LLM-Enhanced mmWave Data Synthesis for Environment-Adaptive Human Activity Recognition

Chengyi Liu, M.S. in Artificial Intelligence

FACULTY MENTOR: Yucheng Xie, Ph.D.

Katz School of Science and Health

Background

- Human Activity Recognition (HAR) enables a wide range of applications, including human-computer interaction, elderly fall detection, smart fitness coaching, and VR/AR body motion tracking.
- Recently, millimeter wave (mmWave)-based HAR has emerged as a promising alternative to vision-based HAR or wearable-based HAR.
- Large language models (LLMs) have demonstrated strong potential in generating rich textual content and exhibiting basic forms of logical reasoning.

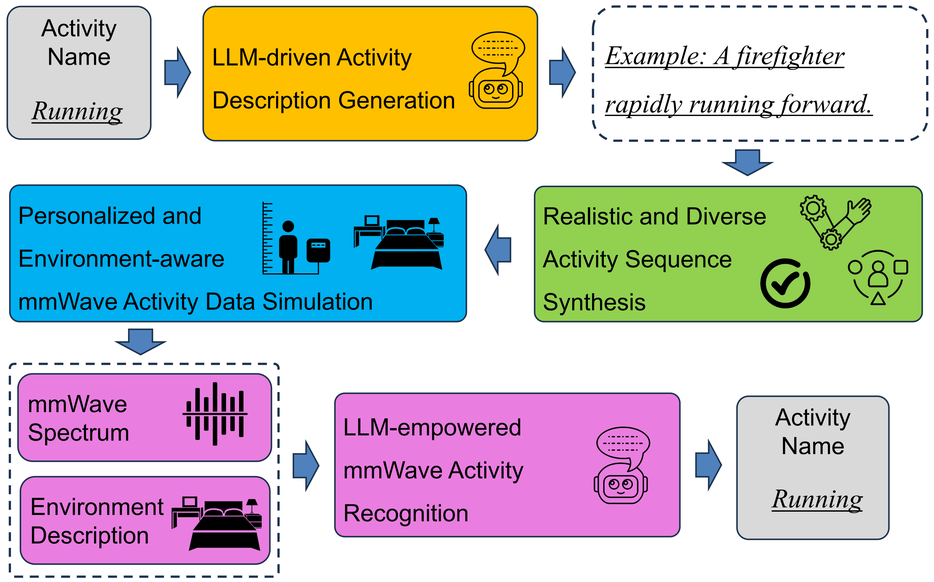
Motivation and Contribution

- The advancement and deployment of mmWave-based HAR are constrained by the high cost associated with collecting large-scale, labeled mmWave datasets from diverse environments.
- We propose a mmWave activity data generation framework that integrates LLM-driven motion description generation with ray-tracing simulation informed by environmental effects and human body characteristics.
- We leverage LLM adaptation and inference to enable environment-adaptive mmWave-based HAR in diverse real-world scenarios.

Challenges

- Automatically generating diverse, structured textual descriptions of common human activities with controllable attributes.
- Synthesizing biomechanically realistic, accurate, and diverse motion sequences from text descriptions.
- Modeling human body characteristics and environmental multipath effects for accurate mmWave activity data simulation.
- Achieving robust mmWave-based human activity recognition across different environments.

System Overview



Methodology

- **LLM-driven Activity Description Generation:** A large language model (LLM) automatically generates diverse and structured textual descriptions of human activities, incorporating controllable motion attributes (e.g., speed, orientation), without requiring manual prompt engineering.
- **Realistic and Diverse Activity Sequence Synthesis:** Human activity sequences are generated from textual descriptions using a motion sequence synthesis model and are subsequently filtered based on biomechanical plausibility, activity relevance, and diversity to ensure high-quality generation.
- **Personalized and Environment-aware mmWave Activity Data Simulation:** Realistic mmWave activity data are generated by constructing personalized 3D human meshes with the Skinned Multi-Person Linear (SMPL) model and simulating the effects of human body characteristics and surrounding environments on mmWave signal propagation via ray tracing.
- **LLM-empowered mmWave Activity Recognition:** Leveraging mmWave data and textual environment descriptions to enable robust activity recognition across different environments via large language model adaptation and inference.

Results

- **Preliminary Results of LLM-driven Activity Description Generation:**
  - ✓ 28 common human activities are defined as the target action set.
  - ✓ For each activity, ChatGPT 4o is used to generate 50 diverse textual descriptions.
  - ✓ Each generated description incorporates speed and orientation attributes to capture motion variations in execution.

Table 1: A set of descriptions generated for the activity "walk".
1. A woman calmly walks down a straight path.
2. A man leisurely walking forward.
3. A girl bravely walks in a straight line.
4. A teenager reluctantly begins to walk.
5. A firefighter rapidly walking forward.
6. .....

- **Preliminary Results of Realistic and Diverse Activity Sequence Synthesis:**
  - ✓ A motion sequence synthesis model is developed to generate skeleton data from each activity description.
  - ✓ An Inverse Kinematics method is developed to filter out descriptions that violate human anatomical constraints.
  - ✓ A motion filtering method is designed to remove inaccurate or irrelevant skeleton sequences.
  - ✓ Two metrics (i.e., Motion Diversity Score and Motion Attribute Coverage Rate) are defined to encourage greater variation among the generated skeleton data.

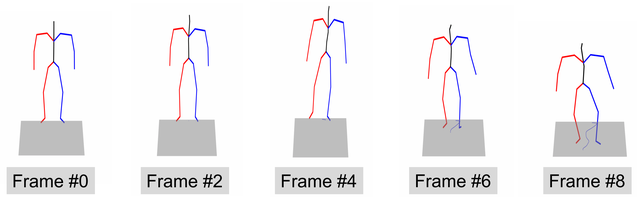


Figure 1: Keyframes from a synthesized motion sequence of a woman calmly walking down a straight path.