King Abdullah University
of Science and Technology

Angle-of-Arrival-Based Gesture Recognition Using Ultrasonic Multi-Frequency Signals

Authors: Hui Chen, Tarig Ballal, Mohamed Saad, Tareq Y. Al-Naffouri

Speaker: Hui Chen
University: King Abdullah University of Science and Technology (KAUST)

Human Gestures

King Abdullah University of Science and Technology

A form of non-verbal communication:

- Conveying information
- Expressing emotions
- Language for disabled people

Human-machine Interactive Products

King Abdullah University
of Science and Technology

Camera

(Microsoft Kinect)

Lasers & IMU

(HTC Vive)

Accelerometer

(Synertial Glove)

Ultrasound

(Ellipticlabs)

IR cameras and infrared LEDs

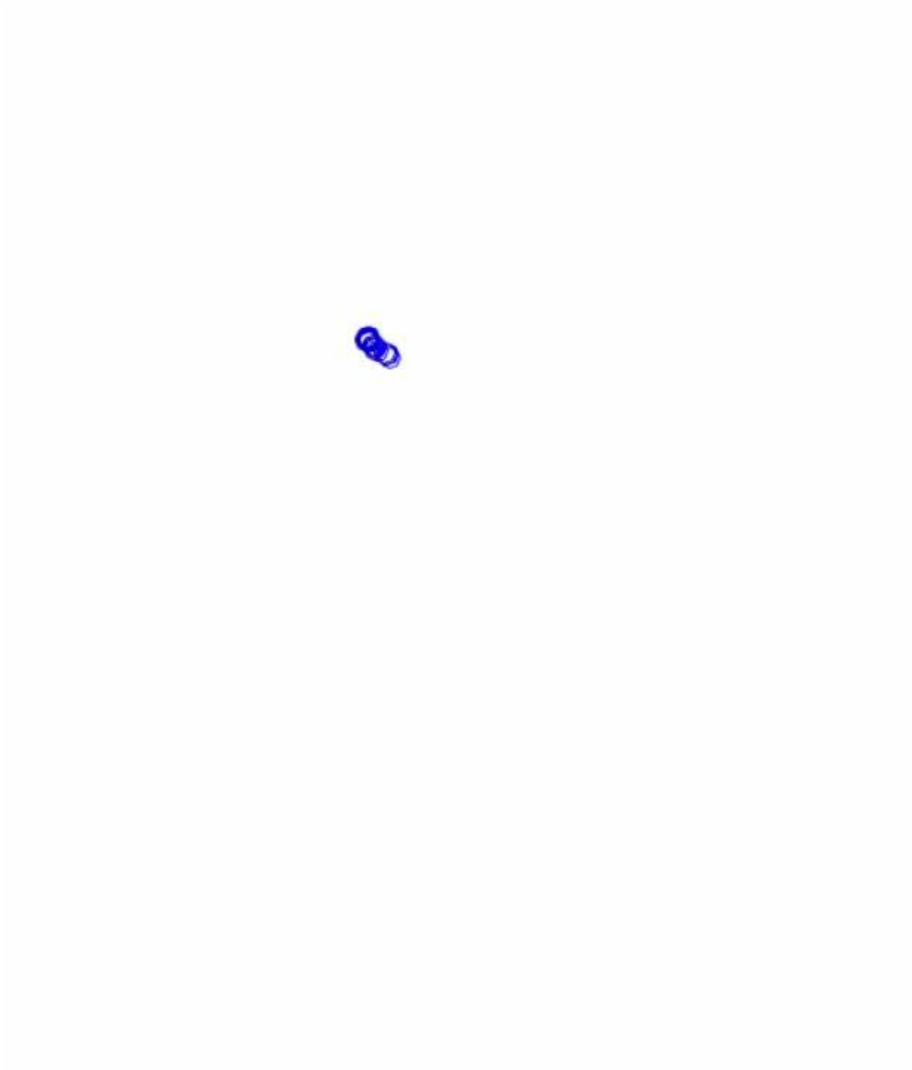
(Leap Motion)

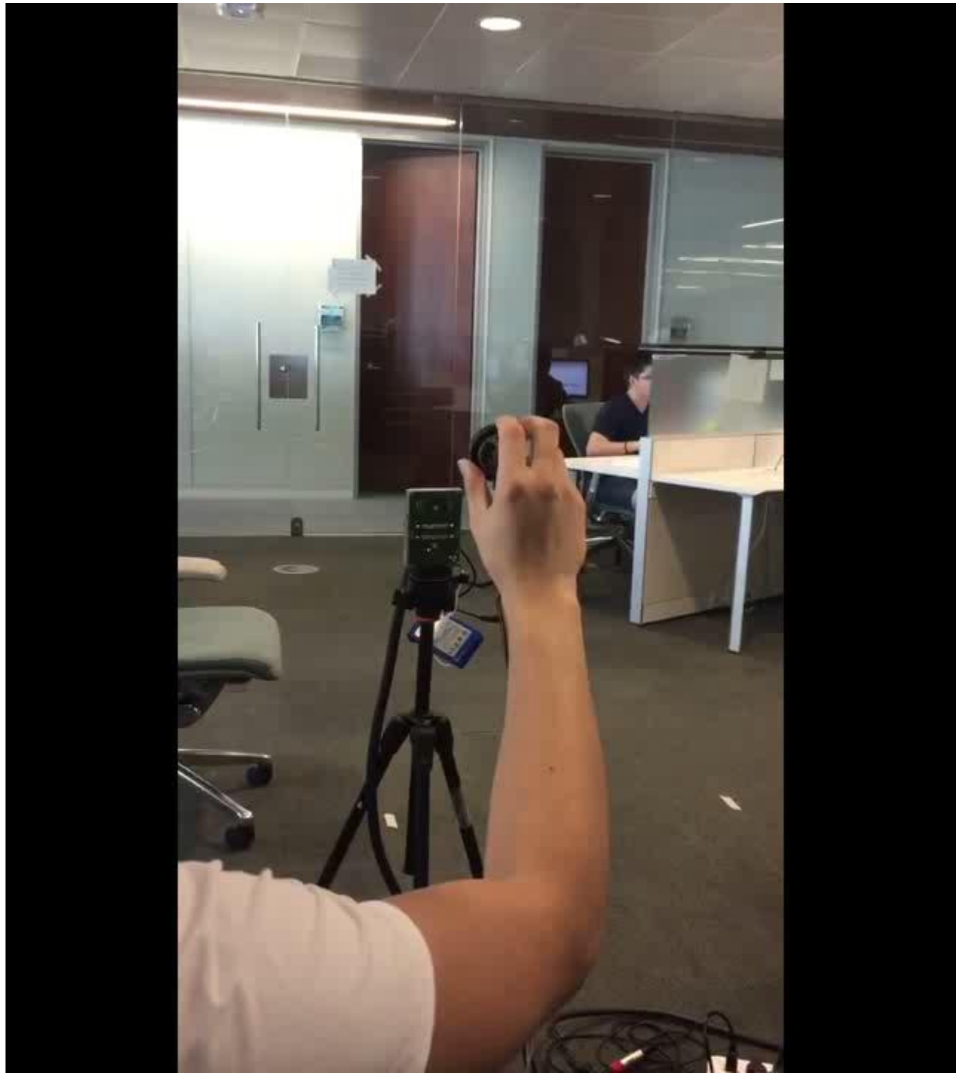
A Demo of Writing "KAUST"

Write 'K', 'A', 'U', 'S', 'T'
continuously within a window.

A Demo of Writing "KAUST"

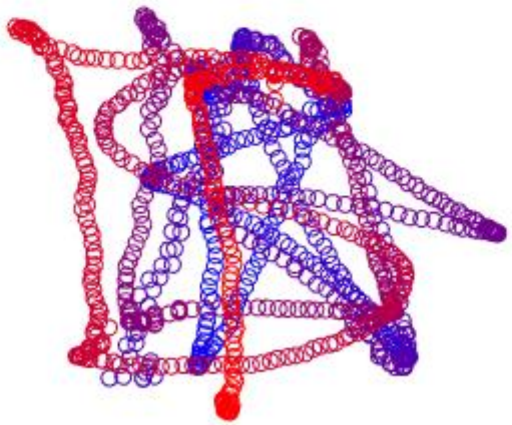
King Abdullah University
of Science and Technology





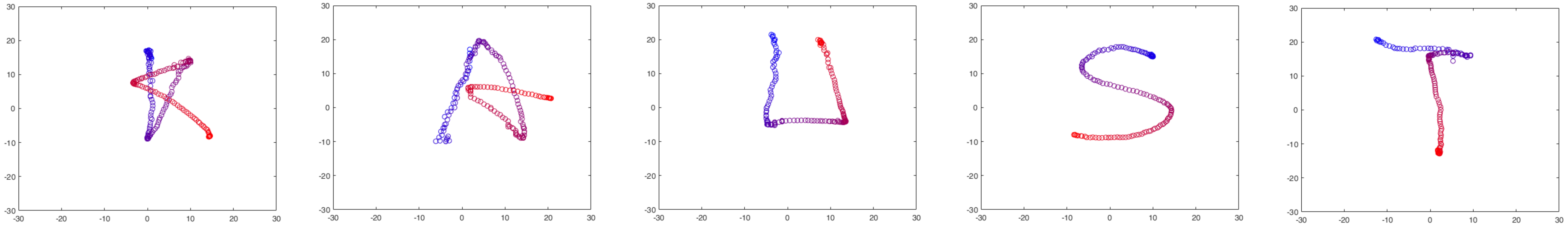
A Demo of Writing “KAUST”

King Abdullah University of Science and Technology



Captured hand motion with proposed system

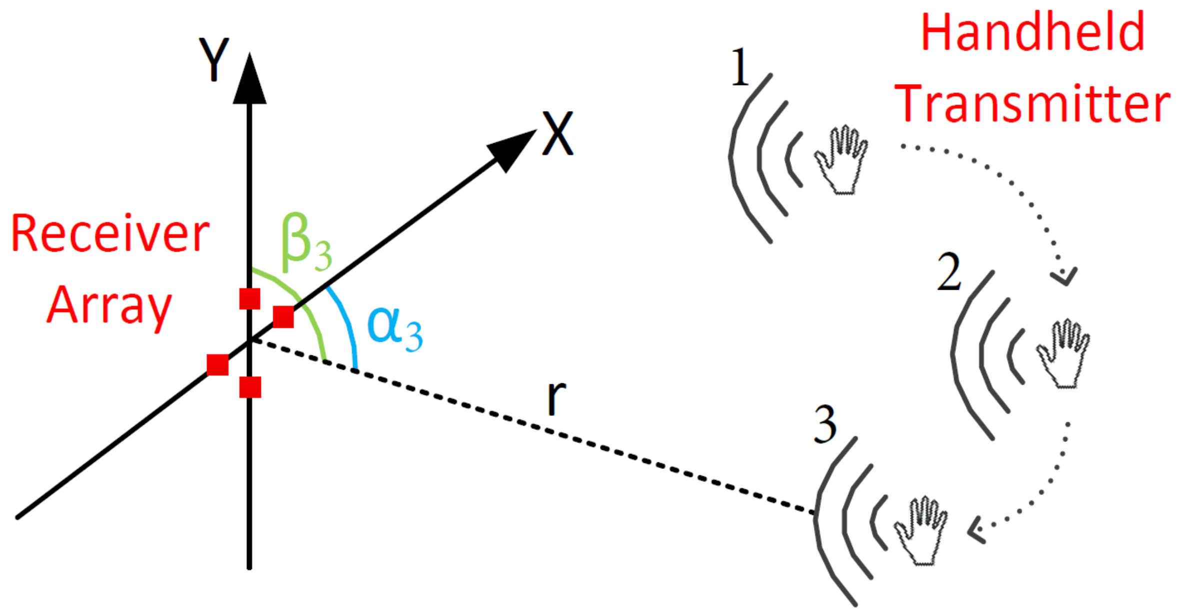
Start from blue to red

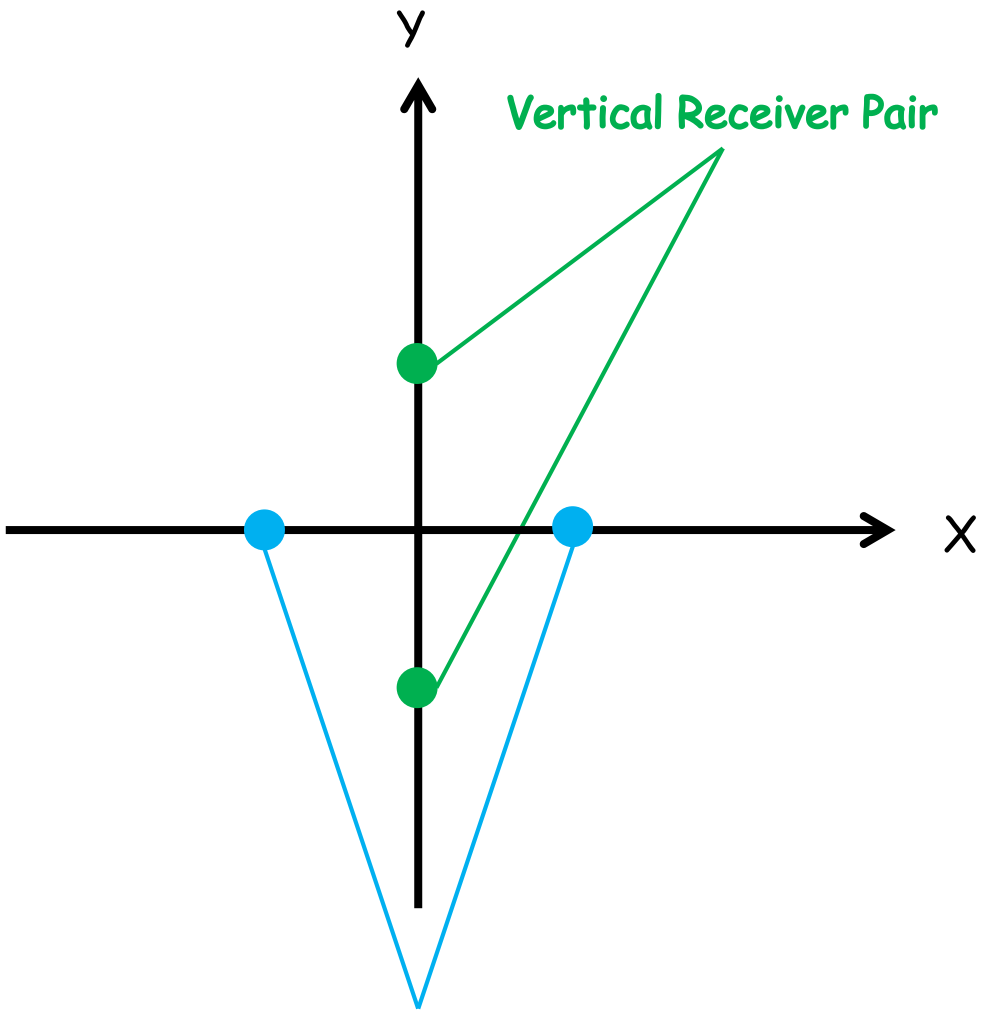


Extracted Segments

How it works?

Angle information → Classify





Proposed System

King Abdullah University of Science and Technology

1 AOA estimation:

Ultrasonic signals are transmitted from the transmitter. AoA algorithm is used to obtain the angle information for each sample.

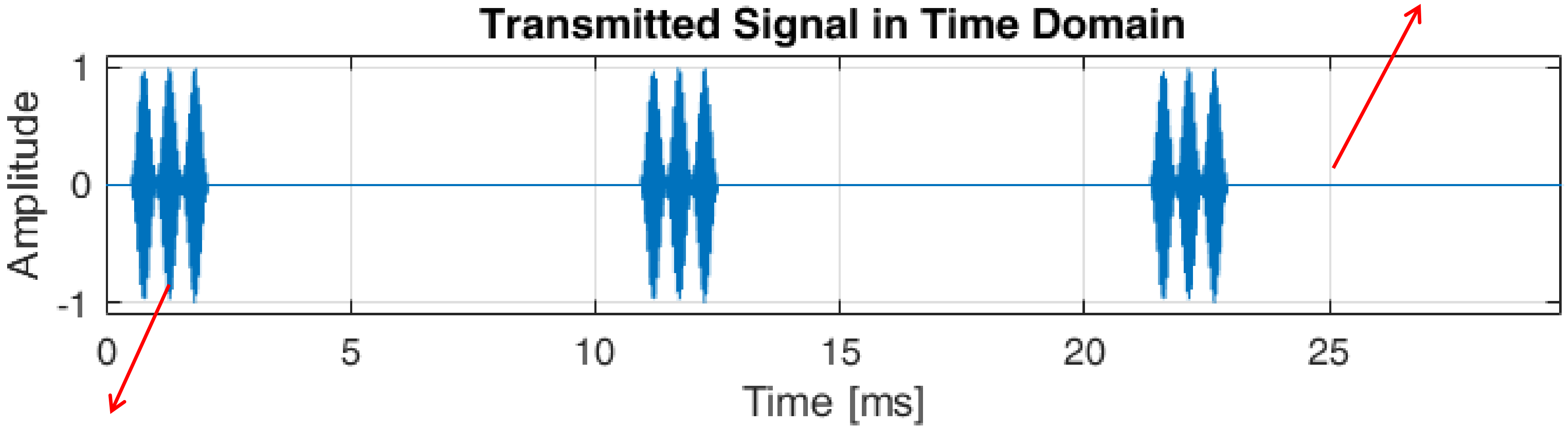
2 Classification:

Estimated angle data are processed and sent to classifier to recognition this gesture.

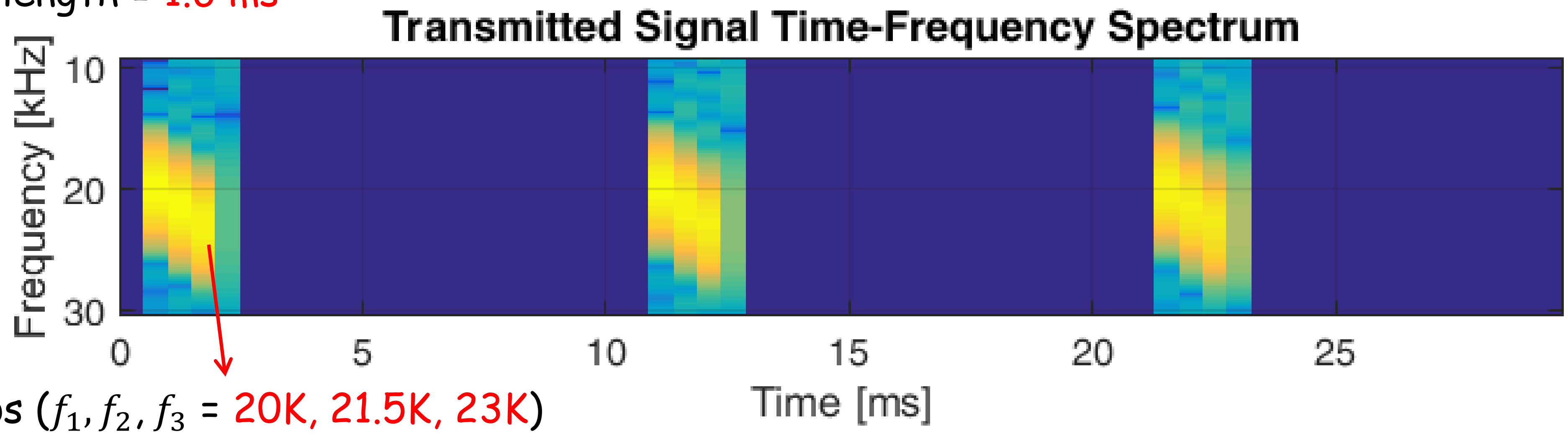
Signal Design

King Abdullah University of Science and Technology

Measurements/Frame per second: 100 (Hz)



Pulse length = 1.6 ms

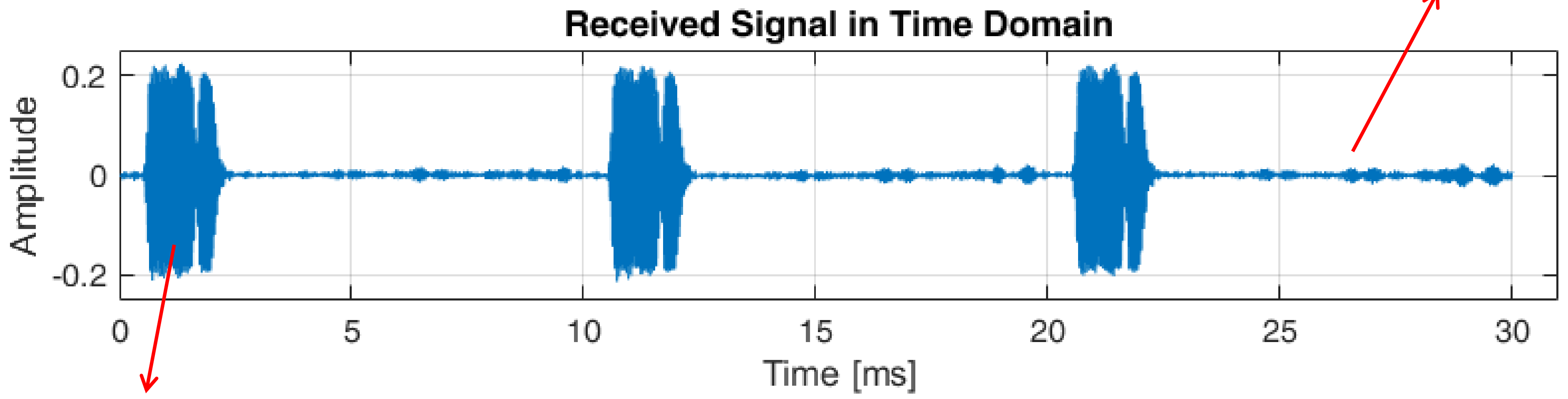


3 hops (f_1, f_2, f_3 = 20K, 21.5K, 23K)

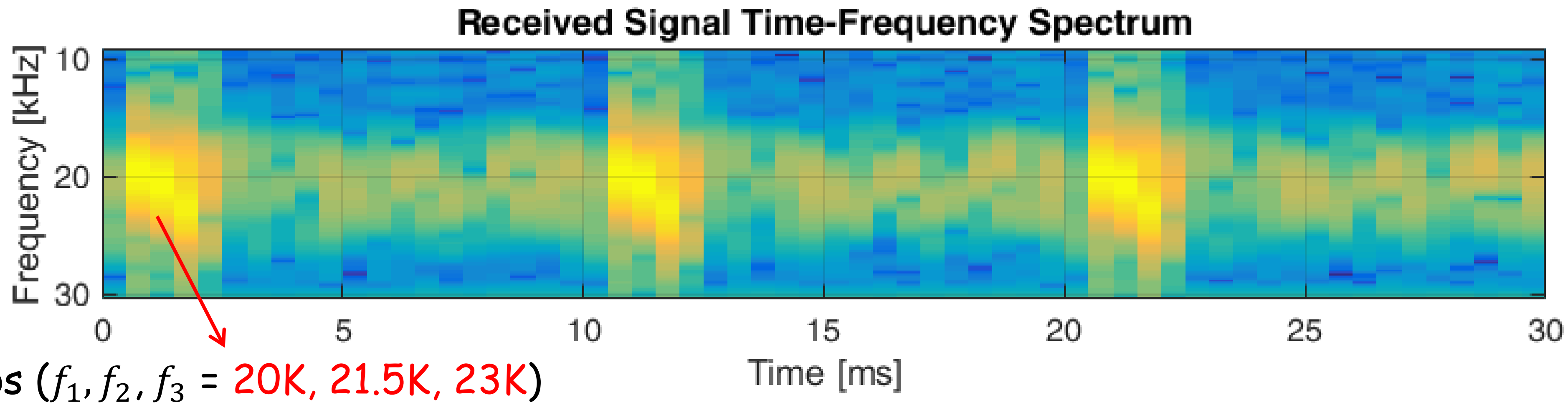
Signal Design

King Abdullah University of Science and Technology

Measurements/Frame per second: 100 (Hz)

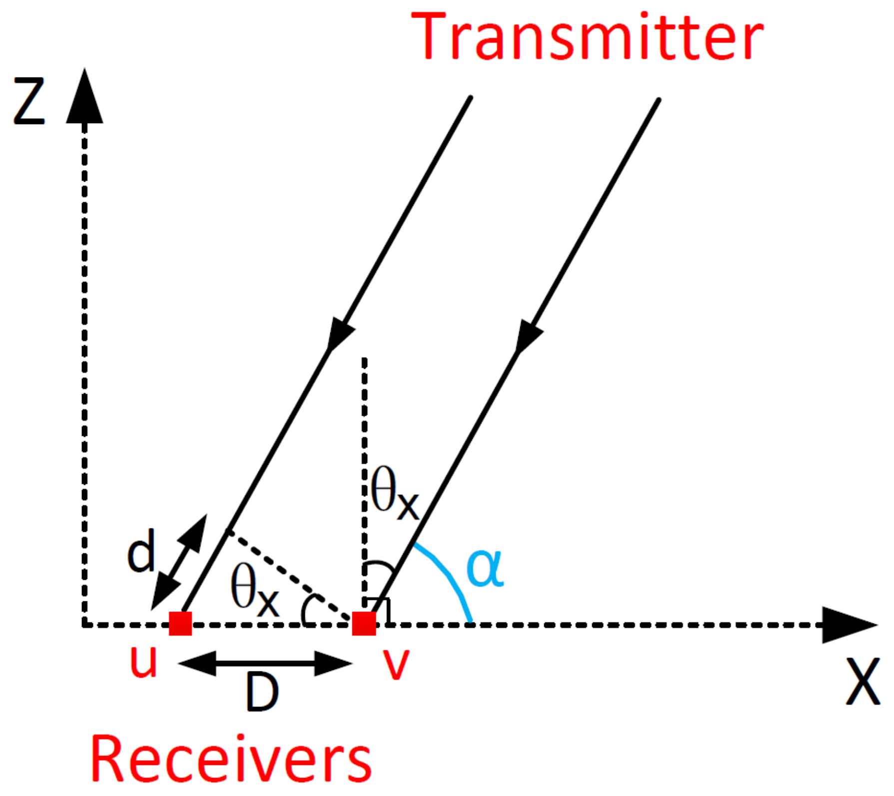


Pulse length = 1.6 ms



3 hops (f_1, f_2, f_3 = 20K, 21.5K, 23K)

Far-field Model of AOA



Far-field Model of
Horizontal Receivers

$$\boxed{sin\theta_x = \frac{d}{D} = \frac{\Delta\varphi}{2\pi f} \cdot \frac{c}{D}}$$

Given:

θ_x:	Horizontal angle
d:	difference of arrival
$\Delta\varphi$:	phase difference
F:	frequency of current signal
C:	Sound speed
D:	Sensor distance

How to get:

Horizontal angle θ_x and vertical angle θ_y

Proposed Angle of Arrival Estimation Algorithm

King Abdullah University of Science and Technology

1 Estimated phase difference $\hat{\psi}_{x,i}$ for carrier frequencies f_1, f_2, f_3 using cross power spectrum:

$$\boxed{\hat{\psi}_{x,i}} = \operatorname{ang}(Y_u(f_i) \cdot Y_v^*(f_i)) = \hat{\phi}_{x,i} - 2\pi N_{x,i}, \qquad (1)$$

$Y_u(f_i)$ is the DFT (Discrete Fourier Transform) of the received signals at sensor u;

$(\cdot)^*$ indicates the complex conjugate operation;

$\hat{\phi}_{x,i}$ is the real phase difference;

$N_{x,i}$ is an integer which makes estimated phase difference $\hat{\psi}_{x,i}$ within (-π, π];

2 Calculate real phase difference $\hat{\phi}_{x,i}$ for $\theta_x = \theta$ (eg: θ = -90°) for all the 3 frequencies:

$$\boxed{\frac{2\pi f_i D \sin(\theta)}{c}}$$

$$\sin(\hat{\theta}_x) = \frac{d}{D} = \frac{\boxed{\hat{\phi}_{x,i}}\, c}{2\pi f_i D}, \qquad (2)$$

Proposed Angle of Arrival Estimation Algorithm

King Abdullah University of Science and Technology

3 Obtain wrapped real phase difference $\tilde{\psi}_{x,i}$ from real phase difference $\hat{\phi}_{x,i}$;

$$\boxed{\tilde{\psi}_{x,i}(\theta)} = \text{wrap}(\tilde{\phi}_{x,i}(\theta)) = \text{wrap}\left(\boxed{\frac{2\pi f_i D \sin(\theta)}{c}}\right), \qquad (3)$$

4 Calculate the error between estimated phase difference and at all frequencies (i = 1,2,3);

$$\sum_{<i>} (|\hat{\psi}_{x,i} - \tilde{\psi}_{x,i}(\theta)|)$$

5 Repeat 2,3 for all the angles θ_x = -90 : 1 : 90 and find the θ gives min error as the estimated angle.

$$\boxed{\hat{\theta}_x} = \arg\min_{\theta} \sum_{<i>} (|\hat{\psi}_{x,i} - \tilde{\psi}_{x,i}(\theta)|), \qquad (4)$$

AOA Estimation Summary

King Abdullah University of Science and Technology

1 Estimated phase difference $\hat{\psi}_{x,i}$

$$\boxed{\hat{\psi}_{x,i}} = \mathrm{ang}(Y_u(f_i) \cdot Y_v^*(f_i)) = \hat{\phi}_{x,i} - 2\pi N_{x,i}$$

2 Calculate real phase difference $\hat{\phi}_{x,i}$ for θ = -90° : 1 : 90 at all the 3 frequencies:

3 Obtain wrapped real phase difference $\tilde{\psi}_{x,i}$

$$\boxed{\tilde{\psi}_{x,i}(\theta)} = \mathrm{wrap}(\tilde{\phi}_{x,i}(\theta)) = \mathrm{wrap}\left(\boxed{\frac{2\pi f_i D \sin(\theta)}{c}}\right)$$

4 Obtain absolute error (difference) summation of $\hat{\psi}_{x,i}$ and $\tilde{\psi}_{x,i}$;

5 Find the θ gives min error as the estimated angle.

$$\boxed{\hat{\theta}_x} = \arg\min_{\theta} \sum_{<i>} (|\hat{\psi}_{x,i} - \tilde{\psi}_{x,i}(\theta)|)$$

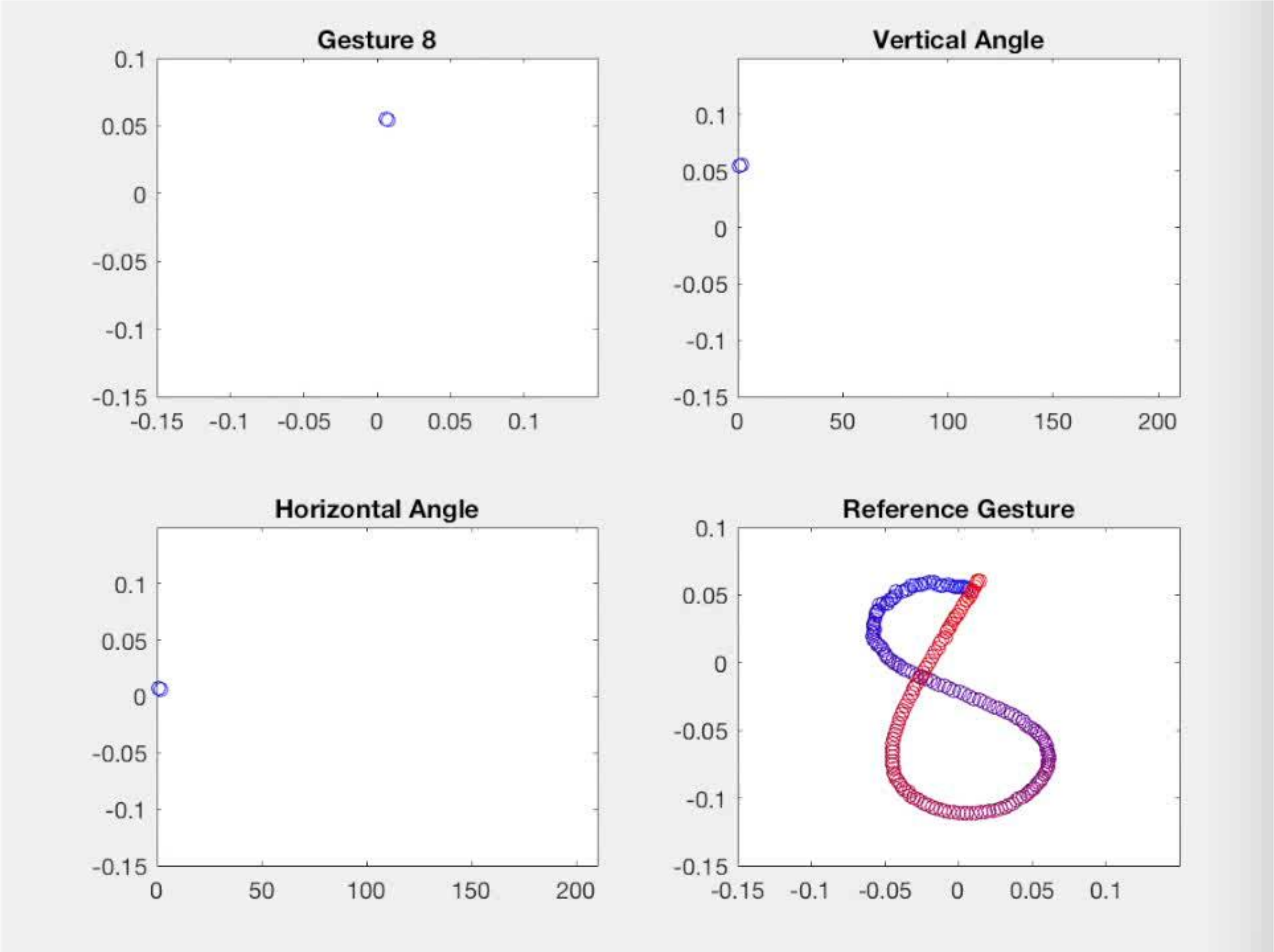
Classifiers

King Abdullah University of Science and Technology

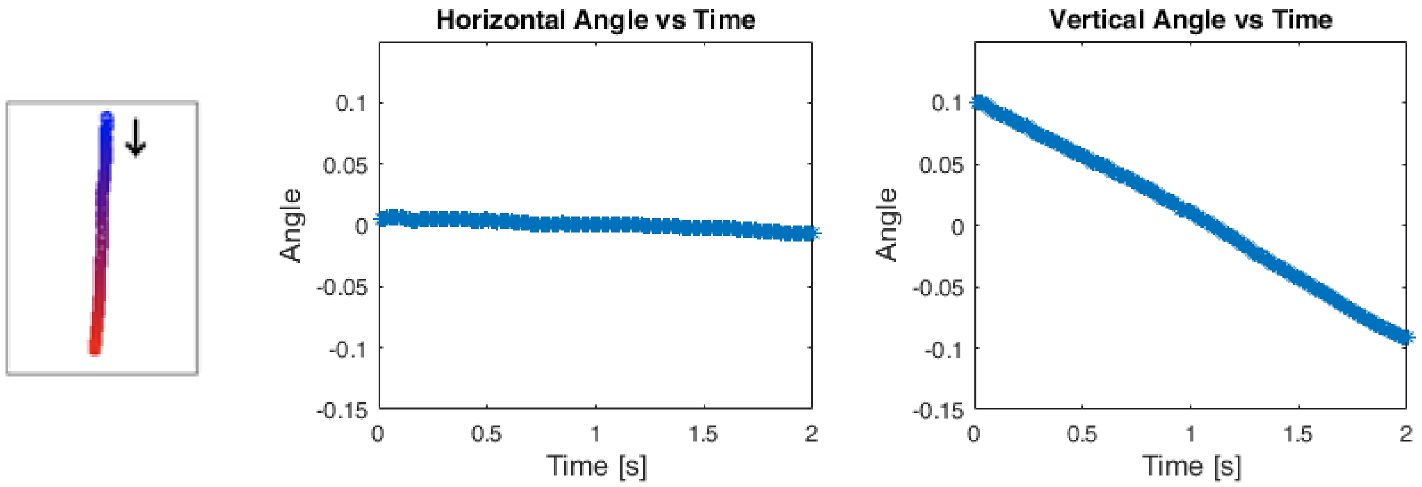
Proposed classifier	Benchmark Classifier
Redundant Dictionary Matching; (RD classifier) ▪ No training procedure;	Neural Network Model Trained by: (NN classifiers) ▪ Train the model with angle data ▪ Train the model with constructed image ▪ MNIST database

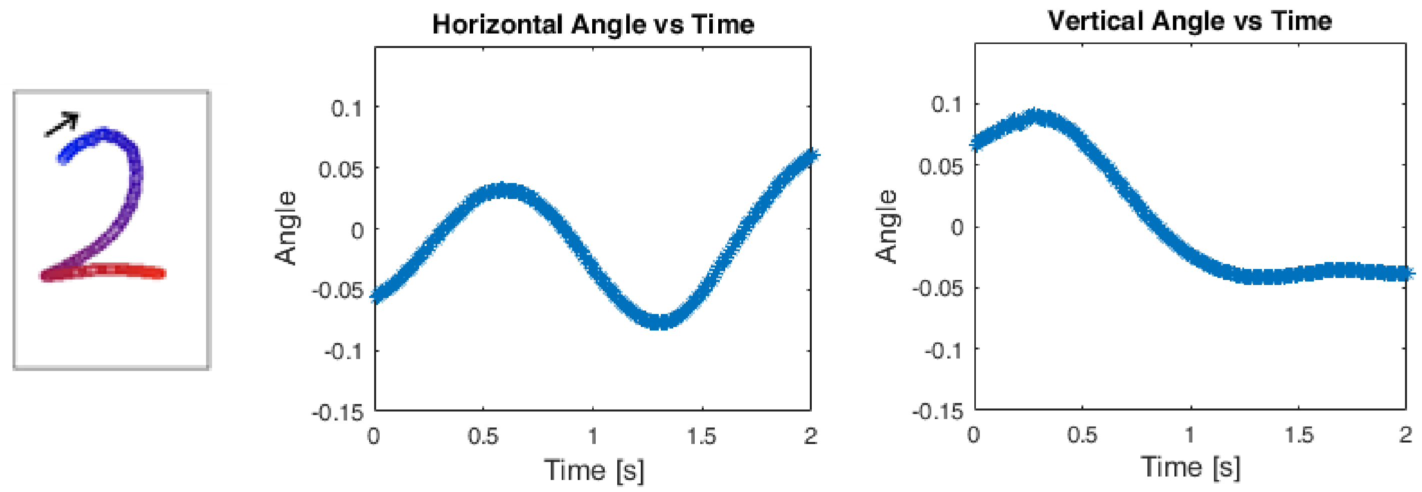
4 classifiers

A Demo of Angle vs. Time



Template Dictionary





Same applies for 3, 4, 5 ~ 9, 0

Horizontal angle vector

+

Vertical angle vector

+

Normalize

=

Template dictionary

Template Dictionary

King Abdullah University of Science and Technology

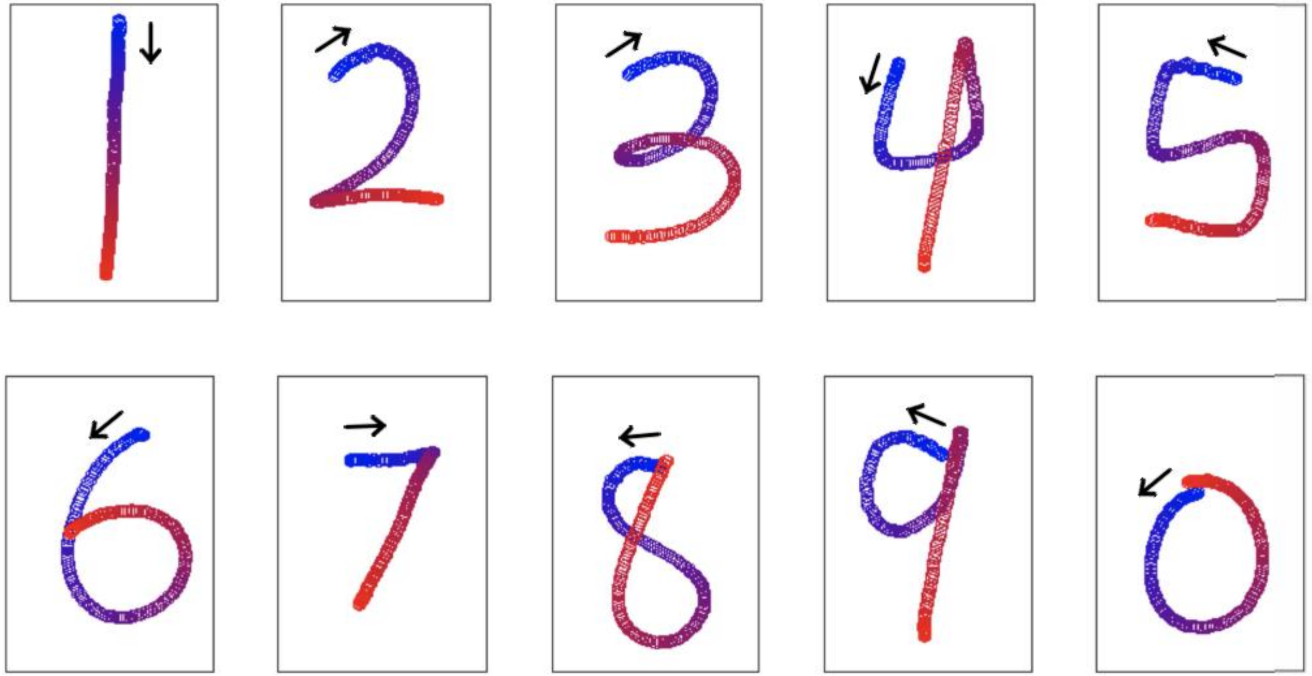
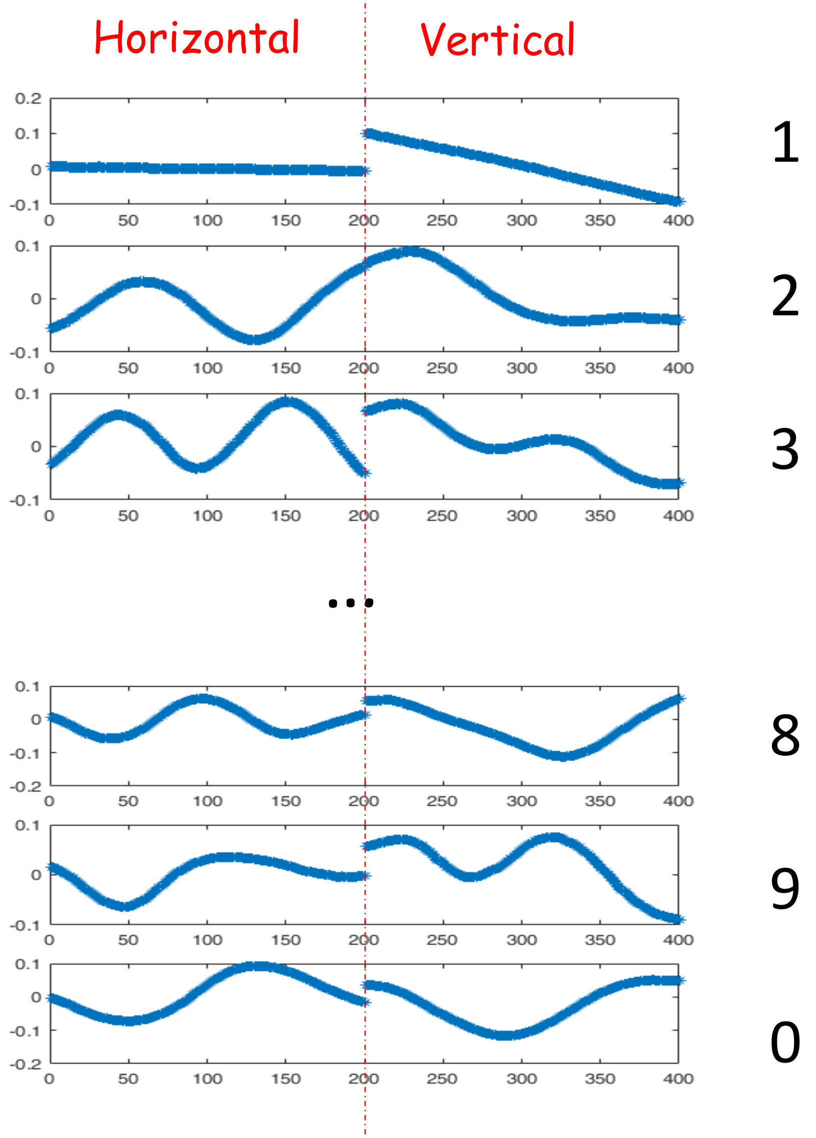
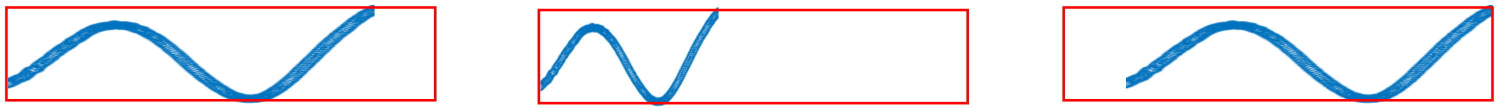


Fig. 4. Templates of the Gestures



Redundant Dictionary

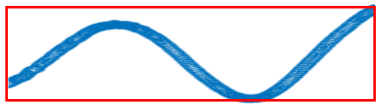
Users have different habits:



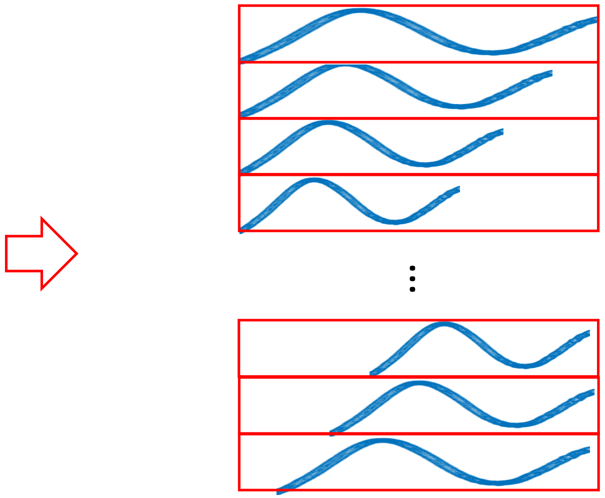
Slow | Faster | Slow delayed

Writing Speed → Shrink

Delay → Shift



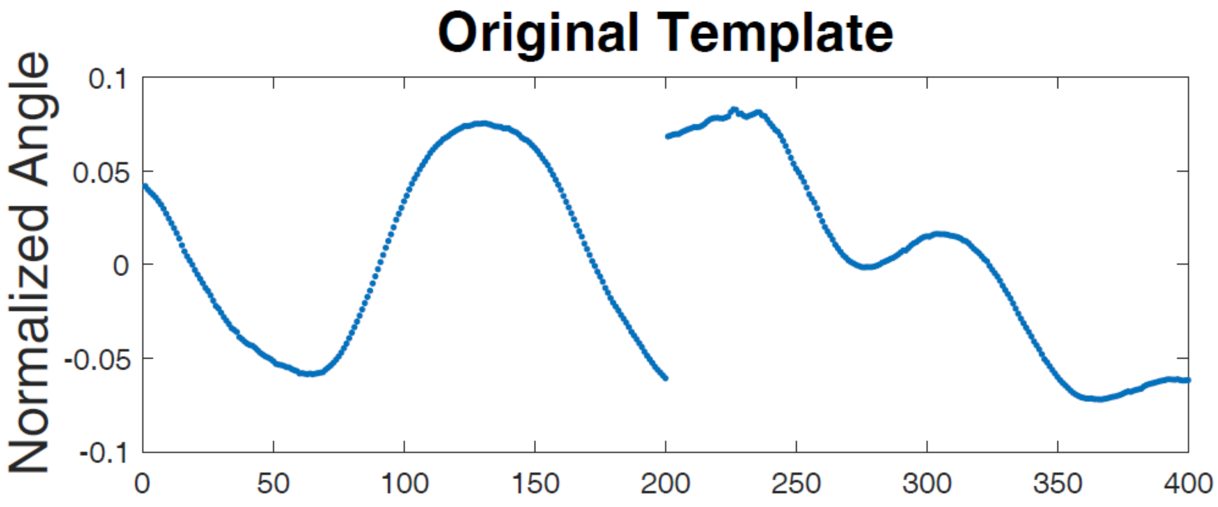
Template dictionary

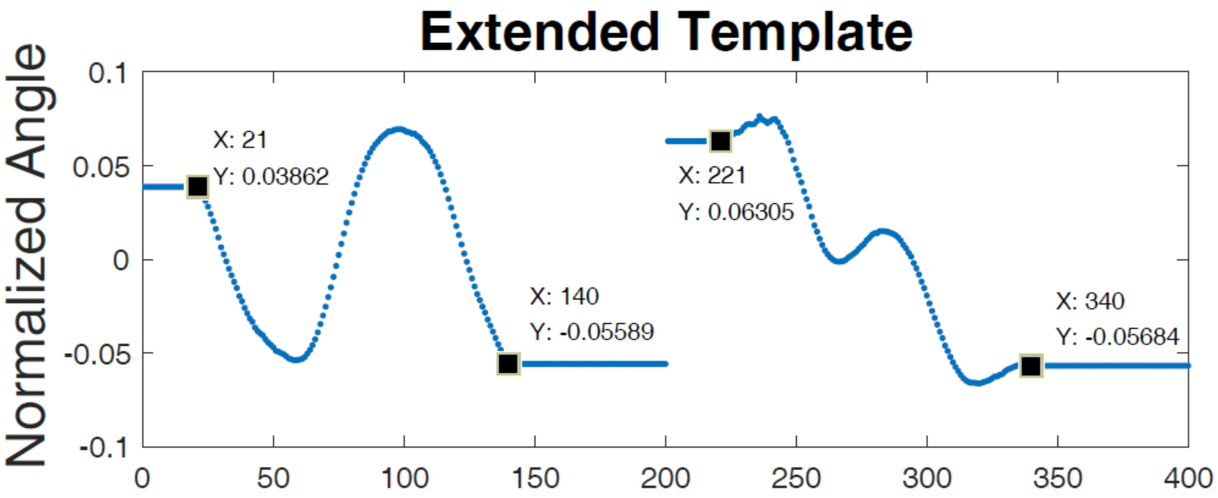


Redundant Dictionary

Redundant Dictionary—One Example

King Abdullah University of Science and Technology





Shrink sample = 20; Shift sample = 20;

200 samples → 1 start point

180 samples → 2 start points

...

100 samples → 6 start points

Extend to 21 templates

Dictionary A_r: 400 by 210

RD Classifier—Classify

King Abdullah University of Science and Technology

$$r = A_r^T g. \tag{6}$$

r: vector of inner products;

A_r: redundant dictionary matrix;

g: normalized angle information using linear interpolation to fit the size of the dictionary.

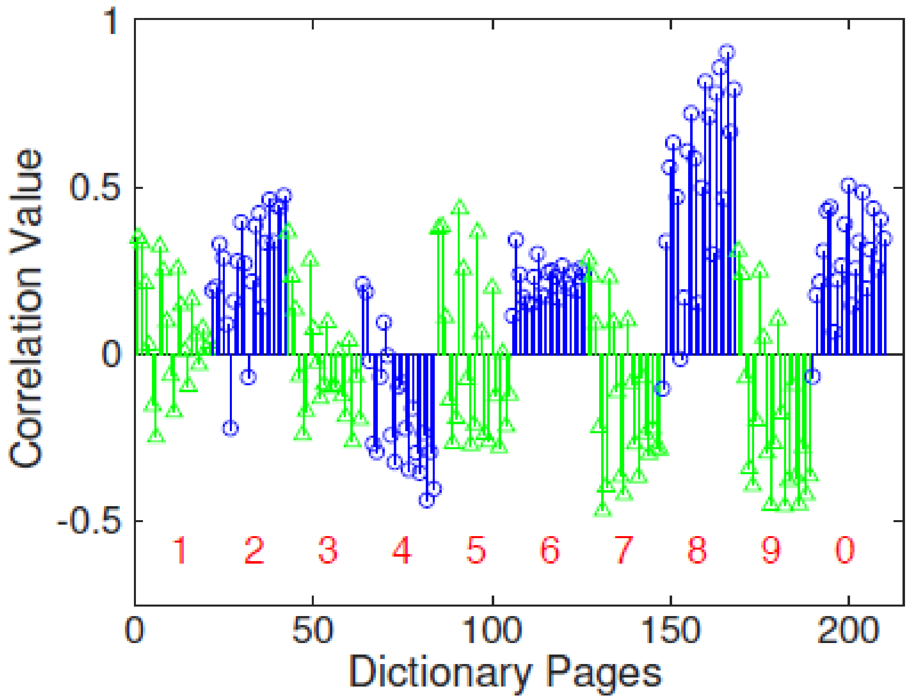
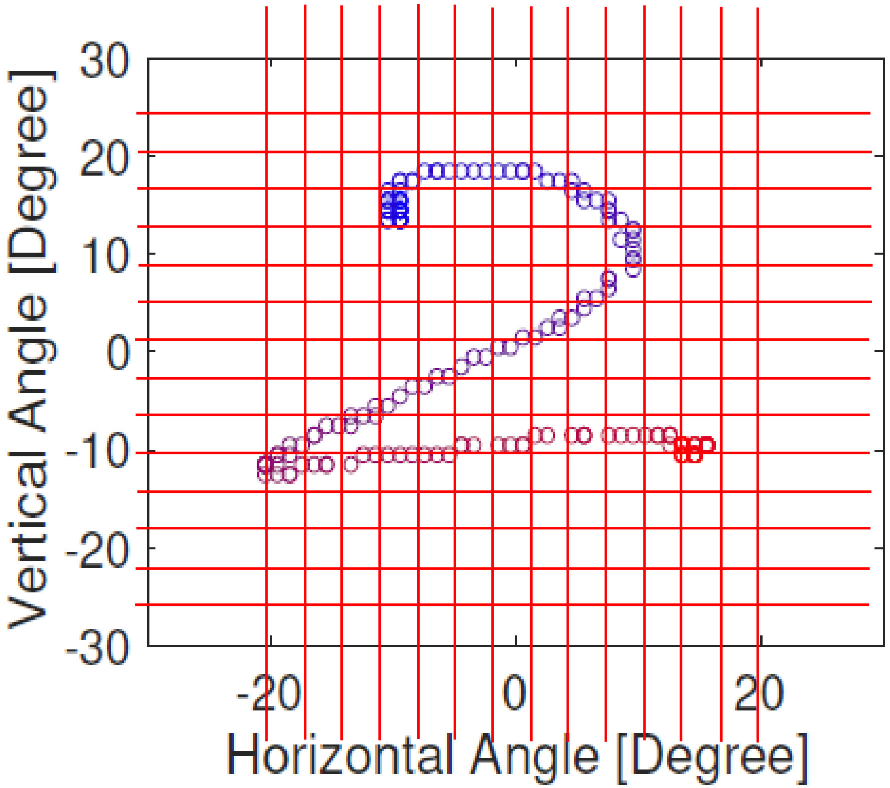


Fig. 9. An Example of Dictionary Matching for Gesture '8'

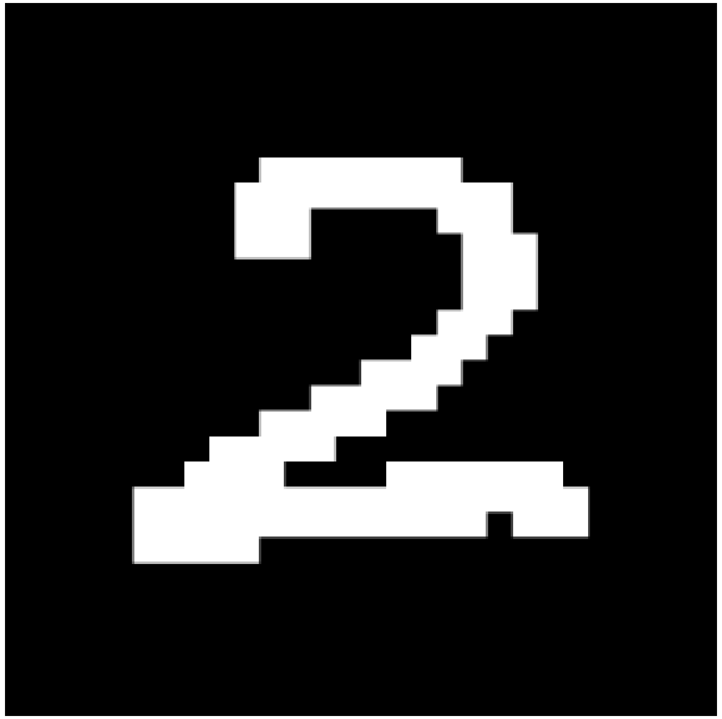
NN Classifier—Image Constructing

King Abdullah University of Science and Technology

Grid into a 28 by 28 matrix



(a) Gesture '2'



(b) Binarized Image

NN Classifier—training data sets

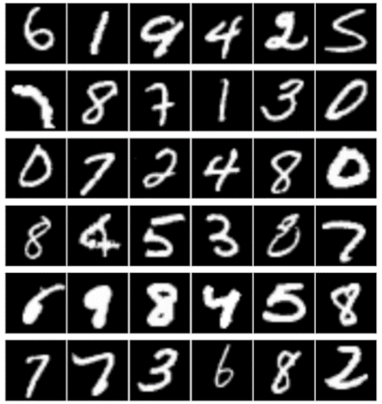
Model: Stacked auto-encoder model from Matlab Neural Network Toolbox.

Training data sets:

1 Train the model with angle data; (1x400 vector)

2 Binarized MNIST database; (28x28 image)

3 Train the model with constructed images; (28x28 image)



MNIST database

System Setup

Handhold ultrasound transmitter

$\downarrow \vec{Z} \rightarrow$

Perform gesture

Receiver array

Computer

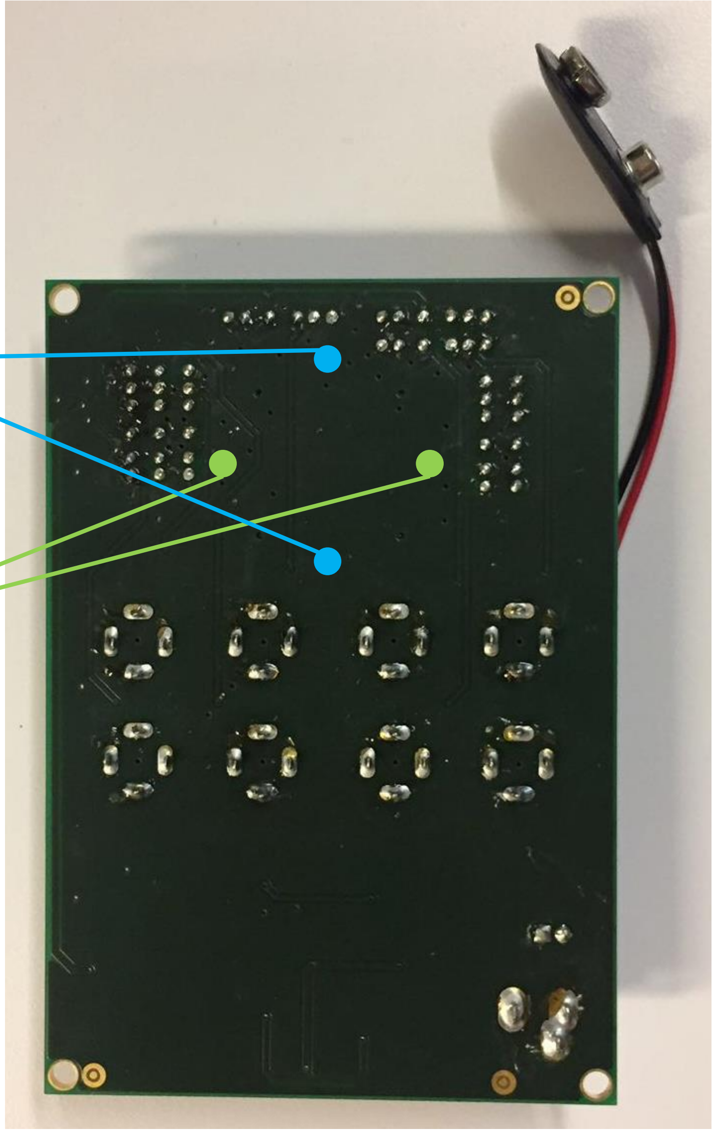
Classify the gesture

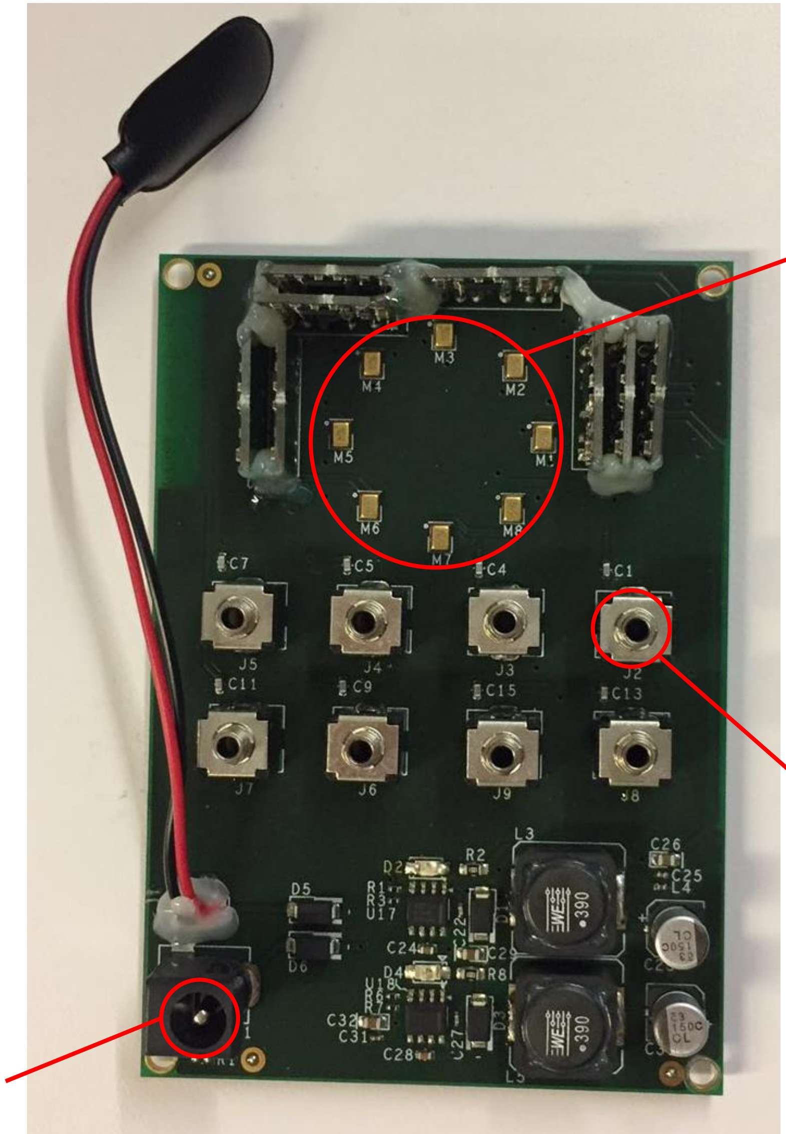
Receiver Array

King Abdullah University of Science and Technology

Vertical pair

Horizontal pair





Circular Receiver Array

Connector

Power

Angle Accuracy Test Result

King Abdullah University of Science and Technology

- Receiver is placed 1 m from the ground;
- Data is collected from 6 points with Vertical angle = 0 (transmitter is the same height with the receiver)

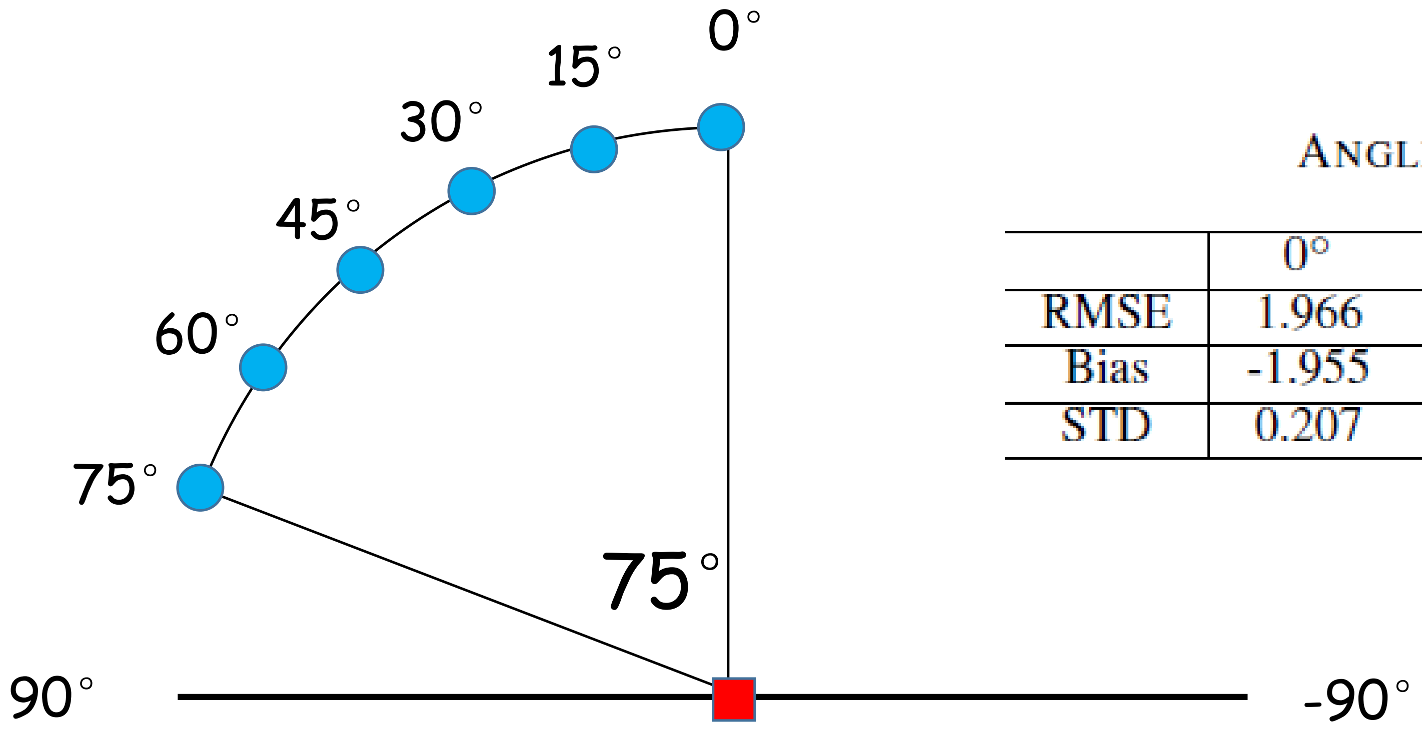
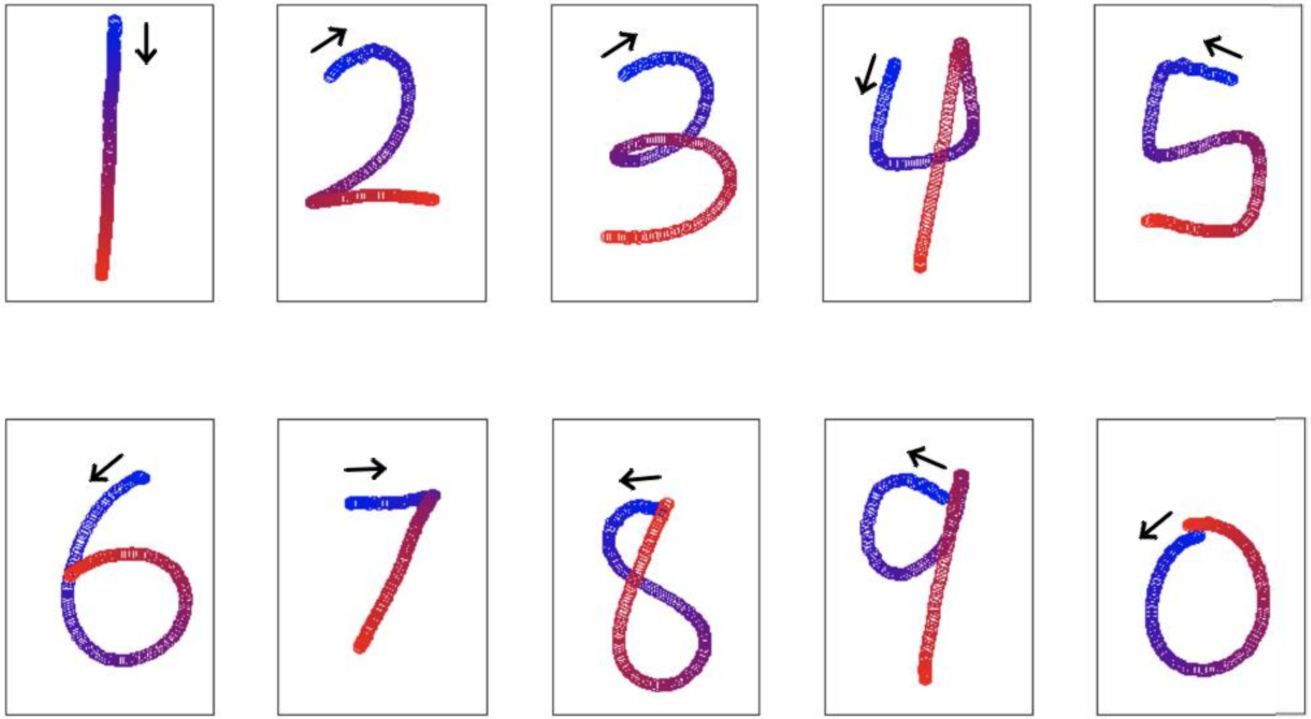


TABLE I
ANGLE OF ARRIVAL ACCURACY TEST

	0°	15°	30°	45°	60°	75°
RMSE	1.966	1.925	1.691	2.520	0.604	1.102
Bias	-1.955	-1.895	-1.620	-2.470	-0.145	0.155
STD	0.207	0.339	0.487	0.500	0.588	1.094

Classifier Performance

King Abdullah University of Science and Technology



Gesture Templates

Data collection:

10 volunteers

10 gestures

Repeated for 10 times

1000 gesture data set

Classifier Performance

King Abdullah University of Science and Technology

Classifier	Accuracy
Redundant Dictionary Matching	95.5%
Neural Network Model	
▪ Train the model with angle data	94.4%
▪ MNIST database	66.5%
▪ Train the model with constructed image	91.1%

Result of Dictionary-based Classifier

TABLE II
CONFUSION MATRIX OF DICTIONARY-BASED CLASSIFIER

Actual Gesture	Classified Gestures (95.5% of 1000 gestures)									
	‘1’	‘2’	‘3’	‘4’	‘5’	‘6’	‘7’	‘8’	‘9’	‘0’
‘1’	96						3	1		
‘2’	4	91					5			
‘3’	5	1	91				3			
‘4’				81		5			9	5
‘5’	3				95			1	1	
‘6’						97				3
‘7’	9						91			
‘8’							1	99		
‘9’					1		1		97	1
‘0’						1				99

Summary

King Abdullah University of Science and Technology

1 A prototype of gesture recognition system;

2 AOA algorithm to obtain 2D angle information from the gesture;

3 Training-free Redundant dictionary classifier;

4 95.5% accuracy.

Current & Future Work

King Abdullah University
of Science and Technology

1 Letter Recognition (√)

2 Hidden Markov Model Classifier (√)

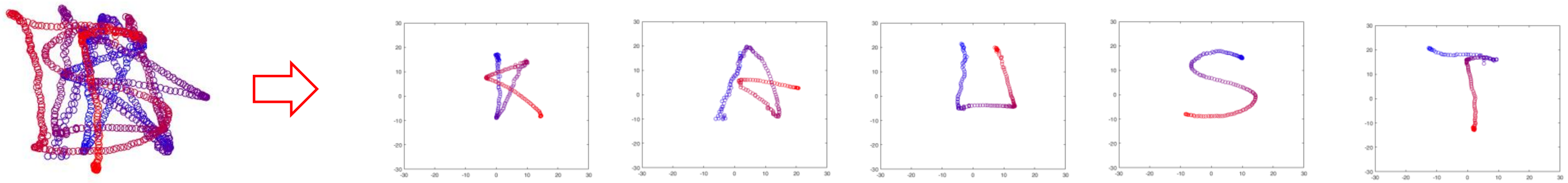
3 Decision Tree (√)

4 Word Recognition

Recent Letter Classification Results

1 With decision tree
- Redundant Dictionary: 96.09%
- Neural network: 95.45%
- Hidden Markov Model: 96.22%

2 Without decision tree
- Redundant Dictionary: 89.39%
- Neural network: 94.29%
- Hidden Markov Model: 95.45%



King Abdullah University
of Science and Technology

Thanks for your time

Speaker: Hui Chen
Email: hui.chen@kaust.edu.sa
Date: 2017-08-25