WiHRR: CSI based English Alphabet Handwriting Recognition

Aktar Hossen  
Taiyuan University of Technology  
China

Feng Xiufang  
Taiyuan University of Technology  
China

Xiong XiaoQiao  
Taiyuan University of Technology  
China

Zhang Xin  
Taiyuan University of Technology  
China

ABSTRACT
Human Computer Interaction has been an important and vital subject of research since last few years and new and innovative techniques are being implemented in this field with each passing day. This research paper tries to improve existing samples of Channel frequency Response (CFR) that is obtained from Wi-Fi devices in the form of channel state information (CSI). In this research article, handwriting is used as an input but character recognition is not done on the paper but in the air. It discusses all the important procedures which are undertaken to accomplish this task which include fetching the input, pre-processing, feature extraction and classification of characters using machine learning algorithms. Current wireless devices can measure the interference caused by handwritten English alphabet in the range of signal coverage during the transmission of wireless signals in their own channels, and these measured interference s can be applied to handwritten recognition. Compared with traditional behavior recognition using cameras or wearable sensor devices, handwritten recognition using Wi-Fi wireless signals has more unexpected benefits: first, the wireless signals do not need LOS (line of sight) monitoring; second, no need to wear redundant sensor devices; and, at the same time, no need to wear redundant sensor devices. Wireless signals can be deployed in various scenarios without dead corners, such as home, office, public area and so on. Compared with the camera, the behavior recognition based on wireless signal is not affected by the weather.

Keywords
CFR, CSI, Handwriting Recognition, pre-processing, Noise Reduction

1. INTRODUCTION
Since 2010, researchers have modified firmware so that a sample version of CFR (Channel Frequency Response) can also be obtained on ordinary Wi-Fi devices in the form of Channel State Information (CSI). Recently, handwriting recognition approaches has been widely applied to HCI (Human-Computer Interface) applications based on CSI. Handwriting is the basic input mechanism of human-computer interaction. Handwriting recognition has two main research directions: one is to recognize characters written or printed on paper, the other is to recognize characters or symbols written in the air.

Handwritten recognition has a wide range of applications in daily life, such as smart home, human-computer interaction, game world, etc. For example, we can input text to the device through aerial writing, design virtual paper for writing, and even simplify the communication between blind and deaf people.

Many researchers have proposed a variety of systems to recognize writing gestures, which are mainly divided into four categories: vision-based systems, sensor-based systems, hardware-based systems and Wi-Fi-based systems. Vision-based systems use computer vision technology to recognize gestures. However, due to the limitations of LOS and brightness, vision-based systems cannot work in dark environments. In sensor-based systems, we can get high-precision handwriting recognition by wearing special sensors, but it is very inconvenient for users to wear sensors. Existing hardware-based systems mainly use dedicated hardware to recognize gestures. It is unrealistic to deploy multiple sensor nodes and dedicated RF devices to achieve handwriting recognition. Recently, the system based on Wi-Fi utilizes RSSI and CSI from commercial Wi-Fi devices for handwriting recognition. It has many advantages, such as no need to carry equipment, wall scanning, privacy protection, night work and low cost, etc. It provides an opportunity to solve these limitations.

In the past two decades, a large number of applications using signal intensity information to perceive environmental characteristics have emerged, among which RSSI (Received Signal Intensity Indicator) is the most representative one. Because RSSI reflects the quality of channel to a certain extent, many wireless communication technologies such as RFID, FM, GSM, Wi-Fi, ZigBee and so on can be used. RSSI information is acquired on the terminal device, which greatly reduces the cost of our research. In the field of wireless perception, the universality of RSSI makes it widely used to detect the presence and location of indoor people, recognize gestures, human activities and voice, calculate the number of people, and identify human users.

To sum up, for handwriting recognition, the computer is used to observe and learn handwritten actions, and the computer is used to realize the classification and recognition of handwritten actions. This application has broad application prospects and potential value in many fields such as smart home, human-computer interaction, virtual reality and so on. It is also a new frontier research direction with important scientific significance in the world.

2. LITERATURE REVIEW
In recent years, hand gesture recognition has attracted a growing interest due to its applications in many different fields, such as human-computer interaction, robotics, computer gaming, or handwriting recognition in the air and so on. Handwriting recognition in the air gradually is becoming a hot research field in recent year based on CSI and RSSI. The strength of RSSI reflects the quality of the channel, but in indoor environment, RSSI has no longer monotonously decrease with the increase of propagation distance due to small-scale shadow fading caused by signal multiple path propagation, thus limiting the ranging accuracy. Multiple path propagation can also lead to RSSI amplitude fluctuation.
which can also lead to wrong fingerprint matching of wireless signals when locating. In order to characterize multiple path propagation, the channel impulse response (CIR) is usually used to model the wireless channel. Precise measurement of CIR/CFR usually requires professional channel analysis tools. Since 2010, by modifying firmware, a sampled CFR can also be obtained in the form of channel state information (CSI). CSI is a physical-layer information that are available in many commodity network interface cards, such as Intel 5300 and Atheros 9390. It describes the multi-path fading, Doppler shift and scattering effect in the specific scenario and characterizes the quality of the communication link. The MIMO-OFDM (Multiple-Input Multiple-Output/Orthogonal Frequency Division Multiplexing) system in the standard of IEEE 802.11/n consists of multiple transmit-receive TX-RX antenna pairs to increase the capacity of the communication link for subcarrier scale information. The MIMO-OFDM modulation scheme is introduced to ensure finer-grained estimated physical-layer information for the CSI. Recently, a number of researchers have developed diverse technologies for inertial-sensing-based HCI methods such as activity recognition, Hand gesture recognition, Handwriting recognition and motion tracking. Different researcher used different method to recognize hand gesture. Yanmin Zhu and his team proposed Vision Based Hand Gesture Recognition. Vision based approaches do not require the user to wear anything (naked hands). Instead, video camera(s) are used to capture the images of hands, which are then processed and analyzed using computer vision techniques. This type of hand gesture recognition is simple, natural and convenient for users. They involved three important phases: detection and segmentation, tracking, and classification. The detection and segmentation part is to detect hands and segment the corresponding. They used tracking algorithm called TLD (Tracking-Learning-Detection). There was some problem also, such as complex background, partial-occlusion, background disturbance, object reappearance, illumination change and running in real time. This model is so costly. This is the problem of this model. [1] Recent studies have shown that hidden Markov model (HMM) approaches is effective at increasing the recognition rate of the inertial sensing-based handwriting and gesture recognition. [8][9] [10][11]. They have used 6-DOF [10] Support Vector Machine [8][9] and window-based [11] approach to identify those data segments which contain handwriting. There are other usability issues of air-writing from user feedback. The box-writing style appears to be easy to learn, but it needs some practice to write with the specified stroke order. In our current system, writing with different stroke orders can cause errors in recognition, especially for shorter words. This is the basic limitation of those paper,super pixel-based hand gesture recognition system based on a novel super pixel earth mover’s distance metric, together with Kinect depth camera. The depth and skeleton information from Kinect are effectively utilized to produce marker less hand extraction. They faced false matching problem in gesture recognition [14].

Giulio Marin and his team from University of Padova published a paper WITH LEAP MOTION AND KINETIC DEVICES. A set of relevant features is extracted from the data produced by the sensor and fed into a SVM classifier in order to recognize the performed gestures. The hand direction estimation from the Leap Motion software that causes an unreliable matching between the fingertip of these gestures and the corresponding angular region [13] Wenjun Zeng presented a study to recognize handwriting in the air, where he used Kinect sensor. He said, The Kinect effect is transforming human-machine interaction in multiple industries. “Writing in the air with hands” is one such exciting example. [12], [2] presents the first attempt at fusing data from inertial and vision depth sensors within the framework of a hidden Markov model for the application of hand gesture recognition.A Team from IEEE proposed an inertial-sensor-based digital pen (inertial pen) and its associated dynamic time warping (DTW)-based recognition algorithm for handwriting and gesture recognition. The accuracy for recognizing eight gestures reached 96.84% and 100% for user-independent and user-dependent recognition, respectively. [3] Shahad Vikram and his team from University of California presented a method of online character recognition, using a data-driven approach. Their method utilizes similarity search technique on multi-dimensional time series, and they used dynamic time warping (DTW) algorithm to identify characters in these time series. [5][6] proposes an accelerometer-based pen device for online handwriting recognition applications. The accelerometer-based pen device consists of a triaxial accelerometer, a microcontroller, and an RF wireless transmission module. Researcher used dynamic time warping (DTW) algorithm to recognize handwriting on wireless-inertial-measurement-unit (WIMU) which uses hand motion analysis to segment hand motion data from a WIMU device that incorporates magnetic, angular rate, and gravity sensors (MARG). Dynamic time warping (DTW) recognition algorithm is used to recognize handwriting in real-time. [7] Toshio Asano and Sachio Honda from Hiroshima Institute of Technology, Japan proposed – A visual interface system that recognizes handwriting of Japanese katakana characters in the air: a light emission diode (LED) pen and a TV camera are used to capture the LED light trajectory, and the movements of the light are converted into direction codes connecting the slant of the handwriting character. In the experiments, we attained a recognition rate of 92.9% for the single-camera system. [15] “WiG” presents a study a device free gesture recognition system based solely on Commercial Off-The-Shelf (COTS) Wi-Fi infrastructures and devices. WiG used CSI data preprocessing, Birge-Massart filter, LOF anomaly detection, features extraction and SVM classification to get good accuracy rate. WiG can achieve an average recognition accuracy of 92% in line-of-sight scenario and an average accuracy of 88% in the none-line-of-sight scenario. [17] “WiDraw” the first hand motion tracking solution that can be enabled on existing mobile devices using only a software patch. WiDraw leverages physical layer information and multiple antennas on commodity devices to track the detailed trajectory of the user’s hand in both LOS and NLOS scenarios, without requiring the user to touch the device or hold any hardware. [19] [20] proposed LLAP, a device-free gesture tracking scheme that can be deployed on existing mobile devices as software, without any hardware modification. We use speakers and microphones that already exist on most mobile devices to perform device-free tracking of a hand/finger. The key idea is to use acoustic phase to get fine-grained movement direction and movement distance, the measurements Jeen-Shing Wang and his research mate Fang-Chen Chuang developed a pen-type portable device and a trajectory recognition algorithm. The pen-type portable device consists of a triaxial accelerometer, a microcontroller, and an RF wireless transmission module. The overall handwritten digit recognition rate was 98%, and the gesture recognition rate was also 98.75%. researcher faced digits overlapped problem. [21] To recognize whole Home gesture Quan Pu and his team presented a novel gesture recognition system that leverages wireless signals to enable whole-home sensing and recognition of human gestures. Since wireless signals do not require line-of-sight and can traverse walls. Further, it
achieves this goal without requiring instrumentation of the human body with sensing devices. They implemented a proof-of-concept prototype of WiSee using USRP-N210s and evaluate it in both an office environment and a two-bedroom apartment. They achieved accuracy of 94%. WiSee leverages MIMO to cancel the signal from the interfering human, it suffers from the near-far problem that is typical to interference cancellation systems. This was the limitation of this paper. [22]. WiFall uses LOF-based (local outlier factor) and SVM (SVM) Support Vector Machine (SVM) anomaly detection algorithm, judging the occurrence of behavior. It mainly uses seven eigenvalues, including the normalized standard deviation (STD) of CSI. The deviation of signal intensity, time range of motion, absolute deviation of median (MAD), quartile distance (IR), signal entropy and speed of signal change w.[23] presented three different gesture recognition models which are capable of recognizing seven hand gestures, i.e., up, down, left, right, tick, circle, and cross, based on the input signals from MEMS 3-axes accelerometers. The accelerations of a hand in motion in three perpendicular directions are detected by three accelerometers respectively and transmitted to a PC via Bluetooth wireless protocol. show that the best of the three models discussed in this paper achieves an overall recognition accuracy of 95.6%. There are some obstacles or error to recognize handwriting in the air. Many researchers have researched in this field and they faced same problems to recognize handwriting.Li Sun and Dimitrios Koutsoukoulos from university at buffalo said on their paper that prototype achieves satisfactory track high accuracy when the hand is within 2 feet from the receiver. If the hand moves more than 2 feet from the receiver the tracking accuracy will reduce. [19]. This is a large error, another limitation of the system is that, the proposed approach assumes that the acceleration and angular velocity of a hand motion decrease when a user begins and finishes handwriting for spotting handwriting and no handwriting data, which requires users to slow their handwriting speed for fraction of second after each letter in the word to recognize handwritten letter spotted during segmentation process. [7] The CSI data provided by the commercial Wi-Fi network card will inevitably be noisy. To recognize keys with CSI values, noise must be removed from CSI time series. The experimental observation shows that the waveform shapes of different keys are quite different. For this reason, the CSI time series is first fed into the low-pass filter to remove the high-frequency noise [25]. So, noise of commercial Wi-Fi network card is another major problem to recognize handwriting. There is another small error in this filed that prevent to achieve high accuracy rate.

3. METHODOLOGY

3.1 Experimental Set up
Aerial handwriting refers to the writing of language or words by gestures in free space. Many researchers use different methods to study aerial handwriting recognition. In recent years, handwriting recognition based on CSI has been widely used in human-computer interaction. Main processes undertaken in this research include:

- WiHRR is a wireless system, so a common commercial router with three transmit antennas will be used to build the system and act as the transmitter (AP).
- WiHHR also includes another hardware element: a desktop or laptop with Intel 5300 network card as the receiver (DP);

- CSI with finer granularity information will be used to realize aerial handwritten English letter recognition.
- Python is used to analyze CSI data using Google Colab as IDE and machine learning algorithms as classifiers.
- In order to obtain good denoising effect, the wavelet transform filter is selected to denoise CSI data.
- Selection of classifier: select support vector machine (SVM) to classify handwritten actions;
- Laboratories, offices, cafeterias and open spaces will be used for experiments to improve the performance of WiHRR systems.

3.2 Procedure of Task accomplishment
Three way setup is accomplished to perform detailed analysis of handwritten characters in the air.

Firstly, the interference signal is processed. Because of the reflection of static objects such as walls and the interference signal directly from the transmitter to the receiver, how to remove the interference signal is a problem we need to solve.

Secondly, after receiving the signal reflected by the target user, it is necessary to analyze, process and extract appropriate signal features.

Finally, the appropriate classifier is selected for classification. Aerial handwriting recognition technology can be classified from many perspectives. From the perspective of classification mechanism, it can be divided into two categories: fingerprint-based mapping and geometric mapping. From the signal characteristics, there are five main types: energy characteristics, channel state characteristics, frequency characteristics, time characteristics and angle characteristics. In this paper, channel state characteristics and frequency characteristics are selected as the analysis objects of channel state information. After extracting these features, support vector machine is used for further classification.

The ultimate goal of this paper is to achieve the recognition of handwritten English letters in the air, and to achieve good recognition accuracy in both visual and non-visual scenes.

3.3 Tasks Accomplished

3.3.1 Data Collection and Pre-Processing
Firstly, we collect the original CSI information. Because the subtle movements of human body will disturb the signal, our work is to extract these fluctuation characteristics as far as possible and reduce the interference of another external environment.

3.3.2 Noise Removal
Noise removal is the most critical part of recognition and classification, so the selection of appropriate filters and filtering methods is particularly critical. At present, Butterworth filter and PCA denoising are more popular. In addition, there are CSI analysis and processing based on wavelet transform. Wavelet transform can directly play a filtering effect, providing pre-processing for feature extraction, reducing the correlation between different features. Faced with CSI signals with complex signal components, the “zoom” characteristic of wavelet transform ensures high frequency in low frequency band of signal analysis. Rate resolution and low time resolution, as well as low frequency resolution and high time resolution in high frequency band.
Therefore, it is helpful to analyze the local characteristics of CSI signals such as falls by focusing. Because CSI signals belong to typical non-stationary signals, the analysis and processing of CSI signals are the main application fields of wavelet analysis technology. Here we choose the wavelet transform which is most suitable for our CSI signal characteristics.

3.3.3 Feature Extraction
We extract the amplitude, phase, the second component of PCA, the third component and the frequency characteristics of the data from the pre-processed CSI information.

3.3.4 Time Frequency Domain Analysis
Through time-frequency domain analysis, the signal is transformed from time domain to frequency domain, which is convenient for subsequent determination.

3.3.5 Classification
Neural Network is used to construct the classification model. Finally, the classification results of different features are compared and a conclusion is drawn.

4. IMPLEMENTATION AND RESULTS
Principal Component Analysis (PCA) has been the main processing algorithm used in this experiment for feature extraction and Hidden Markov Models are used for the purpose of classification. PCA is also used for the purpose of noise removal in this data-set. In our experiment, the whole system is composed of a TP-Link router which serves as the Access Point (AP) and a laptop which serves as the Detect Point (DP). Both of them are off-the-shelf commercial devices. The router is TP-Link WRD6300. The laptop is Dell M411R with 4GB RAM, equipped with Intel 5300 802.11n NIC and installed with OS Ubuntu 12.04 LTS. The AP and DP have two and three external omnidirectional antennas, respectively. The CSI values is provided by the modified driver of the NIC, which is developed by [22]. Our experiment is conducted in a university laboratory, which is 7.5m in length and 4.5m in width. We choose five right-handed students (two females and three males) aged 20-23 as volunteers to collect the CSI data. All letter are written with the index finger on or above the appointed area of the white board, which is located beside the TX-RX pair. As when Hidden Markov models are applied, we see that as number of samples are increasing so is the overall accuracy of the result.

We used a sentence “The five boxing wizards jump quickly.”. This sentence has one quality that it contains all 26 letters and picked up 8 different people who could write it in the air in order to detect the accuracy. Apart from that we have also calculated the results of individual accuracy of alphabets with different number of samples. Table 1 shows the exact accuracies of each alphabet when number of samples are 30, 70 and 100 while Figure 1, Figure 2 and Figure 3 show its graphic representation.

Table 1. Accuracy of all alphabets with PCA with variety in the number of samples.

<table>
<thead>
<tr>
<th>Alphabet</th>
<th>Accuracy when 30 Samples Taken</th>
<th>Accuracy when 70 Samples Taken</th>
<th>Accuracy when 100 Samples Taken</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>70</td>
<td>80</td>
<td>85</td>
</tr>
<tr>
<td>B</td>
<td>66</td>
<td>77</td>
<td>81</td>
</tr>
<tr>
<td>C</td>
<td>59</td>
<td>68</td>
<td>72</td>
</tr>
<tr>
<td>D</td>
<td>78</td>
<td>81</td>
<td>89</td>
</tr>
<tr>
<td>E</td>
<td>70</td>
<td>80</td>
<td>85</td>
</tr>
<tr>
<td>F</td>
<td>66</td>
<td>77</td>
<td>81</td>
</tr>
<tr>
<td>G</td>
<td>59</td>
<td>68</td>
<td>72</td>
</tr>
<tr>
<td>H</td>
<td>78</td>
<td>81</td>
<td>89</td>
</tr>
<tr>
<td>I</td>
<td>70</td>
<td>80</td>
<td>85</td>
</tr>
<tr>
<td>J</td>
<td>66</td>
<td>77</td>
<td>81</td>
</tr>
<tr>
<td>K</td>
<td>59</td>
<td>68</td>
<td>72</td>
</tr>
<tr>
<td>L</td>
<td>78</td>
<td>81</td>
<td>89</td>
</tr>
<tr>
<td>M</td>
<td>70</td>
<td>80</td>
<td>85</td>
</tr>
<tr>
<td>N</td>
<td>66</td>
<td>77</td>
<td>81</td>
</tr>
<tr>
<td>O</td>
<td>59</td>
<td>68</td>
<td>72</td>
</tr>
<tr>
<td>P</td>
<td>78</td>
<td>81</td>
<td>89</td>
</tr>
<tr>
<td>Q</td>
<td>70</td>
<td>80</td>
<td>85</td>
</tr>
<tr>
<td>R</td>
<td>66</td>
<td>77</td>
<td>81</td>
</tr>
<tr>
<td>S</td>
<td>59</td>
<td>68</td>
<td>72</td>
</tr>
<tr>
<td>T</td>
<td>78</td>
<td>81</td>
<td>89</td>
</tr>
<tr>
<td>U</td>
<td>70</td>
<td>80</td>
<td>85</td>
</tr>
<tr>
<td>V</td>
<td>66</td>
<td>77</td>
<td>81</td>
</tr>
<tr>
<td>W</td>
<td>59</td>
<td>68</td>
<td>72</td>
</tr>
<tr>
<td>X</td>
<td>78</td>
<td>81</td>
<td>89</td>
</tr>
<tr>
<td>Y</td>
<td>59</td>
<td>68</td>
<td>72</td>
</tr>
<tr>
<td>Z</td>
<td>78</td>
<td>81</td>
<td>89</td>
</tr>
</tbody>
</table>
Figure 1: Accuracy of all Alphabets with PCA when 30 samples are taken

Figure 2: Accuracy of all Alphabets with PCA when 70 samples taken
It is clearly evident from tables and graphs that more number of samples are tested, more will be the accuracy but there can be certain exceptions for example when samples are increased from 30 to 70, accuracy of character recognition decreases from 78% to 69%. This has occurred due to overloading. It is a phenomenon that takes place when data input exceeds the capacity of the processor but there are multiple ways to tackle that. Apart from this we performed one more type of testing. We used a sentence that contained all 26 alphabetical letters and arranged 8 users to give their input in the air. Their accuracy measures are stored in Table 2 and Figure 4.

Table 2. Accuracy measures of 8 users on a real-time sentence containing 26 alphabets

<table>
<thead>
<tr>
<th>User</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>U1</td>
<td>0.76</td>
</tr>
<tr>
<td>U2</td>
<td>0.79</td>
</tr>
<tr>
<td>U3</td>
<td>0.67</td>
</tr>
<tr>
<td>U4</td>
<td>0.76</td>
</tr>
<tr>
<td>U5</td>
<td>0.65</td>
</tr>
<tr>
<td>U6</td>
<td>0.68</td>
</tr>
<tr>
<td>U7</td>
<td>0.61</td>
</tr>
<tr>
<td>U8</td>
<td>0.79</td>
</tr>
</tbody>
</table>

5. CONCLUSION

The contributions of this paper include a novel handwriting recognition framework without any specialized sensors. We collect the data and then denoise them using Principal Component algorithm, in which we remove the noises. Our main intention here is to recognize 26 capital letters with the help of 30, 70 and 100 samples and calculate the accuracy in each case. Hidden Markov model is used to achieve the purpose of classification in which we use a sentence which can possibly contain all the 26 capitalized English letters. We used 8 different people and calculated the accuracy graphs.

In the future, capital as well as small letters can also be included to achieve the purpose of handwriting recognition and many more classifiers can also be used such as SVM, DNN, CNN, RNN etc. It can also be implemented on other languages such as Arabic, Chinese etc.
REFERENCES

[26] Yuxi Wang, Student Member, IEEE, Kaishun Wu, Member, IEEE, and Lionel M. Ni. WiFall: Device-free Fall Detection by Wireless Networks.IEEE,2014
[28] Zhangjie Fu, Jiashuang Xu, Zhuangdi Zhu, Alex X. Liu and Xingming Sun. Writing in the Air with WiFi Signals for Virtual Reality Devices.IEEE,2018
[31] Zhangjie Fu, Jiashuang Xu, Zhuangdi Zhu, Alex X. Liu and Xingming Sun. Writing in the Air with WiFi Signals for Virtual Reality Devices.IEEE,2018


IfCA™: www.ijcaonline.org 37