A survey on WiFi Channel State Information (CSI) utilization in Human Activity Recognition

Working Paper · May 2016

DOI: 10.13140/RG.2.1.4383.5121

Mustafa Aljumaily
University of Tennessee

A survey on WiFi Channel State Information (CSI) utilization in Human Activity Recognition

Mustafa Aljumaily

Department of Electrical Engineering and Computer Science

The University of Tennessee, Knoxville

mlatief@vols.utk.edu

Abstract

Human activity recognition and classification has been considered a core technology that enables variety of applications such as the health care of elder and sick people, smart homes, fitness tracking, and building surveillance. However, accurate detection, recognition, and classification of human activities is still a big challenge that attracts a lot of research efforts due to the problems related to the human body parts and the difficulty in detecting their actions accurately, human clothes and their negative affect on the detection accuracy, and the surrounding environment conditions. Vision based human activity analysis using computer vision and machine learning as the original solution is still having its limitations that are related to the inability to detect whatever happening behind the walls or in the dark places and the uncomfortable feeling of people with cameras everywhere. Recent advances in the wireless technology gave new solutions to tackle these problems as it has been proven that the movement of human body parts will affect the channel state information (CSI) of wireless signals in the indoor environments. The advantages of this technique is the ability to work in dark, non-line of sight detection and working without affecting human daily life activities like the cameras do in the vision-based systems. This survey paper provides an overview of the basics and applications of wireless CSI and the recent research efforts to utilize the it in track, detect, and classify human's actions and activities in the indoor environments. This would help readers to have an overview of this topic and the existing applications and solutions presented of the wireless CSI in the human activity recognition problem.

1 Introduction

Human Computer Interaction (HCI) is a broad topic that includes a lot of fields of interest and one of them is the human activity and gesture recognition [21]. Traditional solutions that use vision-based human activity analysis in an attempt to understand the movements of human body using computer vision and machine learning techniques have their limitations [19]. Some of these limitations and challenges that are still facing researchers include: (a) Body parts or big size obstacles may cause partial obstruction that prevents accurate detection; (b) An action, observed from different viewpoints, has different appearances; (c) Clothing, especially long skirts, may lead to apparent differences in body parts and their proportional sizes; (d) The start-time and end-time points of an action are sometimes hard to detect accurately; (e) Dynamic backgrounds may make it difficult to locate and observe actions for vision based systems; (f) Smoke-filled, dim, or dark rooms may make it hard to observe actions; (g) People may feel uncomfortable with a camera overhead, especially in a bathroom [19]. WiFi based systems have been proved to give solutions for such problems especially with the recent advances in the wireless communications and the utilization of the CSI of the wireless signals.

The purpose of this paper is to give the readers a general overview on the research efforts on the utilization of the Channel State Information (CSI) of the wireless signals in the different applications of the human activity and gesture recognition starting from the keystroke recognition and ending with the general activity recognition. It surveys the recent applications of the CSI and classifies them based on the type of applications and their benefits.

The rest of this paper is organized as follows. Section II provides some background information about the meaning of the CSI and its emerging applications. Section III discusses research topics on the utilization of the CSI in human activity recognition. Section IV concludes the paper.

2 Background

2.1 WiFi Channel State Information (CSI)

CSI refers to channel properties in wireless communications [1]. CSI describes how a signal propagates from the transmitter to the receiver, and reveals a set of channel measurements depicting the amplitudes and phases of every subcarrier as shown in equation (1):

$$H(f_k) = \left||H(f_k)|\right| e^{jsin(\angle H)} \qquad (1)$$

Where H(fk) is the CSI value at the subcarrier with central frequency of fk, and $\angle H$ is the phase. In general, the receiver evaluates and quantitates CSI, then makes feedback to the sender (a time-division duplex system often needs reverse evaluation). In real application, CSI can be divided into instantaneous CSI and statistical CSI. Also, there is another classification for the CSI as the CSI in the transmitter (SCIT) and the CSI at the receiver (CSIR).

WiFi NICs continuously monitor variations in the wireless channel using CSI, which characterizes the frequency response of the wireless channel [1]. Let X(f, t) and Y(f, t) be the frequency domain representations of transmitted and received signals, respectively, with carrier frequency f. The two signals are related by the expression:

Y (f, t) = H(f, t)× X(f, t), where H(f, t) is the complex valued channel frequency response (CFR) for carrier frequency f measured at time t. CSI measurements basically contains these CFR values. Let N_{Tx} and N_{Rx} represent the number of transmitting and receiving antennas, respectively. As CSI is measured on 30 selected OFDM subcarriers for a received 802.11 frame, each CSI measurement contains 30 matrices with dimensions $N_{Tx} \times N_{Rx}$. Each entry in any matrix is a CFR value between an antenna pair at a certain OFDM subcarrier frequency at a particular time. Therefore, the time-series of CFR values for a given antenna pair and OFDM subcarrier is called CSI stream. Thus, there are $30 \times N_{Tx} \times N_{Rx}$ CSI streams in a time-series of CSI values.

2.2 Applications of WiFi CSI

Channel State Information (CSI) of the wireless signals are available in many commercial devices like the Intel 5300 [2] and the Atheros 9390 network interface cards (NIC) [3]. These information have a large spectrum of applications and utilizations in many fields like the use of CSI in the human activity recognition [4-8] (which is the main focus of this paper). CSI also used in indoor localization [3,9], human activity of falling [4], the detection of the presence of a human in a room or building [8] and even in the application of counting the number of people in a crowd [7]. Another types of activity recognition applications include recognizing the spoken words by using a specialized directional antennas to obtain the necessary CSI variations caused by lips movements during speaking [5]. CSI can also be used to identify in home daily activities using the E-eye system as example [6]. Wikey [10], Wi-Vi [11], WiHear [5], and WiDraw [13] are all just other examples of the versatile applications that can make use of the CSI.

3 Research Topics of WiFi CSI Utilization in Human Activity Recognition

3.1 General Human Activity Recognition.

Human Computer Interaction (HCI) is increasing very fast both in quality and in quantity in our daily life. The old solutions to control computing devices and smart appliances in the smart environments (smart home, smart offices …etc.) included using cameras [14] that track human movements and try to classify and recognize these movements by comparing the collected data (pictures) to an already stored patterns. The other traditional way of recognizing human activities included Radar [15] or wearable devices (like sensors or smart watches) [16, 17]. However, camera based approaches have the fundamental limitations of requiring line of sight (LoS) with enough lighting and the potential privacy breaching of humans. On the other hand, Low cost 60 GHz radar solutions have limited operation range of just tens of centimeters [15]. Also, wearable sensors based approaches are inconvenient sometimes because the users have to wear them everywhere and they are not convenient especially for elders who wear these sensors to monitor their health status but usually forget to wear them. The WiFi CSI usage for human activity recognition has the advantages over the other ways of activity recognition in different ways as it does not require lighting like the camera based approaches, it provides better coverage as it can operate through walls, it preserves user privacy (unlike the cameras), and it does not require users to carry any devices as it relies on the WiFi signals reflected by humans [18].

[18] Uses two commercial off-the-shelf (COTS) WiFi devices to recognize the human activities as shown in figure (1):

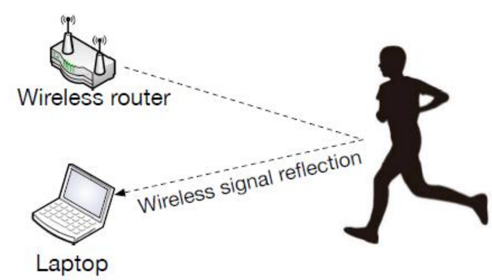

Figure (1)

The laptop is used for continuous receiving signals and the router is used for continuous sending signals. The system is used to detect any human activity happening within the range of these two devices using two theoretical underpinnings: CSI-speed model and CSI-activity model.

[19] Utilized the off-the-shelf Intel 5300 Network Interface Cards (NICs) and a modified driver to reveal a group of sampled versions of Channel Frequency Responses (CFRs) within the WiFi bandwidth to upper layers in the format of

the CSI. Using an open CSI Tool [2] built on the Intel WiFi Wireless Link 5300 802.11n and the MIMO radios with open source Linux wireless drivers, they used that tool to gather CSIs. An $Ntx \times Nrx \times 30$ matrix is taken as the data structure of the CSI, where the third dimension is across 30 subcarriers in the Orthogonal Frequency Division Multiplexing (OFDM) channel. In an Intel 5300 NIC, there is only one transmitting terminal and three receiving ends, so it is a 1×3 MIMO system. They aggregate 30 subcarriers' CSI values into one single value by their average for each MIMO plot. Their work included training stage, followed by features selection and classification and recognition process for the system before using it. Then the real application process included filtering the collected CSI data, pattern segmentation (both online and offline segmentation), feature extraction, and finally activity classification.

[20] Presents device-free location-oriented activity identification at home through the use of existing WiFi Access Points (AP) and WiFi devices (e.g., desktops, thermostats, refrigerators, smart TVs, laptops) and call it E-eyes. Their low-cost system takes advantage of the abundance of WiFi links between such devices and the increasingly fine-grained channel state information that can be extracted from such links. It examines channel features and can uniquely identify both in-place activities and walking movements across a home by comparing them against signal profiles. Signal profiles construction can be semi-supervised and the profiles can be adaptively updated to accommodate the movement of the mobile devices and day-to-day signal calibration.

E-eyes system idea includes using the existing channel state information (CSI) provided by IEEE 802.11n devices and relatively few wireless links, such as those to existing in-home WiFi devices as shown in the figure (2):

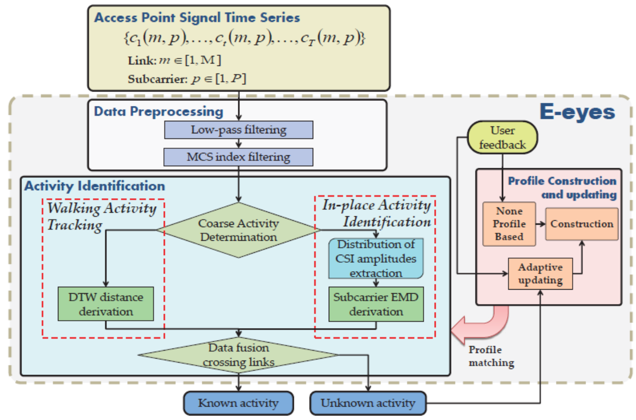

Figure (2) [20].

3.2 Keystroke Recognition Using WiFi Signals

One of the most recent applications of the WiFi information is to recognize the keystrokes of human using two (COTS) devices which are a laptop and a router as shown in figure (3).

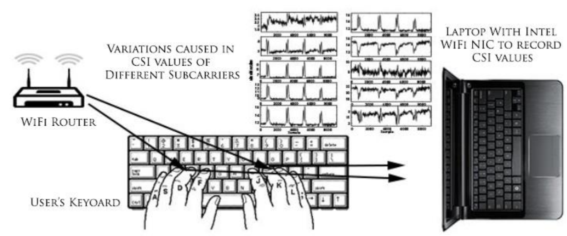

Figure (3).

[10] Shows that WiFi signals can be exploited to recognize keystrokes. The idea behind that is the fact that while typing a certain key, the hands and the fingers of a user move in a unique formation and direction that generates a unique pattern in the time series of the channel state information (CSI) values that they call it the CSI-waveform for that key. The keystrokes of each key introduce relative unique multi-path distortions in WiFi signals and this uniqueness can be exploited to recognize keystrokes. Due to the high data rates supported by modern WiFi devices, WiFi cards provide enough CSI values within the duration of a keystroke to construct a

high resolution CSI-waveform for each keystroke.

In this system, the sender (the router) continuously emits signals and the receiver (the laptop) continuously receives signals. When a human subject types in a keyboard, on the WiFi signal receiver end, WiKey recognizes the typed keys based on how the CSI value changes. CSI values quantify the aggregate effect of wireless phenomena such as fading, multi-paths, and Doppler shift on the wireless signals in a given environment. When the environment changes, such as a key is being pressed, the impact of these wireless phenomena on the wireless signals change, resulting in unique changes in the CSI values.

Other similar systems include [22] which leverage the CSI and RSS information from COTS WiFi devices to classify four arm gestures - push, pull, lever and punch. The fundamental difference between this scheme and the keystroke recognition system is that this scheme extract coarse grained features from the CSI values provided by the COTS WiFi NIC to perform this task while the earlier scheme refines these CSI to capture fine grained variations in the wireless channel for recognizing keystrokes.

3.3 See through Walls with WiFi

The other type of applications of the wireless signals information include seeing through the walls and one example of it is the Wi-Vi [11]. This system is a wireless device that captures moving objects behind a wall. It leverages the ubiquity of WiFi chipsets to make through wall imaging relatively low-power, low-cost, low-bandwidth, and accessible to average users. To this end, Wi-Vi uses WiFi OFDM signals in the ISM band (at 2.4 GHz) and typical WiFi hardware. Wi-Vi is essentially a 3-antenna MIMO device: two of the antennas are used for transmitting and one is used for receiving. It also employs directional antennas to focus the energy toward the wall or room of interest. Its design incorporates two main components:

1) The first component eliminates the flash reflected off the wall by performing MIMO nulling.

2) The second component tracks a moving object by treating the object itself as an antenna array using a technique called Inverse Synthetic Aperture Radar (ISAR).

Wi-Vi can be used in one of two modes, depending on the user’s choice. In mode 1, it can be used to image moving objects behind a wall and track them. In mode 2, on the other hand, Wi-Vi functions as a gesture-based interface from behind a wall that enables humans to compose messages and send them to the Wi-Vi receiver.

The other example of such system is the system proposed in [12] that is also aiming to see behind the walls but this system required both the transmitter and a reference receiver to be inside the imaged room. Furthermore, the reference receiver in the room has to be connected to the same clock as the receiver outside the room which are limitations that the system in [11] does not have.

3.4 Reading human lips using WiFi signals

Some recent research efforts have been focusing of answering the question: can WiFi signals hear us?

And the system in [5] WiHear (WiFi Hearing), explores the potential of using WiFi signals to hear people talk and transmit the talking information to the detector at the same time. They suggested the following potential applications for the proposed system:

1) WiHear introduces a new way to hear people talks without deploying any acoustic sensors. Further, it still works well even when the surrounding environment is noisy.

2) WiHear will bring a new interactive interface between human and devices, which enables devices to sense and recognize more complicated human behaviors (e.g. mood) with negligible cost.

3) WiHear can help millions of disabled people to conduct simple commands to devices with only mouth motions instead of complicated and inconvenient body movements.

WiHear locates the mouth of an individual, and then recognizes his words by monitoring the signal reflections from his mouth. By recognizing mouth moving patterns, WiHear can extract talking information the same way as lip reading. Thus, WiHear introduces a micro-motion detection scheme that can differentiate among different mouth movements with different letters as shown in figure (4) [5].

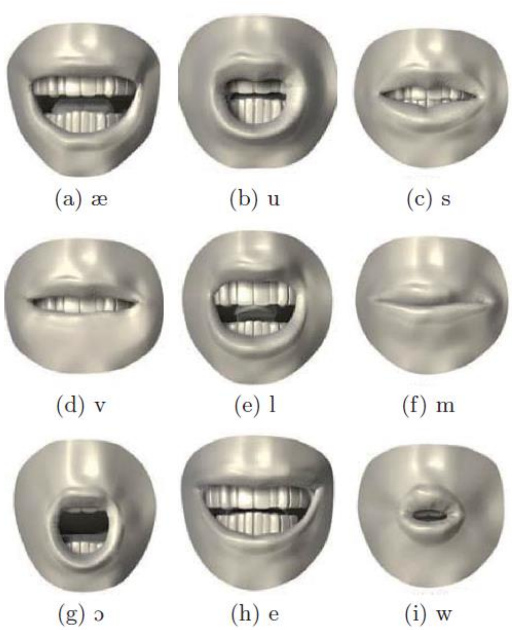

Figure (4): Illustration of vowels and consonants that WiHear can detect and recognize [5].

WiHear framework includes filtering, partial multipath removal, wavelet transform, segmentation, feature extraction, and finally classification and error correction as shown in the figure (5) below [5]:

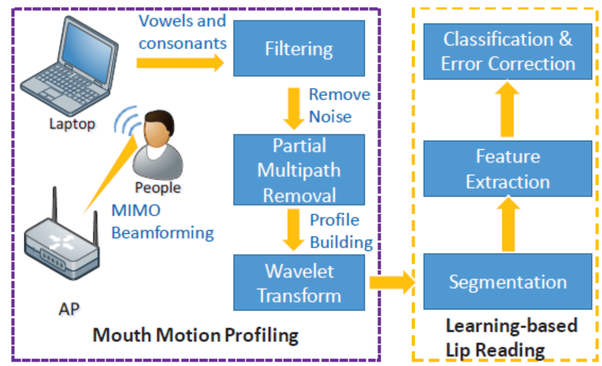

Figure (5)

The other examples of such application was introduced in [27] that leverages radio reflections from human bodies to detect, track, and recognize human body parts motions.

3.5 Hands-Free Drawing in the Air using WiFi Signals

Unlike the Kinect [23] and Leap motion [24], the system proposed in [13] does not need special hardware to track the human hands movements to draw in the air. WiDraw in [13] leverage WiFi signals from commodity mobile devices to enable hands-free drawing in the air. This system introduced 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 Line-Of-Sight (LOS) and Non-LOS scenarios, without requiring the user to touch the device or hold any hardware. By using *WiDraw*, a user can draw arbitrary lines, curves, or even alphabetical characters, simply by using hand motions in the air as shown in figure (6).

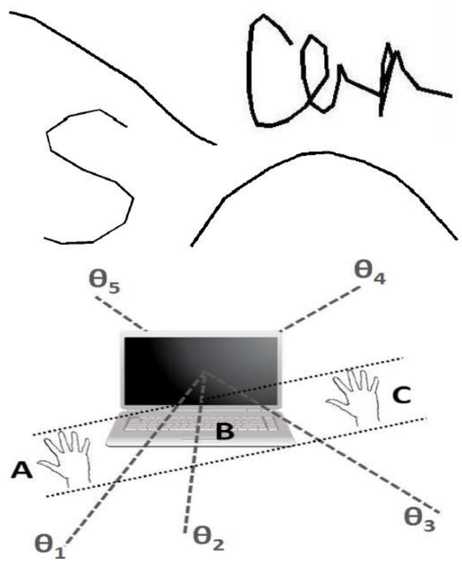

Figure (6): WiDraw's estimated trajectory when the user drew a straight line, a circular curve, the letter "S", and the word "can" and the user's hand perturbs the signal strength of several AoAs along its trajectory.

WiDraw utilizes the Angle-of-Arrival (AoA) values of incoming wireless signals at the mobile device to track the user's hand trajectory.

Other Existing techniques such as MUSIC [25] can utilize the CSI from multiple antennas to estimate the angles at which a wireless signal from a transmitter arrives at the receiver

Also, WiFi Gestures [22] and WiGest [26] are other works that perform WiFi-based gesture recognition using RSSI and CSI information from COTS devices. However, they also rely on a priori learning, and as a result, they can only classify a few simple gestures that they were originally trained on.

4 Conclusion

This paper presents a comprehensive survey on the utilization of the Channel State Information (CSI) of the wireless signals in human activity and gesture recognition. We introduce the CSI basic concept and the research topics of its utilization in many indoor applications. The references are selected based on our limited knowledge of this topic and according to the requirements of the Qualifying Exam.

References

[1] Rappaport, T.S. Wireless Communications, Principles and Practice; Prentice Hall: Bergen, NJ, USA, 1996.

[2] D. Halperin, W. Hu, A. Sheth, and D. Wetherall. Tool release: Gathering 802.11n traces with channel state information. ACM SIGCOMM CCR, 41(1):53, 2011.

[3] S. Sen, J. Lee, K.-H. Kim, and P. Congdon. Avoiding multipath to revive inbuilding WiFi localization. In Proceeding of ACM MobiSys, pages 249–262, 2013.

[4] C. Han, K. Wu, Y. Wang, and L. M. Ni. Wifall: Device-free fall detection by wireless networks. In Proceedings of IEEE INFOCOM, pages 271–279, 2014.

[5] G. Wang, Y. Zou, Z. Zhou, K. Wu, and L. M. Ni. We can hear you with WiFi! In Proceedings of ACM MobiCom, 2014.

[6] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu. E-eyes: In-home device-free activity identification using fine-grained WiFi signatures. In Proceedings of ACM MobiCom, 2014.

[7] W. Xi, J. Zhao, X.-Y. Li, K. Zhao, S. Tang, X. Liu, and Z. Jiang. Electronic frog eye:

Counting crowd using WiFi. In Proceedings of IEEE INFOCOM, 2014.

[8] Z. Zhou, Z. Yang, C. Wu, L. Shangguan, and Y. Liu. Towards omnidirectional passive human detection. In Proceedings of IEEE INFOCOM, pages 3057–3065, 2013.

[9] Z. Yang, Z. Zhou, and Y. Liu. From RSSI to CSI: Indoor localization via channel response. ACM Computing Surveys, 46(2):25, 2013.

[10] Ali, Kamran, Alex Xiao Liu, Wei Wang, and Muhammad Shahzad. "Keystroke Recognition Using WiFi Signals." Proceedings of the 21st Annual International Conference on Mobile Computing and Networking - MobiCom '15 (2015).

[11] Adib, F., & Katabi, D. See through walls with WiFi! Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM - SIGCOMM '13. 2013.

[12] K. Chetty, G. Smith, and K. Woodbridge. Through-the-wall sensing of personnel using passive bistatic WiFi radar at standoff distances. IEEE Trans. Geoscience and Remote Sensing, 2012.

[13] Sun, L., Sen, S., Koutsonikolas, D., & Kim, K. (2015). WiDraw: Enabling Hands-Free Drawing in the Air on Commodity WiFi Devices. Proceedings of the 21st Annual International Conference on Mobile Computing and Networking - MobiCom '15. 2015.

[14] J. K. Aggarwal and S. R. Michael. Human activity analysis: A review. ACM Computing Surveys, 43(3), 2011.

[15] Google project Soli.

https://www.youtube.com/watch?v=0QNiZfSsPc0.

[16] E. Ertin, N. Stohs, S. Kumar, A. Raij, M. al'Absi, and S. Shah. AutoSense: unobtrusively wearable sensor suite for inferring the onset, causality, and consequences of stress in the field. In Proceedings of ACM Sensys, 2011.

[17] K. Yatani and K. N. Truong. Bodyscope: a wearable acoustic sensor for activity recognition. In Proceedings of ACM UbiComp, pages 341–350, 2012.

[18] Wang, Wei, Alex X. Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. "Understanding and Modeling of WiFi Signal Based Human Activity Recognition." Proceedings of the 21st Annual International Conference on Mobile Computing and Networking - MobiCom '15 (2015).

[19] Wang, Yi, Xinli Jiang, Rongyu Cao, and Xiyang Wang. "Robust Indoor Human Activity Recognition Using Wireless Signals." Sensors 15.7 (2015): 17195-7208.

[20] Wang, Yan, Jian Liu, Yingying Chen, Marco Gruteser, Jie Yang, and Hongbo Liu. "E-eyes." Proceedings of the 20th Annual International Conference on Mobile Computing and Networking - MobiCom '14 (2014).

[21] He, Wenfeng, Kaishun Wu, Yongpan Zou, and Zhong Ming. "WiG: WiFi-Based Gesture Recognition System." 2015 24th International Conference on Computer Communication and Networks (ICCCN) (2015).

[22] Rajalakshmi Nandakumar, Bryce Kellogg, and Shyamnath Gollakota. WiFi gesture recognition on existing devices. arXiv preprint arXiv:1411.5394, 2014.

[23] Microsoft kinect.

http://www.microsoft.com/en-us/kinectforwindows/.

[24] Leap motion, inc. leap motion: Mac & pc gesture controller for game, design and more.

https://www.leapmotion.com/.

[25] Ralph O Schmidt. Multiple emitter location and signal parameter estimation. IEEE Transactions on Antennas and Propagation, AP-34(3):276–280, 1986.

[26] Heba Abdelnasser, Moustafa Youssef, and Khaled Harras. Wigest: A ubiquitous wifi-based gesture recognition system. In *Proceedings of IEEE INFOCOM*, 2015.

[27] M. Scholz, S. Sigg, H. R. Schmidtke, and M. Beigl. Challenges for device-free radio-based activity recognition. In Workshop on Context Systems, Design, Evaluation and Optimisation, 2011.