SGRS: A Sequential Gesture Recognition System using COTS RFID

Bo Chen, Qian Zhang, Run Zhao, Dong Li, Dong Wang

chenbo.31@sjtu.edu.cn

Shanghai Jiao Tong University

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

Outline

上海交通大学
SHANGHAI JIAO TONG UNIVERSITY

1

PART ONE

Motivation

Gesture Recognition Applications

Motion gaming

Smart home

VR or AR controller

Sign language

Gesture Recognition Technologies

Computer Vision

- Line of Sight
- Sensitive to light conditions
- Privacy concerns

Wearable Sensor

- Inconvenient
- Poor scalability

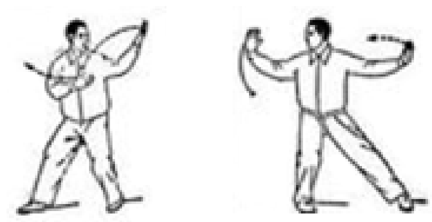
Device-free RF

- Cannot differentiate multiple parts

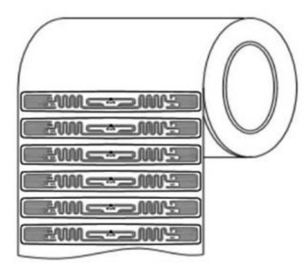
Hard to identify sequential gestures

Sequential Gesture Recognition System

- **Identify sequential gestures, which are composed of a series of temporally-related simple actions in order.**



- **A sequential gesture recognition system with COTS RFID devices.**
  - Battery-free
  - Scalable
  - Non-specific



Sequential Gesture Recognition System

- **Experiment subjects**

Traffic command gestures of Chinese traffic police

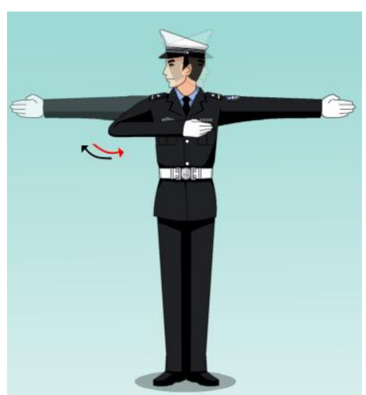
stop

change lane

left turn pending

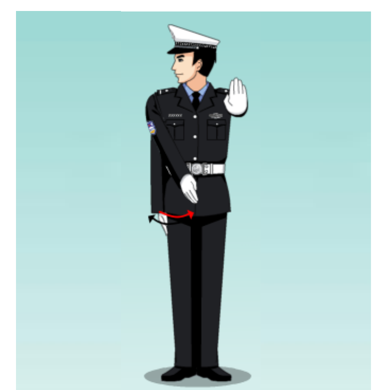
slow down

pull over



go straight

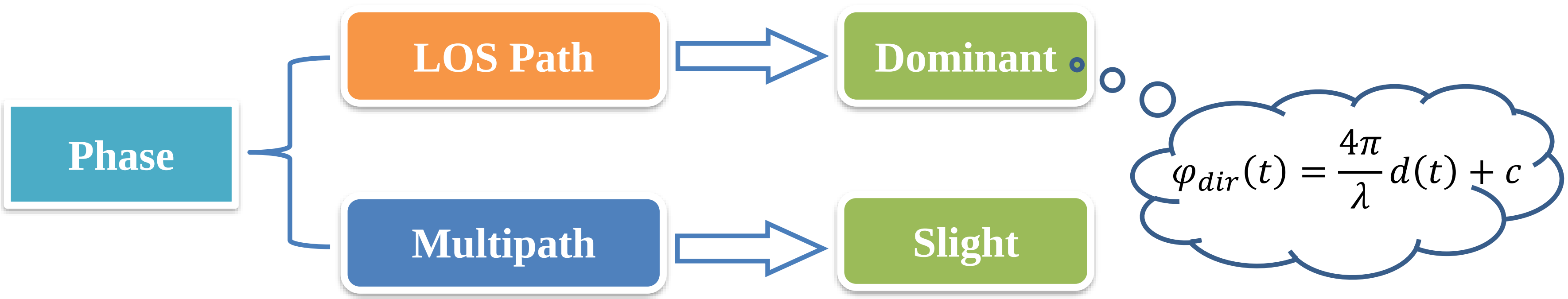
turn left



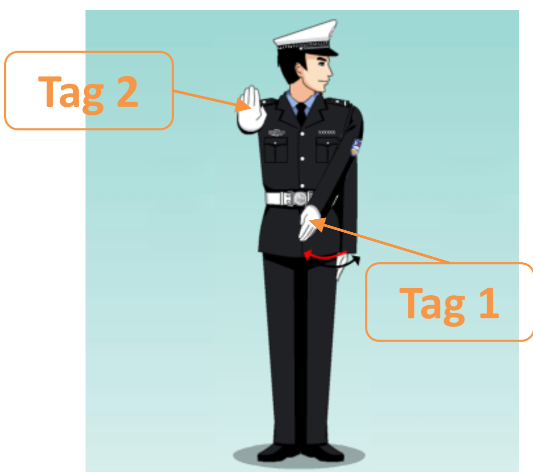
turn right

Key Insight

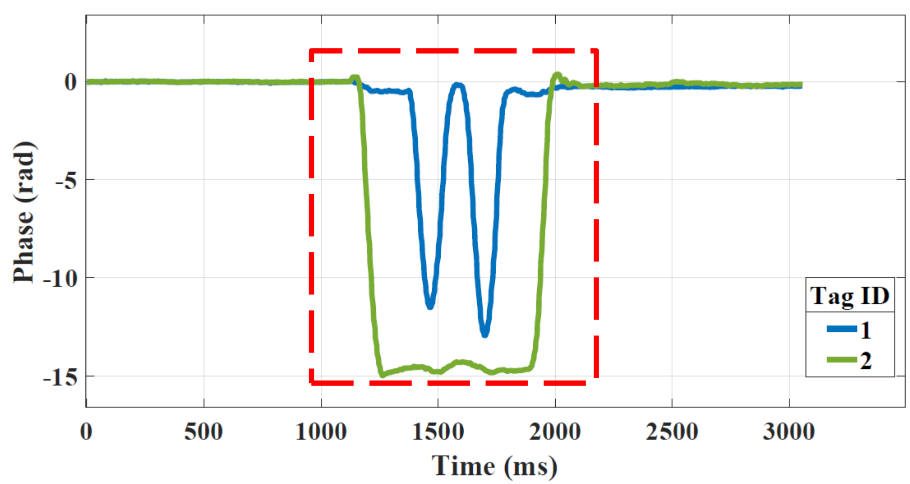
- **RFID phase is affected slightly by the multipath but dominated by the time-varying distance between the tag and the antenna.**



- **Fine-grained signal phase is capable of perceiving sequential gestures.**



turn left



上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

2

PART TWO

Challenges

Challenges and Solutions

C1

Describe temporally-related actions of sequential gesture

- The temporal relation of the same body part
- Combined expression of different body parts

Extract features for each time window

S1

- Extract time-frequency domain features for each tag in each sliding window
- Combine the features of multiple tags into a feature vector

Challenges and Solutions

C2

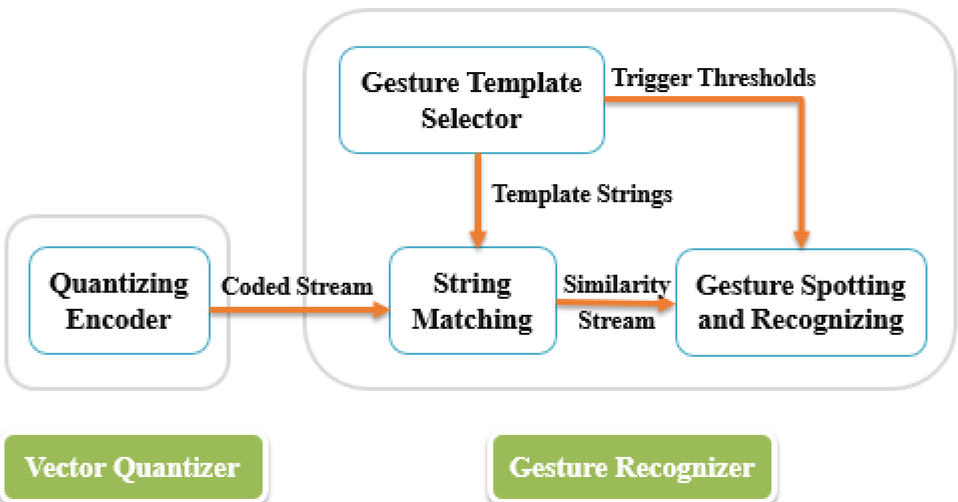
Fast and precise recognition

- Important for an instant human-computer interaction interface

Vector quantization and string matching

S2

- Vector quantization for mapping the high-dimensional feature vectors into a discrete subspace of lower dimension
- Fast and efficient string matching algorithm for recognition

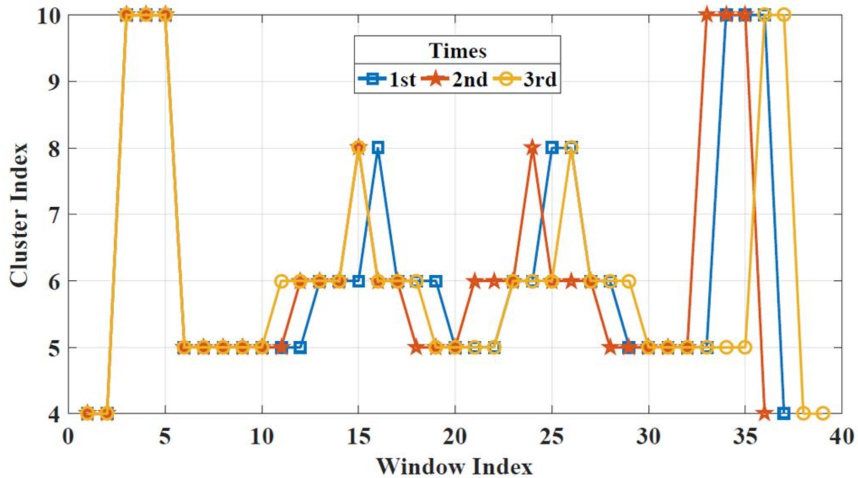


Challenges and Solutions

C3

Individual diversity

- Users with different gesture durations
- Not identical even for the same person



Improved edit distance

S3

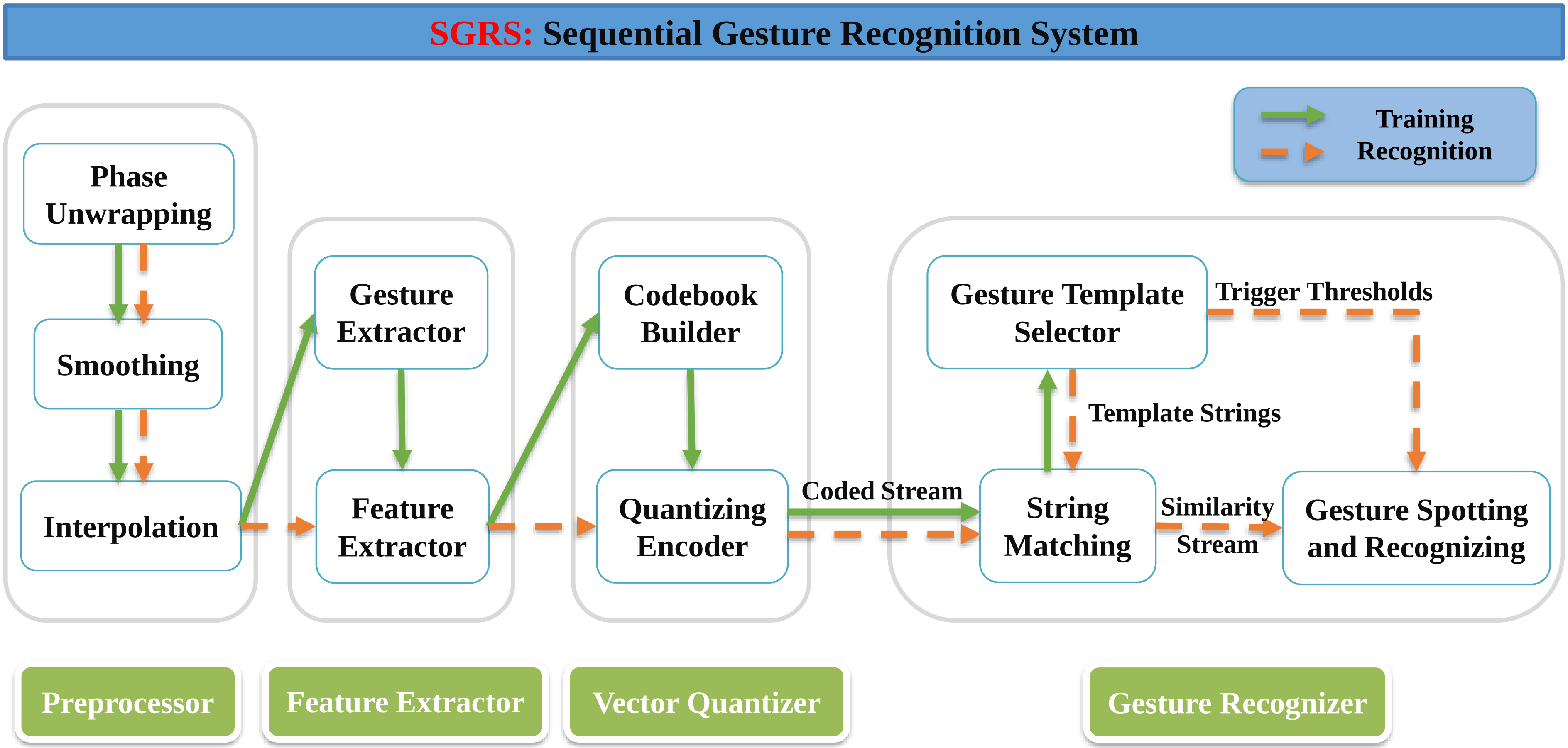
- Improved edit distance algorithm for identifying gestures and suppressing individual diversity
- Increase intra-class similarity and reduce inter-class similarity

3

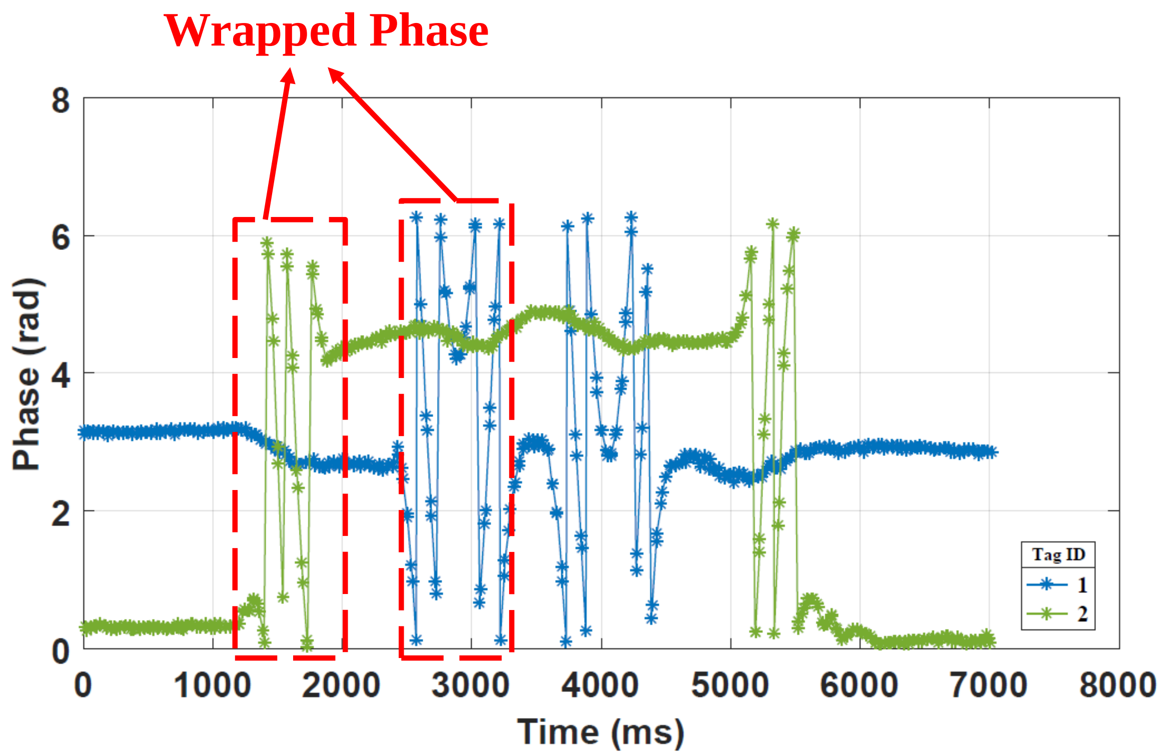
PART THREE

System Architecture

System Overview



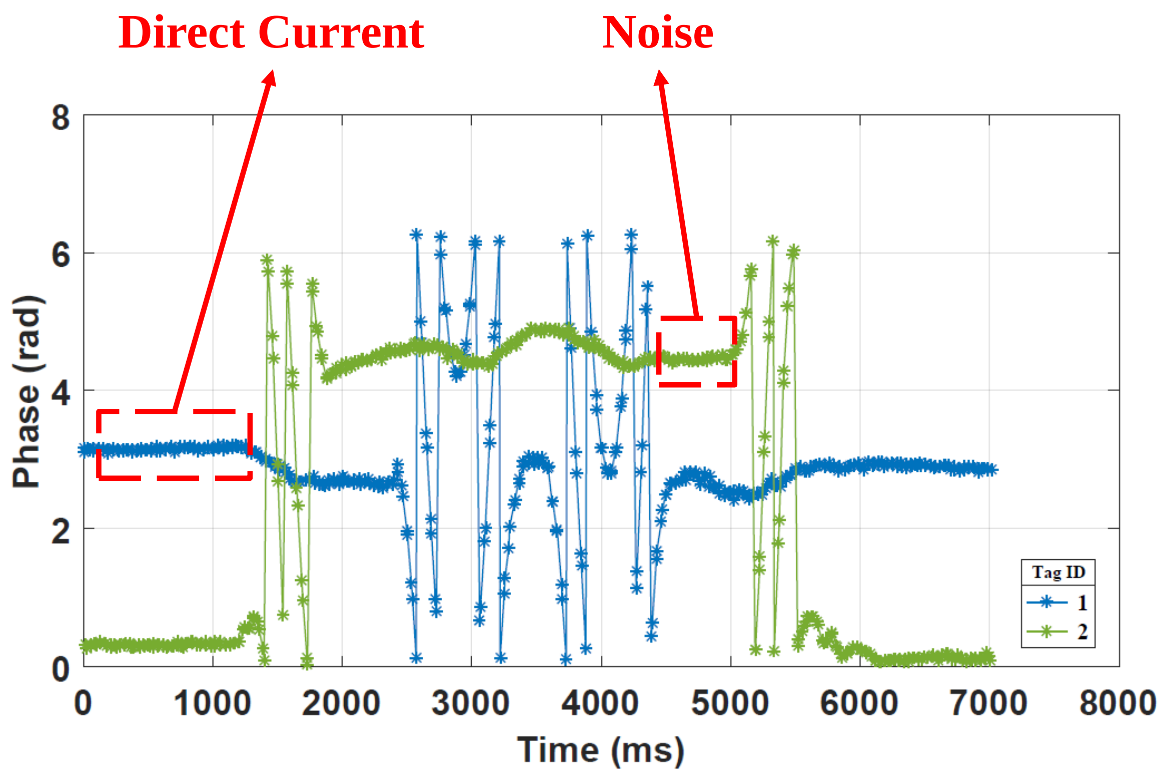
Preprocessor



Phase unwrapping

$$p_i = \begin{cases} p_i - 2 * \pi, & p_i - p_{i-1} \geq \pi \\ p_i, & |p_i - p_{i-1}| < \pi \\ p_i + 2 * \pi, & p_i - p_{i-1} \leq -\pi \end{cases}$$

Preprocessor



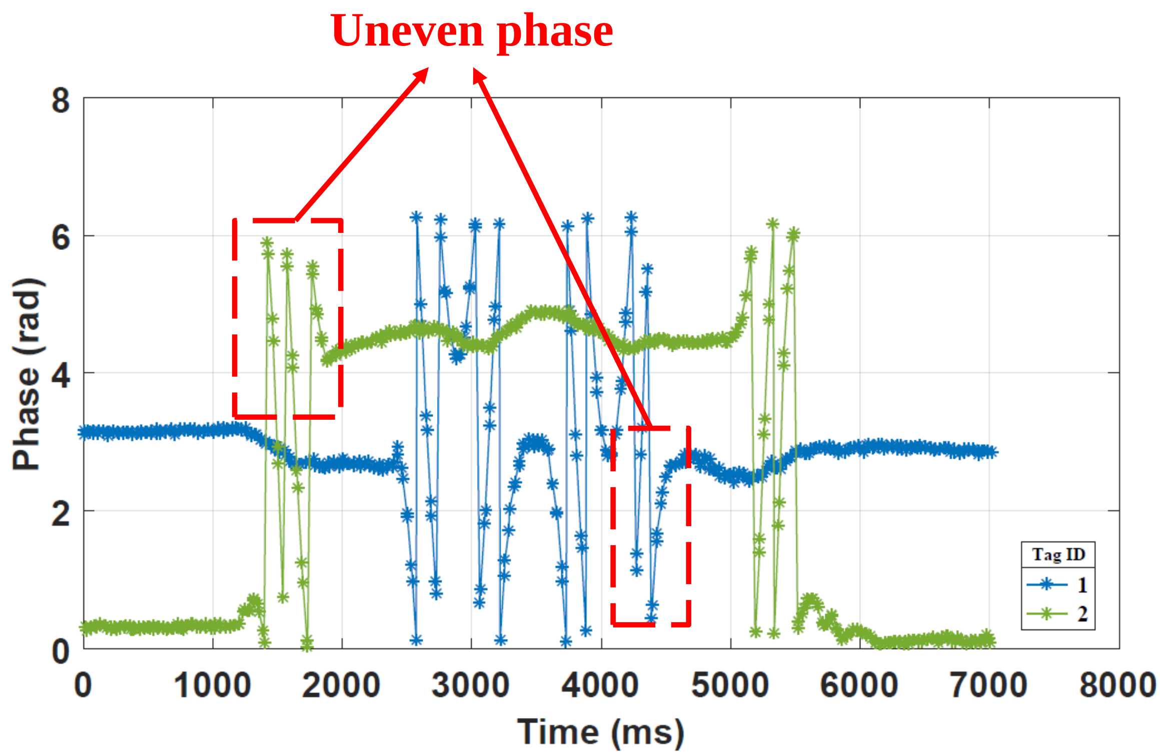
Phase unwrapping

Smoothing

$$p_i = \begin{cases} p_i - 2 * \pi, & p_i - p_{i-1} \geq \pi \\ p_i, & |p_i - p_{i-1}| < \pi \\ p_i + 2 * \pi, & p_i - p_{i-1} \leq -\pi \end{cases}$$

Hampel identifier
Weighted moving average filter
DC removal

Preprocessor



Phase unwrapping

$$p_i = \begin{cases} p_i - 2*\pi, & p_i - p_{i-1} \geq \pi \\ p_i, & |p_i - p_{i-1}| < \pi \\ p_i + 2*\pi, & p_i - p_{i-1} \leq -\pi \end{cases}$$

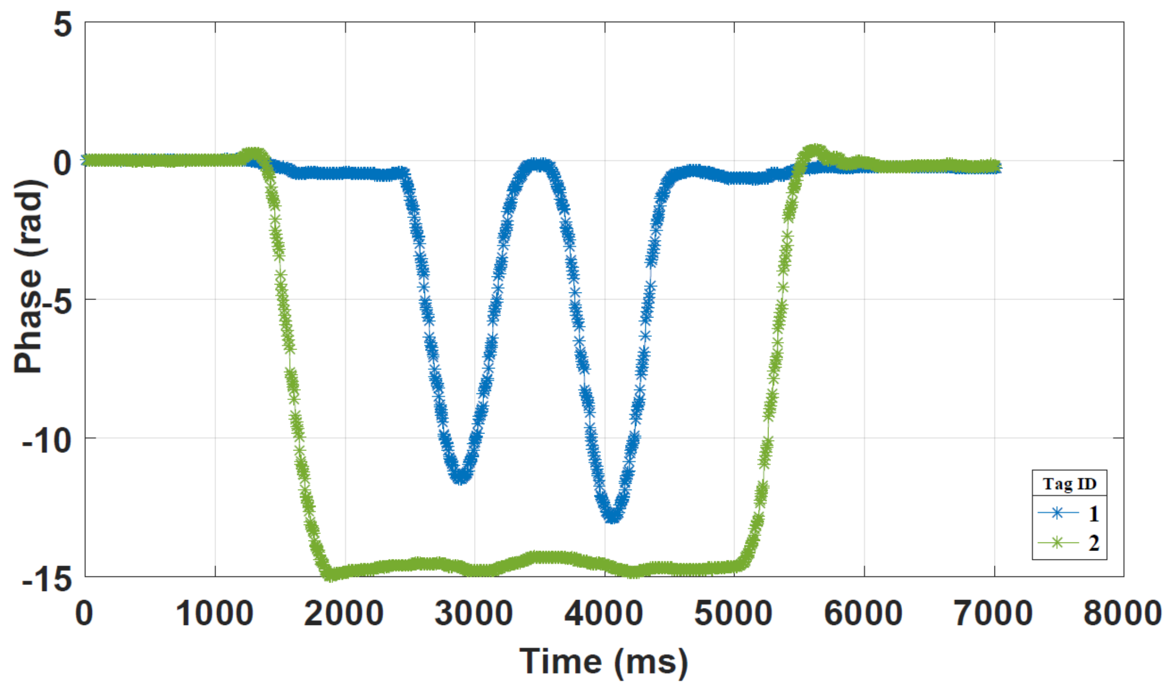
Smoothing

Hampel identifier
Weighted moving average filter
DC removal

Interpolation

Linear interpolation

Preprocessor



Phase unwrapping

$$p_i = \begin{cases} p_i - 2*\pi, & p_i - p_{i-1} \geq \pi \\ p_i, & |p_i - p_{i-1}| < \pi \\ p_i + 2*\pi, & p_i - p_{i-1} \leq -\pi \end{cases}$$

Smoothing

Hampel identifier
Weighted moving average filter
DC removal

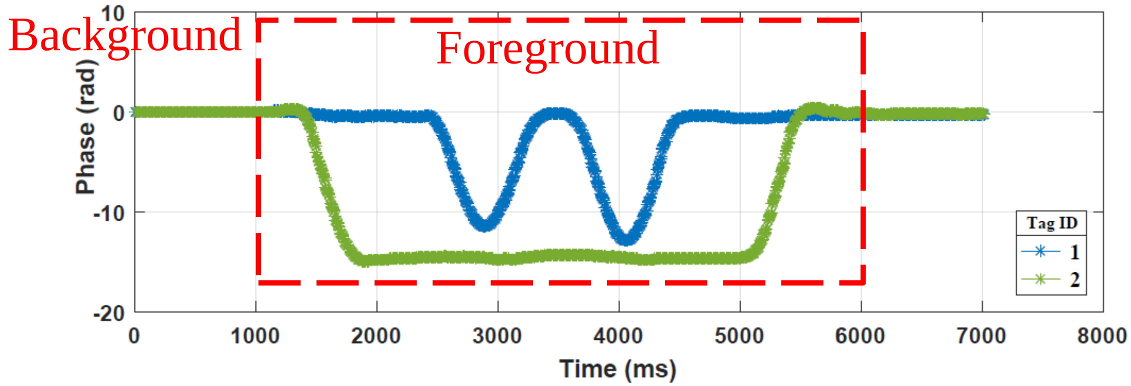
Interpolation

Linear interpolation

Feature Extractor

- **Gesture Extraction (offline training)**

Before feature extraction, the signal fragment corresponding to the gesture is extracted by Foreground detection method.



- **Feature Extraction**

For each sliding window, extract time-frequency domain features of each tag's phase signal into a feature vector.

Feature	Description
Mean	Average phase
Standard deviation	Phase fluctuation
Peak-to-peak amplitude	Magnitude of phase discretization
Standard deviation of PSD	Energy strength fluctuation

Vector Quantizer

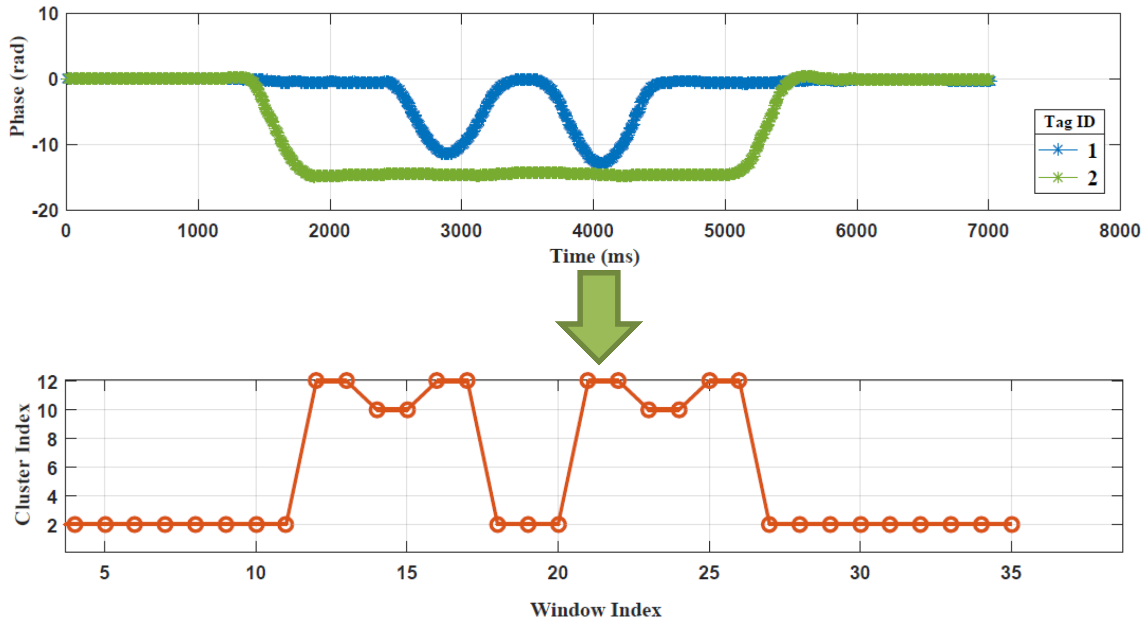
- **Codebook Builder (offline training)**

Cluster the feature vectors by k-means algorithm and use the category number (cluster index) i, centroid vector c_i, mean μ_i and standard deviation σ_i as an item in the codebook.

$$< i, c_i, \mu_i, \sigma_i >$$

- **Quantizing Encoder**

The feature vectors of a series of windows are encoded into a coded stream according to the codebook.

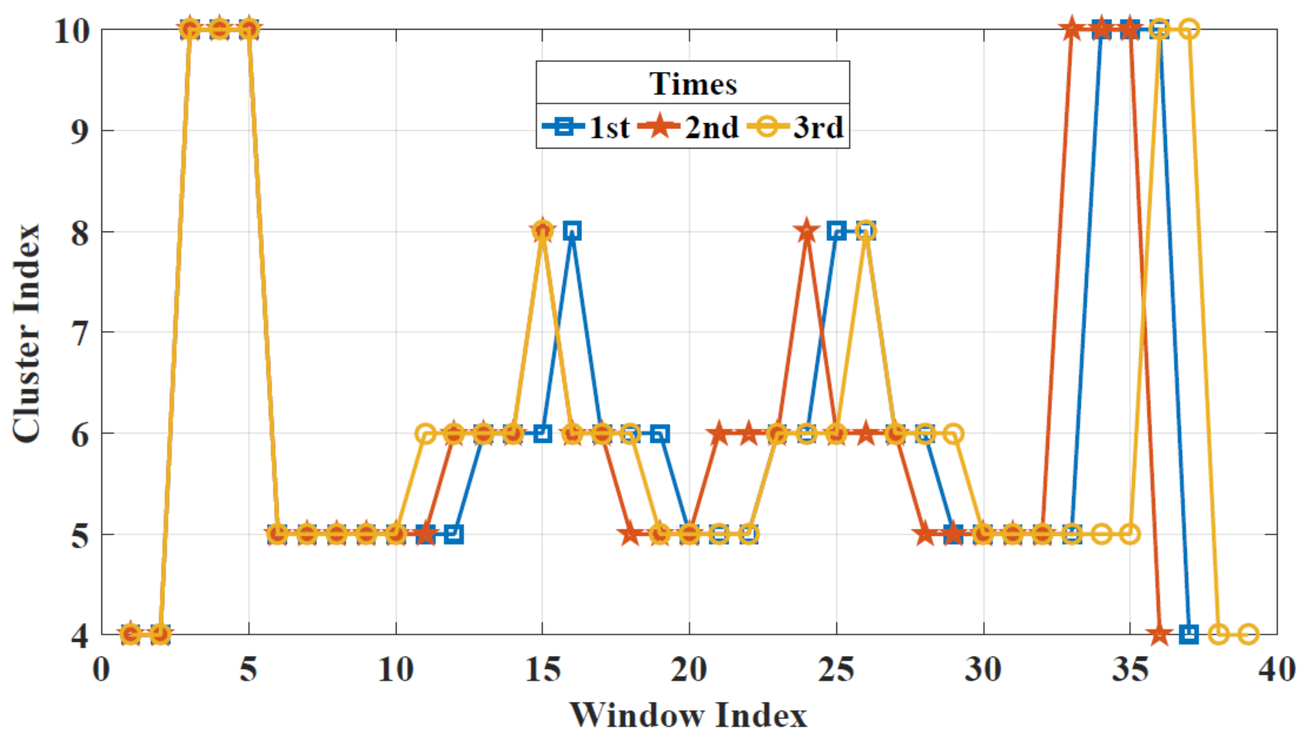


Gesture Recognizer

- **Edit Distance**

Calculate the distance between two strings in terms of the minimum number of edit operations (insertion, deletion and substitution) needed for transforming one to another.

$$D(i,j) = min[D(i-1,j)+p, D(i,j-1)+q, D(i-1,j-1)+r]$$



Gesture Recognizer

- **Edit Distance**

Calculate the distance between two strings in terms of the minimum number of edit operations (insertion, deletion and substitution) needed for transforming one to another.

$$D(i,j) \ = \ min[D(i-1,j)+p, D(i,j-1)+q, D(i-1,j-1)+r]$$

- **Improved Edit Distance**

1. Reduce the cost of edit operations to e(e < 1) when the current symbol is the same as the previous one.
2. The similarity between S1 and S2 :

$$T(S_1, S_2) = 1 - \frac{D(m,n)+C(m,n)}{\max(m,n)}$$

$$C(m,n) = \mathrm{m} - \sum_{i=1}^{m} 1_{S_2}\big(S_1(i)\big) + \mathrm{n} - \sum_{j=1}^{n} 1_{S_1}\big(S_2(j)\big)$$

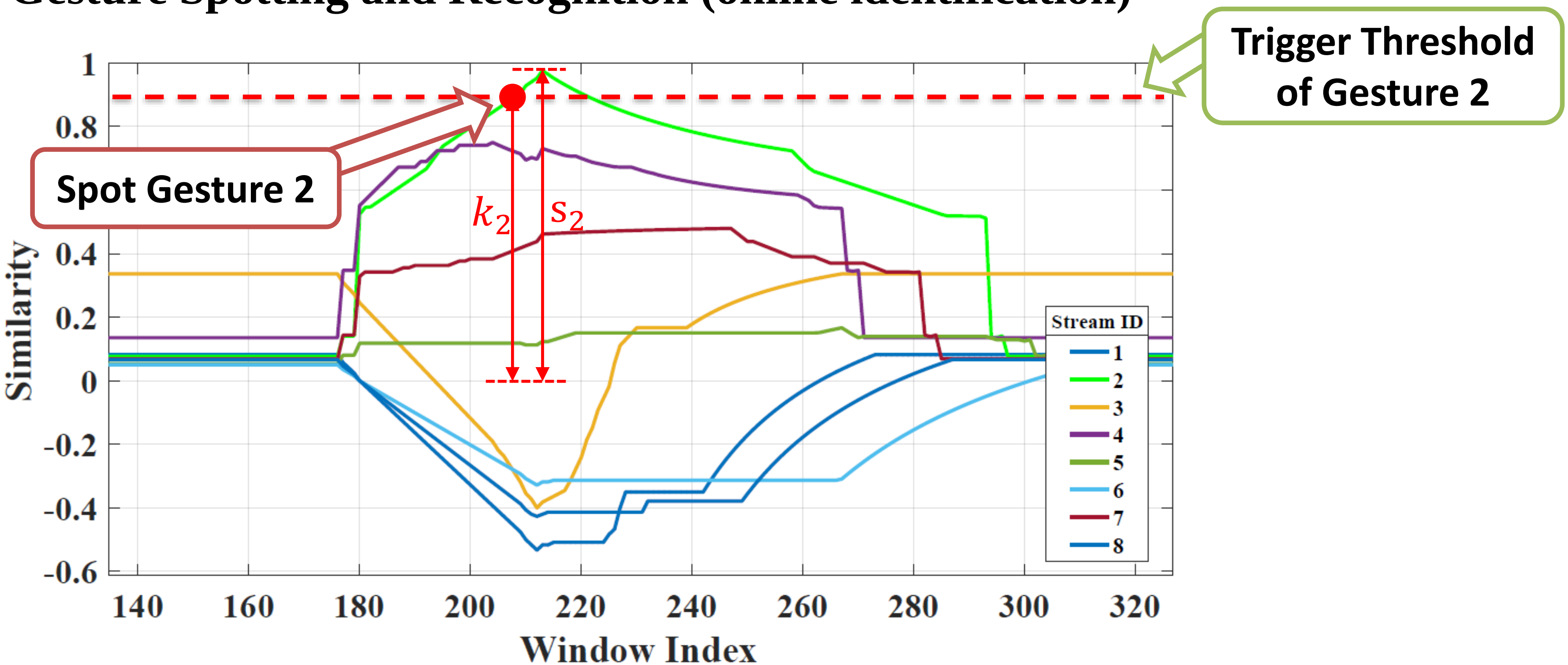
$$1_A(x) \coloneqq \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

Gesture Recognizer

- **Gesture Template Selector (offline training)**

  Choose a template for each sequential gesture and generate the corresponding trigger threshold k_g.

- **Gesture Spotting and Recognition (online identification)**



When collision happens, select the highest confidence level $cl_i = \frac{s_i}{k_i}$

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

4

PART FOUR

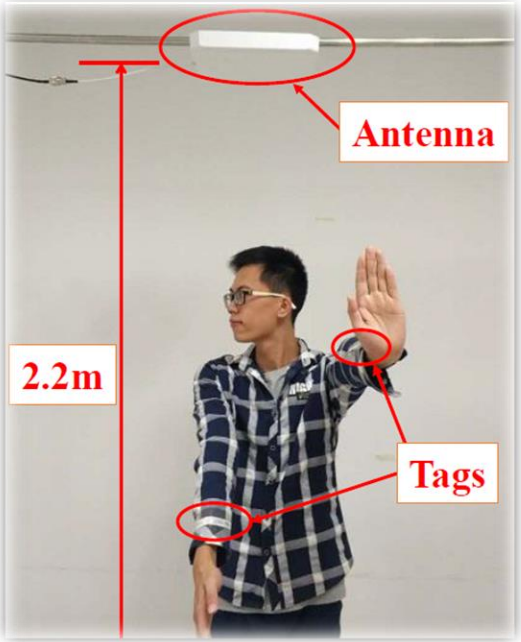
Implementation and Evaluation

Implementation

- **Hardware**

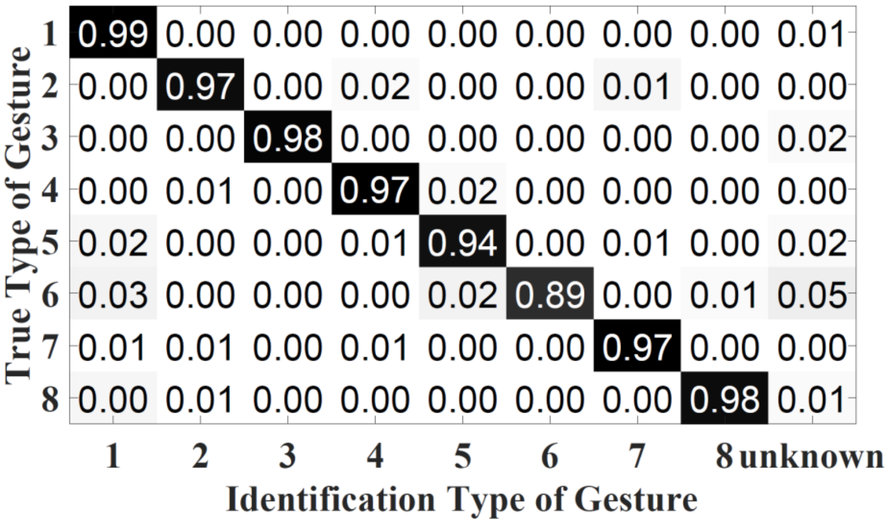
ImpinJ R420 reader	Alien AZ-9640 tags	Laird S9028PCR antenna

- **Setup**

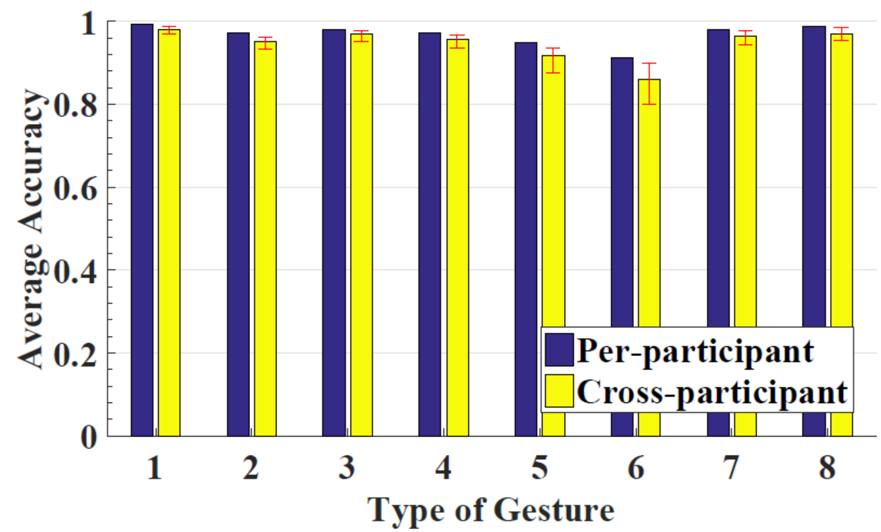


Evaluation

Recognition Accuracy Analysis



Effect of Suppressing Individual Diversity

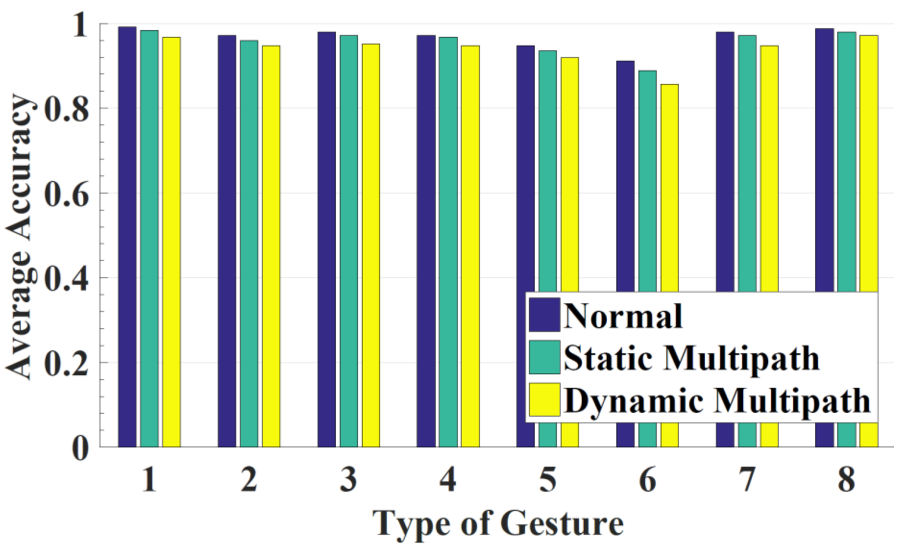


- The average recognition accuracy is 96.2%
- The accuracy of per-participant validation is 96.8%

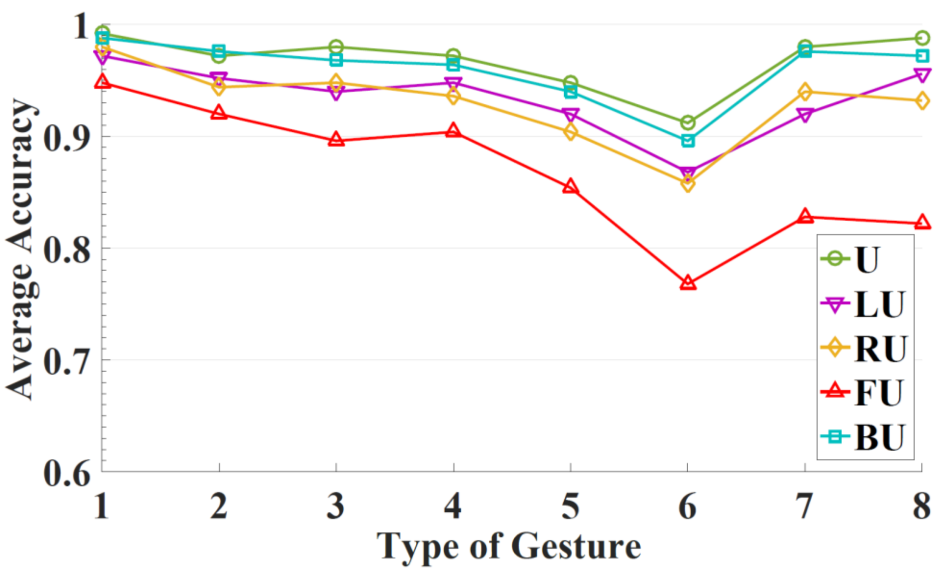
  The accuracy of cross-participant validation is 94.6%

Evaluation

Effect of Resisting Multipath



Effect of Diverse Positions



上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

5

PART FIVE

Conclusion

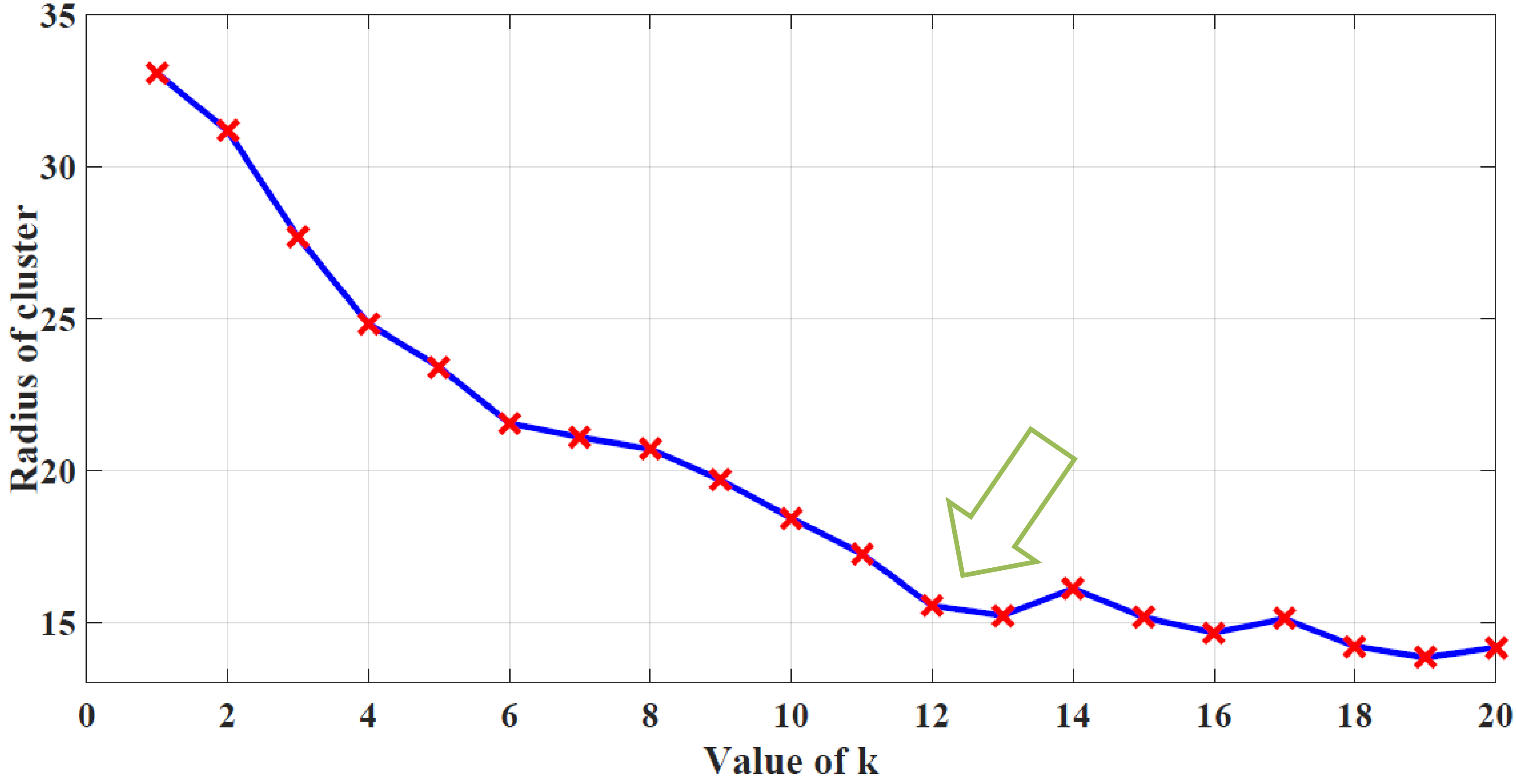
Conclusion

- **Propose the design, implementation and evaluation of SGRS, a sequential gesture recognition system**
  - **Leverage the RFID phase information to perceive the sequential gesture**
  - **Incorporate the vector quantization and improved edit distance to identify sequential gesture**
- **Implemented purely based on COTS RFID devices**
- **Achieve an average recognition accuracy of 96.2% and demonstrate the robustness and feasibility of SGRS**

Q & A

SGRS: A Sequential Gesture Recognition System using COTS RFID

K-means


35
30
25
20
15
Radius of cluster
0
2
4
6
8
10
12
14
16
18
20
Value of k