

Article

An Improved SVM-Based Air-to-Ground Communication Scenario Identification Method Using Channel Characteristics

Guyue Zhu ¹, Yuanjian Liu ^{1,*}, Kai Mao ², Jingyi Zhang ¹, Boyu Hua ² and Shuangde Li ¹

¹ College of Electronic and Optical Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210023, China; 2020020221@njupt.edu.cn (G.Z.); 1221024930@njupt.edu.cn (J.Z.); 2016020104@njupt.edu.cn (S.L.)

² The Key Laboratory of Dynamic Cognitive System of Electromagnetic Spectrum Space, College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China; maokai@nuaa.edu.cn (K.M.); byhua@nuaa.edu.cn (B.H.)

* Correspondence: liuyj@njupt.edu.cn

Abstract: Scenario identification plays an important role in assisting unmanned aerial vehicle (UAV) cognitive communications. Based on the scenario-dependent channel characteristics, a support vector machine (SVM)-based air-to-ground (A2G) scenario identification model is proposed. In the proposed model, the height of the UAV is also used as a feature to improve the identification accuracy. On the basis, an improved scenario identification method is developed including dataset acquisition, identification model training, and height-integrated model feedback. The shooting and bouncing ray/image (SBR/IM) method is used to obtain the datasets of channel characteristics, i.e., root-mean-square delay spread (RMS-DS), *K* factor, and angle spread (AS) under five typical scenarios: over-sea, suburban, urban, dense urban and high-rise urban. SBR/IM is a symmetry-based ray tracing (RT) simulation method. After the identification model is trained, a height-integrated feedback scheme is used to increase the identification performance. The simulation results show that the identification accuracy of improved method is about 14% higher than the method without height feature, which reaches nearly 100% under the over-sea and suburban and over 80% in urban, dense urban, and high-rise urban.

Keywords: scenario identification; air-to-ground (A2G) communication; channel characteristics; support vector machine (SVM); unmanned aerial vehicle (UAV)



Citation: Zhu, G.; Liu Y.; Mao, K.; Zhang, J.; Hua, B.; Li, S. An Improved SVM-Based Air-to-Ground Communication Scenario Identification Method Using Channel Characteristics. *Symmetry* **2022**, *14*, 1038. <https://doi.org/10.3390/sym14051038>

Academic Editor: Jan Awrejcewicz

Received: 16 April 2022

Accepted: 17 May 2022

Published: 19 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Unmanned aerial vehicle (UAV) communication can be used to provide effective and efficient telecommunication facilities and service with low cost [1], and is expected to play an important role in the air-space-earth-sea integrated sixth generation (6G) communication network [2,3]. The authors in [4] proposed an UAV integrated HetNet for smart dense urban, and the improved methods of the UAV integrated HetNet were proposed in [5–8]. Even the deployment of UAV on Mars was proposed in [9]. However, due to the mobility and flexibility, the UAV would experience dynamic communication scenarios and the channel characteristics are proven to be scenario-dependent [10,11]. Therefore, it is necessary to identify the communication scenarios and further optimize the communication system by adopting different channel models. Different from the traditional terrestrial communication, the air-to-ground (A2G) communication shows more obvious 3D scattering space and height variance. According to the accurate identification model, the communication scenarios can be identified in real-time. Furthermore, the scenario-dependent channel models, transmission modes, and optimization algorithms can be used to improve the performance and reliability of A2G communication systems. Therefore, it is vital to propose an efficient and accurate scenario identification method for A2G to assist the UAV cognitive communication.

To date, scenario identification has attracted a lot of attentions since it is a key technology of communication perception integration. The authors in [12] pointed out that there are mainly two scenario identification methods, i.e., visual-inspection-based identification and machine learning (ML)-integrated identification. For example, the authors in [13] identified line-of-sight (LOS)/non-line-of-sight (NLOS) scenarios manually based on the video stream. A visual inspection identification method for the high-speed railway communication system was proposed in [14] based on the geographic information system (GIS). However, visual inspection identification method is laborious and has poor robustness. Therefore, the ML-integrated identification model has been widely applied, where the image/video and channel characteristics are normally used. For example, the authors in [15] used the satellite image with labeled pixel to identify different scenarios by using convolutional neural network (CNN) method. However, the camera is used to record the images or videos in these methods, which would achieve poor performance in the weather with poor visibility.

Recently, the channel characteristics have also been widely used to identify the communication scenarios. For example, the authors in [16–19] identified LOS/NLOS scenarios by using channel characteristics such as receive power, K -factor and delay spread based on ML method like SVM and back propagation neural network (BPNN). However, these methods did not involve the identification of specific geometric scenarios. A CNN-based identification method was proposed in [20] to identify vehicle-to-vehicle (V2V) scenarios according to the channel characteristics. The authors in [21] used measured channel characteristics to identify vehicular communication scenarios by using BPNN. A multi-feature fusion and deep neural network combined method was proposed in [22] to identify the scenarios of high-speed train communication. Furthermore, the authors in [23] proposed an identification method based on the measured channel data and weighted K -nearest neighbor technology for indoor scenario identification. A BPNN-based scenario identification method was proposed in [24] to identify vehicular scenarios according to the channel characteristics. However, these research works were mainly aimed at the terrestrial communication scenarios, which cannot be applied in the A2G scenario identification directly. Moreover, the channel characteristics of A2G communication are normally height-dependent. It is a valuable identification feature but is not considered in the aforementioned works.

To fill these gaps, the main novelties of this paper are summarized as follows:

(1) A SVM-based scenario identification model for A2G communication is proposed by using the channel characteristics. In this model, the height of the UAV is also used as a new feature to improve the identification accuracy.

(2) An improved identification method is developed including dataset acquisition, identification model training, and height-integrated model feedback. The datasets of channel characteristics obtained from the RT simulation are used to train the identification model. Then the height-integrated feedback scheme is introduced into the trained model to improve the identification accuracy.

(3) The proposed scenario identification method is verified by the obtained validation dataset. The performances of different set of channel characteristics for scenario identification are compared and analyzed. The confusion matrix is also used to evaluate the performance of proposed scenario identification method.

The remainder of this paper is organized as follows. Section 2 proposes a channel-characteristic-based scenario identification model. Section 3 develops an improved SVM-based scenario identification method. In Section 4, the performance of proposed scenario identification method is validated and analyzed. Finally, Section 5 draws the conclusions.

2. Channel-Characteristic-Based Scenario Identification Model

The A2G communications normally experience diverse and time-varying scenarios as shown in Figure 1 [25–27]. If these scenarios can be accurately identified, the corresponding channel models can be adopted to optimize the A2G communication systems. Considering that channel characteristics are scenario-dependent, the scenarios can be identified

according to the relationship between channel characteristics and corresponding scenario. However, it is difficult to manually find the complicated inner relationship. It should be mentioned that the UAV channel characteristics including path loss (PL), root mean square delay spread (RMS-DS), K factor, the mean angle of arrival (AOA), and Angular Spread (AS) with flight height can be accurately extracted from the cellular signals [28,29]. Therefore, we propose a channel-characteristic-based scenario identification model combined with SVM method.

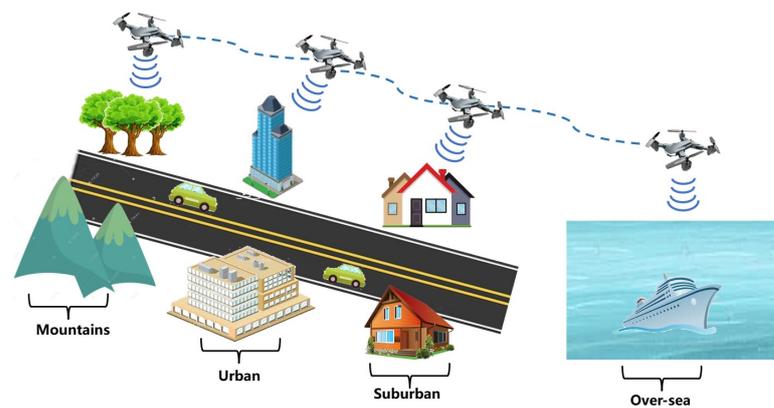


Figure 1. UAV-aided A2G communication scenarios.

The SVM has been proven to be an effective classification and regression model, which is widely used in pattern recognition, nonlinear regression and so on. In the SVM method, hyperplane is used to separate different classification of data, where support vectors represent different data points with approximate distance to the hyperplane. The optimization approach is normally used to find the optimal hyperplane by maximizing the sum of the distances between the hyperplane and support vectors. As shown in Figure 2, the blue and yellow points are two types of data samples, the solid line represents hyperplane and the dotted lines represent support vectors. The hyperplane equation can be expressed as $\mathbf{w} \cdot \mathbf{x} + b = 0$, where \mathbf{w} is the normal vector and b is displacement. The data on both sides of the hyperplane are defined as positive samples and negative samples, i.e., $\mathbf{w} \cdot \mathbf{x} + b$ equals to 1 and -1 , respectively, Ref. [30]. To obtain the optimal hyperplane, we need to find the maximum value of $\frac{2}{\|\mathbf{w}\|}$.

Assuming that the labeled indexes are denoted as \mathbf{I} , and the channel characteristics of target scenario are denoted as \mathbf{C} , the proposed scenario identification model can be expressed as

$$\begin{cases} \mathbf{w} \cdot \mathbf{C} + b \geq 1, \mathbf{I} = 1 \\ \mathbf{w} \cdot \mathbf{C} + b \leq -1, \mathbf{I} = -1 \end{cases} \quad (1)$$

Considering that the channel characteristics of A2G communication are normally height-dependent, which is a valuable feature for scenario identification. In this paper, the height of UAV is introduced into the scenario identification model as

$$\begin{cases} \mathbf{w} \cdot \mathbf{C}'(h \pm \Delta h) + b \geq 1, \mathbf{I} = 1 \\ \mathbf{w} \cdot \mathbf{C}'(h \pm \Delta h) + b \leq -1, \mathbf{I} = -1 \end{cases} \quad (2)$$

where \mathbf{C}' is the channel characteristic after height adjustment, h is the height before adjustment and Δh is the scale for the height change.

Moreover, the kernel function is also crucial to determine the hyperplanes in the training process. The typical kernel functions include logistic regression (LR), linear and Gaussian kernel function [31]. When the number of channel characteristic is large, LR or Linear kernel function is more appropriate. Otherwise, the Gaussian kernel function can be applied. In this paper, since small number of channel characteristics is considered, the

Gaussian kernel function is used to train the SVM. The Gaussian kernel function can be expressed as

$$K(\mathbf{x}_m, \mathbf{x}_n) = \exp(-\gamma \|\mathbf{x}_m - \mathbf{x}_n\|^2) \quad (3)$$

where $\mathbf{x}_m, \mathbf{x}_n$ mean two datapoint and γ is a hyper parameter represents Gaussian filter width.

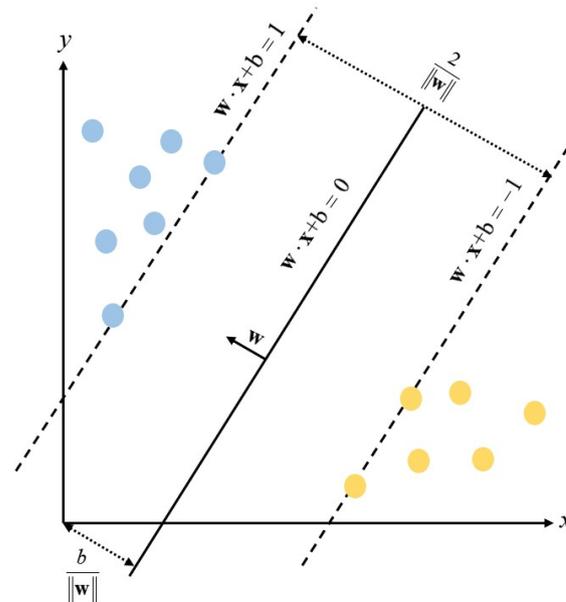


Figure 2. Basic idea of SVM classification.

3. Improved Identification Method

3.1. RT-Based Acquisition of Channel Characteristics

Ray tracing (RT) is an accurate channel parameter calculation method based on field superposition principle and uniform theory of diffraction theory. The shooting and bouncing ray/image (SBR/IM) method is a symmetry-based RT method and it is applied to obtain channel characteristics in this paper due to its high accuracy and low computational complexity. The SBR/IM method includes four steps: scatterer reconstruction, ray decomposition, ray tracking and channel characteristics calculation.

Due to the high complexity of the realistic scenario, a scatterer reconstruction method is used in this paper to reduce the complexity and computational time of the SBR/IM method. The reconstructed scenario is composed of reconstructed the surface of the terrain and buildings. Firstly, the flat surface of terrain can be reconstructed by regular triangle faces and the non-flat surface can be approximated reconstructed by irregular triangle facets. Then, the surface of buildings can also be reconstructed by regular triangle facets due to they are assumed to be rectangles. Finally, material of surfaces will be configured after all scatterer shapes are reconstructed.

The ray decomposition method is as shown in Figure 3a, a regular icosahedron is placed inside a wave-front ball. The wave-front ball is divided to generate the ray tubes, where each tube is represented by a ray from the center of the vertex. Considering that the section of the ray tube increases with the increasing distance, which would affect the accuracy of ray tracking. Therefore, the wavefront division of the regular dodecahedron is necessary as shown in Figure 3b. Each equilateral triangle is uniformly divided into N wave fronts. The expression of N is defined as $N = 20 \times n^2$, where n is the number of ray tube splits. In Figure 3b, the number of ray tube splits is set as $n = 4$.

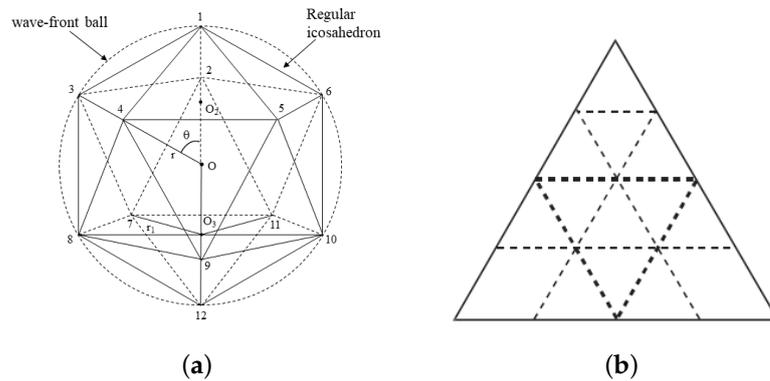


Figure 3. (a) Icosahedron determination and (b) wavefront division for ray decomposition.

For the ray tracking, the reflection, scattering, and diffraction propagations are tracked by using geometric optics theory, where the intersection operation is a key step. The intersection operation is used to determine whether the ray intersects with the triangle facets of scatterers, and then to determine the reflection or scattering points. It is also necessary to determine whether diffraction occurs. Taking the receiver as a receiving sphere, the signal is considered being able to arrive at the receiver if the ray intersects with the sphere.

The path to the receiver can be divided into the direct ray or reflected ray. For the direct ray, the electric intensity and power gain can be calculated by

$$E^{Direct} = E_0 \cdot \frac{e^{-jkd(t)}}{d(t)} \tag{4}$$

$$P^{Direct} = 10\log_{10}\left(G^{rx}G^{tx}\left(\frac{\lambda_0}{4\pi}\right)^2\left|\frac{E^{Direct}}{E_0}\right|\right) \tag{5}$$

where E_0 is the electric intensity of the initial ray tube, $k = 1/\lambda_0$ is the wave number, and $d(t)$ is the distance between the transmitter (TX) and receiver (RX), P^{direct} is the power of the direct ray, and G^{rx} and G^{tx} are antenna gain of the receiver and transmitter, respectively.

The electric intensity of the reflected ray can be calculated as

$$E^{Reflect} = E_0 \cdot \{\Pi\bar{R}_i\} \cdot \{\Pi\bar{T}_i\} \cdot \{\Pi e^{-r_i a_i}\} \cdot SF \tag{6}$$

$$P^{Reflect} = 10\log_{10}\left(G^{rx}G^{tx}\left(\frac{\lambda_0}{4\pi}\right)^2\left|\frac{E^{Reflect}}{E_0}\right|\right) \tag{7}$$

where E_0 is the electric intensity of the initial ray tube, $\{\Pi\bar{R}_i\}$ and $\{\Pi\bar{T}_i\}$ are union vector of reflection and refraction coefficients for the entire ray path, respectively, r_i and a_i are the phase shift and power attenuation of the ray from the reference position, respectively, SF is diffusion factor, $P^{Reflect}$ is the power of the reflected ray, G^{rx} and G^{tx} are antenna gain of the receiver and transmitter.

In this paper, the inter-path delays and angles are calculated in a deterministic method. Denote the location of signal source as \mathbf{Q} and the location vector of RX as \mathbf{R} . The adjacent reflection points are set as the centroid and the position vector is denoted as $\bar{\mathbf{P}}_n$. The delay of m -th ray is defined as adding the intra-path delay offset $\Delta\tau_{n,m}$ to the mean ray delay $\bar{\tau}_n(t)$ as

$$\tau_{n,m}(t) = \bar{\tau}_n(t) + \Delta\tau_{n,m} \tag{8}$$

where $\bar{\tau}_n(t)$ can be expressed as

$$\bar{\tau}_n(t) = \frac{\|\mathbf{Q} - \bar{\mathbf{P}}_n\|_2 + \|\bar{\mathbf{P}}_n - \mathbf{R}\|_2}{c} \quad (9)$$

where c is the speed of light.

The azimuth AOA $\alpha^{rx}_{n,m}(t)$ and elevation AOA $\beta^{rx}_{n,m}(t)$ of m th ray can be obtained as the same way. Furthermore, the mean ray angle $\bar{\alpha}^{rx}_n(t)$ and $\bar{\beta}^{rx}_n(t)$ can be expressed as (10) and (11), respectively.

$$\bar{\alpha}^{rx}_n(t) = \begin{cases} \arccos\left(\frac{|\bar{\mathbf{P}}_n^x - \mathbf{R}^x|}{\sqrt{|\bar{\mathbf{P}}_n^x - \mathbf{R}^x|^2 + |\bar{\mathbf{P}}_n^y - \mathbf{R}^y|^2}}\right), & \bar{\mathbf{P}}_n^x - \mathbf{R}^x \geq 0 \\ \pi - \arccos\left(\frac{|\bar{\mathbf{P}}_n^x - \mathbf{R}^x|}{\sqrt{|\bar{\mathbf{P}}_n^x - \mathbf{R}^x|^2 + |\bar{\mathbf{P}}_n^y - \mathbf{R}^y|^2}}\right), & \bar{\mathbf{P}}_n^x - \mathbf{R}^x < 0 \end{cases} \quad (10)$$

$$\bar{\beta}^{rx}_n(t) = \arcsin\left(\frac{|\bar{\mathbf{P}}_n^z - \mathbf{R}^z|}{\|\bar{\mathbf{P}}_n - \mathbf{R}\|_2}\right) \quad (11)$$

where $(\cdot)^x$, $(\cdot)^y$ and $(\cdot)^z$ represent the x , y , z component, respectively. After the channel parameter are obtained, we can further calculate the channel characteristics.

The RMS-DS is used to describe the channel dispersion phenomenon in the delay domain which is caused by the small-scale fading of multipath effect. The definition of RMS-DS can be expressed as

$$\sigma_\tau = \sqrt{\frac{\sum_{l=1}^L (\tau_l - \bar{\tau}) P_l}{P_R}} \quad (12)$$

where τ_l and P_l are the delay and power of each propagation path, respectively, P_R represents the received power, L represents the number of paths, and $\bar{\tau}$ represents the mean delay which can be further expressed as

$$\bar{\tau} = \frac{\sum_{l=1}^L P_l \tau_l}{P_R} \quad (13)$$

The K factor, also known as the Rice factor, is the power ratio of the dominating multipath component to the summation of other multipath components. The dominating multipath component usually is a direct path in A2G communication. The K factor can be expressed as:

$$K = \frac{P_m}{\sum_{l=1, l \neq m}^L P_l} \quad (14)$$

where P_m is the power of the dominating multipath component.

In the angle domain, the multipath effect will cause angle spread. The AS is used to characterize the magnitude of the angular dispersion, which can be expressed as

$$\sigma_\theta = \sqrt{\frac{\sum_{l=1}^L (\theta_l - \bar{\theta}) P_l}{P_R}} \quad (15)$$

where $\bar{\theta}$ can be expressed as:

$$\bar{\theta} = \frac{\sum_{l=1}^L P_l \theta_l}{P_R} \quad (16)$$

It should be mentioned that the AS of azimuth angle of arrival (AAOA), azimuth angle of departure (AAOD), elevation angle of arrival (EAOA), and elevation angle of departure (EAOD), respectively.

3.2. Training Dataset of Height-Dependent Channel Characteristics

Different from the scenario identification of terrestrial communications, the influence of UAV height should be considered in the scenario identification of A2G communications. In this section, the channel characteristics at different height are obtained and analyzed by using the calculation method in Section 3.1 under five typical A2G communication scenarios including over-sea, suburban and urban scenarios. The urban scenarios are further divided into urban, dense urban and high-rise urban.

The simulation carrier frequency is set as 2.6 GHz, and the transmitting power is set as 20 dBm. Both the transmitters and receivers are equipped with a half-wave dipole antenna with vertical polarization. We set ten groups of transmitters and receivers in each scenario. In each group, the transmitter is placed at the height of 1.7 m, and the receivers are placed at the height between 10 m to 210 m with 2 m intervals. Moreover, we place three receivers at each height, so each scenario has 3000 receivers.

To obtain the needed digital map, the standardized scenario models recommended by International Telecommunication Union-Radiocommunication Sector (ITU-R) are adopted in this paper [32]. The scenario models are related with three parameters α_0 , β_0 , γ_0 , which are defined as follows

- α_0 indicates the proportion of the building area to the total area;
- β_0 indicates the average number of buildings per unit area (buildings/km²);
- γ_0 indicates the height of the building according to the Rayleigh distribution, where h can be calculated by

$$P(h) = \frac{h}{\gamma_0^2} \exp\left(-\frac{h^2}{2\gamma_0^2}\right) \quad (17)$$

The detailed simulation parameters and the parameters of scenario models for the suburban and three urban scenarios are summarized in Table 1. Note that there are no buildings in the over-sea scenario, we just reconstruct several ships for simplicity. The digital maps of five reconstructed scenarios are shown in Figure 4, and the average heights of the scatterers in each scenario are 4 m, 9.44 m, 18.83 m, 24.31 m and 64.43 m. The size of all digital maps is 1 km × 1 km. The material of ship is defined as metal, and the material of buildings is concrete, and the land is defined as dry ground.

Table 1. Simulation parameters.

Parameter	Value
Scenario	Over-sea, Suburban, urban, dense urban, high-rise urban
Area ration of scatters	0.005, 0.1, 0.3, 0.5, 0.5
Number of scatters (km ²)	50, 750, 500, 300, 300
Height of scatters	4 m, Rayleigh distribution of 8, 15, 20, 50
Size of scatters	20 m × 5 m, 11.6 m × 11.6 m, 24.5 m × 24.5 m, 40.5 m × 40.5 m, 40.5 m × 40.5 m
Frequency	2.4 GHz
Bandwidth	100 MHz
Antenna type	Half-wave dipole antenna
Transmitting power	20 dBm
Height of TX	1.7 m
Number of TX	10
Height of RX	Between 10 m to 210 m with 2 m intervals
Number of RX	3000

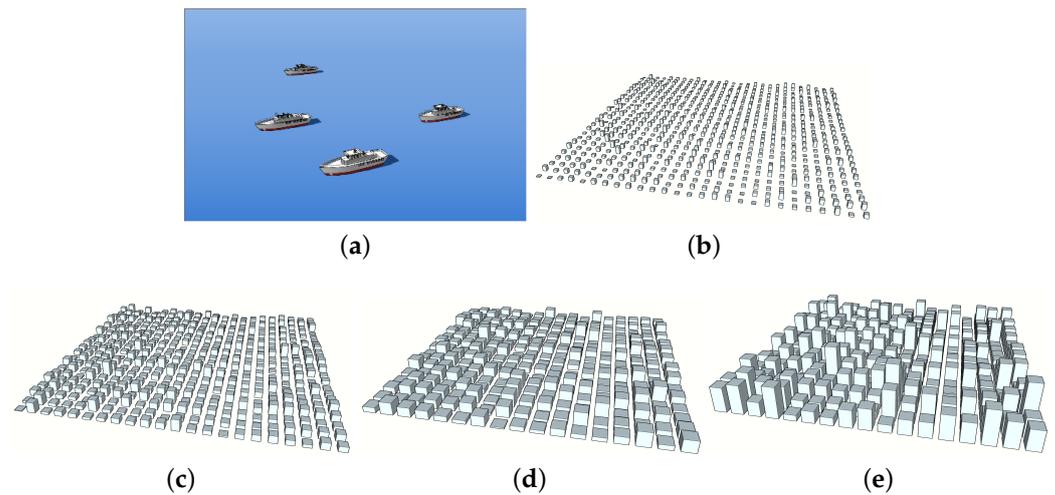


Figure 4. Digital maps of (a) over-sea, (b) suburban, (c) urban, (d) dense urban, and (e) high-rise urban scenario.

Based on the calculation method in Section 3.1, the channel characteristics are calculated at different height. For simplicity, only the RMS-DS and AS of AAOA of one group are shown in Figures 5 and 6, respectively. The y -axis of Figures 5 and 6 are the height of the UAV, and there are 30 data points per height. The x -axis of Figures 5 and 6 are the index of each data point, ranging in size from 1 to 30. Furthermore, there are significant changes around a certain height, which are marked with red lines in the figure, respectively. This is due to the difference in the average height of scatters of scenarios, which leads to different channel characteristics distribution with height. Furthermore, the exact mean values of different channel characteristics under different scenarios are presented in Table 2. It can be found that the channel characteristics, i.e., RMS-DS, K factor, ASs vary a lot under different scenarios and at different height, which makes it possible to identify the scenarios based on the channel characteristics.

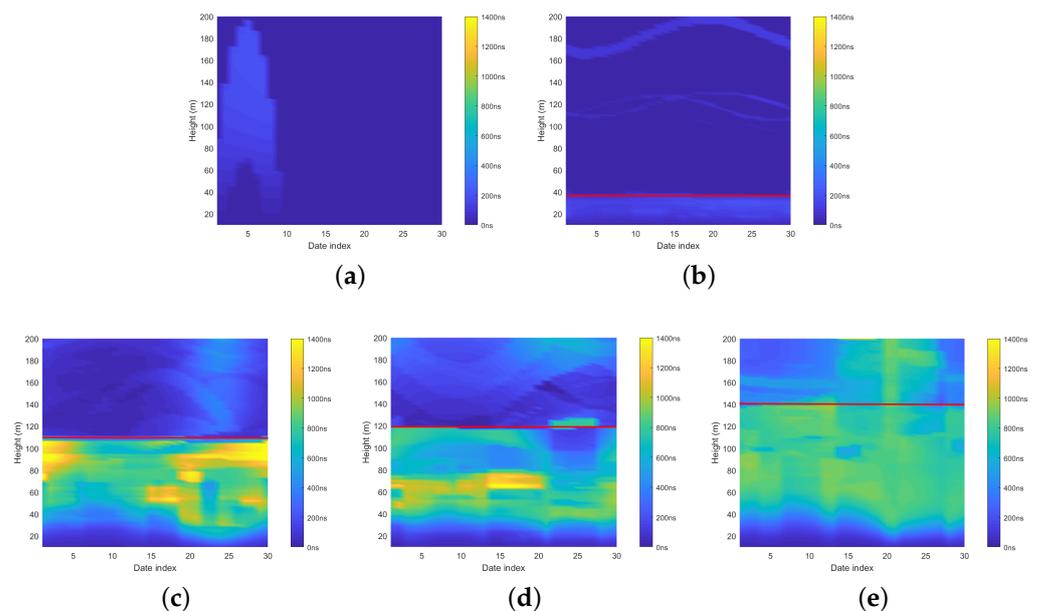


Figure 5. RMS-DS of (a) over-sea, (b) suburban, (c) urban, (d) dense urban, and (e) high-rise urban scenario.

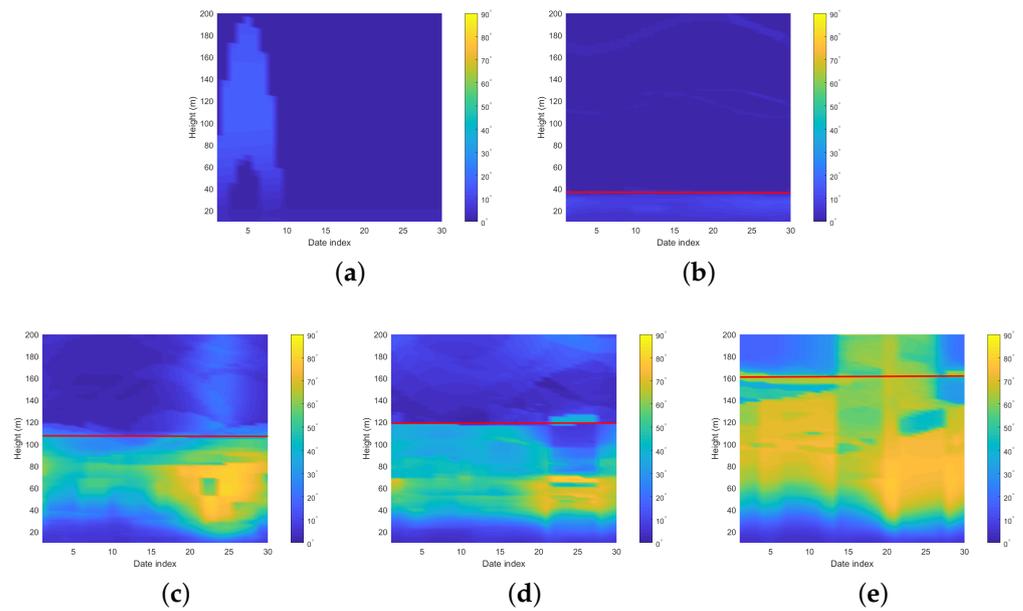


Figure 6. AAOO of (a) over-sea, (b) suburban, (c) urban, (d) dense urban, and (e) high-rise urban scenario.

Table 2. Mean average of parameters of scenarios.

Scenario	σ_{τ} (ns)	K	σ_{AOAA}°	σ_{EAOA}°	σ_{AAOD}°	σ_{EAOD}°
Over-sea	4.6127	6.6838	0.9793	2.1590	2.1805	76.53
Suburban	102.3164	6.3169	0.7497	1.059	1.065	77.9331
Unbran	127.4503	5.3237	16.1668	18.1029	21.7193	68.5246
Dense Urban	143.4544	5.2331	14.2822	18.3465	19.0873	63.8681
Highrise Urban	125.8363	5.3187	37.6931	31.4001	43.3032	43.0069

Based on the above discussion, in this paper we use the channel characteristics, i.e., RMS-DS, K factor, ASs and the height of the UAV as the identification features. The dataset of i -th scenario is denoted as $\mathbf{x}_i = (\sigma_{\tau,i}, \sigma_{AAOA,i}, \sigma_{AAOD,i}, \sigma_{EAOA,i}, \sigma_{EAOD,i}, K_i, h_i, l_i)$, where $l_i \in \{1, 2, \dots, L\}$ denotes the scenario label and L is the number of scenarios.

3.3. Height-Integrated Scenario Identification Method

The proposed height-integrated scenario identification method is shown in Figure 7. It includes three steps, i.e., dataset acquisition and preprocessing, identification model training, and height-integrated model feedback. The details are shown as follows.

Step 1: Data acquisition and preprocessing.

Based on the calculation method in Sections 3.1 and 3.2, the datasets of height-dependent channel characteristics are obtained. Note that the weight of different channel characteristic is different. In order to achieve better identification performance, it is required to preprocess the dataset before training. In this paper, we normalize the data by using z-scores and it can be expressed as

$$x_i^j = \frac{x_i^j - \bar{x}^j}{\sigma_{x^j}} \quad (18)$$

where \bar{x}^j and σ_{x^j} are the mean and variance of the input j -th dimension data, respectively.

Furthermore, dimensionality reduction for high-dimensional input data can prevent the method from slipping into the local optimum and improve the training performance. Therefore, the principal component analysis (PCA) is adopted for dimensionality reduction. The core idea of PCA is to use orthogonal transformation to replace a set of potentially re-

lated variables with a set of linearly unrelated principal components [33]. Low-dimensional principal components can be generated by reasonable selecting eigenvalues. The Gaussian Kernel function in Formula (3) is employed as the core of PCA in this paper.

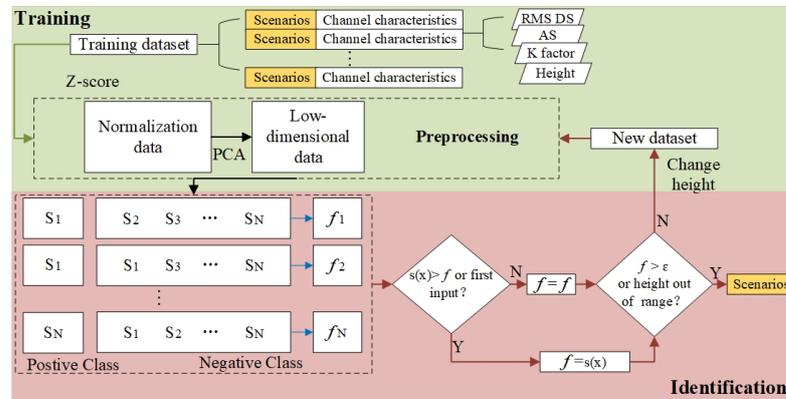


Figure 7. The basic idea of SVM classification.

Step 2: Identification model training.

Then the datasets are divided into two parts, i.e., training data set and testing data set in the proportion of 3:2 as shown in Table 3. Although SVM is a binary classifier, we can use a decomposition methods of multi-class SVM by reconstructing a multi-class classifier from binary SVM-based classifier. For j -th binary SVM classification, it takes the scenario with j -th label as positive class and the rest of others as negative class, where $1 \leq j \leq N$. The results of new samples are determined by combing the labels which is predicted from all the SVM classifiers. Assuming that the multiple binary SVM classifiers are f_1, f_2, \dots, f_N , the final identification result of a sample x is determined by

$$s(x) = \arg \max_i f_i(x) \tag{19}$$

where s is scenario label and “arg max” is to find the scenario label with the highest probability of the multiple binary SVM classifiers.

Table 3. Validation layout parameters.

Data Sets	Datapoints Number
Total data sets	3000/3000/3000/3000/3000
Training data	1800/1800/1800/1800/1800
Testing data	1200/1200/1200/1200/1200

Step 3: Height-integrated model feedback.

After obtaining the trained scenario identification model, the new channel characteristics can be input into the model for scenario identification. If the posterior probability f is greater than the threshold ϵ , the scenario label will be output directly, otherwise the height of the UAV is changed slightly to get a new dataset and repeat the identification procedure. It should be mentioned that whether the height of the UAV is raised or lowered depends on the first scenario label. When the posterior probability is greater than the threshold or the height of the UAV exceeds the height limitation, the model outputs the predicted label of the scenario. In the scenario identification method, the threshold ϵ should be properly determined.

4. Simulation Results and Validation

The selection of identification features is essential to the accuracy of the scenario identification method [34,35]. To validate the rationality of channel characteristics selected

in the paper, we compared the performance of three typical groups of channel characteristics in this section as shown in Table 4. They are Feature Set 1 (K factor + RMS-DS), Feature Set 2 (K factor + RMS-DS + AS), Feature Set 3 (K factor + RMS-DS + AS + path loss). It is found that Feature Set 1 shows the lowest identification accuracy, Feature set 2 achieves a higher identification accuracy of more than 80%. However, the number of identification features is not the more the better. For example, Feature set 3 adds an extra feature Path loss, but the accuracy reduced because the path loss strongly depends on the unknown communication distance. Therefore, the Feature set 2 used in this paper is a rational choice.

Table 4. Accuracy of Proposed Identification Model Under Different Features.

Scenario	Over-Sea	Suburban	Urban	Dense Urban	High-Rise Urban
K factor + RMS DS	100%	100%	79%	70%	85%
K factor + RMS DS + AS	100%	100%	84%	80%	97%
PL + K factor + RMS DS + AS	100%	100%	80%	77%	96%

Moreover, the confusion matrix, also known as the error matrix, is used to validate the effectiveness of proposed identification method. Each row in the matrix represents the final predicted category of the model, and each column represents the actual label of the test set data. The confusion matrices of the method without height factor and the proposed method with height factor are shown in Figure 8a,b, respectively. In this paper, we set $\epsilon = 0.8$. It can be found that the identification accuracy of proposed method under over-sea and suburban scenarios reaches 100%. The identification accuracy of proposed method under the urban, dense urban and high-rise urban scenarios increases 14%, 52%, and 2% than the method without height factor, and the overall identification accuracy increases by 14% from 77% to 91%. Although the channel characteristics of urban scenarios are very similar, the performance of proposed identification method is still better than that of the method without height factor.

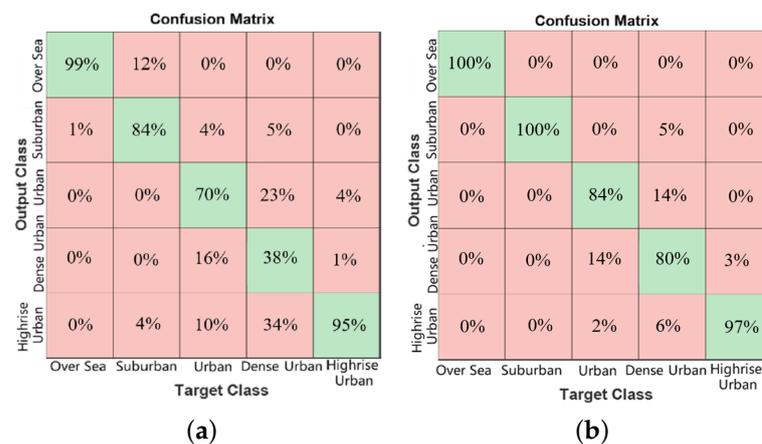
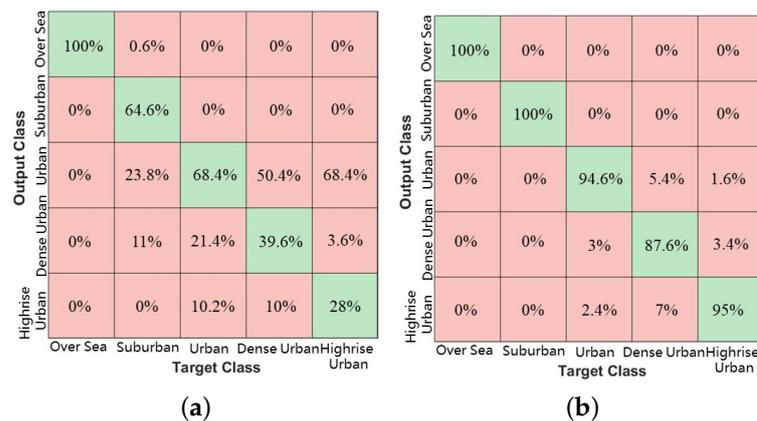


Figure 8. Confusion matrices of the method (a) without height factor and (b) with height factor.

Considering the channel characteristics in this paper are also sensitive to the transmitter height, we perform more simulations under different transmitter height to obtain the training data. The transmitter height ranges from 1.7 m to 6 m for each scenario. The layout parameters of the training dataset are the same as in Table 3, and the layout parameters of the testing data are shown in Table 5. The confusion matrices of the identification model without/with transmitter height variance are compared as shown in Figure 9a,b, respectively. It can be found that the accuracy of the identification model is improved by considering the transmitter height.

Table 5. Validation layout parameters with different transmit heights.

Data Sets	Datapoints Number
1.7 m	100/100/100/100/100
3 m	100/100/100/100/100
4 m	100/100/100/100/100
5 m	100/100/100/100/100
6 m	100/100/100/100/100

**Figure 9.** Confusion matrices of (a) the method only considering the transmitter height of 1.7 m and (b) the method considering the transmitter height from 1.7 m to 6 m.

5. Conclusions

In this paper, a channel-characteristic-based scenario identification model for the A2G communication has been proposed by using SVM. An improved scenario identification method including dataset acquisition and preprocessing, identification model training, and height-integrated model feedback has been developed as well. The datasets of channel characteristics, i.e., RMS-DS, K factor and AS under over-sea, suburban, urban, dense urban and high-rise urban scenarios have been used to train the identification model and validate the proposed identification method. The simulation and validation results have demonstrated that the selection of identification features in this paper is rational, and the identification accuracy of the improved method increases by 14% than the method without height factor. In the future, we will try more different group of identification features and apply it on more scenarios to improve the accuracy and generality of proposed identification method.

Author Contributions: Conceptualization, G.Z. and Y.L.; investigation, G.Z., J.Z.; software, G.Z.; writing—original draft preparation, G.Z.; writing—review and editing, Y.L., K.M., B.H. and S.L.; funding acquisition, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National natural science foundation of China (Grant No. 61871232), in part by the Postgraduate Research & Practice Innovation Program of Jiangsu Province in 2020 under Grant No. SJCX20_0247.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The codes used for the numerical computations are available upon request.

Acknowledgments: The authors are thankful to anonymous reviewers for their instructive reviewing of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

UAV	Unmanned aerial vehicle
SVM	Support vector machine
A2G	Air-to-ground
RMS-DS	Root-mean-square delay spread
AS	Angle spread
RT	Ray tracing
6G	Sixth generation
ML	Machine Learning
V2V	Vehicle-to-vehicle
LOS	Line-of-sight
NLOS	Non-line-of-sight
GIS	Geographic information system
CNN	Convolutional Neural Network
BPNN	Back Propagation Neural Network
LR	Logistic regression
SBR/IM	Shooting and bouncing ray/image
AAOA	Angle spread of azimuth angle of arrival
AAOD	Angle spread of azimuth angle of arrival
EAOA	Angle spread of elevation angle of arrival
EAOD	Angle spread of elevation angle of arrival
ITU-R	International Telecommunication Union-Radiocommunication Sector
PCA	Principal component analysis

References

1. Alsamhi, S.H.; Gupta, S.K.; Rajput, N.S.; Saket, R.K. Network Architectures Exploiting Multiple Tethered Balloon Constellations for coverage extension. In Proceedings of the 6th International Conference on Advances in Engineering Sciences and Applied Mathematics (ICAESAM), Kuala Lumpur, Malaysia, 21–22 December 2016; pp. 24–29.
2. Giordani, M.; Polese, M.; Mezzavilla, M.; Rangan, S.; Zorzi, M. Toward 6G Networks: Use Cases and Technologies. *IEEE Commun. Mag.* **2020**, *58*, 55–61. [\[CrossRef\]](#)
3. Wang, C.X.; Huang, J.; Wang, H.; Gao, X.; You, X.H.; Hao, Y. 6G Wireless Channel Measurements and Models: Trends and Challenges. *IEEE Veh. Technol. Mag.* **2020**, *15*, 22–32. [\[CrossRef\]](#)
4. Gupta, A.; Gupta, S.K.; Rashid, M.; Khan, A.; Manjul, M. Unmanned aerial vehicles integrated HetNet for smart dense urban area. *Trans. Emerg. Telecommun. Technol.* **2020**, *9*, 1–22. [\[CrossRef\]](#)
5. Gupta, A.; Sundhan, S.; Gupta, S.K.; Alsamhi, S.H.; Rashid, M. Collaboration of UAV & HetNet for better QoS: A Comparative Study. *Int. J. Commun. Syst.* **2020**, *5*, 309–333.
6. Rashid, A.; Sharma, D.; Lone, T.A.; Gupta, S.; Gupta, S.K. Secure Communication in UAV Assisted HetNets: A Proposed Model. In *Security, Privacy and Anonymity in Computation Communication and Storage (SpaCCS)*; Springer: Charm, Switzerland, 2019; pp. 427–440.
7. Gupta, A.; Sundhan, S.; Alsamhi, S.H.; Gupta, S.K. Review for capacity and coverage improvement in aerially controlled heterogeneous network. In Proceedings of the International Conference on Optical and Wireless Technologies, Jaipur, India, 10–11 February 2018; pp. 365–376.
8. Gupta, A.; Gupta, S.K. UAV Aided Fog Network (UAFN): A Proposal Framework for Better QoS. In Proceedings of the 2nd International Conference on Computing and Information Technology (ICCIT), Tabuk, Saudi Arabia, 25–27 January 2022; pp. 265–270.
9. Sharma, M.; Gupta, A.; Gupta, S.K.; Alsamhi, S.H. Survey on Unmanned Aerial Vehicle for Mars Exploration. *Drones* **2021**, *6*, 4. [\[CrossRef\]](#)
10. Khawaja, W.; Guvenc, I.; Matolak, D.W.; Fiebig, U.C.; Schneckenberger, N. A Survey of Air-to-Ground Propagation Channel Modeling for Unmanned Aerial Vehicles. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 2361–2391. [\[CrossRef\]](#)
11. Zeng, Y.; Zhang, R.; Lim, T.J. Wireless Communications with Unmanned Aerial Vehicles: Opportunities and Challenges. *IEEE Commun. Mag.* **2016**, *54*, 36–42. [\[CrossRef\]](#)
12. Huang, C.; He, R.; Ai, B.; Molisch, A.F.; Lau, B.K.; Haneda, K.; Liu, B.; Wang, C.X.; Yang, M.; Oestges, C.; et al. Artificial intelligence enabled radio propagation for communications-Part II: Scenario identification and channel modeling. *IEEE Trans. Antennas Propag.* **2022**, in press. [\[CrossRef\]](#)
13. He, R.; Renaudin, O.; Kolmonen, V.; Haneda, K.; Zhong, Z.; Ai, B.; Hubert, S.; Oestges, C. Vehicle-to-vehicle radio channel characterization in crossroad scenarios. *IEEE Trans. Veh. Technol.* **2016**, *65*, 5850–5861. [\[CrossRef\]](#)

14. Zhong, Z.; Ai, B.; Hong, W.; Ding, J. Study on the Propagation Scenario Classification of the High-speed Railway GSM-R System based on GIS. In Proceedings of the 5th IET International Conference on Wireless, Mobile and Multimedia Networks (ICWMMN 2013), Beijing, China, 12–15 June 2013; pp. 306–310.
15. Morocho-Cayamcela, M.E.; Maier, M.; Lim, W. Breaking Wireless Propagation Environmental Uncertainty with Deep Learning. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 5075–5087. [[CrossRef](#)]
16. Greenberg, E.; Bar, A.; Klodzh, E. LOS Classification of UAV-to-Ground Links in Built-Up Areas. In Proceedings of the IEEE International Conference on Microwaves, Antennas, Communications and Electronic Systems (COMCAS 2019), Tel Aviv, Israel, 4–5 November 2019; pp. 1–5.
17. Zhang, J.; Salmi, J.; Lohan, E.S. Analysis of kurtosis-based LOS/NLOS identification using indoor MIMO channel measurement. *IEEE Trans. Veh. Technol.* **2013**, *62*, 2871–2874. [[CrossRef](#)]
18. Tang, P.; Zhang, J.; Molisch, A.F.; Smith, P.J.; Shafi, M.; Tian, L. Estimation of the K-factor for temporal fading from single-snapshot wideband Measurements. *IEEE Trans. Veh. Technol.* **2019**, *68*, 49–63. [[CrossRef](#)]
19. Huang, C.; Molisch, A.F.; He, R.; Wang, R.; Zhong, Z. Machine learning-enabled LOS/NLOS identification for MIMO systems in dynamic environments. *IEEE Trans. Wirel. Commun.* **2020**, *27*, 1292–1304. [[CrossRef](#)]
20. Li, H.; Li, Y.; Zhou, S.; Wang, J. Wireless Channel Feature Extraction via GMM and CNN in the Tomographic Channel Model. *J. Commun. Inf. Netw.* **2017**, *2*, 41–51. [[CrossRef](#)]
21. Yang, M.; Ai, B.; He, R.; Shen, C.; Wen, M.; Huang, C.; Li, J.; Ma, Z.; Chen, L.; Li, X.; Zhong, Z. Machine-Learning-Based Scenario Identification Using Channel Characteristics in Intelligent Vehicular Communications. *IEEE Trans. Intell. Transp.* **2021**, *22*, 3961–3974. [[CrossRef](#)]
22. Zhou, T.; Wang, Y.; Wang, C.X.; Salous, S.; Liu, L.; Tao, C. Multi-Feature Fusion Based Recognition and Relevance Analysis of Propagation Scenes for High-Speed Railway Channels. *IEEE Trans. Veh. Technol.* **2020**, *69*, 8107–8118. [[CrossRef](#)]
23. AlHajri, M.I.; Ali, N.T.; Shubair, R.M. Classification of Indoor Environments for IoT Applications: A Machine Learning Approach. *IEEE Antennas Wirel. Propag.* **2028**, *27*, 2164–2168. [[CrossRef](#)]
24. Yang, M.; Ai, B.; He, R.; Chen, L.; Li, X.; Li, J.; Zhang, B.; Huang, C.; Zhong, Z. A Cluster-Based Three-Dimensional Channel Model for Vehicle-to-Vehicle Communications. *IEEE Trans. Veh. Technol.* **2019**, *68*, 5208–5220. [[CrossRef](#)]
25. Matolak, D.W.; Sun, R. Air-ground channel characterization for unmanned aircraft systems: The over-freshwater setting. In Proceedings of the Integrated Communications, Navigation and Surveillance Conference (ICNS), Herndon, VA, USA, 8–10 April 2014; pp. 1–5.
26. Sun, R.; Matolak, D.W. Air-Ground Channel Characterization for Unmanned Aircraft Systems—Part II: Hilly & Mountainous Settings. *IEEE Trans. Veh. Technol.* **2017**, *66*, 1913–1925. [[CrossRef](#)]
27. Matolak, D.W.; Sun, R. Air-Ground Channel Characterization for Unmanned Aircraft Systems—Part III: The Suburban and Near-Urban Environments. *IEEE Trans. Veh. Technol.* **2017**, *66*, 6607–6618. [[CrossRef](#)]
28. Rodríguez-Piñeiro, J.; Domínguez-Bolaño, T.; Cai, X.; Huang, Z.; Yin, X. Air-to-ground channel characterization for low-height UAVs in realistic network deployments. *IEEE Trans. Antennas Propag.* **2021**, *69*, 992–1006. [[CrossRef](#)]
29. Izydorczyk, T.; Tavares, F.M.L.; Berardinelli, G.; Bucur, M.C.; Mogensen, P.E. Angular distribution of cellular signals for UAVs in urban and rural scenarios. In Proceedings of the European Conference on Antenna and Propagation (EUCAP), Krakow, Poland, 31 March–5 April 2019; pp. 1–5.
30. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
31. Ponte, P.; Melko, R.G. Kernel methods for interpretable machine learning of order parameters. *Phys. Rev. B* **2017**, *96*, 205146–205153. [[CrossRef](#)]
32. ITU-R Recommendation P.1410-2 Propagation Data and Prediction Methods for The Design of Terrestrial Broadband Millimetric Aadio Access Systems. Available online: <https://www.itu.int/rec/R-REC-P.1410-2-200304-S/en> (accessed on 15-04-2003).
33. Lever, J.; Krzywinski, M.; Altman, N. Points of significance: principal component analysis. *Nat. Methods* **2017**, *14*, 641–642. [[CrossRef](#)]
34. Lyu, W.; Li, Y.; Liu, Z.; Huang, C.; He, R. A target recognition-based NLOS identification algorithm. In Proceedings of the IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting (APS/USNC-URSI 2019), Atlanta, GA, USA, 7–12 July 2019; pp. 2093–2094.
35. Huang, C.; Molisch, A.F.; Wang, R.; Tang, P.; He, R.; Zhong, Z. Angular Information-Based NLOS/LOS Identification for Vehicle to Vehicle MIMO System. In Proceedings of the IEEE International Conference on Communications Workshops (ICC Workshops 2019), Shanghai, China, 20–24 May 2019; pp. 1–6.