and network parameters to improve predictive performance.



- **Human Pose Estimation (HPS)**

- ❑ **HPS by DE - MediaPipe's Pose model**

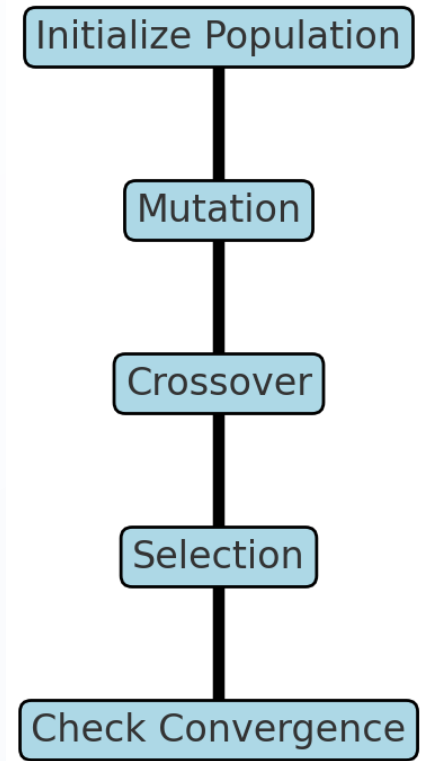
- **Start MediaPipe Pose:** MediaPipe Pose starts with specific parameters, such as model complexity, which determines the accuracy and computational intensity of the model, and the minimum detection confidence, which filters out less certain detections.
- **Video Frame Processing:** For each frame of the video, the frame is captured and converted from BGR to RGB color format because MediaPipe requires RGB input. This conversion is critical for accurate color representation and landmark detection.
- **Pose Landmark Detection:** MediaPipe processes the RGB frame to detect human pose landmarks. It uses a **pre-trained model** based on convolutional neural networks (CNNs) to predict the positions of key body joints, such as elbows, wrists, shoulders, hips, etc. Each detected landmark includes x, y coordinates normalized to the frame size and a z-coordinate representing the depth.
- **Extraction of Joint Positions:** From the detected landmarks, the script extracts the position of each joint. This collection of joint positions forms the basis of the pose that will be analyzed and optimized. These positions are initially approximated based on the output from the MediaPipe model.



- **Human Pose Estimation (HPS)**

- ❑ **HPS by DE - MediaPipe's Pose model**

- **Joint Position Collection:** Initially, for each video frame processed by MediaPipe, the script collects the 2D coordinates (x, y) and depth (z) of each detected joint from the pose landmarks.
- **DE Optimization of Each Joint:** For each joint, DE is applied to optimize its position by minimizing the error between the detected position from the video and a predicted optimal position. This process adjusts each joint's coordinates to potentially more accurate locations, improving the overall pose estimation.
- **Calculation of Rotation and Offset:** After DE optimizes the joint positions, the script calculates the rotation angles (pitch, yaw, roll) and the offset for each joint relative to its parent in the hierarchy. This step is crucial for understanding the spatial relationships and movement dynamics between joints, which is essential for applications that require detailed biomechanical analysis or realistic animations.



GitHub Link

<https://github.com/SeyedMuhammadHosseinMousavi/Real-Time-Differential-Evolution-Human-Pose-Estimation>

