# Data Pre-processing method of Ground Penetrating Radar based on HE-R2M

Jiahe Chen [1, *], Yi Li [2]

[1] School of Electronic Information and Automation, Tianjin University of Science and Technology Tianjin, China, 300222

[2] School of Computer and Communication Engineering, University of Science and Technology Beijing Beijing, China, 100083

* Corresponding Author Email: 18622986475@163.com

**Abstract.** Ground Penetrating Radar (GPR), as a non-destructive detection technology, has been widely used in infrastructure construction, tunneling, agriculture and forestry. It has the advantages of fast detection speed, high accuracy and non-destructive detection. However, GPR is often affected by the interference of subsurface media when facing complex geological environments, resulting in low contrast, high noise and clutter interference in the acquired B-scan images, which seriously affects the detection of subsurface structures and imaging performance. In order to improve the quality of images acquired by GPR in mixed soil media, this paper proposes HE-R2M (Histogram Equalization with Remapping and Morphology) image enhancement algorithm for preprocessing B-Scan images with respect to B-scan hyperbolic features of underground targets. The algorithm is based on the histogram equalization algorithm to initially enhance the image, and for the B-Scan image features, combined with the histogram interval remapping method and morphological processing, it can successfully suppress the clutter features in the B-Scan image while retaining the effective information of the hyperbola. The effectiveness and robustness of the algorithm are verified through simulation and real experiments. The proposed HE-R2M image preprocessing algorithm provides a potential idea for the application expansion of GPR, and offers an effective solution to improve the performance of underground structure detection and imaging.

**Keywords:** GPR, Image enhancement, Two-dimensional image processing.

## 1. Introduction

In recent decades, Ground Penetrating Radar (GPR) has been rapidly developed as an effective means of underground target detection [1]. It has the advantages of fast detection speed, high accuracy, low detection cost, wide detection range, non-destructive testing, and has been widely used in many fields, such as urban infrastructure construction, traffic bridge and tunnel engineering, mining survey, and archaeological excavation. [2]. With the acceleration of urbanization and the rapid growth of urban transport demand, the number of metal objects such as pipes and steel structures buried underground is also increasing [3]. Therefore, in order to ensure the steady progress of urban development, the accurate detection of underground metal targets has become one of the crucial construction links.

B-scan image is a common type of imagery used in GPR to show cross-sectional images of underground targets. It is a two-dimensional image that typically shows the profile of a radar signal between a subsurface target and the surface of the ground. B-scan images can help researchers and engineers to identify objects such as underground pipelines, reinforced structures, and underground caves, and to assess their locations, dimensions, and depths. However, the quality and signal-to-noise ratio of the echo data acquired by GPR are often poor due to the complexity of the structure of the subsurface medium and the interference of external random noise, which largely limits the detection resolution and the final demodulation of B-scan images.

To improve the quality of B-scan images, many pre-processing methods have been proposed and validated. The most traditional B-scan image pre-processing method is the detection of target hyperbola based on Hough transform technique [4]. The accuracy of target recognition in B-scan

images is improved by accurate localization of the hyperbolic region. Least squares method, as another traditional image pre-processing algorithm [5], has also been proposed and improved for the search and differentiation of 2-fold curves in B-scan images. However, the least squares related algorithm is not applicable to multi-curve detection in the absence of the help of image segmentation algorithms because complete noise removal is difficult to achieve in real-world environments. In recent years, with the rapid development of machine learning methods, related GPR image preprocessing algorithms have also been proposed. the VJ algorithm was firstly used to extract the target region in GPR data [6]. Secondly, SVM algorithm was also applied to identify linear and hyperbolic features of subsurface objects [7-8]. The advantage of machine learning is that when data from the same detection region is used to train the framework, the trained model can be reused. For machine learning, most of the input features need to be recognized by humans and the classification results depend on the quality of the features. Deep learning artificial neural networks can skip the feature recognition step and learn the feature representation of buried objects directly from the image. Convolutional neural networks, as the simplest and most effective method of artificial neural networks, are applied to extract meaningful features from GPR and accurately classify complex features contained in GPR images [9-10]. Although all of the above methods can provide higher recognition accuracy for the recognition of hyperbolas in GPR images, there is no method to improve the quality and recognizability of hyperbolas in GPR images from the perspective of image enhancement.

In this study, Histogram Equalization with Remapping and Morphology (HE-R2M) B-scan image pre-processing algorithm is proposed and experimentally verified, which is based on the histogram equalization algorithm, and integrates histogram interval remapping for underground target B-scan hyperbola features and morphological operations to suppress the B-Scan image clutter without damaging the hyperbolic information. In addition, the contrast between the hyperbolic information and the background is enhanced, which facilitates the subsequent data processing and the inversion of the detected targets. The proposed HE-R2M image pre-processing algorithm provides a new idea for the expansion of the application of GPR technology and an effective solution to improve the performance of underground structure detection and imaging.

## 2. Fundamental principle

### 2.1. Histogram equalization algorithm

Histogram equalization is essentially a balanced distribution of the number of grey levels in the image, i.e. the number of grey levels is not concentrated in part of the grey levels above, when the image histogram is completely uniform distribution, the entropy of the image at this time is the maximum (i.e. the probability distribution of each grey level is balanced), and the contrast of the image is the maximum, which in turn enhances the observability of the image. Histogram equalization enhances the contrast of an image and involves transforming the intensity values of the image to distribute them over a larger range of values, resulting in a better visualization of the image. The process includes calculating a histogram of the image representing the distribution of pixel values, and then applying a transformation to the image to expand the histogram more uniformly. Specifically, the steps of histogram equalization are as follows:

1) Count the histogram of the original image. The frequency of each pixel value is obtained by counting the number of occurrences of each pixel value and then normalizing it.

2) Calculate the Cumulative Distribution Function (CDF). The CDF is an integral over the Frequency Distribution Function (PDF), which represents the probability of occurrence of each pixel value in the original image. The CDF can be obtained by cumulatively calculating the PDF. For grey scale values, the CDF is calculated as follows:

$$CDF(i) = \sum_{j=0}^{i} P(j) \quad (1)$$

Where, denotes the frequency of occurrence of pixels in the image with grey value of.

3). Calculate the equalized pixel values. Each pixel value in the original image is mapped to a new pixel value such that the equalized histogram is approximated as a uniformly distributed histogram. This mapping function can be calculated by the following equation:

$$H(i) = round(\frac{CDF(i)-CDF(min)}{(M\times N-1)}) \times (L-1) \quad (2)$$

Where, denotes the mapped pixel value, and denotes the width and height of the image, respectively, denotes the range of pixel values, and denotes the minimum pixel value in the original image.

4) The pixel values in the original image are replaced with the mapped pixel values.

Histogram equalization essentially reduces the grey level of an image in exchange for an increase in contrast. During the equalization process, the less frequent grey levels in the original histogram are classified into very few grey levels or in one grey level and hence cannot be improved. If the details of the image formed from these grey levels are important, local histogram equalization is required.

### 2.2. Histogram interval remapping algorithm

Histogram interval remapping is an effective technique for enhancing the contrast of an image by modifying the histogram to adjust the contrast of the image, especially if the original image is characterized by low contrast, the grey scale range of the image can be adjusted to obtain a portion of the image that has a high and effective information content.

The histogram of the image is a graphical representation of the frequency distribution of grey levels in the image. Remapping the histogram intervals involves dividing the grey levels into a number of intervals and then adjusting the number of pixels in each interval to achieve the desired contrast. The basic idea is to map the intensity values in the original image to a new range of intensities. This is achieved by remapping the cumulative distribution function from the original image to the desired cumulative distribution function. The result is an image with better contrast and better detail visualization.

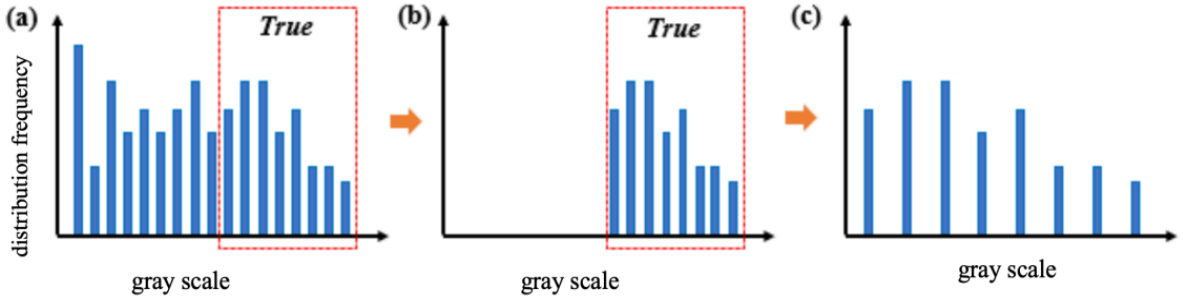



**Fig. 1** Schematic diagram of the principle of histogram interval remapping.

The principle of histogram remapping is shown in Figure 1. Firstly, the histogram of the image after preliminary enhancement is obtained using the histogram equalization algorithm. Then, according to the actual needs of different application scenarios, the appropriate True interval, i.e., the effective grey level information histogram interval, is selected. If the grey value of the original image is set to and the grey value of the enhanced image is set to, the True interval is denoted by assuming that the minimum value of the image is and the maximum value is:

$$(X_{\min}, X_{\max}) \in True \quad (3)$$

Then based on the selected True interval, the entire histogram interval is remapped.

According to:

$$\frac{X-X_{\min}}{X_{\max}-X} = \frac{Y-0}{255-Y} \quad (4)$$

The simplification can be obtained:

$$Y = \frac{255(X - X_{\min})}{X_{\max} - X_{\min}} \tag{5}$$

After the adjustment of this remapping formula, the invalid information such as background noise is all set to 0, and the information higher than the maximum value is all set to 255, and then the histogram range of the True interval is remapped to the whole grey scale range of 0-255, thus enhancing the contrast of the effective information of the image while suppressing the background noise.

### 2.3. Morphological Processing Algorithm

Mathematical Morphology is a set-based operation where the set of shapes includes linear, square and circular structures etc. and noise is also one of the elements of the set. Morphological operations are simply the division of an image into a number of basic shapes, and these basic shapes are then analyzed one by one for image processing. The basic operations are erosion, dilation etc. The basic idea of the erosion operation is to reduce the image inward, which can reduce or even eliminate the details in the image, and can be used to smooth the image and eliminate noise; expansion is the inverse operation of erosion, and the basic idea is to expand outward along the edge of the image, which can be used to eliminate the voids in the image. The combination of the two operations is called morphological open and morphological closed operations. Morphological open operation is expressed as "erosion, then expansion", which can eliminate grey and dark spots in the image that are smaller than the defined structural elements, while morphological closed operation is the inverse of the open operation, which is expressed as "expansion, then erosion", which can eliminate the grey and dark spots in the image that are smaller than the defined structural elements. The morphological closure operation is the inverse of the open operation, which is expressed as "expand and then erode", eliminating the bright spots of structural elements in the image that are smaller than the definition.

## 3. HE-R2M algorithm

Aiming at the characteristics of B-Scan images with more noise, more influence of clutter and complex background, and the HE-R2M image preprocessing method with both image enhancement and noise filtering is designed by the tandem combination of histogram equalization, histogram grey scale interval remapping and morphological processing algorithms.

The flow of the proposed HE-R2M-based GPR image preprocessing algorithm is shown in Fig. 2, where the original B-Scan image acquired by GPR is first filtered by Gaussian filtering to remove the Gaussian white noise caused by the environment or interference during signal acquisition and transmission. After acquiring the initial representation image, the B-Scan image is initially enhanced using the HE algorithm to perform full grey scale histogram equalization. In this process, clutter and noise signals are also enhanced indiscriminately. Then, based on the histogram distribution information, a 10% threshold is selected for the valid information and the True interval containing the valid hyperbolic information is adaptively determined. The HE-RM algorithm based on histogram interval remapping grey scale correction filters out irrelevant information such as disorder and noise without affecting the valid information of the hyperbola. Finally, the morphological open operation based on 6*6 square structural elements is combined into the HE-RM algorithm to filter the deep noise from the enhanced image. So far, after HE-R2M processing, the B-scan image acquired by GPR achieves the enhancement of hyperbolic information under the premise of noise reduction, and obtains clear hyperbolic feature maps, which is convenient for subsequent process.

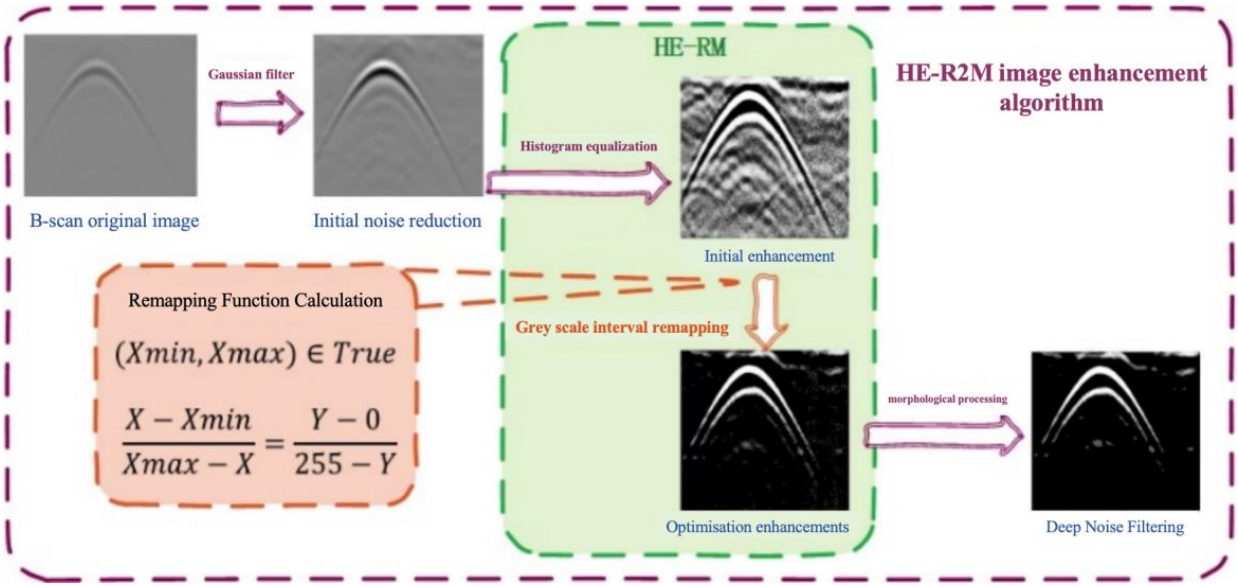



**Fig. 2** Flowchart of HE-RM image enhancement algorithm.

In particular, the True interval range is crucial for the enhancement effect. Usually, the grey level of the hyperbola carrying effective target information tends to be distributed in the last 10% of the effective target information range, and the remaining grey level information is mostly irrelevant environment information and noise information. In the case of unknown materials, after obtaining the grey level histogram distribution, the statistical range is calculated, and the last 10% is taken as the effective information for True interval selection. On a variety of commonly used materials True intervals were judged on the measured data, in the experimental process, after many verifications, to determine the histogram True intervals of some of the materials as shown in Table 1.

**Table 1.** Selection of target True intervals for different materials.

| Material Name | $X_{min}$ | $X_{max}$ |
|---|---|---|
| Metal | 0.9 | 1.0 |
| Plastic | 0.7 | 1.0 |
| PVC | 0.72 | 1.0 |
| Wood | 0.75 | 1.0 |
| Water | 0.8 | 1.0 |
| Air | 0.8 | 1.0 |

## 4. Quality Evaluation Metrics

Since the algorithm sets most of the pixels in the original image to 0 after the histogram interval remapping, i.e., the HE-R2M-enhanced image cannot be used as a reference to calculate the mean signal to noise ratio (PSNR), structural similarity (SSIM) and other metrics commonly used in image quality evaluation. For the effective information of B-Scan image detected by GPR, i.e., hyperbola, the grey values are concentrated in the last 30% of the grey scale, so the Threshold signal to noise ratio (TSNR) is defined as a metric here. TSNR is defined as a suitable threshold value, the grey scale of pixels is compared with the threshold value, pixels larger than the threshold value are considered as signals, pixels smaller than the threshold value are considered as signals, and pixels larger than the threshold value are considered as signals, and pixels smaller than the threshold value are considered as signals. Pixels larger than this threshold are considered as signal and pixels smaller than this threshold are considered as noise, the mean and standard deviation of the signal and noise are calculated separately and the ratio of the two is calculated as TSNR with the following formula:

$$TSNR = \frac{mean(signal)}{std(noise)} \tag{6}$$

# 5. Analysis of experimental results

## 5.1. Simulation experiment

In this paper, the gprMax 3.0 tool was used to perform simulation experiments using the gprMax toolbox to generate the ground-penetrating radar B-scan image dataset. During the data acquisition process, as there are several different materials in the model, representing air, mixed soil and objects to be measured. In order to correspond to the subsequent real-world experiments, we chose the objects to be tested to be metal, PVC water pipe, plastic, air, bottled water and wood. This is also consistent with the objects contained in the shallow underground space in the real application scenario. The objects are already present as built-in materials in gprMax 3.0 and can be accessed using identifiers. However, it is necessary to define the mixed soil by defining its parameters as $\mu r = 1$, $\sigma^* = 0$ and a relative dielectric constant of 4, i.e. $\varepsilon r = 4$ and conductivity $\sigma = 0$.

The following shows the resultant plots of simulated image enhancement for six different materials. It is more difficult to process compared to that of the homogeneous medium due to the enhancement of background interference, more hyperbolic information being submerged and more blurred information from the top position. However, multi-target images are more difficult to process than single targets due to crosstalk and hyperbolic crossings between targets. Therefore, in this paper, simulated images of different materials in non-uniform media are given and the obtained multi-target results are presented in Fig. 3to Fig. 8.

From the experimental results of multi-target simulation image enhancement, it can be seen that the B-scan image enhanced solely by the histogram equalization algorithm not only improves the effective information of the hyperbola, but also amplifies the background and clutter signals, resulting in the hidden image in the background is not significantly enhanced. Among them, as shown in Figs. 4, 5, and 7, due to the crosstalk between multiple targets, the clutter and the effective hyperbola are enhanced at the same time, leading to the mixing of the background-clutter-valid information, which is not conducive to the subsequent extraction of network features. However, using the HE-R2M algorithm proposed in this paper to improve the B-scan image while improving the effective information of the hyperbola, the clutter signals are also significantly suppressed, and only some of the overly strong reflection information in the original image is slightly retained. In addition, the common high harmonics in multi-target B-scan images are also significantly suppressed. The experimental results show that the hyperbolic effective information of multiple materials is significantly enhanced under the condition of selecting appropriate True intervals, which is conducive to improving the output quality of the subsequent imaging classification network.




**Fig 3.** Air simulation experiment results: (a) original image; (b) HE-enhanced image; (c) HE-R2M-enhanced image.




**Fig 4.** Metal simulation experiment results: (a) original image; (b) HE-enhanced image; (c) HE-R2M-enhanced image.




**Fig 5.** Plastic simulation experiment results: (a) original image; (b) HE-enhanced image; (c) HE-R2M-enhanced image.




**Fig 6.** PVC simulation experiment results: (a) original image; (b) HE enhanced image; (c) HE-R2M enhanced image.




**Fig. 7** Water simulation experiment results: (a) original image; (b) HE-enhanced image; (c) HE-R2M-enhanced image




**Fig. 8** Wood simulation experiment results: (a) original image; (b) HE-enhanced image; (c) HE-R2M-enhanced image.

The TSNR of the original images and the images enhanced by HE-R2M for the above six sets of experiments are calculated respectively, and the results are displayed in the following table2. From the TSNR results, it can be seen that the TSNRs of all the material targets are greatly enhanced, among which the TSNR enhancement is the least obvious due to the similarity of the dielectric constants of plastic and PVC with the medium.

**Table 2.** Simulated image performance cure.

| | Air | Metal | Plastic | PVC | Water | Wood |
|---|---|---|---|---|---|---|
| original image | 9.985 | 9.8952 | 10.0272 | 9.4239 | 8.9335 | 9.5745 |
| HE-R2M | 13.7329 | 13.554 | 11.2596 | 11.1126 | 12.7043 | 12.5747 |

### 5.2. Measurement Experiments

In this paper, an SIR 4000 GPR device manufactured by GSSI was used to conduct the field experiments, the antenna was a model 62000 handheld 2GHz antenna, single wheel, count 22146.

The materials to be tested were 8mm diameter steel bars, hollow cylindrical PVC water pipes, plastic, empty cylindrical plastic bottles (containing air), cylindrical plastic bottles containing water and 4cm diameter cylindrical dry wood. Transparent acrylic experimental chamber containing natural environment to obtain soil was used as an experimental environment for non-uniform media. The measuring range of the equipment was set to 800-900 mm, and an antenna with a frequency of 2 GHz was used to detect the different buried materials separately and obtain the reflected echo B-scan hyperbolic images.

Compared with the simulated data, the measured data is obtained by physical detection in the real environment, the background information of the B-scan image is more complex, and the clutter interference and noise interference are obviously strengthened, in order to better prove the effectiveness of the algorithm, the single-target result plots of the measured data of different material targets in non-uniform scenarios are shown here, as shown in Figs. 9-14. From the experimental results, it can be seen that the hyperbolic information of the B-scan image, which has been simply processed by de-rectified wave, is not obvious and is drowned in the ambient background noise. However, after histogram equalization, although the hyperbolic information is significantly improved and can be clearly observed in the background information, due to the fact that the histogram equalization algorithm is an undistinguished processing of the pixels of the whole image, which makes the background information, clutter interference, and noise interference, in addition to the effective information of the hyperbolic curve, enhanced to a certain extent.

And for the next step of imaging classification processing, it may cause interference and affect the recovery results. After the processing of HE-R2M image enhancement algorithm proposed in this study, the background noise, clutter interference and noise interference are effectively suppressed, and the hyperbolic target is clearly visible. From the combination of simulation and real experiment results, the proposed HE-R2M algorithm can effectively improve the GPR B-scan image.

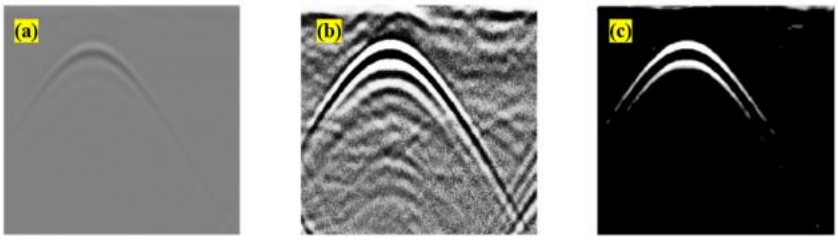



**Fig 9.** Experiment results of Air: (a) original image;
(b) HE-enhanced image; (c) HE-R2M enhanced image.

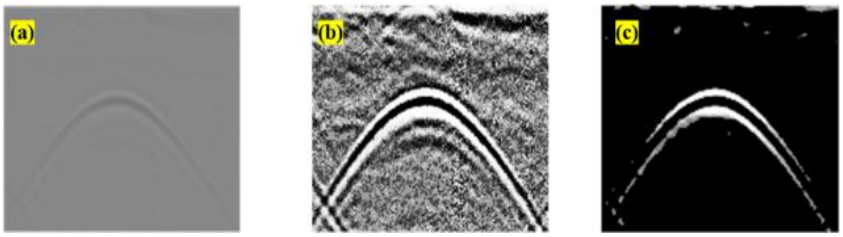



**Fig 10.** Experiment results of Metals: (a) original image;
(b) HE-enhanced image; (c) HE-R2M-enhanced image.

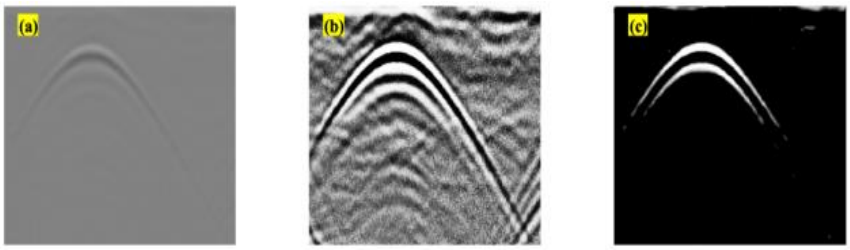



**Fig 11.** Experiment results of Plastic: (a) original image;
(b) HE-enhanced image; (c) HE-R2M-enhanced image.

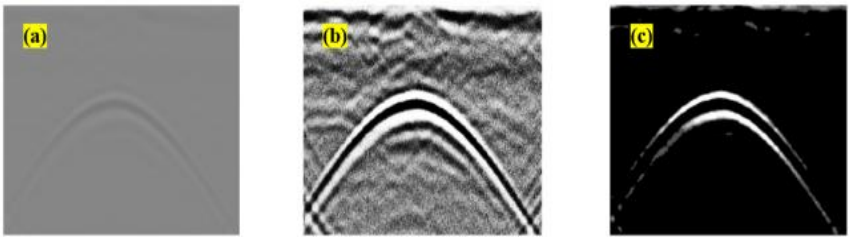



**Fig 12.** Experimental results of PVC: (a) original image; (b) HE-enhanced image; (c) HE-R2M-enhanced image.

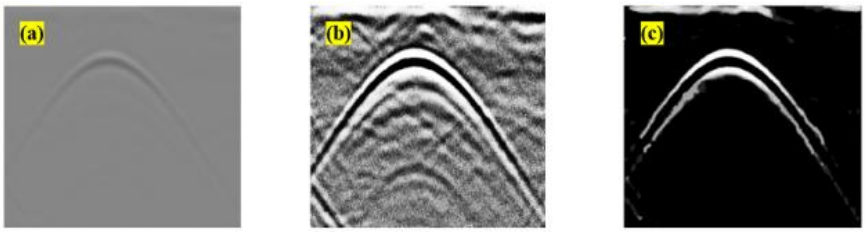



**Fig 13.** Experiment results of Water: (a) original image; (b) HE enhanced image; (c) HE-R2M enhanced image.

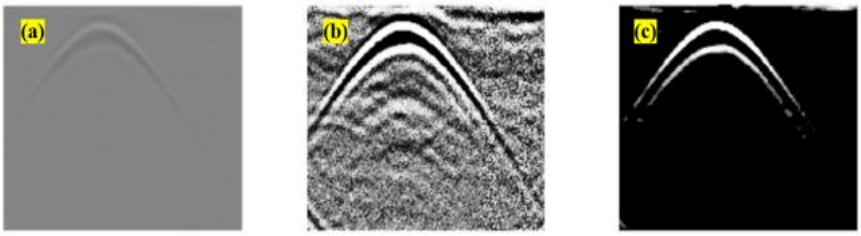



**Fig 14.** Experiment results of Wood: (a) original image; (b) HE-enhanced image; (c) HE-R2M-enhanced image.

The TSNR of the above 6 groups of experiments are calculated separately, and the results are displayed in the table3, which proves that all materials have similar TSNR after HE-R2M enhancement. as shown in Fig. 8~Fig. 13, the visual display of the images also proves that various materials get the same enhancement effect after being processed by HE-R2M algorithm.

**Table 3.** Measured image performance metrics.

| | Air | Metal | Plastic | PVC | Water | Wood |
|---|---|---|---|---|---|---|
| HE-R2M | 13.3048 | 13.8622 | 13.2696 | 13.7369 | 13.5736 | 12.5466 |

## 6. Conclusion

In this paper, for the underground target B-scan hyperbolic features, histogram interval remapping and morphological operations are fused to suppress the B-Scan image clutter without damaging the hyperbolic information. The proposed HE-R2M image pre-processing algorithm flow includes acquiring GPR B-Scan, Gaussian filtering preliminary noise removal, histogram equalization image preliminary enhancement, grey scale interval remapping, and morphological noise filtering. The effectiveness of the proposed algorithm is verified through simulation experiments as well as real experiments, and the proposed algorithm is analyzed by TSNR metrics, which verifies that the proposed algorithm has good performance for different materials with good robustness. The proposed HE-R2M image pre-processing algorithm provides a new idea for the expansion of the application of GPR technology and an effective solution for improving the detection and performance of underground structures.

## References

[1] Hu Rongming, Li Xin, Jingxia, et al. Application of YOLOv7 in ground-penetrating radar B-Scan image interpretation [J]. Surveying and Mapping Bulletin, 2023, (08): 29-33.

[2] REN Wang, YAO Zhen-an, CHEN Long-feng,et al. Analysing the development history of ground-penetrating radar technology [J]. Jiangxi Science, 2024, 42 (01): 100-107.

[3] Lijun Xue. Research on target detection and localisation based on hyperbolic features of ground-penetrating radar B-scan images [D]. Harbin Institute of Technology, 2021.

[4] W. Jiang, Z. Liu, Y.P. Wang, et al. Hough transform-based false point rejection method for passive radar cross-location [J]. Signal Processing, 2023, 39 (11): 1978-1986.

[5] BOOKSTEIN FL. Fitting conie sections to scattereddata [J]. Computer Graphics and Image Processing, 19799(1):56-71.doi:10.1016/0146-664x(79)90082-0

[6] MAAS C and SCHMALZL J.Using pattern recognition toautomatically localize reflection hyperbolas in data fromground penetrating radar [J].Computers &: Geosciences, 2013, 58:116-125.

[7] Li Mangmeng. Research on the metallogenic law of gold ores in Neo-Pacific gneisses in western Liaoning based on SVM [D]. Liaoning University of Engineering and Technology, 2021. DOI: 10.27210/d.cnki.glnju.2021.000673.

[8] SHI Lock, YU Jifeng, CAO Huitao, et al. Reservoir lithology identification based on Gaussian kernel SVM - A case study of Upper Paleozoic clastic rocks in Dongpu Depression [J]. Chinese Science and Technology Paper, 2020, 15 (01): 112-118+136.

[9] Zhang X, Yang WX, Li SB, et al. Analysis of the effect of the use of batch normalisation layer in convolutional neural network on denoising seismic data [J]. Advances in Geophysics, 2024, 39 (01): 183-196.

[10] Cao Maojun, Cui Xinfeng. Intelligent stratigraphic identification method based on one-dimensional convolutional neural network [J]. Computer Technology and Development, 2023, 33 (09): 133-140+148.