Bone-Aware Pose Estimation and Completion: Theoretical Framework

# 1. Introduction

In human pose estimation, the ability to identify specific bones (e.g., left arm, right leg) and infer missing parts is critical for applications in 3D modeling, animation, healthcare, and robotics. This document outlines a theoretical pipeline that builds on MediaPipe's 2D keypoints and extends them into a semantically labeled and biologically plausible 3D pose.

# 2. Key Concepts

- Semantic Bone Labeling: Each bone is defined by a joint pair and given a human-readable label (e.g., 'left\_thigh').

- Visibility Masking: Confidence scores are used to detect missing body parts.

- Symmetry-Based Inference: The body’s natural left-right symmetry allows for prediction of hidden limbs.

- Graph-Based 3D Prediction: Enhanced pose graphs enable reasoning about spatial structure and depth.

# 3. Semantic Bone Mapping

Each bone in the skeleton is labeled by joint indices. This labeling allows downstream models to know which bone is which.

|  |  |
| --- | --- |
| Bone Label | Joint Pair (start, end) |
| left\_upper\_arm | (left\_shoulder, left\_elbow) |
| right\_upper\_arm | (right\_shoulder, right\_elbow) |
| left\_forearm | (left\_elbow, left\_wrist) |
| right\_forearm | (right\_elbow, right\_wrist) |
| left\_thigh | (left\_hip, left\_knee) |
| right\_thigh | (right\_hip, right\_knee) |
| left\_shin | (left\_knee, left\_ankle) |
| right\_shin | (right\_knee, right\_ankle) |
| torso | (neck, spine\_mid) |
| neck\_head | (neck, head) |

# 4. Inference of Missing Bones

If certain keypoints are not visible (due to occlusion or truncation), the corresponding bones can be inferred by leveraging the symmetry of the human body. For example, if the left leg is visible, we can mirror its direction and estimate the position of the right leg.

# 5. Graph-Based Completion Pipeline

The proposed pipeline consists of the following steps:

1. Detect 2D keypoints using MediaPipe.

2. Create a labeled bone graph based on the detected joints.

3. Use visibility masks to determine missing bones.

4. Predict missing bones using mirrored features and known bone ratios.

5. Feed the completed pose into a 3D pose estimation model (e.g., GCN + Transformer).

# 6. Applications and Benefits

- Enhanced anatomical accuracy for pose estimation tasks.  
- Ability to handle occluded or partially visible human bodies.  
- Useful in healthcare, motion capture, AR/VR, and humanoid robotics.