uWave: Accelerometer-based Personalized Gesture Recognition and Its Applications

Jiayang Liu, Zhen Wang, and Lin Zhong
Department. Of Electrical Computer Engineering
Rice University, Houston TX 77005
jiayang@rice.edu, wangzhen127@gmail.com, lzhong@rice.edu

Jehan Wickramasuriya and Venu Vasudevan
Pervasive Platforms & Architectures Lab
Applications & Software Research Center, Motorola Labs
{jehan,venu}@motorola.com

Matthew Knapp
mknapp@wpi.edu
CS 525w Mobile Computing

Introduction to uWave

- **"efficient recognition algorithm"**
  - **Focus on Gestures/Physical Manipulation**
  - **User-dependent Gesture Recognition**
  - **Dynamic Time Warping**
- **Goal: Support efficient personalized gesture recognition on a wide range of devices**

Related Work

- **Computer Vision/Vision Based Techniques**
  - **Translates a “gesture” into “handwriting”**
  - **Fundamentally Limited by Hardware Requirements**
- **Hidden Markov Models**
  - **Require extensive training data to be effective**
  - **Require knowledge of the vocabulary in order to configure the model**

Related Work

- **Dynamic Time Warping (DTW)**
  - **Algorithm for measuring similarity between two sequences which may vary in time or speed**
  - **Allows a computer to find an optimal match between two given sequences with certain restrictions**

Technical Challenges

- **Gesture Recognition lacks a standardized "vocabulary"**
- **Spontaneous interaction requires immediate engagement**

uWave Algorithm Design: Overview

- Premise: "Human gestures can be characterized by the time series of forces applied to the handheld device"
- Template Library
  - Store of one or more time series of known identities for every vocabulary gesture
- Input: Time series of acceleration provided by a three-axis accelerometer

uWave Algorithm Design: Overview

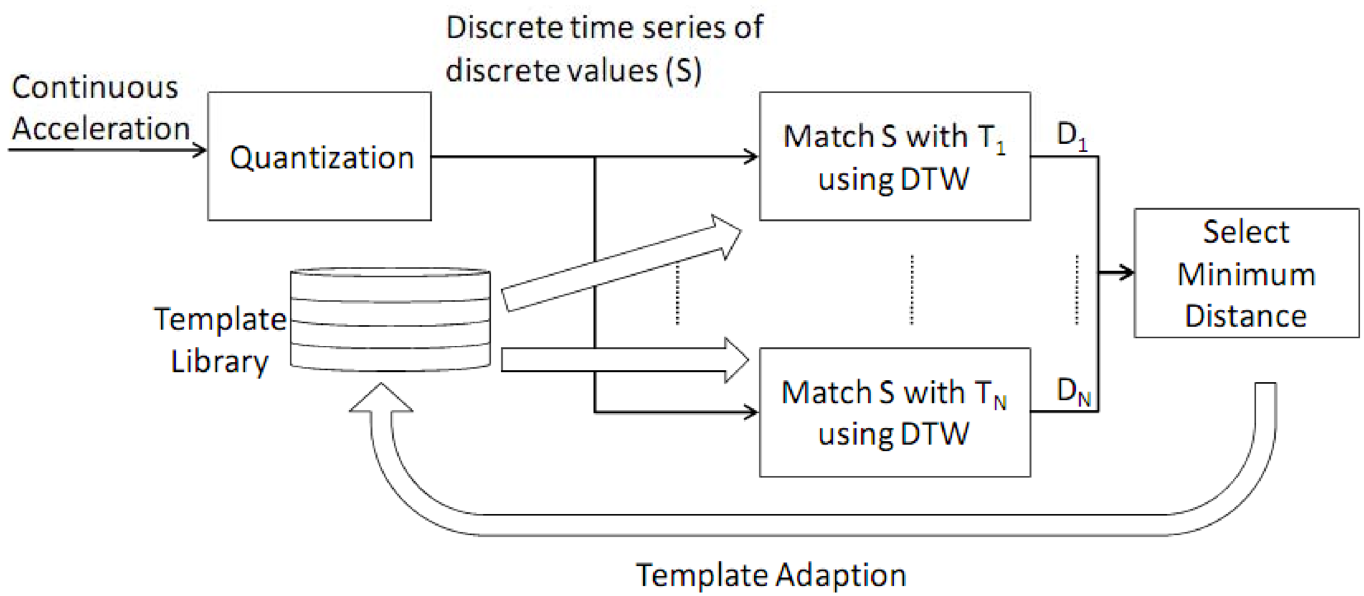


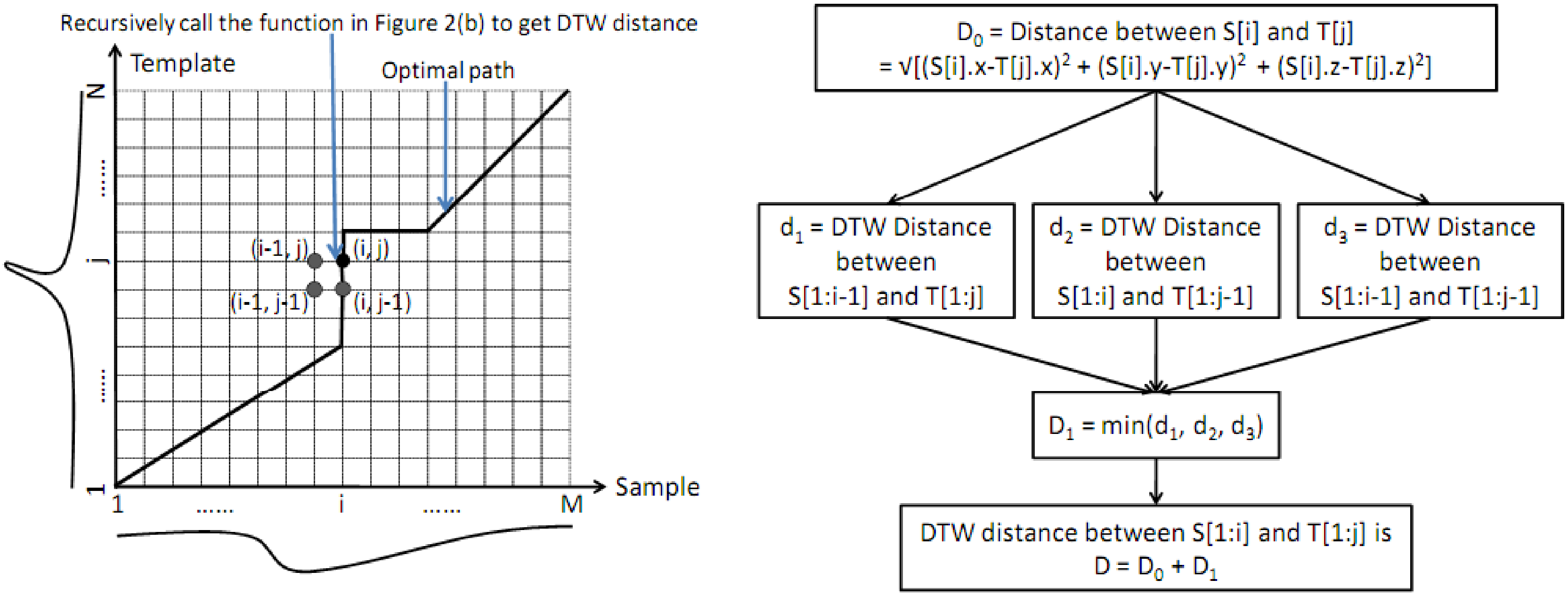
Figure 1: uWave is based on acceleration quantization, template matching with DTW, and template adaptation

uWave Algorithm Design: Quantization

TABLE 1: UWAVE QUANTIZES ACCELERATION DATA IN A NON-LINEAR FASHION BEFORE TEMPLATE MATCHING

Acceleration Data (a)	Converted Value
a > 2g	16
g < a < 2g	11~15 (five levels linearly)
0 < a < g	1~10 (ten levels linearly)
a = 0	0
-g < a < 0	-1~-10 (ten levels linearly)
-2g < a < -g	-11~-15 (five levels linearly)
a < -2g	-16

uWave Algorithm Design: Dynamic Time Warping



(a) Graphic illustration of the recursive algorithm

(b) Algorithm for computing the DTW distance between S[1:i] and T[1:j]

Figure 2: Dynamic Time Warping (DTW) algorithm

uWave Algorithm Design: Template Adaptation

- **Variation between gesture samples by same user**
- **Should adapt templates to accommodate variations**
- **Updating Schemes:**
  - **Positive Update**
  - **Negative Update**

Prototype Implementation

- **Wii remote prototype**
  - Accelerometer range: -3g to 3g
  - Noise below 3.5mg
- **Recognition results returned without perceptible delay on PCs (template library of 8 gestures)**
  - 2ms on Lenovo T60
  - 4ms on T-Mobile MDA Pocket PC
  - 300ms on 16-bit microcontroller in the Rice Orbit sensor

Gesture Vocabulary

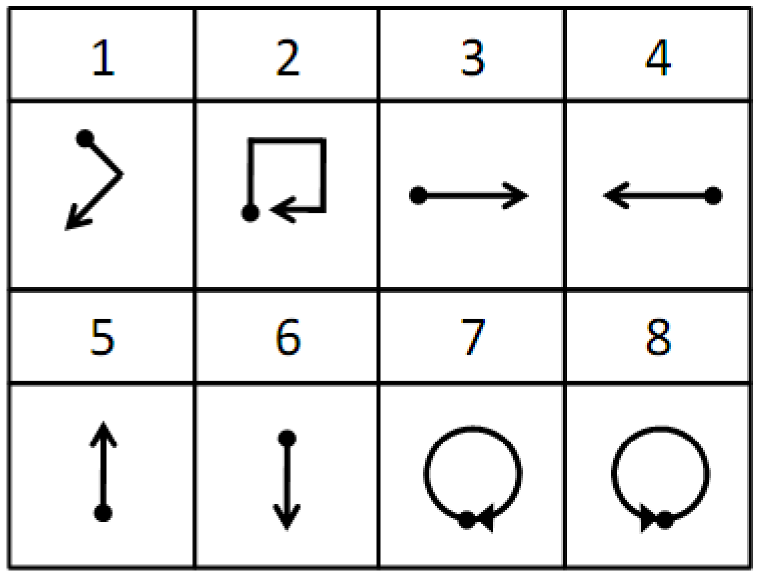


Figure 3: Gesture vocabulary adopted from [6]. The dot denotes the start and the arrow the end

Evaluation: Setup

- Uses the gesture vocabulary from previous slide
- 8 Participants
  - 2 undergraduate, 8 graduate
  - 7 male, 1 female
  - All 20s or early 30s, right handed

Evaluation: Data Collection

- Gestures are collected from 7 days within a period of about 3 weeks
- Each day the participant uses the Wii remote and performs the 8 gestures, 10 times each
- Database at the end consists of 4480 gestures total and 560 for each participant

Evaluation: Recognition without Adaptation

- **Evaluate uWave using the gestures from each subject separately**
- **Use Bootstrapping to improve statistical significance**
- **Use the collected samples to generate 70 tests of uWave**
  - **Produces 70 confusion matrixes**
  - **Averaged into 1 confusion matrix per subject**
  - **Average confusion matrixes of the 8 subjects combined into a final confusion matrix**

Evaluation: Recognition without Adaptation

	>	□	→	←	↑	↓	○	○
>	92.1	0.1	2.4	1.9	0.1	2.9	0.6	0.1
□	1.6	91.6	1.3	1.1	0.7	0.4	2.7	0.6
→	0.5	0	95.9	1.2	0.7	1.7	0	0
←	0.3	0	1.6	96.2	0.7	1.1	0	0.1
↑	0.3	0	1.5	0.6	97.0	0.5	0	0.1
↓	2.4	0	2.4	2.3	1.0	91.7	0.1	0
○	3.4	1.9	2.6	1.7	0.4	0.7	89.2	0
○	1.1	0.6	1.7	0.9	0.8	0.7	0	94.2

	>	□	→	←	↑	↓	○	○
>	98.4	0	0.3	0.4	0	0.4	0.3	0.2
□	0.5	98.3	0.2	0	0.3	0.1	0.4	0.1
→	0.2	0	98.3	0.6	0.1	0.6	0.2	0
←	0.2	0	0.3	98.8	0.3	0.2	0.2	0
↑	0.4	0	0.2	0.4	98.7	0.1	0.2	0
↓	0.7	0	0.6	0.5	0.3	97.7	0.2	0
○	0.5	0.4	0.4	0.1	0.1	0.3	98.1	0.2
○	0.2	0.1	0.1	0.2	0	0	0.2	99.2

Figure 4: Confusion matrixes for the Nokia vocabulary without adaptation. Columns are recognized gestures and rows are the actual identities of input gestures. (Left) Tested with samples from all days (average accuracy is 93.5%); (Right) Tested with samples from the same day as the template (average accuracy is 98.4%)

Evaluation: Recognition without Adaptation

- **Average Accuracy of 93.5%**
  - **Gestures 1,2,6 and 7 have lower accuracy due to similar hand movements**
- **Large variation (9%) among participants**
  - **“The participant with the highest accuracy performed the gestures in larger amplitude and slower speed compared to other participants”**
- **Temporal Compression of the data speeds up recognition by more than 9 times without negatively affecting accuracy**

Evaluation: Recognition without Adaptation

Evaluation Using Samples from the Same Day

- Significantly Higher Accuracy (98.4%) when using only samples from the same day
- Results reported in previous reports may have been overly optimistic
- “The difference between Figure 4 (Left) and Figure 4 (Right) highlights the possible variations for the same gesture from the same user over multiple days and the challenge it poses to recognition.”

Evaluation: Recognition with Adaptation

	>	□	→	←	↑	↓	○	○
>	96.8	0	1.5	0.3	0	1.1	0	0.2
□	0.7	96.4	0.5	0.2	0.2	0.4	1.2	0.5
→	0	0	98.9	0.6	0	0.5	0	0
←	0.2	0	0.3	98.9	0.2	0.5	0	0
↑	0.2	0	0.2	0.1	99.3	0.2	0	0
↓	0.6	0	0.6	0.3	1.7	96.8	0	0
○	0.8	2.0	2.0	0.4	0	0.2	94.6	0
○	1.0	0.4	1.1	0.4	0	0	0	97.1

	>	□	→	←	↑	↓	○	○
>	97.7	0	1.2	0.6	0	0.6	0	0
□	0.6	98.6	0.2	0.1	0	0.1	0.3	0.1
→	0.1	0	99.1	0.4	0.1	0.4	0	0
←	0.1	0	0.4	99.0	0.1	0.4	0	0
↑	0.2	0	0.3	0.1	99.2	0.2	0	0
↓	0.5	0	0.4	0.2	0.5	98.3	0	0.1
○	0.4	0.5	07	0.2	0.1	0.2	98.0	0
○	0.2	0	0.3	0.4	0.1	0.1	0	98.9

Figure 5: Confusion matrixes for the Nokia vocabulary with adaptation, tested with samples from all days. Columns are recognized gestures and rows are the actual identities of input gestures. (Left) Positive Update (average accuracy is 97.4%); (Right) Negative Update (average accuracy is 98.6%)

Evaluation: Recognition with Adaptation

- Produced 7 confusion matrixes for each participants
- Averaged into confusion matrix on previous slide
- Accuracy:
  - Positive Update: 97.4%
  - Negative Update: 98.6%
- Accuracy is much better than without adaptation
  - Close to same day accuracy

uWave-Enhanced Applications: Gesture-based Light-Weight User Authentication

- Prioritizes Ease-of-use over hard security
- Privacy Insensitive
- Enables authentication based on physical manipulation of the device
- Ran studies that showed uWave can recognize user-defined gestures with higher than 99.5% accuracy

uWave-Enhanced Applications: Gesture-based 3D Mobile User Interface

- **Intuitive and Convenient to navigate a 3D interface with 3D hand gestures**
- **Social Networking-based video-sharing service**
- **Rotating Ring Interface**
  - **Employed uWave to navigate the interface**
  - **Uses a series of specific movements such as tilting or slight shaking**

Discussion of uWave

- **Gestures and Time Series of Forces**
  - Diverse opinions on what is a unique gesture
  - Closer to speech than handwriting
- **Challenge of Tilt**
  - uWave uses a single three-axis accelerometer
  - Tilt can change the readings of force applied
  - Opportunity for detecting tilt is limited with a single accelerometer
  - Extra Sensors needed to fully address problem

Discussion of uWave

- **User-Dependent vs. User Independent Recognition**
  - **Much Lower Accuracy for User Independent Recognition (75.4% down from 98.4%)**
  - **No commonly accepted gestures for Interactions**
- **Gesture Vocabulary Selection**
  - **More Complicated Gestures may have higher accuracy**
  - **Number of Complicated Gestures Users can use may be small**

Conclusions

- **Employs a single accelerometer so it can be readily implemented on current devices**
- **Uses DTW to measure similarities between two time series of forces**
- **Tests show uWave achieves 98.6% accuracy with one training sample**
  - **Comparable to HMM-based methods with 12 training samples**
- **Challenges of Variation across Time and Users**

Video Demonstration

- uWave Demonstration

Questions?