Article

Channel Sounding and Scene Classification of Indoor 6G Millimeter Wave Channel Based on Machine Learning

Liang Yin *, Ruonan Yang and Yuliang Yao

Beijing Laboratory of Advanced Information Networks, Beijing University of Posts and Telecommunications, Beijing 100876, China; yang_ruonan@bupt.edu.cn (R.Y.); yaoyuliang@bupt.edu.cn (Y.Y.)
* Correspondence: yinl@bupt.edu.cn; Tel.: +86-010-6228-2760

Abstract: Millimeter wave, especially the high frequency millimeter wave near 100 GHz, is one of the key spectrum resources for the sixth generation (6G) mobile communication, which can be used for precise positioning, imaging and large capacity data transmission. Therefore, high frequency millimeter wave channel sounding is the first step to better understand 6G signal propagation. Because indoor wireless deployment is critical to 6G and different scenes classification can make future radio network optimization easy, we built a 6G indoor millimeter wave channel sounding system using just commercial instruments based on time-domain correlation method. Taking transmission and reception of a typical 93 GHz millimeter wave signal in the W-band as an example, four indoor millimeter wave communication scenes were modeled. Furthermore, we proposed a data-driven supervised machine learning method to extract fingerprint features from different scenes. Then we trained the scene classification model based on these features. Baseband data from receiver was transformed to channel Power Delay Profile (PDP), and then six fingerprint features were extracted for each scene. The decision tree, Support Vector Machine (SVM) and the optimal bagging channel scene classification algorithms were used to train machine learning model, with test accuracies of 94.3%, 86.4% and 96.5% respectively. The results show that the channel fingerprint classification model trained by machine learning method is effective. This method can be used in 6G channel sounding and scene classification to THz in the future.

Keywords: 6G channel sounding; channel scene classification; machine learning; power delay profile

1. Introduction

With limited spectrum resources for the fifth generation (5G) mobile networks, industries urgently need more millimeter wave band [1]. High frequency millimeter wave near 100 GHz to THz is one of the key spectrum resources for the sixth generation (6G) mobile communications system [2]. The wireless signal transmission performance between base stations and mobile stations is mainly determined by wireless channel. 6G will face serious signal occlusion problems, because high-frequency millimeter wave transmits in a straight-line way. Modeling and classifying millimeter wave channel can facilitate cellular communication network design, which is also premise of 6G’s actual deployment. The key to solving this problem is to sound the new high-frequency millimeter wave wirelessly, and then obtain accurate channel impulse response. Wireless channel in different scenes has different characteristics, especially at indoor scenes, which have many obstructions. Therefore, the classification of different wireless transmission scenes is conducive to wireless network optimization [3].

At present, there are two main methods in millimeter wave channel sounding: time domain method and frequency domain method [4].

In references [5–12], time-domain correlation methods were used to measure and model the millimeter wave channel. However, the frequency range of these channel measurements belongs to lower frequency and does not cover the high frequency required by 6G. Reference [5] stated that Beijing Keysight Company carried out a measurement...
of broadband channel characteristics in environment of line-of-Sight (LoS) and non-line-of-Sight (NLoS) with 26 GHz frequency band and 1 GHz bandwidth in an open office. In Reference [6], a 28 GHz channel detector based on an auto rotating horn antenna was proposed. The detector had the ability of time synchronization between the synchronized transmitter and receiver. Relative propagation delay from transmitter to receiver by using the time stamp was recorded by the proposed channel detection system. Reference [7] studied the indoor channel parameters and relevant characteristics by experimental measurements containing LOS and NLOS scenes in a typical corridor at 60 GHz. Reference [8] had measured shadow effects of moving human body in 73 GHz whose transmitter and receiver systems use directional antennas. Based on the double-edged diffraction model, a simple human blocking model was proposed. Reference [9] proposed a modeling method that combined statistical modeling and deterministic modeling with channel attributes extracted from channel data. Combined with the development of big data, Reference [9] also proposed a cluster core based model which integrated the stochastic model and the deterministic model. The complexity of the model was low and the number of cluster cores was limited. There was a physical mapping relationship between cluster cores and real propagation objects. Reference [10] proposed a general wireless channel model and channel detection methods for the 5G Internet of Things (IoT) green wireless communication. By using the perspective of big data mining, intelligent channel detector transformed traditional passive wireless communication scheme into an active and expected wireless communication scheme to achieve efficient and green communication. Reference [11] use channel detection technology to characterize the second-order statistics of indoor Ultra Wide Band (UWB) channels. Reference [12] described usage of circular antennas at both ends for channel detection in urban macro cells with time domain method.

References [13–16] used the frequency domain method to measure and model millimeter wave channel. Because of one platform integration requirement of receiver and transmitter, working range of channel sounding was limited [13]. Reference [14] introduced a commercial channel detector based on Vector Network Analyzer (VNA) for directional radio channel measurement in W-band. Reference [15] discussed the measurement of complex frequency response of millimeter wave indoor wireless channel in coherent broadband frequency domain. Reference [16] proposed a channel impulse response detection and recognition system based on sweep method for 57–64 GHz.

The channel classification in references [17,18] was based on traditional statistical methods. The traditional statistical method will consume huge manpower and time when dealing with massive data. In Reference [17], channel fingerprint features were extracted from channel samples, and corresponding channel fingerprint model was established and verified. In Reference [18], channel average multipath delay characteristic parameters and time domain amplitude envelope characteristics of channel impulse response were extracted, and then discrimination mechanisms of wireless channel in different scenes were given. Both References [9,19] used machine learning algorithms to classify millimeter wave channel scenes. However, the frequency identified in Reference [9] did not cover the high frequency millimeter wave band of 6G service. Reference [19] studied correlation vector machine algorithms for multipath component detection considering channel power delay distributions. However, it only used simulation method, lacking real measured data from actual channel sounding system. Reference [20] used machine learning to identify wireless channel scenarios. However, the approaches it proposed took a lot of computing resources or were less efficient. In addition, it did not identify millimeter wave channel, which was important to 6G service. Reference [21] categorized map-based hybrid channel models using machine learning and shared the database of map-based channel parameters. Reference [22] proposed a Machine Learning (ML)-based time-division duplex scheme in which Channel State Information (CSI) could be obtained by leveraging the temporal channel correlation. Reference [23] shed light on the state of the art in channel prediction and proposed a novel predictor that leveraged the strong time-series prediction capability of deep recurrent neural networks incorporating long short-term memory or
Electronics 2021, 10, 843

Gated recurrent unit. Reference [24] proposed a data-driven Deep Learning (DL) approach to jointly design the pilot signals and channel estimator for wideband massive Multiple-Input Multiple-Output (MIMO) systems. By exploiting the angular-domain compressibility of massive MIMO channels, the conceived DL framework could reliably reconstruct the high-dimensional channels from the underdetermined measurements. Reference [25] evaluated the performance of machine learning method in identifying the channel type in 802.11ac systems. References [21–25] did not cover relatively high millimeter wave frequency or only used simulation data.

Considering the frequency method could not measure channel models of two remote receiving and transmitting places, this paper used time domain channel sounding method. Our improvements compared with studies that already existed in academia are from three points. At first, the frequency range measured by the above methods belong to lower millimeter wave frequency band, which do not cover higher millimeter wave frequency band concerned by 6G service. Secondly, the above methods require a lot of communication expertise and consume a lot of manpower and material resources in channel feature extraction stage. Thirdly, our data is based on real channel scenarios. In our paper, cross-correlation algorithm was used to calculate the channel impulse response between transmitted known data and received channel faded data. This process would generate a lot of baseband data, which provided raw materials for big data modeling and learning. To the best of our knowledge, this paper first proposed a data-driven supervised machine learning model for 6G indoor millimeter wave channel scene classification in W-band.

The main contributions of this article are summarized as follows,

1. We proposed and built the industry’s first commercial off-the-shelf (COTS) hardware-based high-frequency indoor millimeter wave channel sounding system based on time-domain correlation method, which could measure millimeter wave signals in the W-band.
2. We firstly proposed a data-driven supervised machine learning model for 6G indoor millimeter wave scene classification in the W-band.
3. The bagging classifier we firstly proposed is more efficient with 96.5% accuracy.

The article is organized as follows: Section 2 describes the whole process of channel sounding and classification at first. Then it provides the algorithm of channel sounding and Power Delay Profile (PDP) results of measured indoor millimeter wave communication scenes. Finally, it describes extracted features, scene classification algorithms and method verification and evaluation. Section 3 presents the results of classification. The conclusions and discussions are in Section 4.

2. Materials and Methods

2.1. Channel Sounding and Classification

The whole process of channel sounding and classification includes two steps: acquire time-domain I/Q data to get the channel impulse response and then get PDP. The PDP’s mean value, standard deviation, kurtosis, skewness, peak value and relative delay characteristics are extracted as features. Scene classification and recognition model based on these features is trained and finally the unknown scene measured data is tested. The process is shown in Figure 1.
2.2. Channel Sounding

The time-domain correlation method is used directly to calculate the channel impulse response [26]. The pseudorandom sequence has good randomness, and its autocorrelation function is close to the impulse signal. Therefore, pseudorandom sequence is selected as transmitted signal in the channel sounding system. Its autocorrelation function is shown as follows,

\[
r(j) = \begin{cases} 
  1 & j = nN (n = 0, \pm 1, \pm 2, \ldots) \\
  -\frac{1}{N} & j \neq nN 
\end{cases}
\]  

(1)

where \(N\) is the period of the transmitted pseudorandom sequence and \(r(j)\) is the autocorrelation function of the pseudorandom sequence. The received RF signal will be downconverted from RF band to baseband through spectrum analyzer, which will generate baseband complex envelope signal, i.e., baseband I/Q data. Supposed that the complex envelope signal of the transmitter is \(X_{CE}(t)\), then relationship between real RF signal \(X_{RF}(t)\) of the transmitter and the complex envelope signal is

\[
X_{RF}(t) = \text{Re}\left[ X_{CE}(t) \cdot e^{j2\pi f_c t} \right].
\]  

(2)

If this signal is transmitted through a static scattering medium, the received signal is given by

\[
Y_{RF}(t) = \text{Re}\left[ \sum_{i=1}^{N} a_i X_{CE}(t - \tau_i) e^{j2\pi f_c (t - \tau_i)} \right] = \text{Re}\left[ Y_{CE}(t) e^{j2\pi f_c t} \right]
\]  

(3)

where

\[
Y_{CE}(t) = \sum_{i=1}^{N} a_i e^{-j2\pi f_c \tau_i} X_{CE}(t - \tau_i).
\]  

(4)

Complex coefficient \(a_i\) represents magnitude and phase contributions from scatterer \(i\), \(\tau_i\) is its associated propagation time and \(f_c\) is carrier frequency. The block diagram of channel sounding system is shown in Figure 2.
Converting the received RF signal into baseband signal, and then making correlation with the transmitted signal, we can get cross-correlation between transmitted signal and received signal as follows,

\[ R_{xy}(\tau) = E[X_C^*(t)Y_C(t + \tau)]. \]  

(5)

where * means complex conjugate.

Considering correlation function \( R_{xy}(\tau) = \int_{-\infty}^{+\infty} x(u)h(u)du \) and convolution theorem \( y(t) = x(t) \otimes h(t) = \int_{-\infty}^{+\infty} x(t-u)h(u)du \), we get

\[
R_{xy}(\tau) = \int_{-\infty}^{+\infty} h(u)R_x(\tau - u)du = h(\tau) \otimes R_x(\tau) \approx h(\tau) \otimes \delta(\tau) = h(\tau). 
\]  

(6)

Supposed that the stochastic process, described by the impulse response, is wide sense stationary and the amplitudes and phases of different paths are uncorrelated, which means that the autocorrelation function (ACF) of the impulse response disappears for \( \tau_1 \neq \tau_2 \) and shows a delta-like behavior for \( \tau_1 = \tau_2 \), the channel can be defined by the autocorrelation function of the impulse response, which can be simplified to the expression

\[
R_h(\tau, \Delta t) = E[h(\tau,t)h^*(\tau,t + \Delta t)]. 
\]  

(7)

The channel sounding accuracy from above equation depends on sharp autocorrelation shape, so M sequence, Barker code and Linear Frequency Modulation (LFM) signals are used often. The ACF of the impulse response, \( R_h(\tau, \Delta t) \), calculated for \( \Delta t = 0 \), and denoted by \( p_h(\tau) = R_h(\tau) = R_h(\tau,0) \), is called the power delay profile [27].

The measured impulse response is the convolution of the true channel impulse response with the impulse response of the sounder [28]:

\[
h_{meas}(t_i, \tau) = \hat{p}(\tau) \otimes h(t_i, \tau),
\]  

(8)

where the effective sounder impulse response \( \hat{p}(\tau) \) is the convolution of the transmitted pulse shape and the Rx filter impulse response \( \hat{p}(\tau) = p_{TX}(\tau) \otimes p_{RX}(\tau) \).

Therefore, the measured power-delay profile \( D_{measured} \) is convolution between the true power delay profile \( D_{channel} \) with the impulse response of the sounder \( \hat{p}(\tau) \).

\[
D_{measured}(t_i, \tau) \quad = \quad \hat{p}(\tau) \otimes D_{channel}(t_i, \tau).
\]  

(9)

Based on the time-domain correlation method, this paper proposed and built a complete commercial indoor millimeter wave channel measurement system, which could sound the W-band millimeter wave channel. The block diagram of commercial indoor millimeter wave channel sounding system is shown in Figure 3.
Taking the transmission and reception of 93 GHz millimeter wave signal in W-band as an example, M8190A at the transmitting side generates PRBS 11 symbols which are modulated by QPSK. The symbol rate is 150 MHz/s, and a root raised cosine filter with roll off coefficient of 0.35 is adopted. The baseband signal bandwidth is about 100 MHz. The microwave vector signal source E8267D synthesizes two I/Q signals and modulates them to the 15 GHz IF frequency. Then IF signal is mixed with multiplier of LO signal source SMW200A (with the frequency of 13 GHz) and generates a 93 GHz millimeter wave signal at transmitting directional waveguide antenna. After propagation through indoor scenario, the 93 GHz millimeter wave signal is received by the receiving antenna. Received signal is mixed with a down-converter whose local oscillator signal is multiplier of 9 GHz, and then the 15 GHz IF signal is captured. The spectrum analyzer FSW converts it down again to get baseband I/Q signal. Vector signal analyzer (VSA) in FSW67 carries out vector analysis on baseband signal. The parameter configurations of VSA are the same as that of transmitter. The measured distance is about 300 m. The antenna is directional antenna, whose gain is 25 dBi. The commercial indoor millimeter wave channel measurement system based on time-domain correlation method built in lab environment is shown in Figure 4. The distance between transmitter and receiver is shortened in Figure 4 in order to show the composition and structure of the system better [29–39].
which were explained in Table 1. We will use scene numbers in the following parts of our paper.

Table 1. Four different indoor scenes.

| Scene Number | Scene                                           |
|--------------|-------------------------------------------------|
| 1            | An empty water bottle between the Tx and Rx antenna. |
| 2            | Two pieces of A4 paper between the Tx and Rx antenna. |
| 3            | A transparent plastic bag between the Tx and Rx antenna. |
| 4            | A sponge between the Tx and Rx antenna.          |

Figure 5. Indoor channel sounding scene 1.

Figure 6. Indoor channel sounding scene 2.
Figure 7. Indoor channel sounding scene 3.

Figure 8. Indoor channel sounding scene 4.
Each scene was measured 200 times, hence a total of 800 groups of received signals were collected for the four scenes. The power delay profiles of the four scenarios are shown in Figures 9–12 respectively.

Figure 9. Channel sounding PDP of scene 1.

Figure 10. Channel sounding PDP of scene 2.
Figure 11. Channel sounding PDP of scene 3.

Figure 12. Channel sounding PDP of scene 4.

We find that millimeter wave in W-band has a better ability to penetrate A4 paper with less delay and power loss from Table 2. The sponge in scene 4 has a stronger ability to absorb the millimeter wave signal, which has a larger delay and a higher power loss. The delay and power are dependent on the material of the occlusion.
Table 2. Comparison for four scenes.

| Scene Number | Delay/µs | Power/dBm |
|--------------|----------|-----------|
| 1            | 0.06667  | −17.82    |
| 2            | 0.01667  | −17.03    |
| 3            | 1        | −20.69    |
| 4            | 1.567    | −21.66    |

2.3. Channel Classification

A total of 800 groups of power delay profile data were obtained as samples. Six features including mean value, standard deviation, kurtosis, skewness, peak value and relative delay, were extracted from each group of power delay profile samples. Through 10 fold cross validation, 90% of the data set was regarded as the training set and 10% as the validation set. The scene classifier based on these features was trained, and finally we got a classification model. The decision tree, SVM and the integrated bagging classifiers were compared and studied in following parts.

2.3.1. Feature Extraction

Six features of mean value, standard deviation, kurtosis, skewness, peak value and relative delay are extracted from each group of channel power delay profile samples. The mean value of power delay profile $M_\tau$ is

$$M_\tau = E[\tau] = \int_{-\infty}^{\infty} \tau D_{measured}(\tau) d\tau. \quad (10)$$

The standard deviation of power delay profile $S_\tau$ is

$$S_\tau = \sqrt{\int_{-\infty}^{\infty} (\tau - M_\tau)^2 D_{measured}(\tau) d\tau}. \quad (11)$$

The kurtosis of power delay profile is used to measure the flatness of sample data distribution, which is expressed as follows,

$$K = \frac{\frac{1}{N} \sum_{i=1}^{N} (D_{measured}^i - M_\tau)^4}{\left(\frac{1}{N} \sum_{i=1}^{N} (D_{measured}^i - M_\tau)^2\right)^2}. \quad (12)$$

The skewness of power delay profile $S_K$ measures symmetry, which is expressed as follows,

$$S_K = \frac{\frac{1}{N} \sum_{i=1}^{N} (D_{measured}^i - M_\tau)^3}{\left(\frac{1}{N} \sum_{i=1}^{N} (D_{measured}^i - M_\tau)^2\right)^{\frac{3}{2}}}. \quad (13)$$

The transmitter transmitted a pseudorandom sequence. It passed through the transmission path to the receiver. The delay of transmission path can be obtained by comparing the delay difference between transmitter and receiver.

2.3.2. Scene Classification Algorithms

Through 10 fold cross validation, the scene classification model (also called classifier) based on these features was trained. In this paper, decision tree algorithm, SVM algorithm and integrated algorithm bagging algorithm were compared and studied to classify the four scenarios as above.
(1) Decision tree algorithm

The decision tree algorithm classifies the sample instances by arranging them from the root node to a leaf node. Each none leaf node on the tree represents the test of an attribute value, and its branch represents each result of the test. Each leaf node on the tree represents a classification category, and the top node of the tree is the root node [40].

(2) SVM Algorithm

The classification principle of SVM is to distinguish samples by constructing classification hyperplane. The optimal hyperplane is the shortest distance between samples to the optimal hyperplane. The process of finding hyperplane is the process of finding expression,

$$\min \frac{||w||^2}{2}\text{ s.t. } y_i \cdot (w \cdot D_{\text{measured}} + b) \geq 1, i = 1, 2, \cdots, N.$$  \hspace{1cm} (14)

where \( w \) is the slope of the hyperplane and \( b \) is the intercept of the hyperplane. There are two categories. The abscissa and ordinate of each sample of each category are \( h_i \) and \( y_i \). The decision function is

$$f(h) = \text{sgn} \left( \sum_{i=1}^{N} \alpha_i^* y_i \left( D_{\text{measured}} \cdot D_{\text{measured}}^i + b^* \right) \right).$$  \hspace{1cm} (15)

where \( w^* = \sum_{i=1}^{N} \alpha_i^* D_{\text{measured}}^i y_i \).

For multiclassification SVM, the basic idea is to transform multiclassification into two classification. This paper adopts the method of one to one. Suppose that there are \( K \) samples, take two samples from \( K \) samples each time, and it needs to be carried out for a total of \( C^2_K \) times.

(3) Bagging Algorithm

As shown in Figure 13, the specific steps of integrating algorithm bagging algorithm are as follows,

i. Suppose that there are \( n \) samples in set \( S = D_{\text{measured}}^1, D_{\text{measured}}^2, \cdots, D_{\text{measured}}^n \), if we take one sample from set \( S \) every time, put it in a new set \( S^* \) and then put it back, we will take a total of \( n \) times to form a new set \( S^* \). By resampling in this way, it can generate \( T \) training sets \( S^*_1, S^*_2, \cdots, S^*_T \) randomly.

ii. Generate the corresponding decision tree \( C_1, C_2, \cdots, C_T \) by using each training set.

iii. For the sample \( P \) from test set, each decision tree is used to test. Then obtain the corresponding category \( C_1(P), C_2(P), \cdots, C_T(P) \).

iv. By voting, the category with the most output in \( T \) decision trees is taken as the category of sample \( P \) in the test set.

Figure 13. Bagging algorithm.

Bagging algorithm block diagram is shown in Figure 14.
A set \( S = \{ D_1^\text{measured}, D_2^\text{measured}, \ldots, D_n^\text{measured} \} \), which has \( n \) samples.

Initialize \( i = 1, j = 1 \)

Take one sample from \( S \), put it in a new set \( S_i^* \) and then put it back.

\( i < n? \)

Yes: \( i = i + 1 \)

\( j < T? \)

Yes: \( j = j + 1 \)

No: Form a new set \( S_i^* \), which has \( n \) samples.

Generate the corresponding decision tree \( C_j \).

Test sample \( P \) from test set. Obtain \( C_j(P) \).

Voting for obtain the category with the most output in \( T \) decision trees.

End

Figure 14. Bagging algorithm block diagram.

The above algorithms were used to perform 10 fold cross validation on the original data, that is, the eigenvalue samples were randomly divided into 10 copies. One was selected as the validation set each time, the remaining 9 copies were used as the training set for model training and repeated 10 times. The average of 10 test results is used as the estimation of model accuracy.

2.3.3. Method Verification and Evaluation

In order to validate and evaluate algorithms, the predicted results \( f \) need to be compared with the actual annotations \( F \). Define the evaluation metric as a function of \( f \) and \( F \).

\[
\text{score} = \text{metric}(f, F).
\]  

(16)

If the actual class and the predicted class of data are both positive, it is True Positive (TP). If the actual class and the predicted class are both negative, it is True Negative (TN). If the actual class is negative and the predicted class is positive, it is False Positive (FP). Furthermore, if the actual class is positive and the predicted class is negative, it is False Negative (FN). Confusion Matrix is in Table 3.
Table 3. Confusion matrix.

|                  | Actual Positive | Actual Negative |
|------------------|-----------------|-----------------|
| predicted positive | TP              | FP              |
| predicted negative| FN              | TN              |

Therefore, true positive rate (TPR) is,

\[
TPR = \frac{TP}{TP + FN}.
\]  

(17)

And false negative rate (FNR) is,

\[
FNR = \frac{FN}{TP + FN}.
\]  

(18)

3. Results

The classification accuracies of decision tree, SVM and bagging algorithm for four indoor scenes are shown in Table 4, and the bagging algorithm has the highest accuracy.

Table 4. The accuracy of classification algorithm.

| Classification Algorithm | Accuracy |
|--------------------------|----------|
| Decision Tree            | 94.3%    |
| SVM                      | 86.4%    |
| Bagging                  | 96.5%    |

The confusion matrixes of the three models are shown in Figures 15–17. The horizontal axis represents the predicted category, and the vertical axis represents the actual category. The diagonal green box represents the probability of correct classification in the sample scene, and the other represents the probability of wrong classification. Obviously, the bagging model is the best.

Figure 15. Confusion matrix of decision tree.
The selection of bagging algorithm’s training set is random and each round of training set is independent. Each prediction function can be generated in parallel, which can save a lot of time through parallel training. From Table 4, Figures 15–17, the accuracy of bagging algorithm is relatively higher. When we have a large data set, SVM is less efficient. Decision tree is prone to overfitting. Bagging algorithm is more efficient and less prone to overfitting. Therefore, the bagging model is selected to classify and identify these four channel scenarios. The scatter diagrams of model prediction using bagging algorithm are shown in Figures 18–21, in which the circle represents the model prediction is correct and the cross represents the model prediction is error. Different colors represent different channel scenarios.
Figure 18. Scatter diagram of bagging.

Figure 19. Scatter diagram of bagging.

Figure 20. Scatter diagram of bagging.
4. Conclusions and Discussions

This paper proposes and builds a 6G indoor millimeter wave channel sounding system based on the time-domain correlation method. Taking the transmission and reception of a typical 93 GHz millimeter wave signal in the W-band as an example, four indoor scenes are modeled. The collected baseband data is transformed into power delay profile. Furthermore, based on the machine learning method, six fingerprint features from different scenes are trained. For the first time, a data-driven supervised learning model for 6G indoor millimeter wave channel scene classification is proposed in this paper. The decision tree, SVM and the optimal bagging channel scene classification algorithm are implemented, and the actual scene data is classified with the test data, with the accuracy of 94.3%, 86.4% and 96.5% respectively. The results show that the channel fingerprint model trained by machine learning method is effective. This method can be used in 6G channel sounding and scene recognition from 100 GHz to 3 THz in the future.

The method we proposed has three advantages. Firstly, the high frequency indoor millimeter wave channel sounding system we built is the industry’s first COTS hardware-based system, which can be used in 6G channel sounding till to THz in the future. Secondly, the data-driven supervised learning model can classify different high-frequency millimeter wave channel scenarios, which improves the accuracy and efficiency of channel classification. Thirdly, bagging classifier we proposed is more efficient through parallel training and has a 96.5% accuracy. However, it also has disadvantages. When the amount of data is small, the enhancement effect of the bagging classifier is not obvious and the efficiency is relatively not high.

Changes in the environment have an impact on the measurement [41–44]. Indoor propagation environment is complex, and some small environmental changes will affect the wireless signal propagation path and strength distribution. In future study, we will explore the impact of subtle environmental changes on millimeter wave channel measurements.

Author Contributions: Conceptualization, L.Y.; methodology, L.Y. and R.Y.; software, L.Y. and R.Y.; validation, YY; formal analysis, L.Y. and R.Y.; investigation, YY; writing—original draft preparation, L.Y.; writing—review and editing, R.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Youth Program of National Natural Science Foundation of China (Grant No. 61801034) and by National Key R & D Program of China (2018YFB1800802).

Data Availability Statement: The data presented in this study is available on page six of this article.

Acknowledgments: The authors would like to thank State Radio Monitoring Centre for providing instruments used in the experiment.
Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yong, N.; Yong, L.; Jin, D.; Li, S.; Vasilakos, A.V.J.W.N. A Survey of Millimeter Wave (mmWave) Communications for 5G: Opportunities and Challenges. *Wirel. Netw.* **2015**, *21*, 2657–2676.

2. Rappaport, T.S.; Xing, Y.; Kanhere, O.; Ju, S.; Madanayake, A.; Mandal, S.; Alkhatteeb, S.; Trichopoulos, G.C. Wireless Communications and Applications Above 100 GHz: Opportunities and Challenges for 6G and Beyond. *IEEE Access* **2019**, *7*, 78729–78757. [CrossRef]

3. Sipal, V. Impact of the Wireless Channel on the Performance of Ultrawideband Communication Systems. Ph.D. Thesis, University of Oxford, Oxford, UK, 2012.

4. Flikkema, P.G.; Johnson, S.G. A comparison of time- and frequency-domain wireless channel sounding techniques. In *Proceedings of the Southeastcon 96 Bringing Together Education, Science and Technology*, Tampa, FL, USA, 11–14 April 1996.

5. Zhao, X.; Qi, W.; Shu, L.; Wang, M.; Sun, S. Wideband Millimeter-Wave Channel Characterization in an Open Office at 26 GHz. *Wirel. Pers. Commun.* **2017**, *97*, 5059–5075. [CrossRef]

6. Hur, S.; Cho, Y.J.; Lee, J.A.; Kang, N.G.; Park, J.; Benn, H. Synchronous channel sounder using horn antenna and indoor measurements on 28 GHz. In *Proceedings of the IEEE International Black Sea Conference on Communications and Networking*, Odessa, Ukraine, 27–30 May 2014; pp. 83–87.

7. Geng, S.; Liu, S.; Hong, W.; Zhao, X. Mm-wave 60 GHz indoor channel parameters and correlation properties. *Di- anbo Kexue Xuebao/Chin. J. Radio Sci.* **2015**, *30*, 808–813.

8. Maccartney, G.R.; Deng, S.; Sun, S.; Rappaport, T.S. Millimeter-Wave Human Blockage at 73 GHz with a Simple Double Knife-Edge Diffraction Model and Extension for Directional Antennas. In *Proceedings of the 2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*, Montreal, QC, Canada, 18–21 September 2016.

9. Zhang, J. The Interdisciplinary Research of Big Data and Wireless Channel: A Cluster-Nuclei Based Channel Model. *China Commun.* **2016**, *13*, 14–26. [CrossRef]

10. Chen, X.; Liu, S.; Lu, J.; Fan, P.; Letaief, K.B. Smart Channel Sounder for 5G IoT: From Wireless Big Data to Active Communication. *IEEE Access* **2016**, *4*, 8888–8899. [CrossRef]

11. Hayar, A.M.; Knopp, R.; Saadane, R. Subspace Analysis of Indoor UWB Channels. *EURASIP J. Adv. Signal Process.* **2005**. [CrossRef]

12. Sommerkorn, G.; Kaske, M.; Schneider, C.; Hafner, S.; Thoma, R. Full 3D MIMO channel sounding and characterization in an urban macro cell. In *Proceedings of the General Assembly and Scientific Symposium*, Beijing, China, 16–23 August 2014.

13. Zhang, P.; Zhou, Y.; Sun, X.; Wang, H.; Terminal, C.J.M.C. Research on Techniques of Measurement and Modeling for 5G Millimeter Wave Channel. *Mob. Commun.* **2017**. [CrossRef]

14. Martinez-Ingles, M.-T.; Gaillot, D.P.; Pascual-Garcia, J.; Molina-Garcia-Pardo, J.-M.; Rodriguez, J.-V.; Rubio, L.; Juan-Llácer, L. Channel sounding and indoor radio channel characteristics in the W-band. *EURASIP J. Wirel. Commun. Netw.* **2016**, *2016*. [CrossRef]

15. Siamarou, A.G.; Al-Nuaimi, M. A Wideband Frequency-Domain Channel-Sounding System and Delay-Spread Measurements at the License-Free 57- to 64-GHz Band. *IEEE Trans. Instrum. Meas.* **2010**, *59*, 519–526. [CrossRef]

16. Liu, Z.; Li, L.; Ye, F. Feature extraction of wireless mobile channel and the scene discrimination. *J. Shanghai Norm. Univ.* 2018. Available online: http://qktg.shnu.edu.cn/zrb/shsfqszrben/ch/reader/view_abstract.aspx?file_no=20180204&flag=1 (accessed on 6 March 2021). [CrossRef]

17. Chen, Y.; Zheng, Z.W.; Duan, H.; Zhang, M. The Extraction of the Fingerprints and Modeling in the Wireless Communication Channel. *Wirel. Commun. Technol.* **2017**. [CrossRef]

18. Cal-Braz, J.A.; Matos, L.J.; Cataldo, E. The Relevance Vector Machine Applied to the Modeling of Wireless Channels. *IEEE Trans. Antennas Propag.* **2013**, *61*, 6157–6167. [CrossRef]

19. Zhang, J.; Liu, L.; Fan, Y.; Zhuang, L.; Piao, Z. Wireless Channel Propagation Scenarios Identification: A Perspective of Machine Learning. *IEEE Access* **2020**, *8*, 47797–47806. [CrossRef]

20. Lim, Y.C.; Cho, Y.J.; Min, S.S.; Kim, Y.; Valenzuela, R.A. Map-based Millimeter-Wave Channel Models: An Overview, Data for BSG Evaluation and Machine Learning. *IEEE Wirel. Commun.* **2020**, *27*, 54–62. [CrossRef]

21. Yuan, J.; Ngo, H.Q.; Matthaiou, M. Machine Learning-Based Channel Prediction in Massive MIMO with Channel Aging. *IEEE Wirel. Commun.* **2020**, *19*, 2960–2973. [CrossRef]

22. Wei, J.; Schotten, H.D. Deep Learning for Fading Channel Prediction. *IEEE Open J. Commun. Soc.* **2020**, *1*, 320–332.

23. Ma, X.; Gao, Z. Data-Driven Deep Learning to Design Pilot and Channel Estimator for Massive MIMO. *IEEE Trans. Veh. Technol.* **2020**, *69*, 5677–5682. [CrossRef]

24. Kurniawan, E.; Tan, P.H.; Sun, S.; Wang, Y.J.I. Machine Learning-based Channel-Type Identification for IEEE 802.11ac Link Adaptation. In *Proceedings of the 2018 24th Asia-Pacific Conference on Communications (APCC)*, Ningbo, China, 12–14 November 2018.

25. Ribeiro, C.; Gameiro, A. Direct Time-Domain Channel Impulse Response Estimation for OFDM-Based Systems. In *Proceedings of the 2007 IEEE 66th Vehicular Technology Conference*, Baltimore, MD, USA, 30 September–3 October 2007.
27. Fontæn, F.P.; Espiæeira, P.M. *Modeling the Wireless Propagation Channel: A Simulation Approach with MATLAB*; John Wiley and Sons, Ltd.: Hoboken, NJ, USA, 2008.
28. Goldsmith, A. *Wireless Communications*; Cambridge University Press: Cambridge, UK, 2007.
29. Brandli, G.; Dick, M. Alternating-Current Fed Power Supply. U.S. Patent No 4,084,217, 1978.
30. Horowitz, P.; Forster, J.; Linscott, I. The 8-million channel narrowband analyzer. In *The Search for Extraterrestrial Life: Recent Developments*; Cambridge University Press: Cambridge, UK, 1985.
31. Sagnier, P.; Marraffa, L. Parametric Study of Thermal and Chemical Nonequilibrium Nozzle Flow. *AIAA J.* 2015, 29, 334–343.
32. Heiser, A.J. Private Communication. *Private Commun.* 2000, 8, 317–334.
33. Smith, B. An approach to graphs of linear forms. Unpublished work, 1982.
34. McKeeman, W.M. Representation Error for Real Numbers in Binary Computer Arithmetic. *IEEE Trans. Electron. Comput.* 1967, 5, 682–683. [CrossRef]
35. Class, I.E. IEEE Criteria for Class IE Electric Systems for Nuclear Power Generating Stations. *IEEE Std 308-1971 (Rev. IEEE Std 308-1980) 1970, PAS-89, 1365–1372.* [CrossRef]
36. United States of America Standards Institute. *IEEE Standard Letter Symbols for Quantities Used in Electrical Science and Electrical Engineering*: American Society of Mechanical Engineers: New York, NY, USA, 1984.
37. Fardel, R.; Nagel, M.; Nüesch, F.; Lippert, T.; Wokaun, A. Fabrication of organic light-emitting diode pixels by laser-assisted forward transfer. *Appl. Phys. Lett.* 2007, 91, 61103. [CrossRef]
38. Jing, Z.; Tansu, N. Optical Gain and Laser Characteristics of InGaN Quantum Wells on Ternary InGaN Substrates. *IEEE Photonics J.* 2013, 5, 2600111. [CrossRef]
39. Azodolmolky, S.; Perello, J.; Angelou, M.; Agraz, F.; Velasco, L.; Spadaro, S.; Pointurier, Y.; Francescon, A.; Saradhi, C.V.; Kokkinos, P.; et al. Experimental Demonstration of an Impairment Aware Network Planning and Operation Tool for Transparent/Translucent Optical Networks. *J. Light. Technol.* 2011, 29, 439–448. [CrossRef]
40. Cook, D.J.; Krishnan, N.C. *Activity Learning: Discovering, Recognizing, and Predicting Human Behavior from Sensor Data*; John Wiley and Sons, Ltd.: Hoboken, NJ, USA, 2015.
41. Djosic, S.; Stojanovic, I.; Jovanovic, M.; Nikolic, T.; Djordjevic, G.L. Fingerprinting-Assisted UWB-based Localization Technique for Complex Indoor Environments. *Expert Syst. Appl.* 2021, 167, 114188. [CrossRef]
42. Vongkulbhisal, J.; Chantaramolee, B.; Yan, Z.; Mohammed, W.S.; Letters, O.T. A fingerprinting—Based indoor localization system using intensity modulation of light emitting diodes. *Microw. Opt. Technol. Lett.* 2012, 54, 1218–1227. [CrossRef]
43. Stella, M.; Russo, M.; Begušić, D. Location Determination in Indoor Environment based on RSS Fingerprinting and Artificial Neural Network. In Proceedings of the 2007 9th International Conference on Telecommunications, Zagreb, Croatia, 13–15 June 2007.
44. Taok, A.; Kandil, N.; Affes, S. Neural Networks for Fingerprinting-Based Indoor Localization Using Ultra-Wideband. *JCM* 2009, 4, 267–275. [CrossRef]