Thesis for Master’s Degree

Digital Walkie-Talkie Identification scheme based on Sparse Representation with Multiple features

Kiwon Yang

School of Electrical Engineering and Computer Science

Gwangju Institute of Science and Technology

2018

Digital Walkie-Talkie Identification scheme based on Sparse Representation with Multiple features

다중 특징 벡터를 이용한 희소 표현 기반 디지털 무전기 분류체계

Digital Walkie-Talkie Identification scheme based on Sparse Representation with Multiple features

Advisor : Professor Heung-No Lee

By

Kiwon Yang

Graduate: School of Electrical Engineering and Computer Science

Gwangju Institute of Science and Technology

A thesis submitted to the faculty of the Gwangju Institute of Science and Technology in partial fulfillment of the requirements for the degree of Master of Science in the School of Electrical Engineering and Computer Science

Gwangju, Republic of Korea

2018. 12. 26.

Approved by

Professor Heung-No Lee

Committee Chair

Digital Walkie-Talkie Identification scheme based on Sparse Representation with Multiple features

Kiwon Yang

Accepted in partial fulfillment of the requirements for
the degree of Master of science

December 26th 2018

Committee Chair ____________

Prof. Heung-No Lee

Committee Member ____________

Prof. Kiseon Kim

Committee Member ____________

Prof. Byeong Ha Lee

MS/EC
20171079

Kiwon Yang (양기원). Digital Walkie-Talkie Identification Scheme based on Sparse Representation with Multiple features (다중 특징 벡터를 이용한 희소 표현 기반 디지털 무전기 분류체계). School of Electrical Engineering and Computer Science. 2018. 34p. Prof. Heung-No Lee

Abstract

Radio frequency fingerprinting is important in electronic warfare or internet of things network for the security. It gives allies useful information on enemy forces or prevents access of malicious nodes to the system by identifying if the signal is from the specific transmitter. In the past, RF fingerprinting was studied as searching for a unique feature or a simple classification method only. As typical signal parts for RF fingerprinting, the transient signal, the sync signal, and the error in the In-phase/Quadrature domain have been used with the machine learning algorithm such as k-nearest neighbors or support vector machine. From the last year, RF fingerprinting method with multiple features has been researched. However, the studies showing the difference of schemes on using unique feature and multiple features, have not been published. Also, there is no research on RF fingerprinting, which sparse representation based classification (SRC) algorithm is applied to, with multiple features. SRC is a qualified algorithm for the image or signal classification. In this paper, we suggest an effective RF fingerprinting method. The proposed method is to use SRC with multiple features in a signal burst. First, we used the rising transient signal, the falling transient signal, and the sync signal. Then, main lobe is extracted from the each signal. The multiple features are concatenated to use them simultaneously. We show that the classification result of our method recorded accuracy rate over 98% with three features and the result is higher, compared with convolutional neural network method.

Contents

List of Figure

List of Tables

1. Introduction

Classifying radio frequency (RF) transmitters can be useful in the electronic warfare [1] or internet of things (IOT) network [2]. For example, in the electronic warfare, information of enemy's transmitters can be obtained from the classification technique on transmitters. Using the information, efficient tactics can be established. In IOT network, the classification scheme can be employed to prevent the access of counterfeit node to the network. Fig. 1 shows the clear example of the transmitter classification scheme in IOT network.

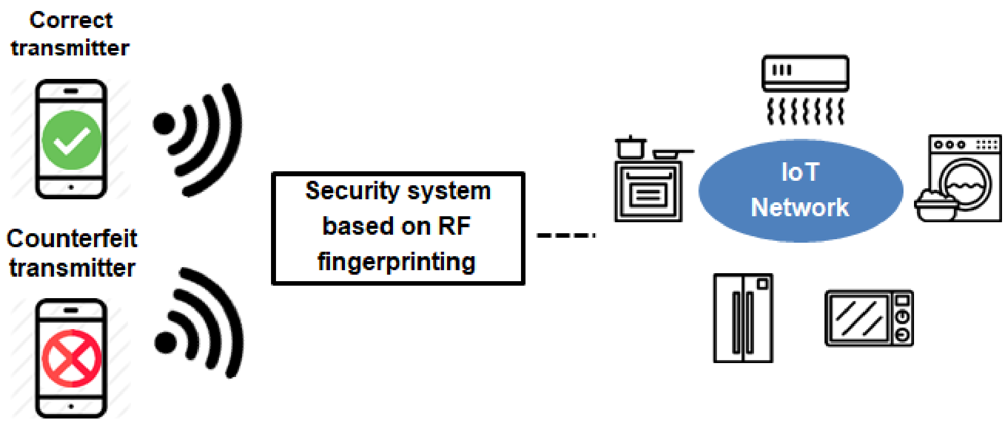


Fig. 1. Example of the transmitter classification scheme in IOT network

There are two nodes : correct node and counterfeit node. Correct node is registered in IOT network for access. Counterfeit node is not registered in IOT network. Thus, it is illegal for the counterfeit node to access to the network. Since media access control (MAC) address or internet protocol (IP) address is used as a node identifier, the counterfeit node can access to the IOT network if it knows the MAC or IP address and the communication standard. If the radio signal from the node (transmitter) has information on classifying the node and the classification scheme is based on the RF signals, it will be difficult to mimic the RF signals perfectly. To make the RF signals totally same, hardware structure and characteristic must be same since the signals are created by some hardware like amplifier and oscillator. Classifying transmitters based on RF signals is called to RF fingerprinting. To do the RF fingerprinting, some parts of the signal must have information on classifying the transmitters. The parts are called as feature. Feature is cultivated from received RF signals. Since feature consists of samples in discrete signal, feature is defined as a sample vector which is cultivated from the received RF signals. Features are known to arise from many sources, including tiny difference in device fabrication process and electronic components [3].

RF fingerprinting has attracted significant attention [4–8]: Patel *et al.* [4] used RF-DNA features which contain information on variance, skewness, and kurtosis, within a preamble response and showed that an ensemble method combining multiple classifiers performs well. Peng *et al*. [5] used four features—differential constellation trace figure, carrier frequency offset, modulation offset, and I/Q offset, where classifications were done by calculating the minimum distance between test data and training data. Yong Qiang *et al.* [6] used mean of instantaneous amplitude of the received signal and the modulation symbol. They found an optimal dimension-reduced matrix that maximizes the quadratic mutual information between the low-dimensional features and the class, and minimizes the classification error. The same authors also investigated an RF fingerprinting scheme based on the low-rank representation of the original data with the robust classifier parameter [7]. Merchant *et al.* [8] used the convolutional neural network for seven commercial Zigbee devices. They collected 1,000 data per a class. In [4-8], RF fingerprinting schemes with multiple features which exhibited high accuracy rate were proposed for Zigbee devices and satellite terminals.

In this paper, we propose a new RF fingerprinting algorithm and a set of three RF features—*rising transient*, *falling transient*, and *sync*—and show the possibility that each feature could provide unique information through our own real-life experiment. The falling transient feature has never been used in RF fingerprinting studies as a unique feature. Our results indicate that the performance of RF fingerprinting improves as each feature is additionally employed. As a classification algorithm, we used sparse representation based classification (SRC) algorithm. Even though SRC is the qualified algorithm in the classification area [9], there are no studies on RF fingerprinting with combination of SRC and multiple features. We show that SRC can be a good platform for RF fingerprinting.

The remainder of this paper is organized as follows :

In Chapter 2, prerequisite knowledges on the proposed experiment will be discussed for understanding. Firstly, we will explain digital mobile radio standard. In this part, a structure of the transmitted signal will be introduced. Then, we will explain what kinds of features have been used in RF fingerprinting area. Following that, sparse representation based classification algorithm and principal component analysis will be explained.

In chapter 3, we will explain how to extract interest signal parts and features for RF fingerprinting. Then, we will demonstrate how to classify the transmitters with unique feature and multiple features. In the multimodal method, we will explain how to use multiple features simultaneously.

In chapter 4, we have two parts : experimental setup and result. In the experimental setup, we will introduce which equipment is used to save the transmitted signals onto the computer. In the experimental result, we will show the accuracy rate of the proposed experiment with a graph and table. Since deep learning is popular nowadays, we will compare the proposed method with convolutional neural network method, referring the study [8].

In chapter 5, we will give summary and future work of this paper.

2. Background

RF fingerprinting is a work to classify transmitters through their RF signal. In RF signal, feature exists. The feature is a sample vector extracted from RF signal. It has unique information on the specific transmitter [10]. For this reason, it is considered as a fingerprint to classify the specific transmitter. The proposed method suggests how to use multiple features with a sparse representation based classification algorithm. To get through to the details on our experiment, we explain basic concepts related in this experiment.

2.1. Digital mobile radio (DMR) Standard

We use eight commercial walkie-talkies for RF fingerprinting. The used walkie-talkies are 2 models. One of them is 'BD-358' and another is 'Sl1M'. They are made by 'Hytera' and 'Motorola', respectively. Each model follows DMR standard [11]. DMR standard is defined in the European Telecommunications Standards Institute (ETSI). It is a protocol for mobile devices on how to modulate each symbol and transmit the RF signals. DMR standard has three main characteristics. First one is 2-slot Time-division multiple access (TDMA) method. TDMA is designed to use the spectrum efficiently. Two nodes can share the same communication channel since each transmitted signal is divided into different time slots. 2-slot TDMA makes signal bursts as Fig 2. We only consider a first signal burst for the proposed experiment.

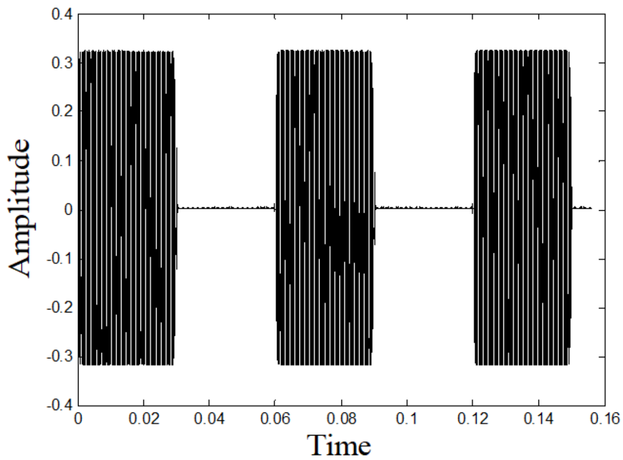


Fig. 2. 2-slot TDMA method forms the burst every 30ms.

Second one is Frequency-shift keying (FSK) modulation method. According to DMR standard document [11], DMR standard uses 4-level FSK to modulate symbols. In 4-level FSK modulation method, information 2 bits constitute a symbol and the symbol is modulated with different frequencies. Table 1 shows the frequency deviations on each symbol.

Table 1. Dibit symbol mapping to 4FSK deviation

Information bits		Symbol	4 FSK deviation
Bit 1	Bit 0		
0	1	+3	+1.944 kHz
0	0	+1	+0.648 kHz
1	0	-1	-0.648 kHz
1	1	-3	-1.944 kHz

Even if each symbol is modulated with the predefined frequency deviation, the modulated frequency is different slightly on each device. It is generated from hardware characteristics as we mentioned. We will explain on why the feature exists in each signal with more details in subsection 2.2. The DMR document [11] mentions that the frequency offset exists up to ± 2 ppm .

The last one is the structure of the burst. The burst consists of a rising transient signal, a steady-state signal, and a falling transient signal in order. Fig. 3 shows the structure of the burst in details.

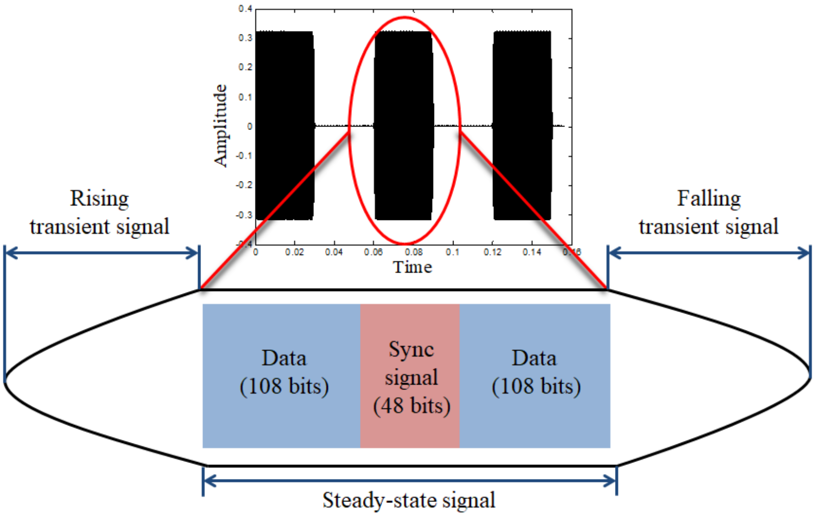


Fig. 3. The structure of the signal burst

At first, the rising transient signal is turned up. The rising transient signal is a growing part from the zero level to a designed level of the RF signal. Subsequent to the rising transient signal, the steady-state signal appears. The steady-state signal has a data and a sync signal. In walkie-talkies, the data signal is a voice. Since the voice could be different whenever push-to-talk of the transmitter is executed, we don’t consider the data part as our interest. The sync signal exists for the synchronization between transmitters and receivers. It is also called as a preamble. The sync signal always has same information regardless of the surrounding environment. Lastly, the falling transient signal follows after the steady-state signal. Contrary to the rising transient signal, the falling transient signal is a decreasing part from a designed level to the zero level of the RF signal. For RF fingerprinting, we consider three signal parts to extract proper features – rising transient signal, falling transient signal, and sync signal. We demonstrate why these signals are considered to extract features in the next subsection.

2.2. Feature

For RF fingerprinting, which features will be used and how the features will be extracted have to be contemplated considerably. A feature is unique characteristic of each transmitter. It arises from many sources, such as device fabrication process and electronic components [12-13]. Generally, typical features have three categories, namely a transient feature, a sync feature, and In-phase and quadrature (I/Q) domain feature.

The transient features are in the transient signals. The transient feature is represented as a variation of the envelope of the transient signals. Fig. 4 shows differences among the transient signals of Bluetooth transmitters [14]. This feature is generated by the different characteristic of each signal amplifier [14]. Even though the different models in Fig. 4 are represented, the same models also have the unique transient feature, which can't be distinguished with the unaided eye.

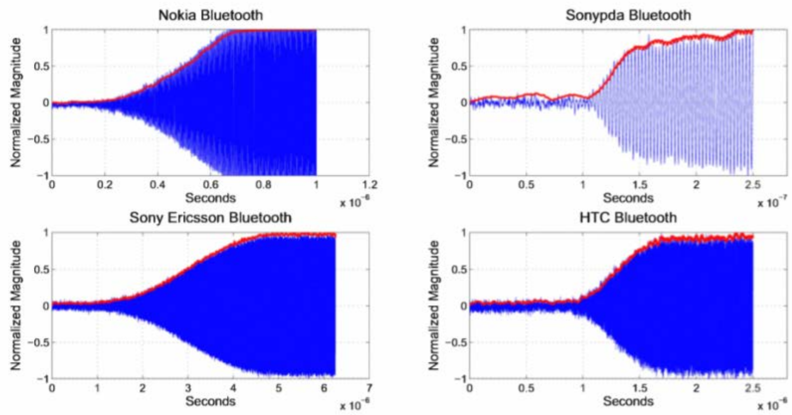


Fig. 4. Transient signals of four Bluetooth transmitters [14]

The sync features are in the preamble. It appears as a carrier frequency offset and modulation offset etc. This feature is generated from the nonlinear property of each oscillator [15]. Fig. 5 shows a measured RF preamble for two different Wi-Fi devices at 2.4 GHz.

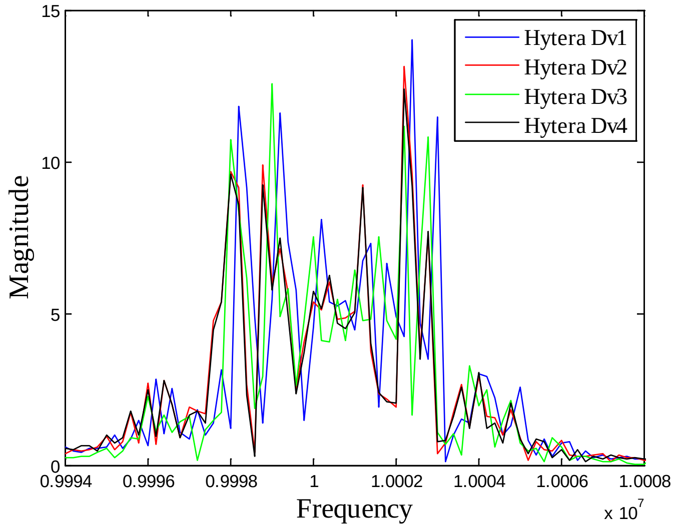


Fig. 5. Measured preambles of four Wi-Fi devices

I/Q domain feature is an error between an ideal magnitude or phase and a measured magnitude or phase. When symbol is modulated, the modulation amplitude, frequency, and phase are fixed on specific symbol. However, there is error in the three components because of different hardware characteristics. In I/Q domain, we can see errors on the amplitude and phase. Fig. 6 represents an error in I/Q domain [16].

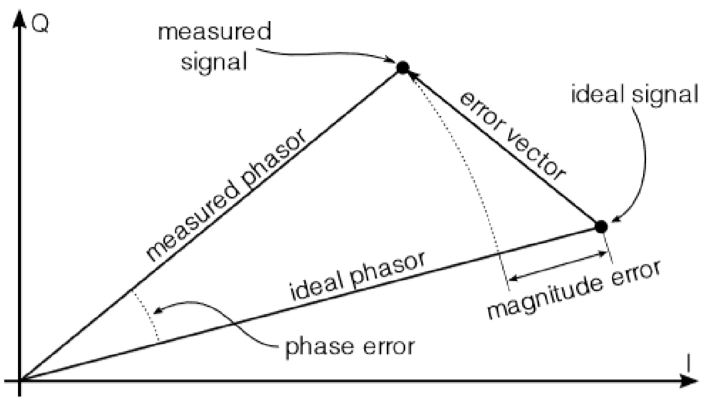


Fig. 6. Error vector in I/Q domain [16]

Also, phase mismatch distortion can be generated. The phase between In-phase and Quadrature should be $\frac{\pi}{2}$. However, the offset could appears with the same reason [5].

The features become fingerprints for classifying transmitters. The features enter the classifier as an input. A classifier outputs a correspoding label on the input.

2.3. Sparse representation-based classification algorithm

Sparse representation based classification algorithm (SRC) is from a compressed sensing theory [17]. SRC is a qualified method for images [9] and signals [18]. For SRC, a sparse unique solution of a linear equation (1) is used,

$$\mathbf{y} = \mathbf{D}\mathbf{s}, \tag{1}$$

where the column vector $\mathbf{y} \in \mathbb{R}^{P\times 1}$ is a test data vector, $\mathbf{D} \in \mathbb{R}^{P\times N}$ is a training data matrix, and $\mathbf{s} \in \mathbb{R}^{N\times 1}$ is a coefficient vector. The column vector in $\mathbf{D}$ is a training data vector. Assuming $M < N$ in Eq. (1), which is underdetermined system, the compressed sensing algorithm finds a sparse unique solution among infinite cases of $\mathbf{s}$. A sparse solution has a least number of components. As a condition to find out if the sparse unique solution exists or not in the system (1), the column vectors in $\mathbf{D}$ should be uncorrelated [19]. In Eq. (1), the condition of uniqueness of sparse solution on L_1 norms is as Eq. (2)

$$\| \mathbf{s}_0 \|_0 < \frac{1}{2}\left(1 + \frac{1}{\mu(\mathbf{D})}\right), \tag{2}$$

where $\mu(\mathbf{D})$ is $\max\limits_{1\le k,\, j\le m,\, k\ne j} \frac{|\mathbf{d}_k^T \mathbf{d}_j|}{\|\mathbf{d}_k\|_2 \|\mathbf{d}_j\|_2}$, $\mathbf{d}_i$ is i th column vector in a matrix $\mathbf{D}$, L_p norm, $\|\mathbf{s}\|_w$, is $\left(\sum_i |s_i|^w\right)^{\frac{1}{w}}$, and s_i is i th component value of $\mathbf{s}$. Thus, the condition to find a sparse solution depends on the mutual correlations between column vectors in a training matrix $\mathbf{D}$. To find a sparse solution, we use a basis pursuit algorithm [20]. Basis pursuit algorithm finds a sparse solution which has the minimum value of L_1 norm of solution as follows :

$$\min_{\mathbf{s}} \|\mathbf{s}\|_1 \quad s.t.\ \mathbf{y} = \mathbf{D}\mathbf{s}. \tag{3}$$

The obtained sparse unique solution $\mathbf{s}$ is divided into L disjoint sub-vectors $\mathbf{s}_l$, for $l = 1,\ldots,L$,

$$\mathbf{s} = \left[\mathbf{s}_1^T, \mathbf{s}_2^T, \ldots, \mathbf{s}_l^T\right]^T, \tag{4}$$

where T is transpose of matrix and L is the number of transmitters (classes). Similarly, $\mathbf{D}$ is also divided into L sub-matrices corresponding to $\mathbf{s}_l$. To output the class of test data, we solve

$$\text{class} = \arg\min_{l} \| \mathbf{y} - \mathbf{D}_l \mathbf{s}_l \|_2 . \tag{5}$$

Fig. 7 represents the modeling of SRC on 2 classes.

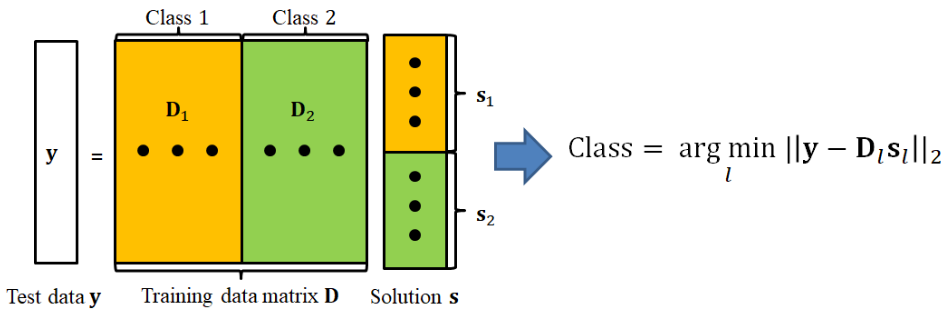


Fig. 7. The model of the sparse representation-based classifier

We assumed that the column vectors in $\mathbf{D}$ are uncorrelated. However, we can't make sure whether the training data vectors are uncorrelated or not. If RF signals become the training data vectors of $\mathbf{D}$ without any signal processing, the solution $\mathbf{s}$ in Eq. (1) is not approximated to be sparse. The main point for ideal SRC is to remove the correlations among the column vectors in $\mathbf{D}$.

2.4. Principal component analysis

Principal component analysis (PCA) [21] is a statistical method to change a set of correlated vectors to uncorrelated vectors. Thus, PCA can be a key technique to remove the correlations among data. Besides, PCA makes the data size reduced. It means that computational operation to find the sparse unique solution in Eq. (1) becomes simpler with PCA. In this subsection, we demonstrate the concept of PCA.

Let us assume that Eq. (1) is a system which passes through PCA. A system before PCA is

$$\mathbf{u} = \mathbf{A}\mathbf{x}. \tag{6}$$

Eq. (6) is similar with Eq. (1). However, $\mathbf{A}$ has correlations among column training vectors unlike $\mathbf{D}$,

$$\mathbf{A} = [\mathbf{a}_1^{(1)}, \ldots, \mathbf{a}_1^{(N)}, \mathbf{a}_L^{(1)}, \ldots, \mathbf{a}_L^{(N)}]^T , \tag{7}$$

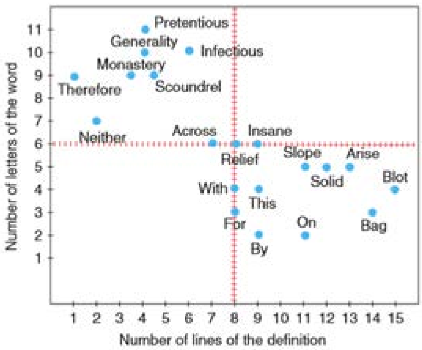
where column vector $\mathbf{a}_n^{(l)} \in \mathbb{R}^{M \times 1}$ is the n th sample vector of l th RF transmitter for $n = 1, \ldots, N$ and $l = 1, \ldots, L$. Let $\mathbf{m} = \frac{1}{L \times N} \sum_{l=1}^{L} \sum_{n=1}^{N} \mathbf{a}_l^n \in \mathbb{R}^{M \times 1}$, which is average vector of columns of $\mathbf{A}$. Then, we perform the eigen-decomposition on

$$(\mathbf{A} - \mathbf{m}\mathbf{1})(\mathbf{A} - \mathbf{m}\mathbf{1})^T = \mathbf{W}\mathbf{\Lambda}\mathbf{W}^T , \tag{8}$$

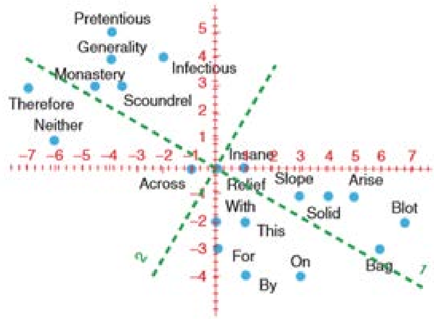
where $\mathbf{A} \in \mathbb{R}^{M \times N}$ is a training matrix, $\mathbf{1}$ is 1s vectors, $\begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{1 \times N}$, $\mathbf{W}$ is a eigenvector matrix of covariance matrix $(\mathbf{A}-\mathbf{m1})(\mathbf{A}-\mathbf{m1})^T$, and $\mathbf{\Lambda}$ is a diagonal matrix on eigenvalues of covariance matrix $(\mathbf{A}-\mathbf{m1})(\mathbf{A}-\mathbf{m1})^T$. The main point of covariance matrix $(\mathbf{A}-\mathbf{m1})(\mathbf{A}-\mathbf{m1})^T$ is that the eigenvectors of the covariance matrix are orthonormal [21], $\mathbf{W}^{\mathrm{T}}\mathbf{W} = \mathbf{I}$. Thus, Eq. (8) can be represented as

$$\mathbf{W}^T (\mathbf{A}-\mathbf{m1})(\mathbf{A}-\mathbf{m1})^T \mathbf{W} = \mathbf{\Lambda}. \tag{9}$$

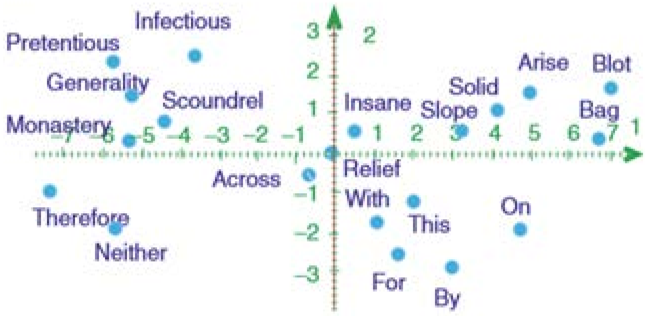
Eq. (9) means that the eigenvalue of the covariance matrix is proportional to the variance of a training matrix $\mathbf{A}$. Fig. 8 shows a meaning of the demonstration.



(a) Original data



(b) New space on variance of data



(C) Projection data onto new space

Fig. 8. Example of principal component analysis

If unique feature exists in different data, more distinct cluster can be formed when the axis of space has the bigger variance. As the eigenvalue matrix $\mathbf{\Lambda}$ is arranged by the size of eigenvalue, the eigenvector matrix $\mathbf{W}$ is also arranged by the eigenvalue. We will define the arranged eigenvector matrix as $\mathbf{V}$. Since the biggest eigenvalue represents the biggest variance of $\mathbf{A}$, the eigenvector corresponding to the biggest eigenvalue represents the most distinct basis vector which can form the distinct cluster on column vectors of $\mathbf{A}$. Thus, we make a new space on the variance of $\mathbf{A}$ using eigenvector matrix $\mathbf{V}$. The undetermined and uncorrelated system, Eq. (1), is completed by mapping the feature matrix $\mathbf{A}$ and test data $\mathbf{u}$ to the new space, $\mathbf{V}$,

$$\mathbf{V}^T(\mathbf{u}-\mathbf{m}\mathbf{1})=\mathbf{V}^T(\mathbf{A}-\mathbf{m}\mathbf{1})\mathbf{s}. \tag{10}$$

Eq. (10) is totally same with Eq. (1). Since the correlations among column data are removed, we can find the sparse unique solution of Eq. (1). Also, the computational operation falls off to find the solution as the size of data is highly reduced.

3. Proposed method

Determining which feature will be selected is important in RF fingerprinting. Even if the same feature is selected, there are a variety of processes to extract the feature. Depending on which feature will be used and how to extract the feature, a classification result can be totally different. If the feature is extracted with the complicated method, the performance could be improved. However, the method may not be applied to other communication protocols if the method requires the detail information of the specific communication protocol. In this section, we deal with the simple method. The performance is pretty increased as just additional features are included

3.1. Extraction of features

To extract features, we should select the corresponding signals; the rising transient signal, the falling transient signal, and steady-state signal. These signals exist in a burst. Fig. 1 shows the raw signal we received. The first burst among generated busts in received signal is our interest in the raw signal. Since the burst width appears during 30ms [11], we can choose the first burst with the corresponding samples. The number of sample of the first burst is 2,880,000.

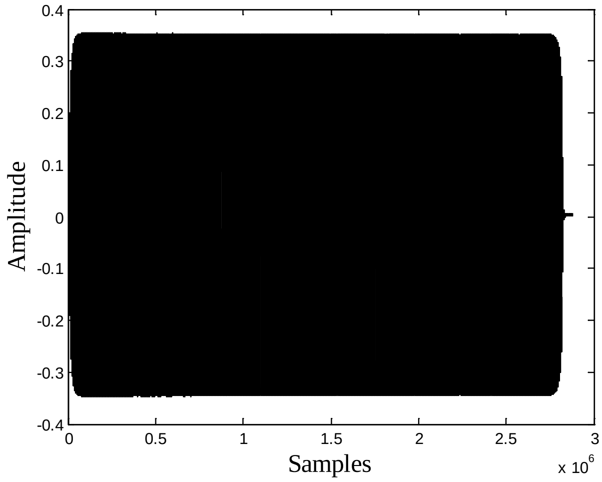


Fig. 9. A burst following DMR standard

Fig. 9 shows the first burst in a raw signal. We now process this burst for our interesting feature.

As we mentioned, the burst has three signals, the rising transient signal, falling transient signal, and steady-state signal, respectively. These signals have unique information which occurs as a result of the inherent nonlinear properties [1]. To extract the interest signal parts, we use a threshold method. The threshold method is performed through time-windowing.

For the rising transient signal, we set the starting point of the time-window to the fastest time when the magnitude of a signal burst exceeds 10% of its maximum. The ending point of the time-window is set to the fastest time when the magnitude of a signal burst exceeds 90% of its maximum. The designed time-window is applied to the burst. Then, the rising transient signal is extracted. To fit all rising transient signals to the same, normalization is implemented. We normalized the rising transient signals by dividing them into the summation of samples. It makes the summation of all samples 1. Fig. 10 shows the extracted rising transient signal.

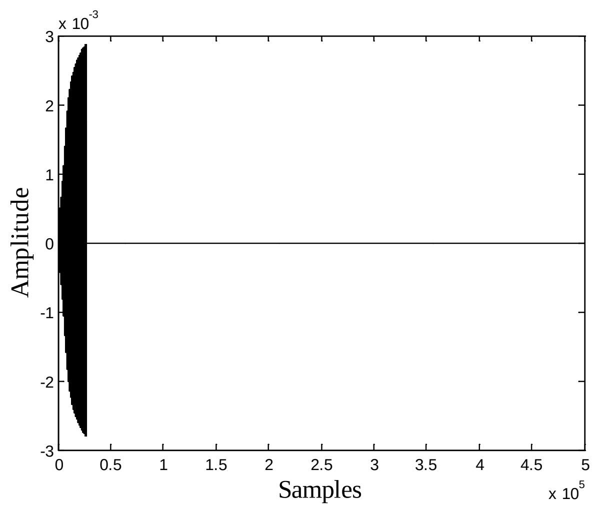


Fig. 10. The extracted rising transient signal

Since the number of time-window for each rising transient signal is different, we set the number of samples to 500,000 and used zero padding after the ending point of each transient signal to make the length of each rising transient signals same. Then, we normalized the extracted rising transient signal. Normalization is used to match the transmitted signal amplitude in the different environment.

Even though we succeed to extract the rising transient signal, it has a problem. The signal length are still many as 500,000 samples. We need to reduce the size of feature for more rapid computation in classification process. We define the feature of target signal as main lobe in spectrum. To convert time signal to spectrum, we perform Fast Fourier Transformation (FFT). Then, we take an operation of absolute value to compare energy and frequency information. Finally, the main lobe of the spectrum is extracted. Fig. 11 represents the main lobe of the rising transient signal.

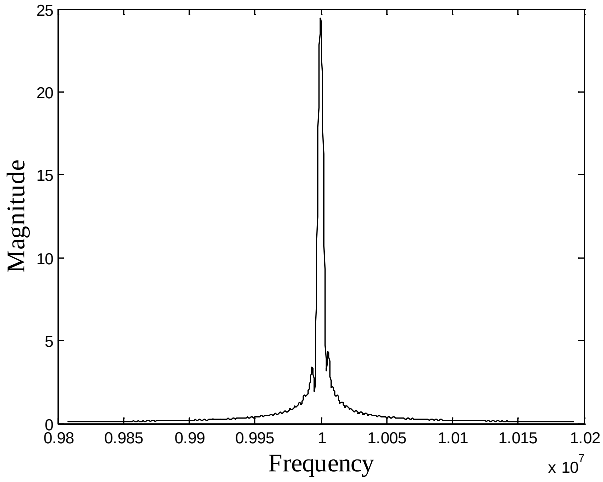


Fig. 11. The main lobe of the rising transient signal

The main lobe is selected by the maximum value. The point which has the maximum value is the center point of the main lobe. The 0.4% bandwidth of the total is extracted with the center point. The numerical value for the selected bandwidth is considered by the DMR standard. This main lobe is used to concatenate the main lobe with other features.

For the falling transient signal, we set the starting point of the time-window to the latest time when the magnitude of a signal burst exceeds 90% of its maximum. The ending point of the time-window is set to the latest time when the magnitude of a signal burst exceeds 10% of its maximum. The designed time-window is applied to the burst. Then, the falling transient signal is extracted in common with the rising transient signal. The normalization we described is applied to the falling transient signal to make all signal strengths same. Fig. 12 shows the falling transient signal.

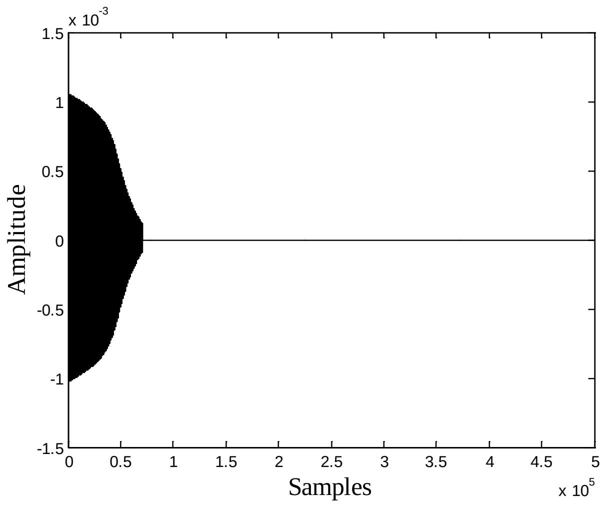


Fig. 12. The extracted falling transient signal

After extracting the falling transient signal, we should make the size of the signal reduced for simple computational operation of classification process. Along with the rising transient signal, the main lobe is selected in the spectrum of the falling transient signal. The process for the main lobe of the falling transient signal is same with the rising transient signal. Fig. 13 represents the main lobe of the falling transient signal.

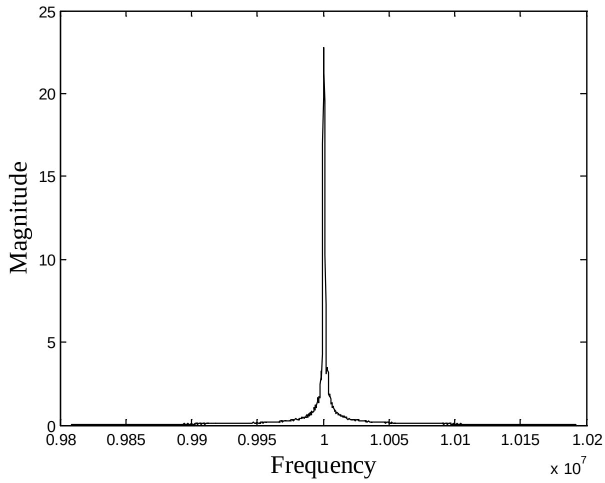


Fig. 13. The main lobe of the falling transient signal

At last, we consider the sync signal. According to the DMR standard [11], the sync signal exists in the center of the burst. It consists of 48 bits. Following the DMR standard, we extracted the sync signal. First, we set the center point of the burst. We selected the center point by choosing other two points which are the last point of the rising transient signal and the starting point of the falling transient signal, respectively. Then, we calculated the number of samples corresponding to 48 bits which is the size of sync signal. Since the sampling rate is 96MHz and the bit rate is 9,600 bits/sec. The number of samples per bit is calculated as below.

$$\text{The number of samples per a bit} = \frac{\text{Sampling rate}}{\text{Bit rate}} \tag{11}$$

The number of samples per a bit in our experiment is 10,000 samples. Thus, the length of the sync signal is 480,000 samples. From the center point, we can get the preamble signal by extracting the samples whose range is 'center point -240,000' to 'center point+239,999'. Fig. 14 shows the extracted preamble signal.

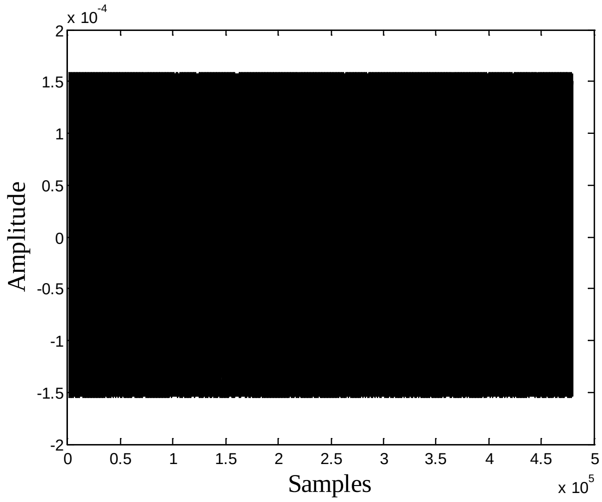


Fig. 14. The extracted sync signal

For similar reasons with transient signals, we should reduce the size of the sync signal. To reduce the size of the sync signal, we extracted the main lobe of the sync signal. The sync signal is changed to the spectrum by FFT. Then, the point having maximum value is set to the center point. The 0.4% bandwidth of the total is extracted with the center point. Fig. 15 shows the main lobe of the sync signal.

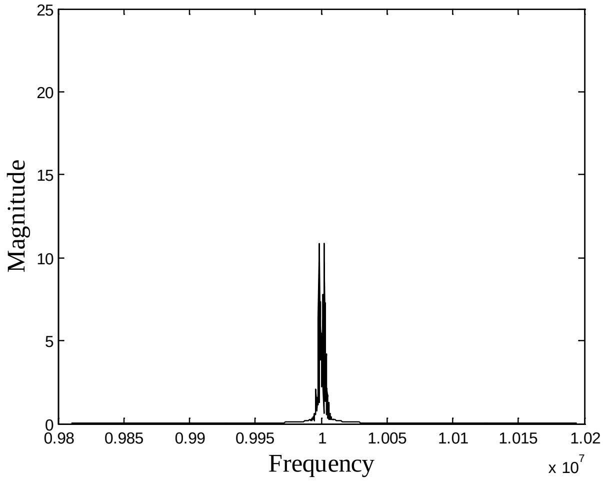


Fig. 15. The main lobe of the sync signal

We can see that the spectrum has two peak points in frequency domain. According to the DMR standard documentation [11], the SYNC (Synchronization) pattern of mobile station is specific as Table 2.

Table 2. SYNC patterns

MS sourced													
Voice	Hex	7	F	7	D	5	D	D	5	7	D	F	D
	Binary	0111	1111	0111	1101	0101	1101	1101	0101	0111	1101	1111	1101
Data	Hex	D	5	D	7	F	7	7	F	D	7	5	7
	Binary	1101	0101	1101	0111	1111	0111	0111	1111	1101	0111	0101	0111
RC sync	Hex	7	7	D	5	5	F	7	D	F	D	7	7
	Binary	0111	0111	1101	0101	0101	1111	0111	1101	1111	1101	0111	0111

The dibit consists of 2 symbols : ‘01’ and ‘11’. Since SYNC patterns have only 2 symbols, two peak points are represented in the spectrum. In case of the walkie-talkie, ‘Data’ part of MS sourced SYNC patterns is applied. The binary bit of SYNC pattern of our walkie-talkies is ‘110101011101011111110111011111111101011101010111’.

Features of each signal are represented in Fig. 11, Fig. 13, and Fig. 15. Using these features simultaneously, we will show that the RF fingerprinting performance is improve.

3.2. Combination of features.

To unite the three features, we concatenated the rising transient feature, the falling transient feature, and the sync feature. Fig. 16 shows the concatenated feature.

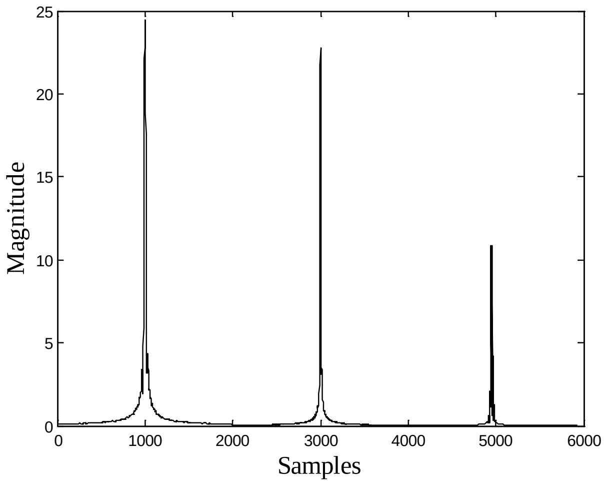


Fig. 16. The concatenated feature

The left main lobe is on the rising transient signal. The center main lobe is on the falling transient signal. The right main lobe is on the sync signal. When the burst is extracted, the number of samples of the burst is about 3,000,000. However, the number of samples of the concatenated feature is 5,920. The size of feature for classification is much smaller than raw signals. Even though the length of feature is reduced, we can’t use this feature as an input for classification since the features could be much correlated.

3.3. Classification

The features are added to the linear Eq. (1) as column vectors. We explained that the column vectors should be less correlated to get a sparse solution of Eq. (1). However, the features could be correlated when signal processing is not employed. To make them uncorrelated, we applied PCA to features. We describe the proposed scheme in singlemodal and multimodal RF fingerprinting.

3.3.1. Singlemodal RF fingerprinting

We consider that one of the rising transient, falling transient, and sync features is used as the sole representative feature of a single transmitter. We make a feature matrix in which the columns are the sample vectors of the feature of candidate RF transmitters. Mathematically, for L RF transmitters (Classes) and N sample vectors of a feature of each RF transmitter, we construct the feature matrix $\mathbf{A}$ as follows

$$\mathbf{A} = [\mathbf{a}_1^{(1)}, \cdots, \mathbf{a}_N^{(1)}, \mathbf{a}_1^{(2)}, \cdots, \mathbf{a}_N^{(L)}], \tag{12}$$

where the column vector $\mathbf{a}_N^{(L)} \in \mathbb{R}^{M\times 1}$ is the n^{th} sample vector of a feature of the l^{th} RF transmitter for $n = 1, \cdots N$ and $l = 1, \cdots, L$. M is the size of the sample vector of the feature of the RF transmitter. We set an unknown sample vector, which is the sample vector of the feature of an unknown RF transmitter, as $\mathbf{u}$. From the PCA operation, $\mathbf{A}$ is mapped to a training matrix $\mathbf{D}$ as follows:

$$\mathbf{D} = \mathbf{V}^T(\mathbf{A} - \mathbf{m1}), \tag{13}$$

where $\mathbf{m} = \frac{\sum_{l=1}^{L}\sum_{n=1}^{N}\mathbf{a}_n^l}{L \times N} \in \mathbb{R}^{M\times 1}$ is an average vector of columns of $\mathbf{A} \in \mathbb{R}^{M\times NL}$, $\mathbf{1} := [1 \quad 1 \quad \cdots \quad 1]$ is the 1 by NL vector of 1s, and $\mathbf{V} \in \mathbb{R}^{M\times P}$ is the arranged eigenvector matrix of $(\mathbf{A} - \mathbf{m1})(\mathbf{A} - \mathbf{m1})^T \in \mathbb{R}^{M\times M}$. As we mentioned, the PCA can remove the correlations among column vectors in a difference matrix $(\mathbf{A} - \mathbf{m1})$. According to the study [22], PCA can improve the performance of SRC. In Eq. (13), the column vectors of $\mathbf{V}$ are arranged according to the eigenvalues in descending order. The dimension of $\mathbf{V}$ can be selected by user as $P \in \{1, \cdots, M\}$.

In common with the feature matrix $\mathbf{A}$, an unknown signal is also processed for the test data. From the PCA operation, the feature of the unknown signal $\mathbf{u}$ is mapped to a test data $\mathbf{y}$ as bellows

$$\mathbf{y} = \mathbf{V}^T(\mathbf{u} - \mathbf{m}). \tag{14}$$

From the test data $\mathbf{y}$ and the training data matrix $\mathbf{D}$, we can find the sparse solution $\mathbf{s}$. Then, the class is output by Eq. (5), the error estimation between the test data and the value of multiplications of the training data matrix with the disjoint solution vector corresponding to the each class.

3.3.2. Multimodal RF fingerprinting

The proposed method is to perform RF fingerprinting with multiple features which exist in the burst signal. Each feature of the burst signal is concatenated. From the concatenation, the multiple features become a unique vector. The proposed model is represented in Fig. 17.

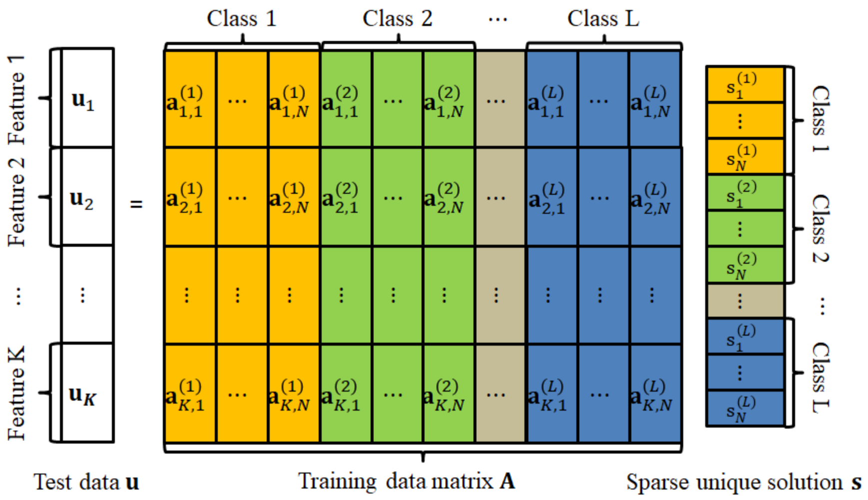


Fig. 17. Feature concatenation in the proposed method

The feature matrices are concatenated in arrow-wise manner, On the combined matrix, the PCA is applied. Let us denote the n^{th} sample vector of the k^{th} feature of l^{th} RF transmitter for $n = 1, \cdots, N$, $l = 1, \cdots, L$, and $k = 1, \cdots, K$, where K is the number of features to be combined through $\mathbf{a}_{k,n}^{(l)} \in \mathbb{R}^{M \times 1}$. The feature matrices are concatenated as follows :

$$\mathbf{A} = \left[\mathbf{A}_1^T \quad \mathbf{A}_2^T \quad \cdots \quad \mathbf{A}_K^T \right]^T, \tag{15}$$

where $\mathbf{a}_{k,n}^{(l)}$ forms the columns of feature matrix $\mathbf{A}_i \in \mathbb{R}^{M \times NL}$, i.e.,

$$\mathbf{A}_k = [\mathbf{a}_{k,1}^{(1)}, \cdots, \mathbf{a}_{k,N}^{(1)}, \mathbf{a}_{k,1}^{(2)}, \cdots, \mathbf{a}_{k,N}^{(L)}]. \tag{16}$$

From Eq. (13), we obtain training data matrix $\mathbf{D}$. In multimodal scheme, $\mathbf{m} = \dfrac{\sum_{l=1}^{L}\sum_{n=1}^{N} \mathbf{a}_n^l}{L \times N} \in \mathbb{R}^{MK \times 1}$ is an average vector of columns of $\mathbf{A} \in \mathbb{R}^{MK \times NL}$, $\mathbf{a}_n^l = [(\mathbf{a}_{1,n}^l)^T \ (\mathbf{a}_{2,n}^l)^T \ \cdots \ (\mathbf{a}_{K,n}^l)^T]^T \in \mathbb{R}^{MK \times 1}$ is a column vector which combines K features, and $\mathbf{V} \in \mathbb{R}^{MK \times P}$ is the arranged eigenvector matrix of $(\mathbf{A} - \mathbf{m})(\mathbf{A} - \mathbf{m})^T \in \mathbb{R}^{MK \times MK}$. The dimension of $\mathbf{V}$ can be selected by user as $P \in \{1, \cdots, MK\}$.

To obtain the test data $\mathbf{y}$ of SRC, we first concatenate the sample vectors of different features of an unknown transmitter $\mathbf{u}_k$ as follows:

$$\mathbf{u} = [\mathbf{u}_1^T \quad \mathbf{u}_2^T \quad \cdots \quad \mathbf{u}_K^T]^T . \tag{17}$$

Finally, $\mathbf{y}$ is obtained by mapping the difference between concatenated vector $\mathbf{u} \in \mathbb{R}^{MK \times 1}$ and $\mathbf{m}$, i.e. $\mathbf{u} - \mathbf{m}$, onto the space with orthonormal basis $\mathbf{V}$ using Eq. (14). Fig. 17 is changed to Eq. (1) which has principal components as training data matrix and test data.

3.4. Overall scheme of the proposed method

We described the detailed process for RF fingerprinting with multiple features. The overall scheme is represented in Fig. 18.

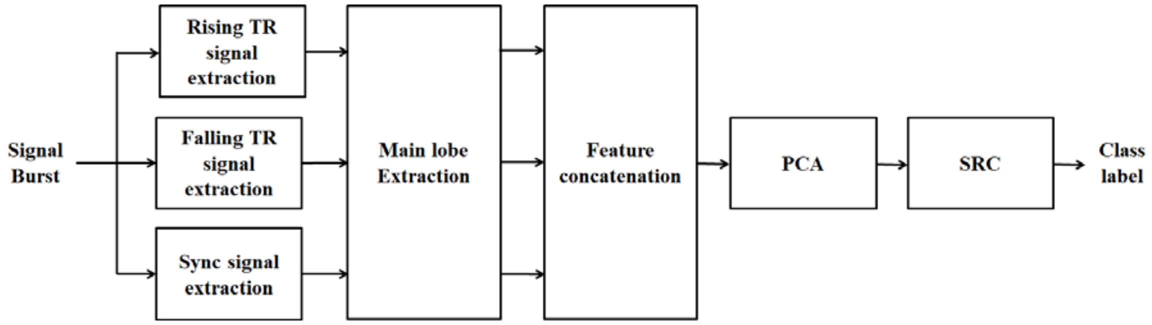


Fig. 18. Multimodal classification algorithm scheme

The order of this experiment as follows.

① The signal burst is extracted.

② The rising transient signal is extracted by the threshold method.

③ The falling transient signal is extracted by the threshold method.

④ The sync signal is extracted by the threshold method and the location in DMR standard.

⑤ The main lobe of each signal is extracted.

⑥ Main lobes of each signal are concatenated.

⑦ PCA is applied to the concatenated vector u and matrix A

⑧ SRC is performed in Eq. (3) and final class is output.

4. Experiment

4.1. Experimental system

We performed RF fingerprinting experiment for 8 walkie-talkies. The used transmitters are as below,

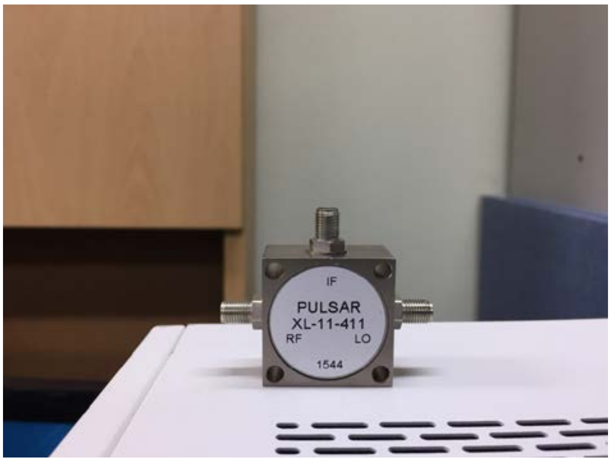
(a) ‘SL1M’, which is made in Motorola (b) ‘BD-358’, which is made in Hytera

Fig. 19. The digital walkie-talkies used in the experiment

In our experiment, 4 models of ‘SL1M’ and 4 models of ‘BD-358’ are used. As we mentioned, even though 2 models are made in the different manufacturers, they follow the same communication standard, which is DMR standard. The channel frequency is around 423.1875 MHz.

For the experiment, we only consider the first received signal burst. To get the signal burst, we set some experimental systems. The equipment for the system is represented in Fig. 20.


(a) SMA male mini car-mounted antenna



(b) XL-11-411 RF mixer


(c) E4438C ESG vector signal generator


(d) IF recording system


(e) The experimental system for RF fingerprinting

Fig. 20. The equipment for acquisition of the signal burst

The range of the received frequency of SMA male mini car-mounted antenna is 400-470 MHz. As the *push-to-talk* function of a transmitter was executed, the transmitted signal was acquired from the antenna. The transmitted RF signal has 423.1875 MHz. The signal is down-converted to 10 MHz using an XL-11-411 RF mixer and an E4438C ESG vector signal generator. Then, the down-converted signal is filtered and sampled from intermediate frequency (IF) recording system. IF recording system has PX14400 operator which functions as a low-pass filter and analog-to-digital converter. The IF signals are sampled at 96 MHz. The sampled signals are saved to a computer and loaded to MATLAB.

To perform the experiment, we captured 50 signals per a digital walkie-talkie. Total 400 signals were saved. The decimation rate is set to 250, considering the bandwidth of the RF signal following the DMR standard and the sampling rate. The rate, 250, means 0.4% bandwidth in total spectrum. The channel bandwidth is 12.5kHz in DMR standard. In the range 0 to over 10MHz, the bandwidth can be extracted without any loss of information on the main lobe when the decimation rate is set to 250. The experiment was conducted in the line-of-sight environment. The distance between transmitter and receiver was about 0.7m. Signal to noise ratio (SNR) was around 35-40dB.

4.2. Experimental result

To evaluate the performance of the proposed classifier, we used a five-fold cross validation technique. It means that five times of experiments are performed on the different training sets per an experiment. Among 50 signals per a class, 10 signals are used for test data and the others are set to the training data. Then, the average of accuracy rate for each test data is output. After the experiment, other 10 data, which are not used as test data, are set to the test data. The others are set to the training data. This cycle is maintained until all of data will be used for test data. The average accuracy rate on the five iterations is output as a final result.

Fig. 21 shows the classification result of four BD-358 and four SL1M walkie-talkies, 1) on the rising transient feature, 2) on the falling transient feature, 3) on the both transient features combined, 4) on the sync feature, and 5) on the all features combined.

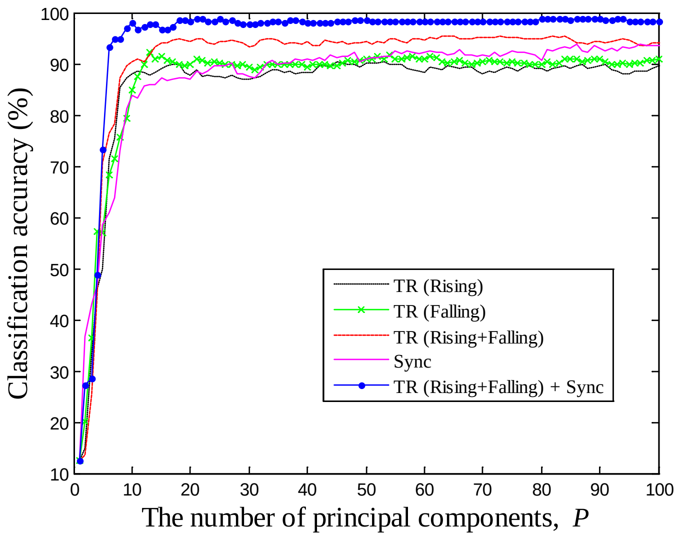


Fig. 21. Classification result of four BD-358 and four SL1M walkie-talkies

The number of principal components means the number of used eigenvectors of $(\mathbf{A}-\mathbf{m1})(\mathbf{A}-\mathbf{m1})^T$ to form the new space on the variance of columns of $\mathbf{A}$. When the rising transient feature is used only, the accuracy rate was 90.5% with 45 principal components. When the falling transient feature is used only, the accuracy rate was 92.25% with 13 principal components. When the sync feature is used only, the accuracy rate was 93.5% with 86 principal components. With the concatenated feature, the accuracy rate is increased. When the rising transient feature and falling transient feature is concatenated, the accuracy rate was 95.5% with 63 principal components. When the three features – rising transient feature, falling transient feature, and sync feature- are concatenated, the accuracy rate was 98.75% with 21 principal components.

The reason why there is no increase anymore when $P > 20$ is that eigenvector of over 20$^{\text{th}}$ eigenvalue of $(\mathbf{A}-\mathbf{m1})(\mathbf{A}-\mathbf{m1})^T$ doesn't have enough information to represent the differences of data. Eq. (18) shows the eigenvalue matrix of $(\mathbf{A}-\mathbf{m1})(\mathbf{A}-\mathbf{m1})^T$.

$$
\Lambda = \begin{bmatrix}
675938 & 0 & 0 & 0 & & & & & & & & & & \cdots \\
0 & 89526 & 0 & 0 \\
0 & 0 & 17728 & 0 \\
0 & 0 & 0 & 9091.8 \\
& & & & 3532.7 \\
& & & & & 3159.7 & & & & & \ddots \\
& & & & & & 2403.1 \\
& & & & & & & 1809 \\
& & & & & & & & 817.9 \\
& & & & & & & & & 733.5 \\
& & & & & & & & & & 577.9 \\
\vdots & & & & & & & & & & & 403.8 \\
& & & & & & & & & & & & 345.2 \\
& & & & & & & & & & & & & 320.8 \\
& & & & & & & & & & & & & & 157.3 \\
& & & & & & & & & & \ddots & & & & & 138.2 \\
& & & & & & & & & & & & & & & & 118.3 \\
& & & & & & & & & & & & & & & & & 71.9 \\
& & & & & & & & & & & & & & & & & & 54.1 \\
& & & & & & & & & & & & & & & & & & & 48.9 \\
& & & & & & & & & & & & & & & & & & & & 37.6 \\
& & & & & & & & & & & & & & & & & & & & & 34.1 \\
& & & & & & & & & & & & & & & & & & & & & & \ddots
\end{bmatrix} \quad (18)
$$

First eigenvalue in Eq. (18) has the biggest value. It means that the corresponding eigenvector on the first eigenvalue represents the basis of the new space on the biggest variance. The 20th eigenvalue has the pretty smaller value than the first. As we mentioned, the eigenvalue is proportional to the variance of the column vectors of a matrix. Thus, there is no information on the differences among column vectors of a matrix as $P > 20$. Table 3 shows the highest accuracy rate of the proposed method.

Table 3. Accuracy rate of the proposed method

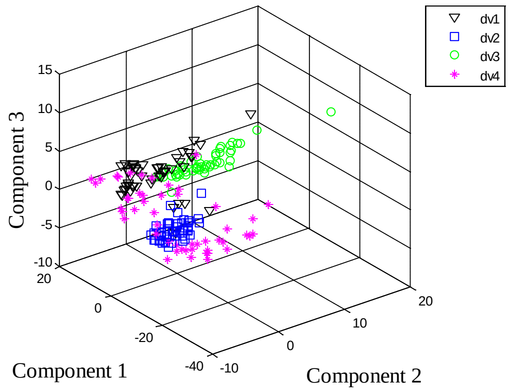
	4 BD-358	4 SL1M	4 BD-358 4 SL1M
	Accuracy rate (Minimum number of PC)		
TR(R)	88% (24)	82% (48)	90.5% (45)
TR(F)	87.5% (45)	90% (12)	92.25% (13)
TR(R + F)	93% (49)	92% (20)	95.5% (63)
Sync	99% (45)	83.5% (22)	93.75% (86)
TR(R + F) + Sync	99% (44)	98.5% (22)	98.75% (21)

R: Rising, F: Falling, PC: Principal components

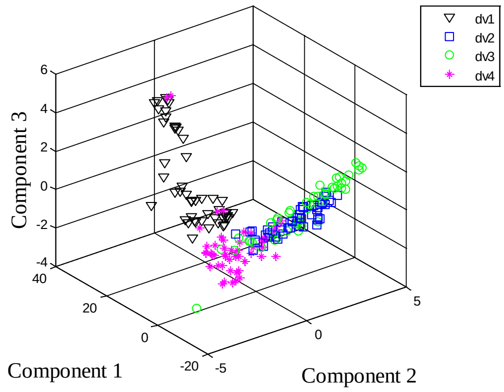
The minimum number of principal components means the minimum number of column vector in $\mathbf{V}$, which is the eigenvector matrix of covariance matrix $(\mathbf{A}-\mathbf{m1})(\mathbf{A}-\mathbf{m1})^T$. It shows that the accuracy rate is getting improved when the additional feature is concatenated.

It is pretty noticeable that the falling transient feature can be used as a unique feature. As we mentioned, the falling transient signal part has not been used in the previous work for RF fingerprinting. According to the Fig. 21 and Table 3, around 90% accuracy rate is recorded when falling transient feature is used as a unique feature. Moreover, the accuracy rate is improved when the falling transient feature is used with the rising transient feature. It shows the possibility that the falling transient feature could have the independent information for RF fingerprinting.

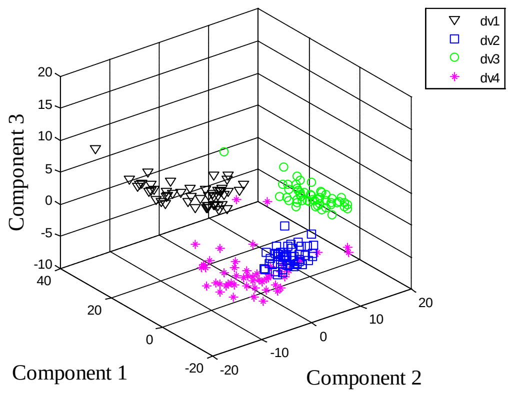
To check how well the cluster for each class is formed, we mapped the principal components of the concatenated feature of each class in ‘SL1M’ model onto the 3 dimensional space. Fig. 22 represents the space.



(a) SL1M transient (rising + falling) features



(b) SL1M sync features



(c) SL1M transient (rising + falling) + sync features

Fig. 22. Feature map of a 3-D principal components space

The meaning of 'Component 1', 'Component 2' *etc* is the P^{th} column vector of the eigenvector matrix $\mathbf{V}$ of the covariance matrix $(\mathbf{A}-\mathbf{m1})(\mathbf{A}-\mathbf{m1})^T$. In Fig. 22, we can see that the clusters are distinctly clustered when the additional feature is concatenated. Because of the distinctly formed clusters, SRC can find the sparse solution well for RF fingerprinting.

To show that fifty sample vectors per a class are enough for the proposed experiment, we did additional experiment. Using thirty sample vectors and forty sample vectors per a class, we did experiment to check how well SRC classify the classes with the less number of sample vectors. First, we performed the test with forty sample vectors per a class. The test proceeds with four cross validation method. After the first additional experiment, we did the second test with thirty sample vectors. In the second test, we used three cross validation method. The result is represented in Fig. 23.

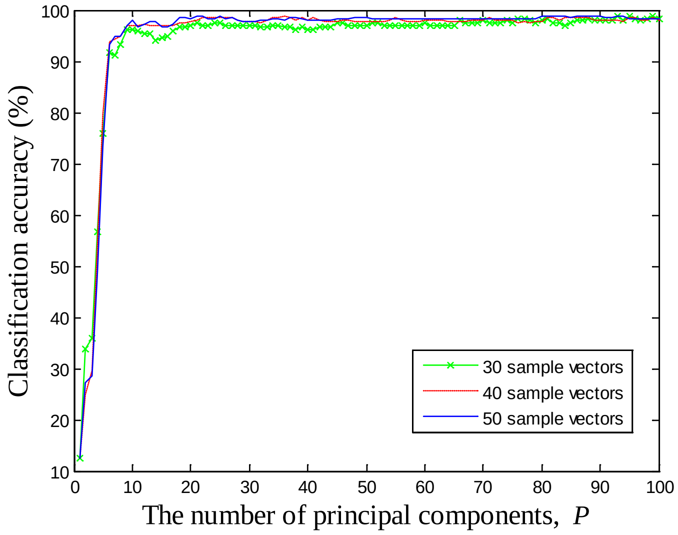


Fig. 23. Classification result when 30 sample vectors, 40 sample vectors, and 50 sample vectors per each class are used

The classification accuracy with the thirty sample vectors is 98.75%. Also, the accuracy with the forty sample vectors is 98.75%. The accuracy rates are same as the fifty sample vectors are used. The difference is the number of principal components. When the thirty sample vectors are used, 98.75% is recorded with 93 principal components. When the forty sample vectors are used, same accuracy is recorded with 36 principal components. Those values are bigger than 21 principal components. Thus, the maximum accuracy can be achieved with the less number of principal components as the number of sample vectors per a class is increased.

As a third additional experiment, we used convolutional neural network (CNN) to compare the result of the method with ours. CNN is widely used for classification or regression as a deep learning method. To construct neural network, we referred to the study [8]. Table 4 shows the structure of layers in CNN.

Table 4. The structure of layers in CNN

Layer	Dimension	Activation
Input	5920×1	-
Convolution 1D	128×19	Exponential linear unit
Max Pooling	2×1	-
Convolution 1D	32×15	Exponential linear unit
Max Pooling	2×1	-
Convolution 1D	16×11	Exponential linear unit
Max Pooling	2×1	-
Fully connected	128	Exponential linear unit
Fully connected	16	Exponential linear unit
Fully connected	7	Softmax

The size of the filter in 1D convolutional layer and pooling layer is same with [8]. However, we set the number of epoch and batch size heuristically from the several experiments. The accuracy rate is determined from five cross validation. Five training data sets are trained separately. As a stride, we set it to 1 on each convolutional layer and pooling layer. Thus, five test data sets enter the five training network we trained, respectively. The accuracy rate of each training data set is represented in Table 5.

Table 5. Accuracy rate of CNN

Data set	Accuracy rate
Test data set 1	78.75%
Test data set 2	90%
Test data set 3	93.75%
Test data set 4	97.5%
Test data set 5	93.75%

The average accuracy of five test data sets is 90.75%. This numerical value is similar with the result of [8] and pretty lower than the proposed method which recorded 98.75%. This comparison shows that SRC can perform RF fingerprinting well with the less number of training data. Deep learning usually requires many training data for the competent result. Since it is difficult to gather many data of military transmitters in electronic warfare, SRC can be useful with the less number of training data sets. Also, SRC is simple to extend the more number of training data and more kinds of features. Thus, the performance of the proposed method could be improved with the many training data sets and features.

5. Conclusion

In this paper, we proposed a multimodal RF fingerprinting scheme based on sparse representation. We showed that the proposed scheme improves accuracy significantly. To use multiple features simultaneously, we concatenated multiple features in a row-wise manner. Then, we applied principal components analysis to the concatenated training data matrix since SRC is sensitive to the mutual correlations between columns of the training data matrix. As features, we used a rising transient feature, a falling transient feature, and a sync feature. Using threshold method, we extracted each signal parts and used main lobe of the parts as a unique feature. With three features, the proposed method recorded 98.75%, which is higher than the convolutional neural network method did. The experiment showed the possibility that the three signal features we have cultivated from RF signals could provide mutually independent information. Especially, it is noticeable that the falling transient feature can be used for RF fingerprinting. Since SRC is simple to extend the more number of training data and kinds of features by concatenating in the row-wise manner, the proposed scheme is efficient for RF fingerprinting in such an electronic warfare or the security of IOT network.

As a future work, we need to do more experiments with a variety of features such as errors in inphase & quadrature (I/Q) domain. Also, phase mismatch distortion, which is generated by the off-beam phase between I/Q, can be useful. Also, we need to use more number of training data. In [4-8], more than five hundred training sample vectors per a class are employed for their experiment. It is meaningful to check how accurate the proposed method will be with many training sample vectors. Lastly, raw data should be used for the experiment since a signal processing like taking operation of absolute value makes information on RF fingerprinting lost. We need to do additional experiments with time signal by Inverse Fast Fourier Transformation of our feature without taking operation of absolute value.

6. Reference

[1] Y. Jia, S. Zhu, and L. Gan, “Specific emitter identification based on the natural measure,” *Entropy*, vol. 19, pp. 117, 2017.

[2] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, “Internet of things for smart cities,” IEEE Internet Things J., vol. 1, no. 1, pp. 22–32, 2014.

[3] X. Wang, P. Hao, and L. Hanzo, “Physical-layer authentication for wireless security enhancement: Current challenges and future developments,” IEEE Commun. Mag., vol. 54, no. 6, pp. 152–158, 2016.

[4] H. J. Patel, M. A. Temple, and R. O. Baldwin, “Improving ZigBee device network authentication using ensemble decision tree classifiers with radio frequency distinct native attribute fingerprinting,” *IEEE Trans. Rel.*, vol. 64, no. 1, pp. 221–233, Mar. 2015.

[5] L. Peng, A. Hu, J. Zhang, Y. Jiang, J. Yu, and Y. Yan, “Design of a Hybrid RF Fingerprint Extraction and Device Classification Scheme,” *IEEE Trans. Internet Things.*, May, 2018 (in press)

[6] Y.Jia, J.Ma, L.Gan, “Combined Optimization of Feature Reduction and Classification for Radiometric Identification,” *IEEE Signal Process. Lett.*, vol. 24, no. 5, pp. 584-588, May, 2017.

[7] Y. Jia, J. Ma, L. Gan, “Radiometric Identification Based on Low-Rank Representation and Minimum Prediction Error Regularization,” *IEEE Commum. Lett.*, vol. 21, no. 8, pp. 1847-1850, Aug, 2017.

[8] K. Merchant, S. Revay, G. Stantchev, and B. Nousain, “Deep learning for Rf device fingerprinting in cognitive communication network,” *IEEE, J.Sel.Topics Signal Process*., vol. 12, no. 1, pp. 160-167, Feb. 2018.

[9] J. Wright, A. Yang, A.Ganesh, S.Sastry, and Y.Ma, “Robust face recognition via sparse representation,” *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 30, no. 2, pp. 210–227, Feb. 2009.

[10] Q. Xu, R. Zheng, W. Saad, and Z. Han, “Device fingerprinting in wireless networks: Challenges and opportunities,” IEEE Commun. Surveys Tuts., vol. 18, no. 1, pp. 94–104, 2016.

[11] ETSI TS 102 361-1 v2.4.1. “Electromagnetic compatibility and Radio spectrum Matters (ERM); Digital Mobile Radio (DMR) Systems; Part 1: DMR Air Interface (AI) protocol,” European Telecommunications Standards Institute, 2016.

[12] Huang, G.; Yuan, Y.; Wang, X.; Huang, Z. Specific Emitter Identification Based on Nonlinear Dynamical Characteristics. Can. J. Electr. Comput. Eng. 2016, 39, 34–41. 21.

[13] Polak, A.C.; Dolatshahi, S.; Goeckel, D.L. Identifying wireless users via transmitter imperfections. IEEE J. Sel. Areas Commun. 2011, 29, 1469–1479.

[14] S. Ur Rehman, K. Sowerby, and C. Coghill, ‘‘RF fingerprint extraction from the energy envelope of an instantaneous transient signal,’’ in Proc. Austral. Commun. Theory Workshop (AusCTW), 2012, pp. 90–95.

[15] I. Kennedy, P. Scanlon, F. Mullany, M. Buddhikot, K. Nolan, and T. Rondeau, “Radio transmitter fingerprinting: A steady state frequency domain approach,” in *Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th*. IEEE, 2008, pp. 1–5.

[16] V. Brik, S. Banerjee, M. Gruteser, and S. Oh, “PARADIS: Physical 802.11 Device Identification with Radiometric Signatures,” Proc. ACM MobiCom, pp. 116-127, Sept. 2008.

[17] D.L.Donoho, “Compressed sensing,” *Information Theory, IEEE Trans. on*, vol. 52, no.4, pp.1289–1306, April 2006.

[18] K. Huang and S. Aviyente, “Sparse Representation for Signal Classification,” *Neural Information Processing Systems*, 2006.

[19] A. M. Bruckstein, D. L. Donoho, and M. Elad, “From sparse solutions of systems of equations to sparse modeling of signals and images,” SIAM Rev., vol. 51, no. 1, pp. 34–81, Feb. 2009.

[20] S. S. Chen, D. L. Donoho, and M. A. Saunders, “Atomic decomposition by basis pursuit,” *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, 2001.

[21] H. Abdi and L. Williams, “Principal component analysis,” *Wiley Interdiscipl. Rev., Comput. Statist.*, vol. 2, no. 4, pp. 433–459, 2010.

[22] M. Zhongli, L. Qianqian, L. Huixin, M. Zhongli, L. Zuoyong, “Image representation based PCA feature for image classification,” *Mechatronics and Automation (ICMA) 2017 IEEE International Conference on*, pp. 1121–1125, 2017.

7. Acknowledgements

입학하던 때가 엊그제 같은데, 벌써 졸업을 앞두게 되었습니다. 처음 연구실에 왔을 때는 우여곡절이 많았지만, 교수님과 연구실 선후배님들의 도움으로 여기까지 올 수 있었습니다. 먼저 연구에 대한 아낌없는 조언과 지원을 해주신 이흥노 교수님께 진심으로 감사의 말씀 드립니다. 교수님께 배운 것들을 항상 되새겨 사회에 도움을 줄 수 있는 공학인이 되도록 노력하겠습니다. 제 대학원 생활과 연구에 대해 많은 도움을 주신 인포넷 연구원 분들 에게도 감사의 말씀 드립니다. 특히 개인적으로 많은 지도를 해주신 주성이형, 재혁이형 정말 감사합니다. 2 년간 이곳에서 동고동락하며 쌓아온 추억들을 잊지 않겠습니다. 아울러, 인포넷 연구원 분들의 앞날에 항상 행복한 일만 있기를 기원하겠습니다. 감사합니다!