An Ultra-Low Energy Human Activity Recognition Accelerator for Wearable Health Applications

Ganapati Bhat*, Yigit Tuncel
Sizhe An, Umit Y. Ogras
Arizona State University

Hyung Gyu Lee
Daegu University

CASES, October 14, 2019
Outline

- Motivation
- Related Work
- Human Activity Recognition Accelerator
  - Baseline HAR Engine
  - Activity-Aware 2-Level HAR Engine
- Low Power Optimizations
- Experimental Results
- Conclusion
15% of the world’s population lives with a disability*

110-190 million people face difficulties in functioning*

**Intl. Parkinson and Movement Disorders Society Task Force on Technology:**

– Low-cost and small form-factor wearable devices offer great potential

– Enabled by advances in low power sensors and processors

Why Human Activity Recognition (HAR)?

- Identify activities, such as walking, sitting, driving, jogging
- First step to solutions for movement disorders

*We have to know what the patient is doing to reach a conclusion*

- HAR can provide valuable insight

**Applications of HAR**
- Patient rehabilitation
- Fall detection
- Physical activity promotion
Challenges of Wearable Health Technology

- **Adaptation & technology** challenges hinder widespread adoption
  - **Comfort:** Awkward to wear or carry a device
  - **Compliance:** Stop using technology due to maintenance
  - **Applications:** No killer applications

- **27% users give up due to charging reqs [1]**
  - Practical solutions must minimize energy

![Graph showing the increase in the number of papers on the use of wearable devices for PD over the years from 2008 to 2018.](image)

- Flexible energy harvesting devices can address these problems

- **Flexible PV-cell**
  - Ambient power is still lower than 10 to 30 mW requirement
  - Mere 40 hrs with 130 mAh battery

[1] Ana Lígia Silva de Lima et al.. *Feasibility of Large-Scale Deployment of Multiple Wearable Sensors in Parkinson’s Disease*. PLOS One 12, 12 (2017), e0189161
Challenges of Wearable Health Technology

- **Adaptation & technology** challenges hinder widespread adoption
  - **Comfort**: Awkward to wear or carry a device
  - **Compliance**: Stop using technology due to maintenance
  - **Applications**: No killer applications

- 27% users give up due to charging reqs [1]

### Low-power accelerators needed to meet energy budget

- Flexible energy harvesting devices can address these problems

### However,

- Ambient power is still lower than 10 to 30 mW requirement
- Mere 40 hrs with 130 mAh battery

---

Our Novel Contributions

- The first integrated full hardware accelerator for HAR
  - Sensor reading to activity classification

- Novel activity-aware design to minimize energy consumption
  - 22.4 µJ per activity (>17 days with 130 mAh battery)

- Post layout evaluation using TSMC 65 nm LP

- Extensive experimental evaluation with 22 users
  - Dataset released to public (https://github.com/gmbhat/human-activity-recognition)

A critical step towards *self-powered* health monitoring devices
### Related Work

<table>
<thead>
<tr>
<th>Ref</th>
<th>Target App.</th>
<th>Technology</th>
<th>Frequency</th>
<th>Voltage</th>
<th>Power</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1] Alan CW Wong et al.</td>
<td>Vital signal monitoring</td>
<td>130 nm</td>
<td>32 kHz or 16 MHz</td>
<td>1.0 V</td>
<td>530 µW</td>
<td>16 mm²</td>
</tr>
<tr>
<td>[2] Yuxuan Luo et al.</td>
<td>Vital signal monitoring</td>
<td>130 nm</td>
<td>1-20 MHz</td>
<td>0.9 V</td>
<td>93–322 µW</td>
<td>6.25 mm²</td>
</tr>
<tr>
<td>[3] Nick Van Helleputte et al.</td>
<td>Signal acquisition</td>
<td>180 nm</td>
<td>1 MHz</td>
<td>1.2 V</td>
<td>191 µW</td>
<td>49 mm²</td>
</tr>
<tr>
<td>[4] Xin Liu et al.</td>
<td>Signal acquisition</td>
<td>180 nm</td>
<td>Up to 2 kHz</td>
<td>1.1 V</td>
<td>88.6 µW</td>
<td>5.45 mm²</td>
</tr>
<tr>
<td>[5] Wouter Bracke et al.</td>
<td>Sensor AFE for physical act.</td>
<td>500 nm</td>
<td>120 Hz</td>
<td>2.7 V - 3.3 V</td>
<td>120 µW</td>
<td>196 mm²</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>HAR</td>
<td>65 nm</td>
</tr>
<tr>
<td>100 kHz</td>
<td></td>
</tr>
<tr>
<td>1.0 V</td>
<td></td>
</tr>
</tbody>
</table>

- **Frequency**
  - 32 kHz or 16 MHz
  - 1-20 MHz
  - 1 MHz
  - Up to 2 kHz
  - 120 Hz
  - 100 kHz
- **Voltage**
  - 1.0 V
  - 0.9 V
  - 1.2 V
  - 1.1 V
  - 2.7 V - 3.3 V
  - 1.0 V
- **Power**
  - 530 µW
  - 93–322 µW
  - 191 µW
  - 88.6 µW
  - 120 µW
  - 45 – 51 µW
- **Area**
  - 16 mm²
  - 6.25 mm²
  - 49 mm²
  - 5.45 mm²
  - 196 mm²
  - 1.35 mm²


Outline

- Motivation
- Related Work
- Human Activity Recognition Accelerator
  - Baseline HAR Engine
  - Activity-Aware 2-Level HAR Engine
- Low Power Optimizations
- Experimental Results
- Conclusion
Baseline HAR Engine Overview

- **Stretch sensor input:** Measures bending of the knee
- **Accelerometer input:** Measures acceleration at ankle
- **Activities**
  - Sit
  - Stand
  - Walk
  - Jump
  - Up/down stairs
  - Lie Down
3-axis accelerometer data

- The most commonly used sensor for activity recognition
- Since it is notoriously known to be noisy, preprocess using 8-point moving average filter

\[ \bar{s}[kT_s] = \frac{1}{8} \sum_{i=-3}^{4} s[(k + i)T_s] \]

where \( T_s \): Sampling time, 
\( \bar{s}[kT_s] \): Averaged sample at time \( kT_s \) 
\( s[kT_s] \): Raw sample at time \( kT_s \)

- InvenSense MPU-9250
- Low pass filter
- Filter applied to 3-axis data
- **3-axis accelerometer data**
  - The most commonly used sensor for activity recognition
  - Since it is notoriously known to be noisy, preprocess using 8-point moving average filter

- **Use a textile-based stretch sensor (first time for HAR)**

\[
\bar{s}[kT_s] = \frac{1}{8} \sum_{i=-3}^{4} s[(k + i)T_s]
\]

where \(T_s\): Sampling time, \(s[kT_s]\), \(\bar{s}[kT_s]\): Raw, averaged sample at time \(kT_s\)
3-axis accelerometer data
- The most commonly used sensor for activity recognition
- Since it is notoriously known to be noisy, preprocess using 8-point moving average filter

Use a textile-based stretch sensor (first time for HAR)

- Stretchsense Stretch Sensor
- Low pass filter

It has much less noise and power consumption since it is passive

\[
\bar{s}_{kT_s} = \frac{1}{8} \sum_{i=0}^{7} s_{[kT_s] + iT_s}
\]

where \( T_s \): Sampling time, \( s_{[kT_s]} \). Raw, averaged sample at time \( kT_s \)
Input Sensor Data – Segmentation

- 3-axis accelerometer data
  - The most commonly used sensor for activity recognition
  - Since it is notoriously known to be noisy, preprocess using 8-point moving average filter

- Use a textile-based stretch sensor (first time for HAR)

- Segment data into windows by detecting local minima in stretch sensor

- 5-pt derivative to define trends in data

- A new segment when the trend changes from
  - Decreasing to Increasing
  - Flat to Increasing
Feature Generation

- Non-uniform samples due to variable segment length
- Down sample and smooth
  - Down sample block standardizes number of samples
  - 64 for accelerometer, 32 for stretch sensor
- 16-bit Neural Network Features

- Statistical Features
  - Variance of $a_x$, $a_y$, $a_z$, $b_{acc}$ and mean of $a_y$
  - Min, max of stretch sensor and window length
Baseline DNN Classifier

- Detailed neural architecture space exploration
- 2 Hidden layers
  - ReLU Activation
- Output layer with 8 neurons
  - Linear activation with $\text{max}$
  - More hardware-friendly compared to softmax
- Operation and optimizations
  - Design a parameterized module
  - Instantiate for hidden and output layers
  - Only one hour required to change from 3 layer to 2 layer network
- 84% of human activities are static (e.g. sit, stand, lie down)
  - We do not need a DNN to classify them
  - At the same time, more complex dynamic activities must be classified accurately

- **Divide the activities into two classes**
  - A simple support vector machine (SVM) to identify static vs dynamic
  - A 2-Layer NN classifier for dynamic activities

---

**Diagram Description**

- Raw Data:
  - Stretch Sensor
  - 3-axis Accelerometer

- Preprocessing:
  - Stretch
  - $a_x$
  - $a_y$
  - $a_z$
  - $b_{acc}$

- Feature Generation:
  - **Down-Sample & Smooth**
    - Statistical Features
    - Static vs. Dynamic?:
      - DWT
      - 64-point FFT
    - DNN Features
    - 2-Layer NN Classifier

- Activity-aware Classifier:
  - SVM
  - Decision Tree Classifier
  - Static Activity
  - Dynamic Activity

---

84% of human activities are static (e.g. sit, stand, lie down)
- We do not need a DNN to classify them
- At the same time, more complex dynamic activities must be classified accurately

Avoids power hungry FFT and DNN blocks for 84% of activities
Features are reused between SVM and decision tree

DWT and FFT calculated only if activity is dynamic
Outline

- Motivation
- Related Work
- Human Activity Recognition Accelerator
  - Baseline HAR Engine
  - Activity-Aware 2-Level HAR Engine
- Low Power Optimizations
- Experimental Results
- Conclusion
Clock and Data Gating

- Human activities are in the order of few Hz
  - Use this information to clock gate unused blocks
Clock and Data Gating

- Human activities are in the order of few Hz
  - Use this information to clock gate unused blocks

- Data dependencies
  - e.g., downsampling depends on segment detection

![Diagram of data processing flow](image)
Clock and Data Gating

- Human activities are in the order of few Hz
  - Use this information to clock gate unused blocks

- Data dependencies
  - e.g., downsampling depends on segment detection

![Diagram of clock and data gating circuit]
- **Insight from wearable applications domain**
  - Data collection and preprocessing have to be always ON
  - Processing blocks can be activated after the data is available

- **Major power savings potential by turning off processing pipeline**

- **Divide logic into two domains**
  - Segmentation, filtering, FIFO in *always-ON domain*
  - Downsampling, feature generation, and NN in *gated domain*

- **Use signal from segmentation to wake up**

---

**Diagram:**

- **Power Domain 1**
  - Filtering
  - FIFO
  - Segmentation
  - PCU

- **Power Domain 2**
  - Stretch DS & Stats
  - Features
  - Classifier

- **Power Domains Signaling:**
  - 1: Global clk
  - 2: Gated clk 1
  - 3: Gated clk 2
  - 4: Gated clk 3
  - 5: Gated clk 4

- **Control Logic:**
  - Sleep signal to PCU
Outline

- Motivation
- Related Work
- Human Activity Recognition Accelerator
  - Baseline HAR Engine
  - Activity-Aware 2-Level HAR Engine
- Low Power Optimizations
- Experimental Results
- Conclusion
Experimental Setup

- **Design tools and hardware technology**
  - TSMC 65 nm LP
  - Cadence Innovus for APR
  - Synopsys PrimeTime for power

- **User studies**
  - Data from 22 users
  - Total of 4740 segments

- **Training data split**
  - 4 users for test
  - 18 users for training
    - 60% train, 20% cross-val, 20% test
  - 37% test data from unseen users

- **Data used in ESWEEK IoMT design contest**
  - 16 teams from 7 countries

- **Presentations on Tuesday**
  - 15th 12 pm to 1pm

- **Data available open source**
- Synthesize at 100 kHz
- Floorplan during APR
  - Optimize to match logic
- Total area = 1.353 mm²
- FFT has the highest area
- Blocks with memory have higher area
  - FIFO for storing samples
  - Neural network
Design Area: 2-Level Engine

- Total area = 1.357 mm$^2$
  - Only 0.3% larger than the baseline design
- Resembles baseline design
  - Processing blocks are common
Accuracy of the Baseline Engine

- **Weight and Activation Quantization to 16-bits**
  \[ \Delta q = \frac{2W_{\text{max}}}{2^{16}} \text{ where } W_{\text{max}}: \text{ Largest weight} \]

- **Confusion matrix for baseline classifier**
  - Greater than 93% accuracy for all activities

<table>
<thead>
<tr>
<th></th>
<th>Jump</th>
<th>Lie Down</th>
<th>Sit</th>
<th>Stand</th>
<th>Walk</th>
<th>Stairs up</th>
<th>Stairs Down</th>
<th>Transition</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jump</td>
<td>442</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>6</td>
<td>97</td>
</tr>
<tr>
<td>Lie down</td>
<td>0</td>
<td>474</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Sit</td>
<td>0</td>
<td>0</td>
<td>665</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>93</td>
</tr>
<tr>
<td>Stand</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>576</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>27</td>
<td>93</td>
</tr>
<tr>
<td>Walk</td>
<td>31</td>
<td>0</td>
<td>1</td>
<td>10</td>
<td>1913</td>
<td>0</td>
<td>10</td>
<td>42</td>
<td>95</td>
</tr>
<tr>
<td>Stairs up</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>101</td>
<td>6</td>
<td>1</td>
<td>93</td>
</tr>
<tr>
<td>Stairs down</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>97</td>
<td>1</td>
<td>97</td>
</tr>
<tr>
<td>Transition</td>
<td>7</td>
<td>2</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>4</td>
<td>0</td>
<td>229</td>
<td>83</td>
</tr>
</tbody>
</table>
99% accuracy in classifying static and dynamic activities

- Accuracy improvement with 2-Level engine

1% to 8% accuracy improvement with only 0.3% larger area
Power Consumption of Baseline Engine

- Always ON modules consume about 14 µW
- FFT has highest power among classification blocks
- Total power consumption of 51 µW
Power Consumption of 2-Level Engine

- Static activities consume 19.5 µW (2.6x reduction)
- Dynamic activities consume 44.6 µW (1.14x reduction)
- 10x improvement compared to embedded solutions
  - Including sensor and communication energy
- 17 day operation using a 130 mAh flexible battery
Peak Power Consumption Benefits

- Our goal is to operate with ambient energy
  - Peak power must be lower than energy harvesting capacity

- More than 80% time spent in static activities
  - Activity-aware engine provides lower peak power

2.6× and 1.1× reduction in peak power for static and dynamic activities

Facilitates operation with ambient energy
Conclusion

- Presented two human activity recognition engines
  - Fully integrated solution from sensor to activity classification
  - Novel activity-aware engine
  - 22.4 µJ per activity using TSMC 65 nm LP
  - Further power savings possible with voltage scaling

- Dataset from 22 users released to public

A critical step towards **self-powered** healthy monitoring devices