Distribution Characteristics of Ground Echo Amplitude and Recognition of Signal Grazing Angle

Guangwei Zhang, Ping Li *, Guolin Li and Ruili Jia

Abstract: With the continuous advancement of electronic technology, terahertz technology has gradually been applied on radar. Since short wavelength causes severe ground clutter, this paper studies the amplitude distribution statistical characteristics of the terahertz radar clutter based on the measured data, and provides technical support for the radar clutter suppression. Clutter distribution is the function of the radar glancing angle. In order to achieve targeted suppression, in this paper, selected axial integral bispectrum (selected AIB) feature is selected as deep belief network (DBN) input to complete the radar glancing angle recognition and the network structure, network training method, robustness are analyzed also. The ground clutter amplitude distribution can follow normal distribution at 0~45° grazing angles. The Weibull distribution and G0 distribution can describe the amplitude probability density function of ground clutter at grazing angles 85° and 65°. The recognition rate of different signal grazing angles can reach 91% on three different terrains. At the same time, the wide applicability of the selected AIB feature is verified. The analysis results of ground clutter amplitude characteristics play an important role in the suppression of radar ground clutter.

Keywords: THz ground clutter; bispectrum; amplitude probability density function (PDF); deep belief network

1. Introduction

The rapid development of radar technology has continuously improved the detectability of radar systems, but the environment has a profound impact on target detection [1–3]. The environmental interference of ground and sea clutter has always been a hot research issue. Research on the characteristics of radar clutter can be traced back to the 1950s. Kaplan first proposed a statistical method to study clutter [4]. Researches about clutter characteristics have always received extensive attention. The key issue of clutter research is how to accurately model the clutter amplitude probability density function. At present, researchers have proposed a variety of models based on topography, vegetation, water content, grazing angle, and signal frequency bands. There are lognormal model [5,6], Gauss Rayleigh model [7,8], symmetric steady-state model [9,10], Weibull model [7,11], K distribution model [5,12], G0 model [13,14]. There is little research on the characteristics of near-field clutter above 100 GHz, which provides a direction for the follow-up development.

High-order spectrum analysis is a cutting-edge research direction in signal processing. Among the high-order spectrum, the bispectrum obtained from the third-order correlation function is the most widely used. Two-dimensional bispectrum features are difficult to be further processed. The integral bispectrum transforms two-dimensional the bispectrum function into a one-dimensional function, which is conducive to the calculation. The radial integral bispectrum (RIB) [15], axial integral bispectrum (AIB) [16], circular integral bispectrum (CIB) [17], surrounding-line integrated bispectrum (SIB) [18] and other methods have been proposed. Professor Zhang Xianda proposed the selected bispectrum [19].
The deep belief network can be used for unsupervised learning and supervised learning. DBN is a probabilistic generative model with a traditional discriminant. The generative model establishes joint distribution between observation data and labels. By training the weights, the entire neural network can generate training labels according to the maximum probability [20].

The amplitude distribution characteristics of ground clutter are the basis of radar clutter suppression and play a key role in target detection. There is little research on the characteristics of ground clutter in the terahertz frequency band [21,22]. This paper is based on measured data to fill this gap. Previous studies mainly focused on targeted interference suppression of signals with large grazing angle (≥60°) signal. Since other glancing angles are not considered, no significant research results were obtained [23].

Previously, for grazing angle recognition, the backscattering coefficient of different terrains was mainly studied, but the backscattering coefficient is a multi-factor function. As the environment changes, the backscattering coefficient is also changing, which does not have wide applicability [24,25]. In this paper, measure THz ground clutter from alpine meadows, alpine dry grassland and alpine swamp and echo signal selected AIB are used as DBN network input to recognize grazing angle, to suppress clutter at different grazing angles. The bispectrum characteristics of ground clutter are studied. The amplitude distribution characteristics of ground echo at different grazing angles are analyzed. The axial integrated bispectrum, radial integrated bispectrum, circular integrated bispectrum, surrounding-line integrated bispectrum and selected axial integrated bispectrum (selected-AIB) are calculated. The deep belief network is used to identify different grazing angle echo.

The paper is organized as follows. The amplitude statistical distribution and bispectrum are recalled and analysed in Section 2. The selected AIB and network optimization are analysed in Section 3. The statistical characteristics of the amplitude are analysed based on the measured data in Section 4. DBN are used to classify and identify grazing angles in Section 5. The research content is summarized in Section 6.

2. Area Target Scattering Model and Bispectrum

2.1. Area Target Scattering Model

The terahertz signal wavelength is far smaller than the size of the area target. According to the electromagnetic scattering theory, the scattering characteristics of the area target already belong to the quasi-optical zone. The interaction between the various parts of the scatterer is very weak, and the scattering is almost a local phenomenon rather than accumulation process. The area integration current comes from the stagnation points and the integration end points. The scattering energy of the area target mainly comes from these scattering points, which are called the scattering center of the area target. The echo characteristics of the area target are determined by multiple scattering centers.

The Geometric Theory Diffraction (GTD) [26] is a classic theory for describing scattering center. The approximate formula of the scattering field can be described as follows:

\[
E_q(f(t), \theta, \gamma) = \sum_{m=1}^{M} A_m(t) S_{qm}(\theta, \gamma)(\frac{f}{f_0}) a_m e^{-j4\pi f(x_m \cos \gamma \cos \theta + y_m \cos \gamma \sin \theta + z_m \sin \gamma)}/c
\]  

(1)

The signal energy mainly comes from M strong scattering center. m is the serial number of the scattering center. \( f_0 \) is the center frequency of the broadband signal, \( q \) is specific polarization combination of transmit and receive. \((x, y, z)\), \( \theta \), \( \gamma \) describe the position of the scattering center and the attitude angle. \( S_{qm} \) is the scattering coefficient factor. \( a_m \) is central parameter. \( f \) is signal radio frequency. \( c \) is speed of light.

The total scattered field of the area target can be simplified as:

\[
E_q(f(t)) = \sum_{m=1}^{M} A_m(t) S_{qm}(\frac{f}{f_0}) a_m e^{-j4\pi f r_m}/c
\]  

(2)
where, \( r_m \) represents the distance from the scattering center to the radar in the direction of the radar’s sight line. \( \left( \frac{f}{f_0} \right)^{a_m} \) reflects the relationship between the scattering intensity of the scattering center and the frequency. The area target scattering can be expressed by the above formula. In order to facilitate the analysis and simplify the mathematical model of the signal, the noise term is omitted, and the signal is considered as a sawtooth chirp signal. Sawtooth FM frequency time relationship is shown in Figure 1.

![Figure 1. Sawtooth FM frequency time relationship.](image)

The transmitted signal frequency is expressed as:

\[
 f_T(t) = \Delta F_M f_M t + \frac{1}{2} \Delta F_M f_M \tau + f_0
\]

(3)

The received signal frequency is expressed as:

\[
 f_R(t) = \Delta F_M f_M t - \frac{1}{2} \Delta F_M f_M \tau + f_0
\]

(4)

\( \Delta F_M \) is the frequency offset. \( \tau \) is the delay time. \( f_M \) is modulation frequency.

The ground echo can be expressed as:

\[
 x_q(t) = E_j(f_R(t)) = \sum_{m=1}^{M} A_m(t) S_{qm} e^{-j4\pi f_0 t / c} (1 + \frac{\Delta F_M f_M t - \frac{1}{4} \Delta F_M f_M \tau}{f_0}) e^{-j4\pi f_R(t) r_m / c}
\]

(5)

\[
 r_m = R_{0m} + v_{0m} t
\]

(6)

\( R_{0m}, v_{0m} \) are the initial distance and velocity of scattering center (considering the case of uniform velocity).

\[
 a_{qm} = A_m(t) S_{qm} e^{-j4\pi f_0 t / c}
\]

(7)

\[
 \omega_m = -j4\pi f_R(t) r_m / c
\]

(8)

For terahertz radar, when the bandwidth is the same, it has smaller relative bandwidth. The above formula can be further simplified as:

\[
 x(t) = \sum_{m=1}^{M} a_{qm} e^{j\omega_m t}
\]

(9)

For the sawtooth chirp system, the echo from area target multiple scattering points is a multiple sinusoidal signals combined model.

Substitute \( a_{qm}, \omega_m \) into the formula:

\[
 x(t) = \sum_{m=1}^{M} A_m(t) S_{qm} \exp \left\{ -j4\pi(f_R(t)) \frac{r_m}{c} \right\}
\]

(10)
The ground echo received by the radar is the vector sum of the backscattered electromagnetic field. The center of the ground scatterer in the radar beam changes with space and time, which will affect the clutter amplitude. Therefore, the description for the amplitude distribution characteristics requires statistical methods.

So far, many echo models have been proposed. The more commonly recognized amplitude probability density distribution function models include the Rayleigh distribution, lognormal distribution, Weibull distribution, K distribution, etc. In recent years, the gamma distribution, and G0 distribution have been proposed.

2.1.1. Rayleigh Distribution

Within the radar resolution range, if the number of scatterers is large, the echo amplitude and phase have random characteristics. In the case of low-resolution radar, Rayleigh distribution is more suitable for describing ground clutter distribution. The echo amplitude follows the Rayleigh distribution. If $x$ is the envelope amplitude, and $\sigma^2$ is power, the probability density function is:

$$f(x) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad x \geq 0$$



2.1.2. Lognormal Distribution

$x$ is the envelope amplitude of the clutter. The probability density function is:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right) \quad x \geq 0$$

$\mu$ — Scale parameter, which represents the median of the distribution.

$\sigma$ — The shape parameter, which represents the skewness of the distribution.

2.1.3. Weibull Distribution

In the case of close detection, that is, severe clutter, the Weibull distribution is more suitable for describing ground clutter distribution. The asymmetry of this distribution is weaker than lognormal distribution. $x$ is the envelope amplitude. The probability density function is:

$$f(x) = \frac{p}{q} \left(\frac{x}{q}\right)^{p-1} \exp\left(-\left(\frac{x}{q}\right)^p\right) \quad x \geq 0$$

$q$ — Scale parameter, which represents the median of the distribution.

$p$ — Shape parameter, which represents the skewness of the distribution.

2.1.4. K Distribution

The clutter distribution models of the Rayleigh distribution, lognormal distribution and Weibull distribution are all based on single point statistics, so they are only suitable for single pulse detection. Their main disadvantage is the lack of time and space correlation. In recent years, the K-distribution mixing model has been introduced.

$$f_k(x) = \frac{2}{a\Gamma(v+1)} \left(\frac{x}{2a}\right)^{v+1} K_v\left(\frac{x}{a}\right) \quad x \geq 0$$

where $x$ is the amplitude and $a(a > 0)$ is the quantization parameter whose change determines the characteristics of the clutter distribution. $v$ is the shape parameter $-1 < v < \infty$, and $K_v$ is the modified Bessel function.
2.1.5. Gamma Distribution

The gamma distribution clutter model probability density function is:

\[ f(x, \mu, L) = \frac{L^x}{\mu^L} \left( \frac{Lx}{\mu} \right)^{L-1} \exp\left( -\frac{Lx}{\mu} \right) \quad x \geq 0 \]  

(16)

\( \mu \) is the scale parameter, \( L \) is the shape parameter.

G0 distribution.
The G0 distribution clutter model probability density function is:

\[ f(x, L, \gamma, \alpha) = \frac{2L^L \Gamma(L - \alpha)}{\Gamma(\alpha)(\gamma + Lx^2)^{L-\alpha}} \quad x \geq 0 \]  

(17)

\( L(L > 0) \) is the equivalent sight and \( \gamma(\gamma > 0) \) is the scale parameter, which is related to the average value of the scattered energy. \( \alpha(\alpha < 0) \) is the shape parameter, which reflects the uniformity of the observation area.

K – root distribution.
The K – root distribution clutter model probability density function is:

\[ f(x, L, v, \mu) = \frac{2L^L x^{L-1}}{\Gamma(L)(v)} \left( \frac{Lx}{\mu} \right)^{L-1} K_{L-1}(2 \sqrt{\frac{Lx}{\mu}}) \quad x \geq 0 \]  

(18)

where \( K_{L-1}(\cdot) \) is the second modified Bessel function and \( \Gamma(\cdot) \) is the gamma function. The K-root model can be regarded as a product model of two independent random variables. One variable follows the gamma distribution with mean \( \mu \) and shape parameter \( L \), and the other variable follows the gamma distribution with mean \( \mu \) and shape parameter \( v \). Gamma distribution, G0 distribution and k-root distribution are more flexible to describe changes in clutter distribution.

2.2. Bispectrum Characteristic

The bispectrum of continuous-time signal \( x(t) \) is defined as:

\[ B(\omega_1, \omega_2) = S_3x(\omega_1, \omega_2) = \sum_{\tau_1 = -\infty}^{\infty} \sum_{\tau_2 = -\infty}^{\infty} |C_3x(\tau_1, \tau_2)| \exp[-i(\omega_1 \tau_1 + \omega_2 \tau_2)] \]  

(19)

where

\[ C_3x(\tau_1, \tau_2) = \int_{-\infty}^{\infty} x^3(t)x(t + \tau_1)x(t + \tau_2)dt \]

\[ = E\{x^3(t)x(t + \tau_1)x(t + \tau_2)\} \]  

(20)

The bispectrum is uniquely defined by its values in the triangular region \( 0 \leq \omega_2 \leq \omega_1 \leq \omega_2 + \omega_1 \leq \pi \). The bispectrum can not only reflect the signal amplitude information but also reflect the phase information. The bispectrum can handle a variety of signal models, such as non-Gaussian, nonlinear, noncausal, nonminimum phase, Gaussian coloured noise and blind signals. In theory, the interference of Gaussian noise is completely suppressed, because the third-order cumulant of one-dimensional Gaussian stationary random signals is zero.

3. Selected AIB and Network Optimization

Some feature points are lost in the RIB calculation process, and these points contain useful information. The AIB method also has disadvantage, and most of the integral bispectrum phase information are lost. \( B_p(a, \theta) \) is the polar coordinate representation of \( B(\omega_1, \omega_2) \) in the CIB. Note that, \( B_p(a, k\pi/2) \) with integer \( k \) provides no phase information. \( a \) is near \( k\pi/2 \), and \( B_p(a, \theta) \) provides a little phase information. Only partial phase and amplitude information features are used in the SIB, and the features are not fully utilized.
The axial integral bispectrum (AIB):

\[
\text{AIB} = \frac{1}{2\pi} \int_{-\infty}^{\infty} B(\omega_1, \omega_2) d\omega_2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} B(\omega_1, \omega_2) d\omega_1
\]

(21)

The radial integral bispectrum (RIB):

\[
\text{RIB}(a) = I(a) = \int_{0}^{1/(1+a)} \frac{1}{1+af_1} B(f_1, af_1) df_1
\]

(22)

The circular integral bispectrum (CIB):

\[
\text{CIB}(a) = \frac{2\pi}{\int_{0}^{\theta} B_p(a, \theta) d\theta}
\]

(23)

The surrounding-line integrated bispectrum (SIB):

\[
\text{SIB}(s) = \sum_{R} B(\omega_1, \omega_2)
\]

(24)

3.1. Selected AIB

A one-dimensional Fisher class [19] is introduced to calculate the signal AIB feature point discriminant degree. The AIB feature points with the high discriminant are selected and re-taken as the new feature vectors. The trivial AIB feature points in the signal are deleted. The dimension of the feature vector is reduced so that the network training time gets shorter, and the application ability of the network is improved. \(k = \{1, 2, \cdots, 512\}\).

The one-dimensional Fisher is expressed as:

\[
Xm^{(ij)}(a) = \frac{\sum_{l=i}^{j} [\text{mean}_k(x_k^{(i)}(a)) - \text{mean}_l(\text{mean}_k(x_k^{(i)}(a)))]^2}{\sum_{l=i}^{j} \text{var}_k(x_k^{(i)}(a))}
\]

(25)

\(\text{mean}_k(x_k^{(i)}(a))\) is the mean (centroid) of all the AIBs in the ith class.

\(\text{var}_k(x_k^{(i)}(a))\) is the variance of all the AIBs in the ith class.

\(\text{mean}_l(\text{mean}_k(x_k^{(i)}(a))\) is the total centroid of all the sample AIBs in all the classes.

\(Xm^{(ij)}(a)\) indicates the discriminant between the ith class and jth class at the a point.

The four grazing angles are needed to calculate \(m_{12}(a), m_{13}(a), m_{14}(a), m_{23}(a), m_{24}(a), m_{34}(a)\).

The feature extraction process:

Third-order cumulant is used the to calculate the signal bispectrum and calculate the signal AIB feature according to the bispectrum. Calculate the discriminant degree of different AIB features points according to [25] and perform standardization processing.

\[
m^{(ij)}(a_n) = \frac{m^{(ij)}(a_n)}{\sqrt{\sum_{K=1}^{M} [m^{(ij)}(a_n)]^2}}
\]

(26)

The first \(Q\) points are taken as signal feature vectors, \(m^{(ij)}(a_1) > m^{(ij)}(a_2) \cdots > m^{(ij)}(a_Q)\). Re-establish the four grazing angle signal feature vectors. The \(Q\) value needs to be adjusted. \(i\) and \(j\) class selected AIB feature vectors are:

\(i\) class selected -AIB: \(\{s_i(a_1), s_i(a_2), \cdots, s_i(a_Q)\}\).

\(j\) class selected -AIB: \(\{s_j(a_1), s_j(a_2), \cdots, s_j(a_Q)\}\).
The selected-AIB is used as the DBN input to train the network and output the prediction result.

3.2. Data Preprocessing

During preprocessing, one-dimensional Fisher is used to extract and amplify important features, while trivial features are deleted and suppressed. The data are normalized to improve the convergence speed of the model and prevent the model gradient exploding. The data are randomly scrambled, and K-fold cross validation is performed so that the data sequence does not affect network training. The advantage is to increase the randomness to avoid regular data deviation or falling into local optimum, which improves the generalization performance of the network.

3.3. Optimization of Deep Belief Network

Considering the training speed of the network, the three-layer network is selected for classification. The second hidden layer activation function is restricted Boltzmann machine (RBM). When the feature vector is mapped to different feature spaces, it retains as much feature information as possible to establish joint distribution between the observation data and the label. The number of hidden units is used as a parameter to be adjusted. The third layer is the output layer, which is a fully connected layer. DBN network is shown in Figure 2.

![DBN network](image)

**Figure 2.** DBN network.

\[ V = [v_1, v_2, \cdots, v_N] \text{ is visible layer state. } A = [a_1, a_2, \cdots, a_N] \text{ is the offset of the visible layer. } H = [h_1, h_2, \cdots, h_M] \text{ is state of the hidden layer. } B = [b_1, b_2, \cdots, b_M] \text{ is bias of the hidden layer. } [W_{ij}]_{NM} \text{ is weight matrix between the visible layer and the hidden layer.} \]

The visible layer is real-valued, and the potential function can be expressed as:

\[
E(V, H) = \sum_{i=1}^{N} \frac{1}{2\sigma_i^2} (v_i - a_i)^2 - VH^T - HB^T \quad (27)
\]

\(\sigma_i\) is the standard deviation of the neuron in the visible layer. By defining the energy function, the probability distribution function can be defined:

\[
P(V) = \frac{\sum_{H} \exp(-E(V, H))}{\sum_{V} \sum_{H} \exp(-E(V, H))} \quad (28)
\]

It has been proven that RBM can approximate probability distribution functions arbitrarily. Use the maximum likelihood estimation method to describe the input data
probability distribution function. In order to simplify the calculation, likelihood function that minimizes the negative is adopted:

$$
\text{Loss}(\theta) = - \sum_{i=1}^{\theta} \log p(x_i | \theta) 
$$

(29)

$\theta = \{W, A, B\}$ is the parameter that needs to be adjusted during training. The training process of the network is the optimization process of parameter $\theta$.

Further processing of the above potential function, get the two important inferences $P(H | V), P(V | H). P(H | V)$ indicates the ability of the data features extracted through the network to represent the original data, and judges whether the extracted hidden features can be used for classification. $P(V | H)$ is the basis of network optimization and adjustment, and provides adjustment directions, and guarantees the smooth progress of the entire network training. According to the visible layer, the state of all neurons are directly calculated. This parallel mechanism can greatly speed up training and inference. The reason why it can be calculated in parallel is that RBM eliminates the dependence of neurons between layers.

The hidden layer training process:

$$
p(H = [1, \ldots, 1]_{1 \times M} | V) = \prod_{j=1}^{M} \frac{1}{1 + \exp(-VW - B)} 
$$

(30)

The visible layer training process:

$$
p(V = [1, \ldots, 1]_{1 \times N} | H) = \prod_{i=1}^{N} \frac{1}{1 + \exp(-HW^T - A)} 
$$

(31)

The training process of the network is actually to search the optimal parameters, and the optimal parameters are:

$$
\theta^* = \text{argmin} \text{loss}(\theta) 
$$

(32)

Use the gradient descent method to update the parameters:

$$
\theta(t + 1) = \theta(t) - \eta \frac{\partial \text{Loss}(\theta)}{\partial \theta} 
$$

(33)

$\eta$ represents the learning rate, which is used to control the speed of learning. Aiming at the Bernoulli restricted Boltzmann machine, after adopting contrastive divergence (CD-K)approximation calculation, various parameters are obtained:

$$
\frac{\partial}{\partial W} \text{Loss}(\theta) = V_k H_k - V_0 H_0 
$$

(34)

$$
\frac{\partial}{\partial A} \text{Loss}(\theta) = V_k - V_0 
$$

(35)

$$
\frac{\partial}{\partial B} \text{Loss}(\theta) = H_k - H_0 
$$

(36)

The training process of DBN can be regarded as the initialization of weight parameters, which overcomes the shortcomings of BP network. Back Propagation-Deep Belief Network (BP-DBN) needs to only perform partial search on the weight parameter space. Compared with forward neural network, BP-DBN convergence time is faster.

4. Feature Extraction and Network Optimization

4.1. Measurement and Test Radar

The test signal frequency is 120 GHz, and the bandwidth is 1 GHz, and the modulation method is the sawtooth chirp signal. Radar is an integrated transceiver system.
Transmitting power is 1 mW. The angle between the centerline of the beam and the ground is defined as the glancing angle. The grazing angle is controlled by the rotatable pan-tilt platform. The beam width is 60°. Modulation period is 946 μs. The sampling rate of the beat frequency signal is 1 MHz. There are 512 points per period. The test schematic is shown in Figure 3. The wavelength is 2.5 mm. The antenna is array antenna, as shown in Figure 4. It features on-chip MMW built-in-self-test (BIST) circuits, a harmonic transceiver, software linearization (SWL) circuits, and a digital interface. This chipset has been tested in a low-cost package, where the antennas are integrated. The vertical height of the test equipment from the ground is 5 m.

The radar transmits 120 GHz signal. The echo reflects from the ground when it touches the ground. After the radar receives the echo, it performs mixing frequency processing and outputs beat frequency signal. The computer is connected to the radar data output terminal to record data. Data processing flow is shown in Figure 5. First, analyze the ground clutter amplitude probability density distribution. The ground clutter amplitude PDF is compared with the normal distribution, Rayleigh distribution, K distribution, Weibull distribution, lognormal distribution, gamma distribution, G0 distribution, and K-root distribution.

Figure 3. Test equipment.

Figure 4. Integrated test radar.

Figure 5. Data processing flow.
4.2. Alpine Meadow Amplitude PDF

The alpine meadow is covered with low vegetation, and the soil is moist. The ground is relatively flat. An alpine meadow test environment is shown in Figure 6.

![Alpine meadow test environment](image)

**Figure 6.** Alpine meadow test environment.

The 200 periodic signals are used to analyse signal amplitude PDF. The per period contains 512 points. Range between maximum amplitude and minimum amplitude is divided into 512 parts. The point number in per part is calculated, and know the measured signal PDF curve. According to the theoretical distribution formulas, theoretical distribution curves can be obtained.

The measured data amplitude PDF contains 512 points. Theoretical distribution curves also contains 512 points. The corresponding points are subtracted, and then difference sums is used as the fitting error. Use fitting error to determine the fit degree between the measured signal PDF curve and theoretical distribution curves.

$s(i)$ is the measured data amplitude PDF curve. $k(i)$ is theoretical distributions curve. The fitting error is expressed as:

$$err = \sum_{i=1}^{512} |s(i) - k(i)|$$  \hspace{1cm} (37)

The goodness of fit is introduced to describe the shape fit degree between the theoretical distribution curve and the ground echo amplitude PDF curve. The statistical parameter of goodness is expressed as the determination coefficient $R^2$. The maximum value $R^2$ is 1, indicating that the shape fit result between the curves is perfect.

$$R^2 = 1 - \frac{\sum_{i=1}^{512} (y_i - \bar{Y})^2}{\sum_{i=1}^{512} |Y_i - \bar{Y}|^2}$$  \hspace{1cm} (38)

Next, the measured echo amplitude PDF curve is compared with theoretical distributions curve. At grazing angle 20°, the ground echo amplitude PDF and theoretical distributions curve are shown in Figure 7. The fitting error and goodness of fit $R^2$ are shown in Table 1.
The first picture is the comparison between the measured data amplitude distribution and the normal distribution. The second picture shows the comparison between the measured data amplitude distribution and other theoretical distributions. Different theoretical distribution curves are distinguished by different curve shapes. Through observation, it is found that the curves degree of fit between the measured data amplitude distribution and the normal distribution is high. Other theoretical amplitude distribution curves and measured data distribution curves reflect the differences in peak, peak position, and curve fluctuations.

Fitting error describes the difference from the digital perspective. Fitting error from measured signal amplitude PDF curve and normal distribution curve is 0.13. It is the smallest among all fitting errors. Fitting error from measured signal amplitude PDF curve and G0 distribution curve is 0.18. $R^2$ describes the difference from the curve shape perspective. Normal distribution curve $R^2$ is 0.97, and it can explain that normal distribution curve shape is closest to measured signal amplitude PDF curve. The suboptimal $R^2$ value is G0 0.95. Through the above analysis, at grazing angle 20°, measured signal amplitude PDF curve can be described as normal distribution.

At grazing angle 45°, the ground echo amplitude PDF and theoretical distributions curve are shown in Figure 8. The fitting error and goodness of fit $R^2$ are shown in Table 2.
The first picture is the comparison between the measured data amplitude distribution and the normal distribution. The second picture shows the comparison between the measured data amplitude distribution and other theoretical distributions. Through observation, it is found that the degree of fit between the measured data amplitude distribution and the normal distribution is high at grazing angle 45°. The four distribution functions of Lognormal, Gamma, G0, K-root four distribution functions gradually approach the measured data amplitude distribution, and the divergence gradually decreases.

Fitting error from measured signal amplitude PDF curve and normal distribution curve is 0.10. It is the smallest among all fitting errors. Fitting error from measured signal amplitude PDF curve and G0 distribution curve is 0.13. Normal distribution and measured signal amplitude PDF curve $R^2$ is 0.98, and it explain that normal distribution curve shape is closest to measured signal amplitude PDF curve. The suboptimal $R^2$ value is G0 0.97. Through the above analysis, at grazing angle 45°, measured signal amplitude PDF curve can be described also as normal distribution.

At grazing angle 65°, the ground echo amplitude PDF and theoretical distributions curve are shown in Figure 9. The fitting error and goodness of fit $R^2$ are shown in Table 3.

![Figure 9. The 65° ground clutter PDF and theoretical curves.](image)

### Table 2. The 45° curve fitting error and $R^2$.

| Distribution | Fitting Error | $R^2$  |
|--------------|---------------|--------|
| Rayleigh     | 0.61          | 0.54   |
| K            | 0.66          | 0.49   |
| Weibull      | 0.33          | 0.83   |
| Lognormal    | 0.18          | 0.95   |
| Gamma        | 0.14          | 0.97   |
| G0           | 0.13          | 0.97   |
| K-root       | 0.15          | 0.97   |
| normal       | 0.10          | 0.98   |

The first picture is the comparison between the measured data amplitude distribution and the normal distribution. It is obvious that the peak value position of the measured

![Table 3. The 65° curve fitting error and $R^2$.](image)
data amplitude distribution curve moves forward with the increase of the grazing angle. The normal distribution is no longer the optimal description curve. Through observation, it is found that the degree of fit between the measured data amplitude distribution and the G0 distribution curve distribution is high.

Fitting error from measured signal amplitude PDF curve and normal distribution curve is 0.15. Fitting error from measured signal amplitude PDF curve and G0 distribution curve is 0.14, same as Weibull distribution. It is the smallest among all fitting errors. Normal distribution curve $R^2$ is 0.94. Weibull distribution and G0 distribution curve $R^2$ value is G0 0.96. Through the above analysis, at grazing angle 65°. In terms of $R^2$ and fitting error, Weibull distribution and G0 distribution has ability to describe the measured signal amplitude PDF curve.

At grazing angle 85°, the ground echo amplitude PDF and theoretical distributions curve are shown in Figure 10. The fitting error and goodness of fit $R^2$ are shown in Table 4.

![Figure 10. The 85° real ground clutter amplitude PDF and theoretical curves.](image)

**Table 4.** The 85° curve fitting error and $R^2$.

| Distribution     | Fitting Error | $R^2$  |
|------------------|---------------|--------|
| Rayleigh         | 0.15          | 0.95   |
| K                | 0.16          | 0.95   |
| Weibull          | 0.12          | 0.96   |
| Lognormal        | 0.20          | 0.91   |
| Gamma            | 0.15          | 0.95   |
| G0               | 0.12          | 0.97   |
| K-root           | 0.17          | 0.93   |
| normal           | 0.14          | 0.94   |

It is observed that the peak value position of the measured data distributions curve further moves forward, and the difference with the normal distribution curve is further enlarged. Other theoretical distribution curves are gradually approaching the measured data distribution curve. Weibull distribution and G0 distribution from peak value, peak position, and curve fluctuation can describe the measurement data distribution well.

Fitting error from measured signal amplitude PDF curve and normal distribution curve is 0.14. Fitting error from measured signal amplitude PDF curve and G0 distribution curve is 0.12. Weibull Fitting error is 0.12. From the digital perspective, G0 distribution curve and Weibull distribution curve are optimal description curve. Normal distribution curve $R^2$ is 0.94. G0 distribution curve $R^2$ is 0.97. Weibull distribution curve $R^2$ is 0.96 and is the basically same as G0 distribution $R^2$. Weibull distribution and G0 distribution curve are close to measured signal amplitude PDF curve, at grazing angle 85°.

5. DBN Classification

Figures 11–14 show the alpine meadow ground echo AIB, RIB, CIB, and SIB when the grazing angles are 20°, 45°, 65°, and 85°. It can be seen directly from the figures that the different grazing angle signals characteristic have obvious differences in figures. The AIB,
RIB, CIB, and SIB are taken as signal features and as DBN input information. The figure abscissa is the number of points. The figure ordinate is normalized value. After the features are normalized, they are used as the network input. Normalization is conducive to the network training and prediction of the network, and avoids over-fitting and under-fitting.

Figure 11. Alpine meadow echo AIB.

Figure 12. Alpine meadow echo CIB.
Four grazing angles ground echo are classified with DBN. There are 60 sets of data for each grazing angle and 240 sets of data for four different grazing angles. The DBN classification result is shown in Table 5.

Table 5. The DBN classification result.

| Distribution | 20°  | 45°  | 65°  | 85°  |
|--------------|------|------|------|------|
| AIB          | 78.33% | 98.33% | 51.67% | 85.00% |
| CIB          | 76.67% | 70.00% | 66.67% | 75.00% |
| RIB          | 73.33% | 41.67% | 65.00% | 35.00% |
| SIB          | 85.00% | 83.00% | 88.33% | 35.00% |
When the AIB is used as the input feature, the grazing angle 45° signal recognition rate reaches 98.33%, the grazing angle 65° signal recognition rate is low 51.67%, and the overall recognition rate reaches 78.33%.

When the CIB is used as the input feature, the four grazing angle recognition rates are basically consistent at approximately 70%. When the RIB is used as the input feature, the network has poor overall recognition rate. When the SIB is used as feature input, the 20°, 45°, and 65° recognition rates reach more than 80%, but the 85° signal recognition rate is very low at approximately 35%.

Through the above analysis, AIB as classification feature has the best overall recognition rate. According to the above analysis results, improve AIB. The AIB has no ability to suppress the trivial feature bispectrum points and interference bispectrum points. As a result, network uses the AIB feature as input and cannot complete the classification well. By introducing one-dimensional Fisher, important feature points are selected. Amplify important bispectrum features and suppress trivial bispectrum points.

Figure 15 shows that the selected AIB amplifies the important points in the data and suppresses the trivial points. The DBN input vector of four grazing angles are very different, which is more conducive to the subsequent classification.

Use the selected AIB and DBN to classify the four grazing angles. The confusion matrix results are shown in Table 6 and Figure 16.

|                | 20°   | 45°   | 65°   | 85°   |
|----------------|-------|-------|-------|-------|
| Alpine meadow  | 98.33%| 95.00%| 100%  | 100%  |
| Alpine dry grassland | 96.67%| 93.33%| 95%   | 96.67%|
| Alpine swamp    | 91.67%| 91.67%| 96.67%| 100%  |

Figure 15. Selected AIB.
Use the selected AIB and DBN to classify the four grazing angles. The confusion matrix results are shown in Table 6 and Figure 16.

(a) alpine meadow
(b) alpine dry grassland
(c) alpine swamp

Figure 16. Selected AIB confusion matrix.

The recognition rate of the four grazing angle signals can reach 95% in alpine meadow. The proposed selected AIB greatly increases the recognition rate of signals with different grazing angles. The processing method of network input data in this paper is of targets in typical alpine areas.

In order to verify the validity and versatility of the features, two different terrains, alpine dry grassland and alpine swamp, were selected for the grazing angle recognition. Figure 16 shows the confusion matrix of the two types of terrain angle recognition. Table 6 shows the result of the grazing angle recognition. Alpine dry grassland grazing angle recognition is more than 93%. Alpine swamp grazing angle recognition is more than 91%. It shows that the selected AIB feature is effective in identifying the glancing angle of typical alpine area targets.

6. Conclusions

In order to suppress the ground clutter of the terahertz radar, the statistical distribution characteristics of ground clutter amplitude under different grazing angles are analysed. Based on the bispectrum, the clutter selected AIB is extracted. Under the short data length, the characteristic information of the signal is retained. Clutter selected AIB is used as DBN input to complete grazing angle recognition. The study finds that the PDF curve of ground clutter amplitude is more in line with the normal distribution curve when the grazing angle is $0^\circ$ ~ $45^\circ$. The Weibull distribution and G0 distribution can describe the PDF curve of ground clutter amplitude at grazing angle $85^\circ$ and $65^\circ$. The selected AIB can identify different grazing angle signals, and the recognition rate is greater than 95% in alpine meadow. The recognition rate of grazing angle is more than 93% in alpine dry grassland. The recognition rate of grazing angle is more than 91% in alpine swamp. It shows that the selected AIB feature is effective in identifying the glancing angle.

Author Contributions: Conceptualization, G.Z.; Data curation, P.L.; Formal analysis, G.L.; Funding acquisition, R.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.
Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the database is being built. When the database is completed, it can be provided to researchers.

Conflicts of Interest: We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled.

References

1. Yu, H.; Liu, N.; Zhang, L.; Li, Q.; Zhang, J.; Tang, S.; Zhao, S. An Interference Suppression Method for Multistatic Radar Based on Noise Subspace Projection. IEEE Sens. J. 2020, 20, 8797–8805. [CrossRef]

2. Asada, M.; Suzuki, S. Terahertz Emitter Using Resonant-Tunneling Diode and Applications. Sensors 2021, 21, 1384. [CrossRef]

3. Sabery, S.M.; Bystrov, A.; Navarro-Cia, M. Study of Low Terahertz Radar Signal Backscattering for Area Identification. Sensors 2021, 21, 2954. [CrossRef]

4. Kaplan, S.; McFall, R. The Statistical Properties of Noise Applied to Radar Range Performance. Proc. IRE 1951, 39, 56–60. [CrossRef]

5. Oliver, C.; Quegan, S. Understanding Synthetic Aperture Radar Images; SciTech Publishing: Norwood, MA, USA, 1998; pp. 128–130.

6. Krylov, V.; Zerubi, J. Generalized gamma mixtures for supervised SAR image classification. Proc. Graph. Con Russ. 2010, 6, 107–110.

7. Moser, G.; Zerubia, J.; Serpico, S.B. SAR amplitude probability density function estimation based on a generalized Gaussian model. IEEE Trans. Image Process. 2006, 15, 1429–1442. [CrossRef]

8. Szajnowski, W. Estimator of log-normal distribution parameters. IEEE Trans. Aerosp. Electron. Systems 1977, 13, 533–536. [CrossRef]

9. Oliver, C.J. Optimum texture estimators for SAR clutter. J. Phys. D Appl. Phys. 1993, 26, 1824–1835. [CrossRef]

10. Stacy, E.W. A Generalization of the Gamma Distribution. Ann. Math. Stat. 1962, 33, 1187–1192. [CrossRef]

11. Li, H.-C.; Hong, W.; Wu, Y.-R.; Fan, P.-Z. On the Empirical-Statistical Modeling of SAR Images with Generalized Gamma Distribution. IEEE J. Sel. Top. Signal. Process. 2011, 5, 386–397. [CrossRef]

12. Nicolas, J. Introduction to second kind statistics: Application of logmoments and log-cumulants to second kind statistics: Application of logmoments and log-cumulants to analysis of radar images. Traitement Du Signal. 2002, 19, 139–167.

13. Wang, Y.; Wang, W.; Zhou, M.; Ren, A.; Tian, Z. Remote Monitoring of Human Vital Signs Based on 77-GHz mm-Wave FMCW Radar. Sensors 2020, 20, 2999. [CrossRef]

14. Feng, J.; Cao, Z.; Pi, Y. Multiphase SAR Image Segmentation with Statistical-Model-Based Active Contours. IEEE Trans. Geosci. Remote Sens. 2013, 51, 4190–4199. [CrossRef]

15. Chandran, V.; Elgar, S.L. Pattern recognition using invariants defined from higher order spectra—One-dimensional inputs. IEEE Trans. Signal. Process. 1993, 41, 205–212. [CrossRef]

16. Tugnait, J. Detection of non-Gaussian signals using integrated polyspectrum. IEEE Trans. Signal. Process. 1994, 42, 3137–3149. [CrossRef]

17. Liao, X.; Bao, Z. Circularly integrated bispectra: Novel shift invariant features for high-resolution radar target recognition. Electron. Lett. 1998, 34, 1879–1880. [CrossRef]

18. Junguo, M. Research on Feature Extraction and Recognition Method of Space Radar Target. Ph.D. Thesis, Beijing, China, 2006.

19. Heo, J.; Jung, M.; Lee, M. FPGA Implementation of an Efficient FFT Processor for FMCW Radar Signal Processing. Sensors 2021, 21, 6443. [CrossRef]

20. Yan, J.; Liu, H.; Jiu, B.; Bao, Z. Power Allocation Algorithm for Target Tracking in Unmodulated Continuous Wave Radar Network. IEEE Sens. J. 2014, 15, 1098–1108. [CrossRef]

21. Ulaby, F.T.; Deventer, T.V.; East, J.R. Millimeter-wave bistatic scattering from ground and vegetation targets. IEEE Trans. Geosci. Remote Sens. 2002, 26, 229–243. [CrossRef]

22. Fung, A.K.; Li, Z.; Chen, K.S. Backscattering from a randomly rough dielectric surface. IEEE Trans. Geosci. Remote Sens. Lett. 1992, 30, 356–369. [CrossRef]

23. Rappaport, T.S.; Deng, S. 73 GHz wideband millimeter-wave foliage and ground reflection measurements and models. In Proceedings of the 2015 IEEE International Conference on Communications (ICC), ICC Workshops, London, UK, 8–12 June 2015.

24. Kim, S.B.; Moghaddam, M.; Tsang, L. Models of L-Band Radar Backscattering Coefficients Over Global Terrain for Soil Moisture Retrieval. IEEE Trans. Geosci. Remote Sens. 2013, 52, 1381–1396. [CrossRef]
25. Chunliang, X.; Yan, C.; Ling, T; Mingquan, J.; Zhengchan, L.; Haiping, L. Measuring the Microwave Backscattering Coefficient of Paddy Rice Using FM-CW Ground-Based Scatterometer. In Proceedings of the 2008 International Workshop on Education Technology and Training & 2008 International Workshop on Geoscience and Remote Sensing, Shanghai, China, 21–22 December 2008; pp. 194–198.

26. Shi, J.; Yin, D. Computational high frequency waves through curved interfaces via the Liouville equation and geometric theory of diffraction. *J. Comput. Physics*. **2008**, *227*, 6106–6139.