

Sensor-based Semantic-level Human Activity Recognition using Temporal Classification



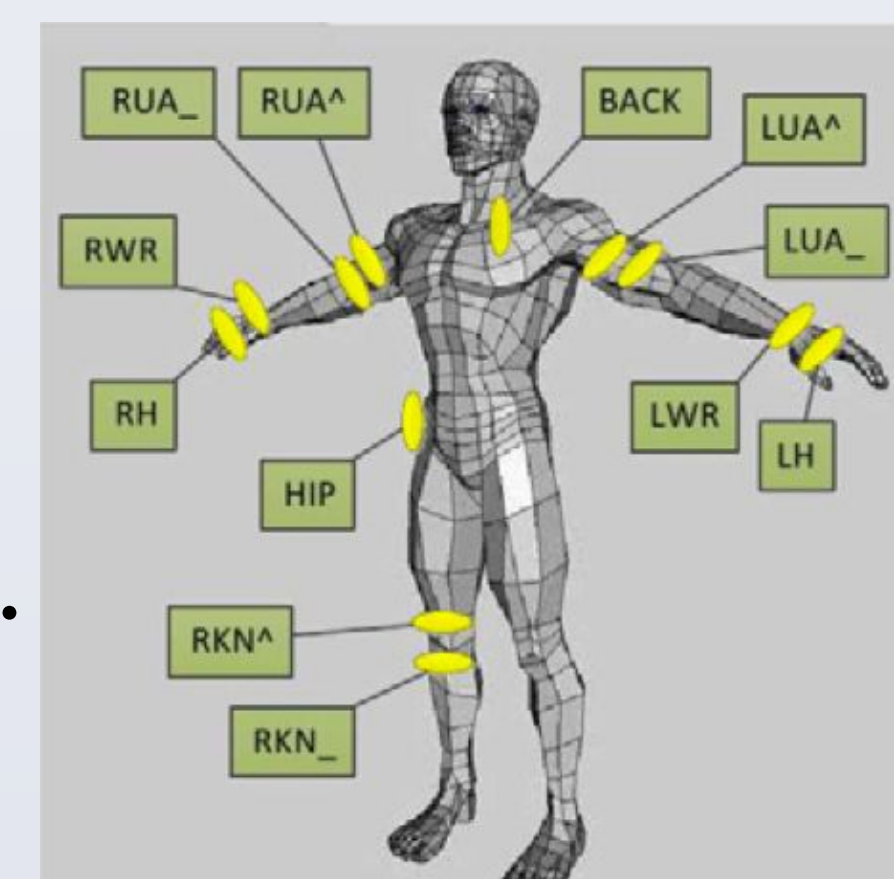
Rui Xu , Chuanwei Ruan , Weixuan Gao

Motivation

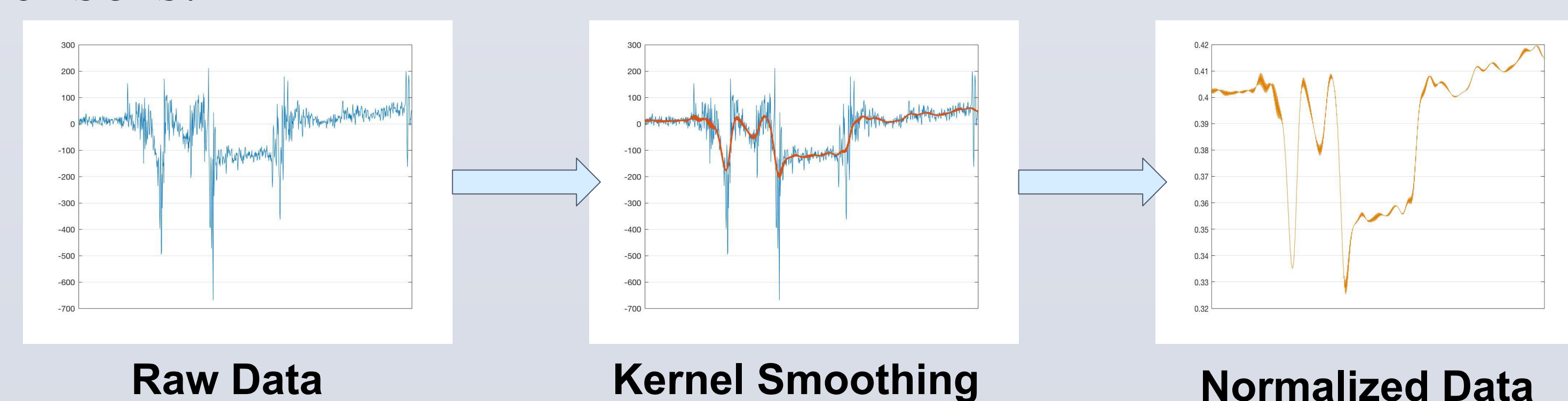
Human activity recognition has always been an active research topic driven by the potential ranging from health care, assistive technologies to manufacturing and smart surveillance system. Most of the effort has been put on recognizing simple action. However, semantic-level activity recognition, which for example could enhance the flexibility and the capability of humanoid robot in a complex context, is not well developed.

Our purpose is to develop a hierarchical approach to recognize semantic-level activities based on sensor-based raw data and low-level actions.

Data



The dataset comprises the reading of motion sensors recorded while users executed typical daily activities. More than 27,000 atomic activities were recorded. There is a total of 242 attributes, including 145 attributes generated from body-worn sensors.



| | Number of Classification | Example |
|--------------|--------------------------|----------------------------|
| Locomotion | 5 | Walk, Sample, Sit |
| Low level | 73 | unlock, stir, close |
| Middle Level | 18 | Open Door1, Drink from cup |
| High Level | 6 | Relaxing, Coffee Time |

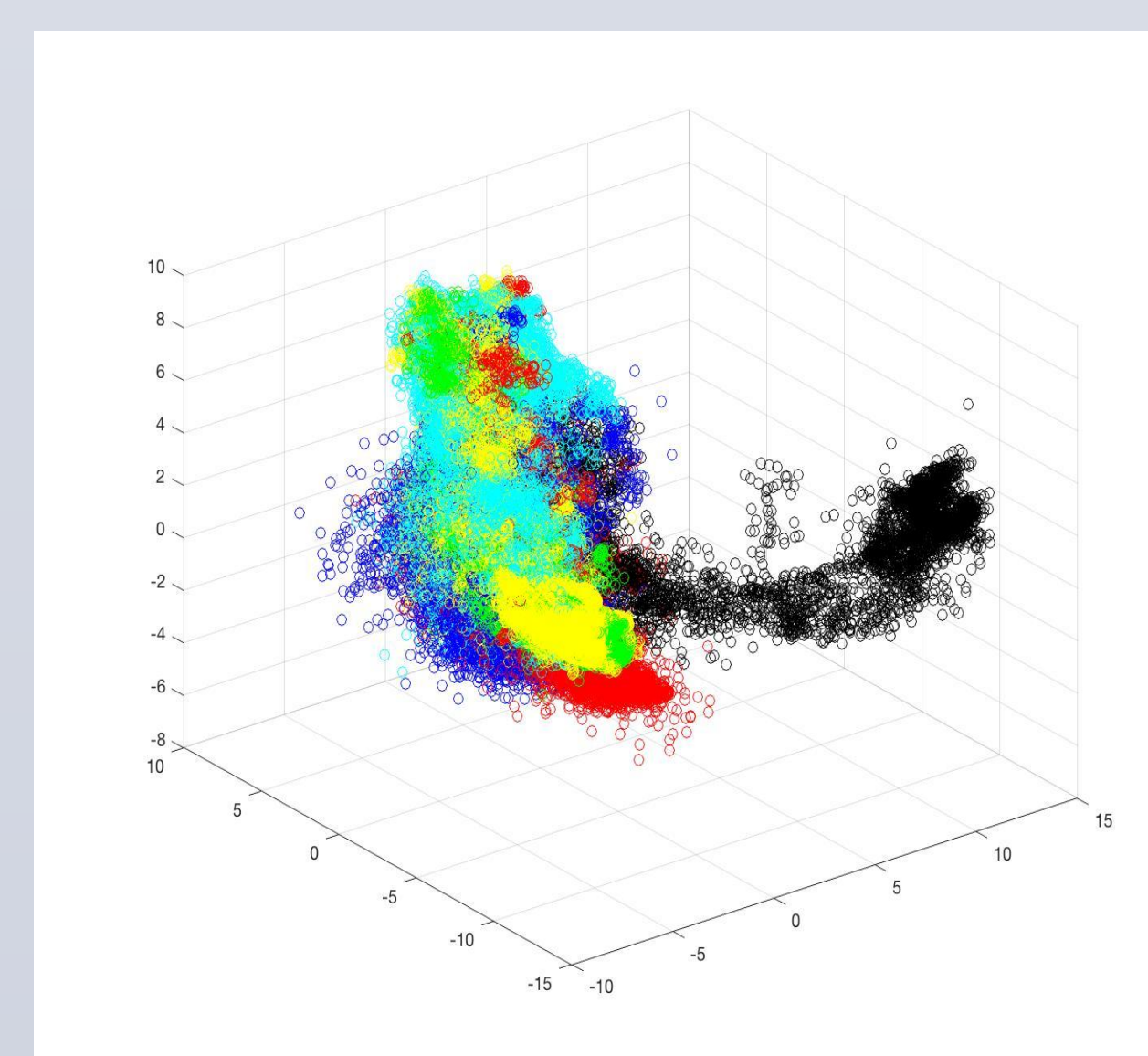
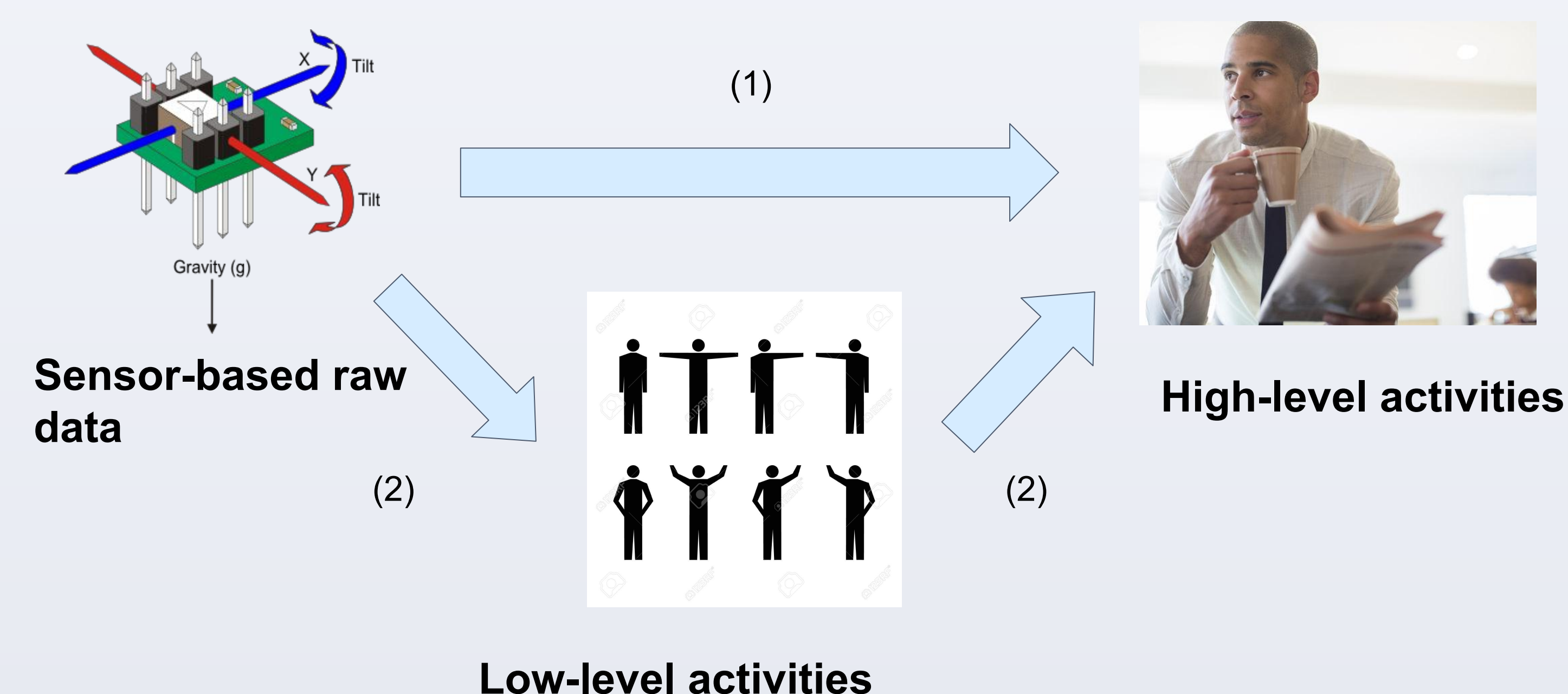


Figure 1:PCA reduction of multi-dimensional space to 3D. Individual examples are color coded according to high-level classification

Methodology

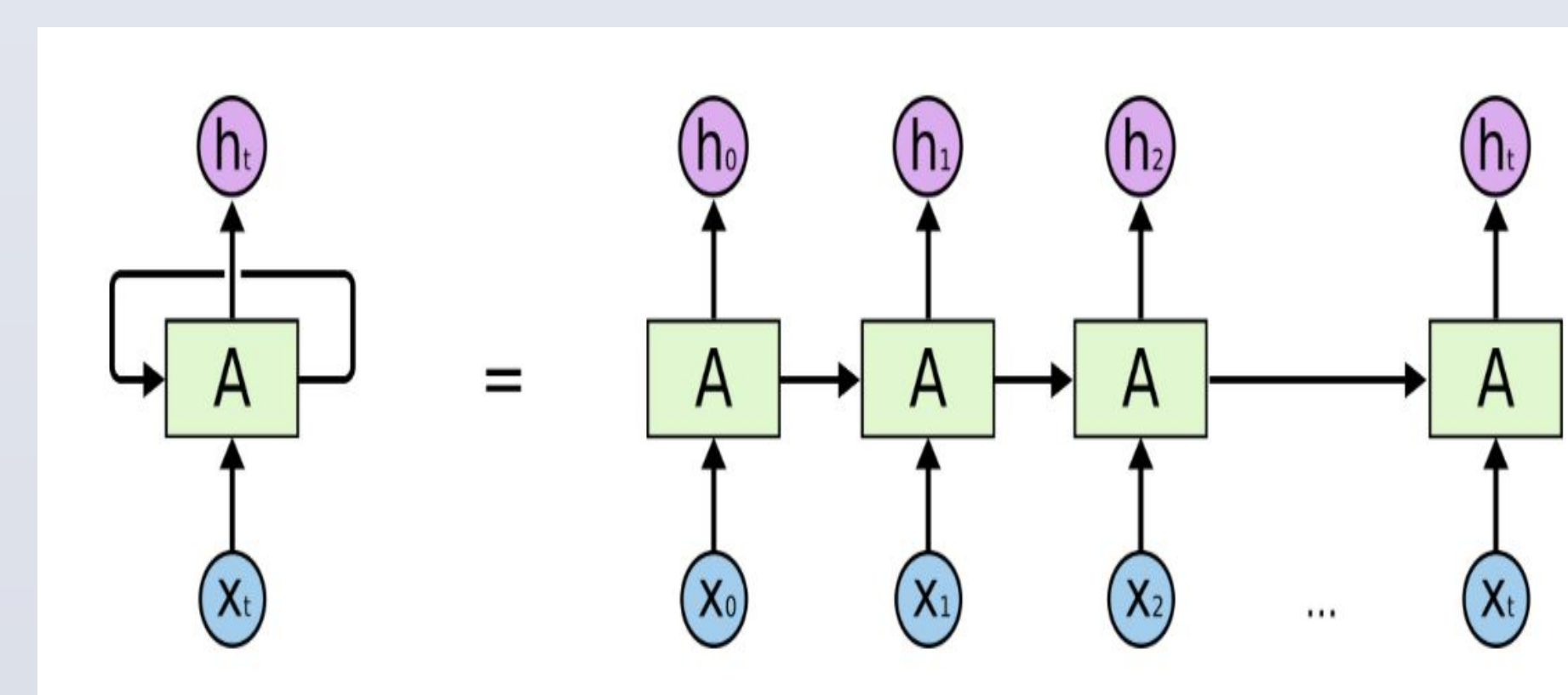
(1) Without Considering Temporal Pattern:



- Baseline is from raw data to high-level activities
- Low-level activities are used as middle layer
- KNN, Boosting, and Randomforest are applied

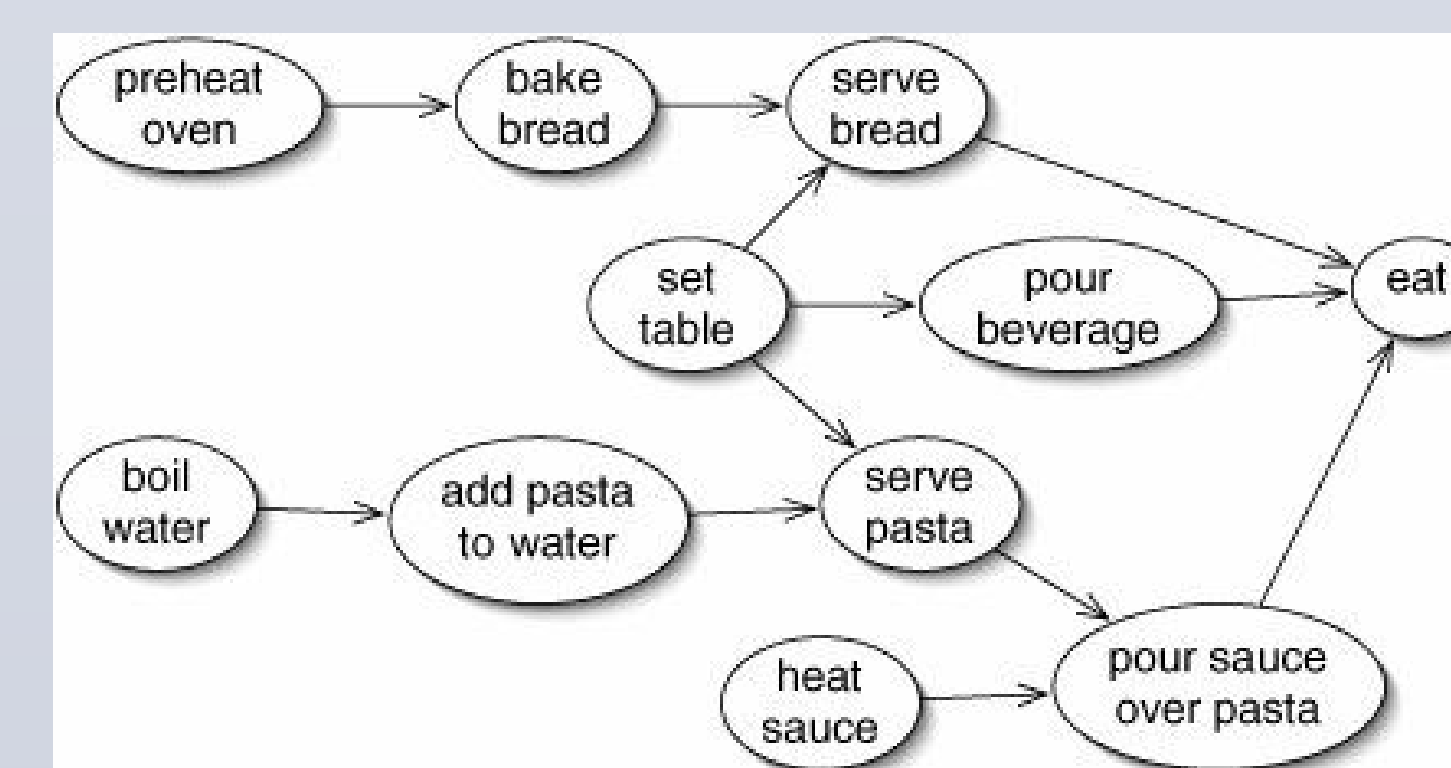
(2) Recurrent Neural Network:

- Different combinations of low-level actions are considered as unique "words"
- Every high-level activity corresponds to a sequence of such "words"
- Apply Long Short Term Memory(LSTM) with softmax layer to predict high-level activities.



(3) Graphical Modeling+LSTM+Multitask Learning

- Building temporal sequential relatedness by using topological sort especially from low → high
- Incorporate raw → high and raw → low → high in LSTM



Results and Analysis

Accuracy(with null class, subject 1):

| | KNN | Boosting Tree | Random Forest |
|----------------------------|--------|---------------|---------------|
| Raw to Locomotion | 86.59% | 92.17% | 94.1% |
| Raw to Low-level Left Arm | 82.66% | 84.78% | 82.5% |
| Raw to Low-level Right Arm | 72.15% | 71.73% | 74.47% |
| Raw to High | 76.40% | 79.72% | 86.01% |
| Low to High | 46.74% | 44.7% | 45% |

- Random Forest generally outperformed other methods
- Too many equidistant points on Low to High using KNN so we add small random noise
- Low level activities info is not put into good use because we failed to extract temporal sequential pattern

Accuracy(without null class, all subjects):

| | LSTM | Boosting Tree | Random Forest |
|-------------|--------|---------------|---------------|
| Low to High | 70.97% | 44.13% | 45.95% |

- Using more data fails to increase the accuracy of Random Forest and KNN due to the high variance between different subjects.
- LSTM significantly outperforms other methods by considering temporal relations.

Future Works

- Improve the current models on low-level activities and locomotions by more sophisticated features engineering and data preprocessing.
- Gather more data and try more appropriate neural network frames. The current poor performance of LSTM are partially due to lack of data and computational resources.
- Build a LSTM models which directly projects the sequences of raw sensor data to the high-level labels.