Analyzing Features for Activity Recognition

Tâm Huynh and Bernt Schiele
Multimodal Interactive Systems
TU Darmstadt
Germany
[huynh,schiele]@mis.tu-darmstadt.de

Activity Recognition using
Wearable Sensors

- Basic Activities (Movements)
  - Walking, Standing, Sitting, Jogging, etc.
  - Clearly defined by motion of body parts
  - Well-suited for recognition with acceleration sensors
Steps in Recognition

- Sensing
  - Conversion from physical inputs to signal data
  - e.g. human motion to acceleration signal

- Feature Extraction
  - Measuring of signal properties useful for classification
  - e.g. mean, variance

- Classification
  - Assignment of the sensed data to a category
  - e.g. “Walking”
Motivation

- Features for Recognition - a common approach:
  - For all activities:
  - ... take a fixed set of features and
  - ... compute them on fixed window lengths

![Example of Acceleration Signal](image)

- Walking
- Jogging

Outline

- Motivation
  - Choosing Features and Window Lengths
- Experimental Setup
  - Hardware
  - Data Set
- Evaluation
  - Features
  - Windows
- Summary & Outlook
Motivation

- Can we do better if we select features and window lengths separately for each activity?

Hardware & Data Set

- Recordings by Intel Research, Seattle
  - Integrated sensor board attached to shoulder
    - Sensors for 3D-acceleration, audio, IR/visible light, temperature, compass, barometric pressure, humidity
  - 200 min of Basic Activities
    - Walking, standing, jogging, skipping, hopping, sitting (in a bus)
Features

- One-dimensional features computed over acceleration signal
  - Mean [e.g. Ravi05, Intille04, Kern03]
  - Variance [e.g. Ravi05, Heinz03, Kern03]
  - Energy [e.g. Ravi05, Intille04]
  - Spectral Entropy [e.g. Intille04]
  - Pairwise correlation b/w axes [e.g. Ravi05, Intille04]
  - FFT coefficients [e.g. Intille04]
- Each feature computed over five different window lengths
  - 0.25, 0.5, 1.0, 2.0 and 4.0 seconds

Data Analysis

- Goal
  - For each activity, analyze combinations of
    - Feature
    - Window length
  - Evaluate relative recognition performance
- Approach: Cluster Analysis
  - Cluster Precision as an indicator for relative recognition performance
Cluster Analysis

- Judge quality of a (feature, window length) combination by looking at how well it clusters activities.
- K-means clustering in feature space.
- For each cluster, compute precision $p$ as indicator for cluster homogeneity.
  
  \[
  \begin{align*}
  \text{Activity 1} & : p_1 = 1, p_2 = 0 \\
  \text{Activity 2} & : p_1 = 0, p_2 = 1 \\
  & : p_1 = 2/3, p_2 = 1/3
  \end{align*}
  \]

- Overall precision $p$ for an activity = weighted average over precision for individual clusters.
  
  $p(\text{activity}) \leq 1$; 1 is best.

Cluster Precision vs. Recognition

- Is cluster precision an indicator for recognition performance?
  - Comparison of cluster precision to results of a simple classifier.
    - Nearest cluster centroid.
    - Variation of distance threshold to obtain ROC curves.
  - Cluster precision seems to correlate with recognition performance:
    - Activity: walking
    - Feature: variance of acceleration
    - Window sizes: 0.25, 0.5, 1.0, 2.0 and 4.0 seconds
Results - Features

• Cluster Precisions of five activities:

- Walking
- Jogging
- Standing
- Skipping
- Sitting (in a Bus)

<table>
<thead>
<tr>
<th>Feature</th>
<th>Jogging</th>
<th>Standing</th>
<th>Skipping</th>
<th>Walking</th>
<th>Sitting (in a Bus)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy</td>
<td>0.1</td>
<td>0.3</td>
<td>0.5</td>
<td>0.7</td>
<td>0.2</td>
</tr>
<tr>
<td>Spectral Entropy</td>
<td>0.2</td>
<td>0.4</td>
<td>0.6</td>
<td>0.8</td>
<td>0.1</td>
</tr>
<tr>
<td>Axis Correlation</td>
<td>0.3</td>
<td>0.5</td>
<td>0.7</td>
<td>0.9</td>
<td>0.2</td>
</tr>
<tr>
<td>FFT Bands</td>
<td>0.1</td>
<td>0.3</td>
<td>0.5</td>
<td>0.7</td>
<td>0.2</td>
</tr>
<tr>
<td>FFT Coefficients</td>
<td>0.4</td>
<td>0.6</td>
<td>0.8</td>
<td>1.0</td>
<td>0.3</td>
</tr>
</tbody>
</table>

- No single 'best' feature for all activities
- FFT features and Variance do consistently well

Results - FFT Features

• Walking: Cluster Precisions for different window lengths of FFT features
Results - FFT Features

- Clear differences b/w high- and low-intensity activities
- Crucial which coefficient/ window length is chosen

Results - Window Lengths

- No single best window length when looking at average
- Significant gains in precision when choosing features and window length individually for each activity

<table>
<thead>
<tr>
<th>Activity</th>
<th>Window Lengths</th>
<th>Cluster Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking</td>
<td>Variance/ 1 sec</td>
<td></td>
</tr>
<tr>
<td>Jogging</td>
<td>FFT coeff. 5+6/2 sec</td>
<td></td>
</tr>
<tr>
<td>Skipping</td>
<td>FFT coeff. 2+3/2 sec</td>
<td></td>
</tr>
<tr>
<td>Sitting</td>
<td>FFT band 3/2 sec</td>
<td></td>
</tr>
<tr>
<td>Standing</td>
<td>FFT coeff. 13+14/0.5 sec</td>
<td></td>
</tr>
</tbody>
</table>

Window lengths: 0.25, 0.5, 1.0, 2.0, 4.0 seconds
Summary

- Analysis of single features/window lengths for basic activities
  - Cluster precision as indicator for recognition performance
- Results
  - Features
    - No single ‘best’ feature for any activity
    - FFT features often perform well
      - But different coefficients for different activities
    - FFT bands might be compromise to individual coefficients
    - Variance does consistently well
    - Mean is less discriminant than variance
  - Window Lengths
    - No single ‘best’ length for all activities
  - For best results, choose an individual (feature, window length)
    combination for each activity

Outlook

- Current and future work
  - Extend approach to more activities
  - Use more sensors and a larger set of features
  - Apply more elaborate classifier scheme (SVM, AdaBoost, ..)
Thanks

... Questions?

Activity Recognition using Wearable Sensors

- Context Awareness plays a central role in ubiquitous computing
- Activity is an integral part of a user’s context
- Wearable Sensors are well-suited for recognizing basic activities
  - E.g. Walking, Standing, Sitting, etc.
  - Acceleration sensors are small enough to be integrated into phones, watches, clothes, etc.
Cluster Analysis

- Judge quality of a (feature, window length) combination by looking at how well it clusters activities
- Calculate precision $p$ as an indicator for how homogeneous the clusters are
  - Best case: $p = 1$
  - Worst case: $p = \text{a priori probability of activity}$

Cluster Precision

- Example: Cluster Precision of Activity 1

\[
p = \frac{8p_1 + 2p_2}{8+2} \approx \frac{3}{4}
\]
Cluster Precision vs. Recognition

- Is cluster precision an indicator for recognition performance?

  Example: Walking

```
(feature, window size):
  - (acceleration FFT coeff. 2+3, 1.0 sec)
  - (acceleration FFT coeff. 1+2, 0.5 sec)
  - (acceleration variance, 1.0 sec)
  - (acceleration FFT bands, 2.0 sec)
  - (acceleration FFT coeff. 7+8, 2.0 sec)
```

Results - Window Lengths

```
Walking
  - 0.25
  - 0.5
  - 1.0
  - 2.0
  - 4.0 seconds

Jogging

Standing

Skipping

Sitting (in a bus)
```