The feature extraction and matching of time-varying fast fading channel

Zhang Yankui[1,a], Wang Daming[2,b], BaBin[3,c]

[1,2,3]The PLA Information Engineering University,Zhengzhou City, 450001, China

[a]834799435@qq.com, [b]wdm123 @aliyun.com, [c]xidianbabin@163.com

Keywords:fingerprint database; feature extraction;Scene RecognitionOFDM

Abstract. Forthe issues of feature extraction and scene recognitionin the complexwireless transmission channel, this paper presents a feature recognition based on a "fingerprint library" algorithm. The algorithm is fully tap the different scenarios, differentiated features within the different regions of the radio channel, for different channel parameters were extracted using the improved KNN algorithm to classify the channel scene. Simulation analysis and experimental data show that the algorithm solves the channel characteristics problemsof wireless channel parameter extraction and channel scene recognition.

Introduction

The wireless communication not only lead to the rapid development of the global economy,but also have a profound impact on human life and social development. In mobile communication,the radio channel between the sender and the receiver is closely related to the surrounding environment, the radio channel in different environments have some differentiating features. Differentiated features wireless channel radio channel called "fingerprint". Radio channel "fingerprint" feature extraction and matching algorithms, that is, extract differentiated featuresof different scenes or differentiation characteristics in different regions of the wireless channel based on a priori model and test data, and then analyzed and summarized the "fingerprint" and gives clear and accurate "mathematical description." How to find or extract these characteristics and apply to optimize the wireless network, it is currently a hot topic.

[1] [1] introduced the varying channel estimation method, but not make presentation and analysis ofhow to extractthe channel characteristic parameters, channel estimation performance is poor, and cannot be used for time varying channel. 错误!未找到引用源。错误!未找到引用源。错误!未找到引用源。 introduced the calculation methodof Dopplershift and Les K factor, but the application environment is slowly fading channel, it is not fit to change thescene ofvarying channel.错误!未找到引用源。proposed a PDP algorithm for time-varying channel parameter extraction, but it is good performance at high SNR, when the SNR is low, the channel parameter extraction failed. 错误!未找到引用源。错误!未找到引用源。错误!未找到引用源。describes the application of KNN classification algorithm, under which the modelsneed sufficiently large sample space, when the sample space is small, the algorithm errors. Based on the analysisabove, this paper put forward an algorithm to identify the channel scene with the extraction oftime-varying channel characteristic parameters andmatch the channel characteristics using the improved KNN algorithm.

Varying channel feature extraction

Take into account of the diversity and complexity of the time-varying wireless channel , this paperconsider the main factors affecting the quality of the channel are: RMS delay spread, varying characteristics of the channel and the Les K factor and other parameters.

Channel impulse response value acquisition. In order to improve the transmission quality of the signal, filteris usually addedto the transmitter and receiver systems. The real impact of thechannel measurementfrom the transceiver and sender filter can be equivalently represented by a function $g(k)$, the result of the channel measurements as follow:

$$r(k)=\sum_{m=0}^{M-1}h(k-m)\cdot g(m) \quad k=0,1,\cdots,K-1 \quad (1)$$

Usually, the characteristic parameter extraction requires real channel responsein some conditions, so the influence of the equivalent filter function need to be eliminatedfirstly. The equivalent filter actually has a low-pass characteristic, and the high-frequency componentwill be filtered out after the convolution of the channel response and theequivalent filter, therefore it cannot simply use the method of de-convolution for channel response recovery. It can be regarded as the input signal, and then the question can be converted to an issue which is known received signal and the channel response, it is fit to usethe method of least squares estimation.

Set a sample number K, in general $K \geq L$, the nth channel impulse response of the samples is:

$\boldsymbol{h}(n) = \left[h_0(n), h_1(n) \cdots h_L(n)\right]^T$ (2)

The equivalent filter is $\boldsymbol{g}(k) = \left[g(k), g(k-1) \cdots g(k-M+1)\right]^T$, in which M is the sample points of transmitting and receiving filter.

The received datamatrix of equivalent filter is defined as follow：

$$r(k,n) = \sum_{m=0}^{M-1} \boldsymbol{h}(k-m,n) \cdot g(m) + u(k,n) \quad (3)$$

among them $k = 0,1,\cdots K-1, n = 0,1,\cdots N-1$.(3)can be abbreviated written as a vector form:

$\boldsymbol{r(n)} = \boldsymbol{Gh_l(n)} + \boldsymbol{u(n)}$ (4)

Channel Estimation equations can be establishedwith the algorithm of least-squares:

$\hat{\boldsymbol{h}}(n) = \arg\min \left\|\boldsymbol{r}(n) - \boldsymbol{Gh}(n)\right\|^2$ (5)

RMS delay spread (RMS-DS) extract. In the mobile radio communication system from the transmitter to the receiver, the signal tends to go through more than one path, which is called multipath propagation, also known as small-scale fading. Small-scale fading have a critical impact on the transmission quality of the signal, so its design determines the choice of transmission technology and digital receivers. Radio channel is mainly composed of small-scale fading, multipath propagation delay spread and Doppler spread.

After the formula(5) gives the channel response value, the equivalence of the TDE, sinusoidal frequency estimation and DOA estimation was taken used to get the multipath delaywith MUSIC spatial spectrum estimation algorithm

$$h(k) = \sum_{l=0}^{L-1} h_l \cdot \delta(k-\tau_l),\ k\text{=}1,0,...,K\text{-}1 \quad (6)$$

Change the channel $\hat{\boldsymbol{h}}(n)$ to the corresponding Fourier transform, and the frequency-domain signal samples, set the sampling frequency f_s, number of samples N, $f_k = -f_S \tau_l / N$. And formula (6) changesto:

$$H(k) = \sum_{l=0}^{L-1} h_l \cdot e^{\mathrm{j}2\pi f_k k},\ k = 0,1,\ldots,K-1 \quad (7)$$

Delay estimation problem thus converts to complex sinusoidal signal frequency estimation. The channel autocorrelation matrix was corresponding to characteristics decomposition, take D eigenvector corresponding the largest eigenvalues span the signal subspace $\hat{U}_S$, and the noise subspace $\hat{U}_N$ was spanned with eigenvectors corresponding the small eigenvalues

$$\hat{U}_S = [V_1, V_2, ..., V_D],\ \hat{U}_N = [V_{D+1}, V_{D+2}, ..., V_L] \quad (8)$$

Autocorrelation matrix can be expressed as:

$$\boldsymbol{R} = \sum_{i=1}^{D} \lambda_i \boldsymbol{V}_i \boldsymbol{V}_i + \sum_{i=D+1}^{L} \lambda_i \boldsymbol{V}_i \boldsymbol{V}_i = \sum_{i=1}^{D} A_i^2 \boldsymbol{e}_i \boldsymbol{e}_i + \delta_N^2 \boldsymbol{I} \quad (9)$$

Among it, $\boldsymbol{e}_i = [1, e^{-\mathrm{j}2\pi f_s \tau_i / N}, e^{-\mathrm{j}2\pi f_s \tau_i / N \cdot 2}, ..., e^{-\mathrm{j}2\pi f_s \tau_i / N \cdot (L\text{-}1)}]^{\mathrm{T}}$, A_i is the constant coefficients, δ_N^2 is the noise power, $\boldsymbol{I}$ is the unit matrix.

It can be proved that $\{\boldsymbol{V}_1, \boldsymbol{V}_2, ..., \boldsymbol{V}_D\}$ and $\{\boldsymbol{e}_1, \boldsymbol{e}_2, ..., \boldsymbol{e}_D\}$ span to the same space, and $\hat{U}_S$ is orthogonalto $\hat{U}_N$, so

$$\boldsymbol{e}_i^{\mathrm{H}} \sum_{j=D+1}^{L} \alpha_l \boldsymbol{V}_j = 0,\ i = 1,2,...,D \quad (10)$$

Therefore, when $\boldsymbol{e} = \boldsymbol{e}_i$

$$\hat{S}_{MUSIC}(f) = \frac{1}{\boldsymbol{e}^{\mathrm{H}} \hat{\boldsymbol{U}}_N \hat{\boldsymbol{U}}_N^{\mathrm{H}} \boldsymbol{e}} \to \infty \quad (11)$$

The Lth multipath delay of the nth sample can be obtained by the MUSIC spectrum:

$$\boldsymbol{\tau}(n) = \left[\tau_0, \tau_1, \cdots, \tau_L\right],\ n = 0,1,...,N\text{-}1 \quad (12)$$

Define the power of the Lth multipath delay of the nth sample, and its expression is calculated:

$$P_l(k,n) = \left\|h_l(k,n)\right\|^2 \quad (13)$$

among them $k = 0,1,\cdots K-1$.So finally can get RMS delay spread expression:

$$\tau_{rms}(n) = \sqrt{\frac{\sum_{k=0}^{k=K-1} P_l(k,n)\tau^2(k,n)}{\sum_{k=0}^{k=K-1} P_l(k,n)} - \left(\frac{\sum_{k=0}^{k=K-1} P_l(k,n)\tau(k,n)}{\sum_{k=0}^{k=K-1} P_l(k,n)}\right)^2} \quad (14)$$

Feature extraction of the time varying channel. In order to distinguish the varying feature of the channel over time, now define the nth received signal energy:

$$P_{r\_sig}(n) = \sum_{k=0}^{K-1} \left|r(n,k)\right|^2 \quad (15)$$

The faster the signal energy changes in the time-domain, the more high frequency components in the frequency domain. To characterize the rate of speed received signal energy varies with time, FFT and normalization process can be used to deal with the received signal energy at different times , which was expressed as formula(16)

$$\boldsymbol{F}_{r\_sig}(n) = \frac{1}{\max[\boldsymbol{F}_{r\_sig}(n)]} FFT(\boldsymbol{P}_{r\_sig}(0), \boldsymbol{P}_{r\_sig}(1), \cdots \boldsymbol{P}_{r\_sig}(N)) \quad (16)$$

In practice, we can calculate the frequency response of the received signal energy, and then calculate the euclidean distance, then matching a given scenewith a certain algorithm and feature.

Feature extraction of Les K factor. Les K factor characterizing the ratio of thethe direct path energy and scattering component energy, the larger K value is ,the stronger the direct path energy is, $K = +\infty$ meansthe direct component is zero, without the presence of LOS, then, $K = -\infty$ means that the direct component is large enough and scattering component can be neglected, i.e. strong LOS scenario. So Les K factor can be used as a measure of the direct path energy of the three scenarios fingerprints.

Les K factor estimations mainly include KS statistical tests, moment estimation, maximum likelihood estimation, a simple, high accuracy and computation method based on the estimated moment, so this paper estimates a moment, two moments estimate.

$\mu_1 = E(\left|h_{n,k,\tau}\right|)$ 、 $\mu_2 = E(\left|h_{n,k,\tau}\right|)$ is the first moment and second moment, the relationship between the μ_1 、 μ_2 is as follows

$$f(K) = \frac{\mu_1^2}{\mu_2} = \frac{\pi e^{-K}}{4(K+1)}\left[(K+1)I_0(\frac{K}{2}) + KI_1(\frac{K}{2})\right]^2 \quad (17)$$

Equation(17)is not closed-form solution with respect to K values, butcan be obtained through the construction of the table look-up approach, which is to construct a function of the value tableof

$f(K)$, solving minimum error corresponding to K values according to the table in this article, the precision value of K Set 0.001. Table value of the curve shown inFig.1:

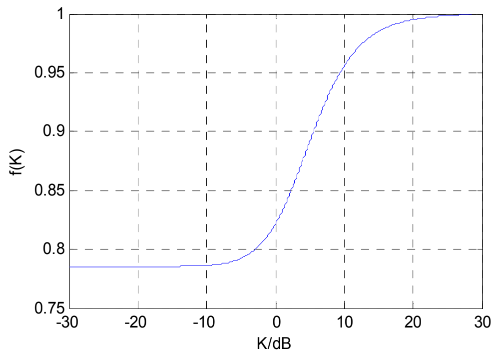

Fig.1Relationship with K value curve

The parameters $\mu_1 = E(|h_{n,k,\tau}|)$ 、 $\mu_2 = E(|h_{n,k,\tau}|^2)$ are the key for K factor estimation, but the actual $h_{n,k}$ needs to be deal withdue to the effect of channel noise, fading and different delays. As the followingsteps

(1) Calculation the channel response power delay profile (PDP):

$$PDP(n,k) = |h_{n,k,\tau}|^2 \quad (18)$$

(2) Identify the PDP effective signal component

Signal acquisition is a pulse signal, the active ingredient means an a signal near the peak of the pulse signal after exclusion of noise arriving through different paths, it is considered invalid signal sampled when there is no arrive signal in the K = 100 sampling points.

(3) μ_1 、 μ_2 are calculated using the look-up table, and the kth sample of Les K factor value calculated after completion of all the statistical sample cumulative distribution function CDF.

Varying channel characteristics match

Based on the analysisabove, the normalized power spectrum $\boldsymbol{F}_{r\_sig}(n)$, delay spread $\tau_{rms}(n)$ and K factor $\boldsymbol{K}_{CDF}(n)$ can be obtained. By extracting characteristic parameter configuration feature vector of each scene, and then regard these feature vectors as the training sample, through a supervised learning algorithm to classify aggregation model.

To complete the classification of the channel characteristics, this papergives the corresponding improvement algorithm according to the studying typical supervised learning algorithm,K nearest neighbor classification algorithm (KNN, k-Nearest Neighbor). K-nearest neighbor classification algorithm is one of themost classic approach data mining classification techniques. The core idea KNN algorithm is that if a sample of K in the feature space most adjacent sample belongs to a category, the sample also falls into this category, and the category having the sample characteristics. The method based solely on the nearest one or several categories of samples to determine the category to be sub-sample belongs in determining the classification decision,andonly related with a very small amount of adjacent samples. Since KNN method is mainly limited by the surrounding adjacent samples, rather than the discrimination class field, so it is more suitable for overlap or class field of more sample set.

KNN classification algorithm is performed in two steps, first established feature space, in this model established feature space using three scenes feature vector composed of 15 samples, due to the cross-border use KNN classification algorithm, so the establishment of the sample feature vectors as follows :

$$\boldsymbol{C}(n) = \left[\boldsymbol{K}_{CDF}^{T}(n), \boldsymbol{F}_{r\_sig}(n)\right]^{T} \quad (19)$$

$\boldsymbol{F}_{r\_sig}(n)$, $\boldsymbol{K}_{CDF}(n)$ of $\boldsymbol{C}(n)$ from the same sample. Therefore, 15 samples constitute a characteristic feature vector space. For any sample, KNN algorithm should select a threshold valuefirstly, then placethe test sample into the feature space, calculate theeuclidean distance between test sample and feature space eigenvectors, when the Euclidean distance is greater than the threshold value, this feature vector will not be considered, the number of feature within the ball of the threshold will be count among the three scenarios, determination test samples belonging to the scene whosenumber of channelis the largest.

Considering the limited training samples in this model, it will result in a miscarriage of justice if simple classification with several recent results of sample. Thus, the algorithm can be improved by dividing the original space into two spaces, establish a K factor of feature space $\boldsymbol{D}_K$ and receive signals normalized power spectral feature space $\boldsymbol{D}_F$, $\boldsymbol{K}_{CDF}(n)\in\boldsymbol{D}_K$, $\boldsymbol{F}_{r\_sig}(n)\in\boldsymbol{D}_K$. In order to take advantage of all the features vectorsof object space, and give different weights accordance to the Euclidean distance, now establish the following evaluation function:

$$W=\sum_{n=1}^{N}\left(\frac{Length(\boldsymbol{K}_{CDF}(n))}{\left\|\boldsymbol{K}_{CDF}(n)-\boldsymbol{K}_{test}\right\|}+\frac{Length(\boldsymbol{F}_{r\_sig}(n))}{\left\|\boldsymbol{F}_{r\_sig}(n)-\boldsymbol{F}_{test}\right\|}\right)\quad(20)$$

$\boldsymbol{K}_{test}$, $\boldsymbol{F}_{test}$ were testing samples of the K-factor and normalized power spectrum of the received signal, $Length(\boldsymbol{K}_{CDF}(n))$ and $Length(\boldsymbol{F}_{r\_sig}(n))$ are the length of two types feature vectors. From the weight function, the farther the distance between the test sample feature vectors,the smaller weights. The maximum weight of the final scene is where the test sample scene.

Performance Simulation and Found

In order to analyzethe main factors affecting thechannel, three known scene of the received data and two scenes to be matched data is given, and then feature extraction and matching measured according to the above method.

Channel Response Analysis. In order to verify the algorithm for channel estimation accuracy, this paper analyzetime domainsignal under three scenariosfirstly, the original received data signal amplitude as shown inFig.2:

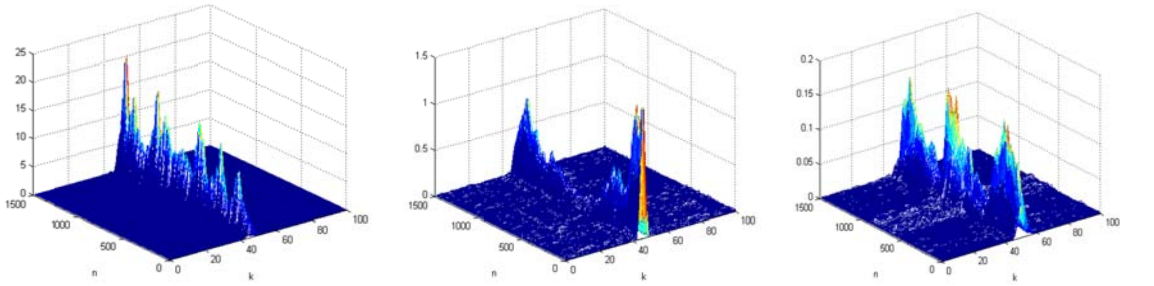

Fig.2The three-dimensional amplitudemap of originally received data signal

After receiving the signal through the least squares estimation, to give impulse response of the channel, and then let the filter as the input signal and convolvewith the estimated channel impulse response, results are shown inFig.3:

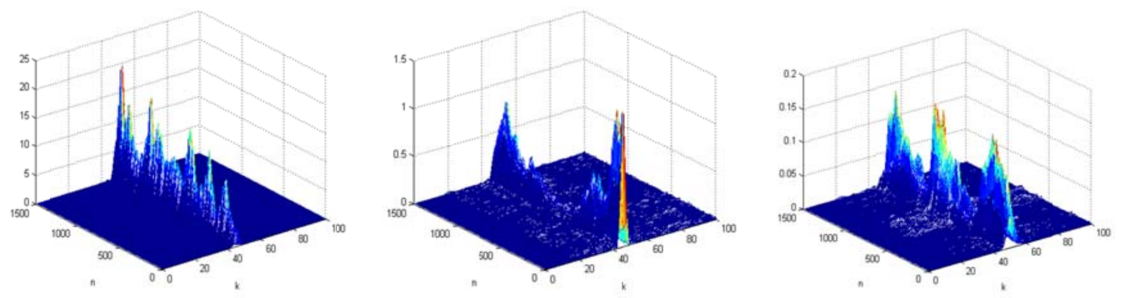

Fig.3 Thethree-dimensional amplitude map of equivalent filter estimated channel

It is shows that the three-dimensional map ofequivalent filter through the channel are almost the same with the original map, it is possible to determine that the estimation method is valid, the error is small by comparison.

Channel variable parameter simulation. For the issue,it can get the three normalized signal energy change over timebased on the given scenes, as shown in Fig.4:

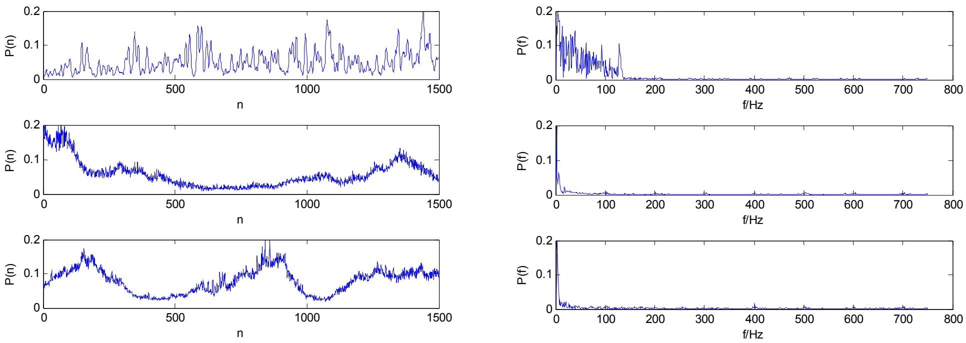

Fig.4 Three scenarios of normalized time-domain diagram and the received signal energy spectrum diagram

Fig.4shows that signal energy change can be fast or slow, because the sender signal is the same within the different samples, so the energy of the received signal can represent the changes of slow and fast characteristics of channel. In order to further describe the change of the channel relationship between speed and signal power spectrum, transform it to frequency domainand can be seen: scene 1 change fast, the corresponding high frequency component is rich. The high frequency components of the scene 2 and 3 are poor. So it can use the received signal power spectrum as a parameter to characterizethe rate of channel change.

The simulation results ofLes K factor. Through a given sample data, the measurement data of all five groups of three scenes Les K factor is as follows:

Tab 1Themean Les K factor of the three scenes

	1	2	3	4	5	average
Scene 1	4.32	4.71	4.56	5.31	5.20	4.86
Scene 2	4.33	4.67	3.84	5.52	5.36	4.74
Scene 3	6.11	6.69	6.23	6.32	6.25	6.32

It can be seen thatLes K factor as an indicator parameter can distinguish three scenes.Scene three can be separated from the other two scenes. The thirdscenes has a rich direct path than the other two scenes, the test scene may be located in relatively open areas.

The improved KNN classification algorithm for identify the channel scene. The conclusion of the two above just researched from the subjective, in order to identify objectivity, it can use the KNN classification algorithm.

To get an accurate determination result, and maximize the use of useful information, it can be calculated weights between the given samples and the build spaceaccording to formula(20), with the following results shown in Tab 2:

Tab 2 Combined KNN algorithm weight of each scene

	Channel scene1	Channel scene2	Channel scene3
Data sample 1 （$i=1$）	2	0.89	0.66
Data sample 2 （$i=2$）	1.43	1.47	1.02

From the improved KNN algorithm, the final result is that: data 1 "Test1ForScene.mat" belongs to category Scene 1, data 2 "Test2ForScene.mat" belongs to Category 2 scenes.

Conclusions

This paperanalysis the characteristics of the radio channel, including time-varying characteristics, the multipath characteristics that direct path energy proportion in the total energy according to extractthe characteristic parameters of the time-varying channel. An improved KNN algorithm is put forwards to match the scene. Simulation and experimental results show that the algorithm can better describe the time-varying channel characteristics, and plays an important role in the modeling of time-varying channel.

Reference

[1] Liu Yanfang. MIMO time-varying channel estimation based on particle filter [D]. Beijing JiaotongUniversity, 2009.

[2] Yang Zhengju. Blind identification algorithm research for time-varying channel [D]. The PLA Information Engineering University, 2012.

[3] He Yujia. Under the environment of high-speed wireless channel measurement and modeling based on LTE key technology research [D]. Beijing JiaotongUniversity, 2013.

[4] TianXuefei. Les high-speed environment factor and the Doppler spectrum measurement research [D]. XidianUniversity, 2011. The DOI: DOI: 10.7666 / d.y. 1865542.

[5] MAZhangyong, YANYongqing, ZHAOChunming, et al. An Improved Channel Estimation Algorithm -based on Estimating Level Crossing Rate for CDMA Receiver [J]. Journal of electronic: English edition, 2003, 2 (2) : 235-238.

[6] Feng Song. Wireless channel measurement parameter extraction algorithm research [D]. XidianUniversity, 2013. The DOI: DOI: 10.7666 / d.D. 365662.

[7] Li Rong, YeShiwei, Shi Zhongzhi, et al.SVM-KNN classifier, a new method for increasing accuracy of SVM classification [J]. Journal of electronics, 2002, 30 (5) : 745-745. The DOI: DOI: 10.3321 / j.i SSN: 0372-0372. 2002.05.035.

[8] BuFanjun. Improved KNN algorithm and its application in text categorization [D]. JiangnanUniversity, 2009. The DOI: DOI: 10.7666 / d.y. 1583945.

[9] Zhang Xiaohui, Li Ying, Wang Huayong, et al. The application characteristics of the polymerization of Chinese text classification improved KNN algorithm [J]. Journal of Northeastern University: natural science edition, 2003, 24 (3) : 229-229. The DOI: DOI: 10.3321 / j.i SSN: 1005-3026.2003.03.007