An Efficient And Adaptive Compression Algorithm Based On DCT Transform For GPR Raw Data

Peng Li$^{1,2,3}$, Xiaojun Liu$^{1,3}$, Guangyou Fang$^{1,2,3}$

1 Key Laboratory of Electromagnetic Radiation and Detection Technology, Chinese Academy of Science, Beijing 100190, China.
2 Institute of Electronics, Chinese Academy of Science, Beijing 100190, China.
3 University of Chinese Academy of Sciences School of Electronic, Electrical and Communication Engineering, Beijing 100049, China

*Corresponding author: Peng Li (lp0741@163.com)

Abstract. Ground Penetrating Radar (GPR) creates large amounts of data which causes serious problems for storage and transmission. In this paper, a novel compression algorithm for GPR raw data based on DCT transform and inter-frame difference is proposed. The algorithm first conducts DCT transform on the GPR raw data, and differentiates the adjacent frame data in DCT domain to reduce the variance and dynamic range. Then the differential data is evenly divided into several blocks, and different bit Lloyd-Max quantizers are selected to quantify different sub-block data through an adaptive bit allocation mechanism, thereby obtaining a higher compression gain. The compression experiments with satellite-borne GPR, aircraft-borne GPR and vehicle-borne GPR raw data confirmed the effectiveness of the compression algorithm.

1. Introduction

Ground Penetrating Radar (GPR) is an important non-destructive detection technique that uses the penetration of electromagnetic waves in low-frequency bands to image shallow and deep-buried targets$^{[1]}$. GPR is widely used in road maintenance, geological engineering, and polar ice thickness detection, etc.. To achieve high-resolution imaging$^{[2]}$, a large bandwidth is required and therefore GPR generates massive raw data, which causes serious problems for data storage and transmission. Especially in deep space exploration such as Mars and Venus, the limited downlink capability of the data channel cannot satisfy the transmission rate of the GPR data. Therefore, the raw data of GPR must be compressed efficiently before transmission.

Usually, the radar echo received by the GPR is sampled discretely and the sampled data of each echo is stored as 1-D frame. The continuity of geological stratification and similarity of composition of adjacent underground areas make the adjacent frame data similar. DCT (discrete cosine transform) can remove correlation and concentrate the signal energy, which further increases the similarity of adjacent frame data in DCT domain$^{[3]}$. By differentiating the adjacent frame data in DCT domain, the variance and dynamic range will be greatly reduced for lower compression distortion.

This paper proposes a novel DCT-FDAQ (inter-frame difference adaptive quantization) algorithm for GPR raw data based on DCT transform and inter-frame difference. And a kind of adaptive bit allocation mechanism required for higher compression gain is also adopted to select Lloyd-Max quantizers with different quantization bits.
2. DCT-FDAQ Algorithm

Figure 1 shows the block scheme of the proposed DCT-FDAQ algorithm.

2.1. DCT transform

DCT transform can remove the correlation and is widely used in various international standards of image and voice compression coding. The proposed DCT-FDAQ algorithm conducts DCT transform of the GPR raw data, which enhances the similarity of the adjacent frame data in DCT domain, so as to effectively reduce the variance and dynamic range by means of inter-frame difference. The DCT and IDCT transform processes are as follows

\[ X(k) = \frac{2}{N} \sum_{n=0}^{N-1} x(n) \cos \left( \frac{2n+1}{N} \pi k \right) \quad (k = 0, 1, 2, \cdots, N-1) \] (1)

\[ x(n) = \frac{2}{N} \sum_{k=0}^{N-1} c(k)X(k) \cos \left( \frac{2n+1}{N} \pi k \right) \quad (n = 0, 1, 2, \cdots, N-1) \] (2)

Where \( c(k) \) denotes the weight coefficient by

\[ c(k) = \begin{cases} 
\frac{1}{\sqrt{2}}, & k = 0 \\
1, & k = 1, 2, \cdots, N-1
\end{cases} \] (3)

2.2. Inter-frame difference

The previous frame data of GPR in DCT domain is subtracted from the adjacent following frame data by

\[ D_i = F_i - F_{i-1} \] (4)

However, it should be noted that in the actually designed DCT-FDAQ algorithm, the inter-frame difference is by

\[ D_i = F_i - \hat{F}_{i-1} \] (5)
Where $D_i$ denotes the differential frame data, $\hat{F}_{i,1}$ denotes the reconstructed frame data (will be introduced in section 2.3.4), and replacing $F_{i,1}$ with $\hat{F}_{i,1}$ in equation (5) is just for avoiding the accumulation of quantization noise.

2.3. Adaptive quantization

2.3.1. Block

In general, the differential frame data $D_i$ contains a large amount of data, and the variance and dynamic range of each data interval vary greatly. In order to save calculating and storage resources, the DCT-FDAQ algorithm uniformly divides the differential frame data into several blocks for quantization in units of sub-block data.

2.3.2. Lloyd-Max quantizer.

The optimal non-uniform quantizer is the Lloyd-Max quantizer\textsuperscript{[4]} under the condition of knowing the signal PDF (probability density function). Figure 2 shows the statistical properties of real-world GPR raw data, the raw data approximately obeys the Gaussian distribution. The DCT is a linear transform and the inter-frame difference can be considered as a linear combination of the raw data, for which the PDF will not be changed.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure2.png}
\caption{Statistical properties of real-world GPR raw data}
\end{figure}

Therefore, the Gaussian Lloyd-Max quantizer is used to quantify the differential frame data in the DCT-FDAQ algorithm, and the quantization process is shown as follows

$$D_i' = LM(D_i)$$  \hspace{1cm} (6)

Where the $f(x) = LM(x)$ denotes the function of Gaussian Lloyd-Max quantizer.

2.3.3. Adaptive bit allocation mechanism.

The adaptive bit allocation mechanism\textsuperscript{[5]} is implemented using the rate distortion function of continuously distributed memoryless Gaussian source. The Gaussian rate distortion function is as follows

$$R(d) = \frac{1}{2} \log \frac{\sigma^2}{d} (0 \leq d \leq \sigma^2)$$  \hspace{1cm} (7)

Where $R(d)$ denotes the bit rate (number of allocated bit) of Lloyd-Max quantizer, $d$ denotes the quantization distortion, and $\sigma^2$ denotes the variance, equation (7) can be expressed as

$$d(R) = \sigma^2 2^{-2R}$$  \hspace{1cm} (8)
Supposing that the differential frame data $D_i$ is divided into $M$ blocks: $B_1, B_2, \cdots, B_M$, $\sigma_i^2$ denotes the variance of $B_i$, $R_i$ denotes the bit number allocated for $B_i$, and $d_i$ denotes the quantization distortion. The average quantization distortion of $D_i$ is as follows

$$
\bar{d} = \frac{1}{M} \sum_{i=1}^{M} d_i = \frac{1}{M} \sum_{i=1}^{M} \sigma_i^2 2^{-2R_i} \quad (i = 1, 2, \cdots, M)
$$

Therefore, the adaptive bit allocation mechanism can be described as: under the condition that the average allocated bit $\bar{R} = \frac{1}{M} \sum_{i=1}^{M} R_i$ is fixed, the average quantization distortion $\bar{d}$ defined by equation (9) is the smallest. The method of Lagrangian multiplier is used to solve this problem, we represent $R_i$ in vector form $R = (R_1, R_2, \cdots, R_M)$, and the objective function is established as follows

$$
R = \arg \min_R L(R, \lambda) = \arg \min_R \left[ \frac{1}{M} \sum_{i=1}^{M} \sigma_i^2 2^{-2R_i} + \lambda \left( \bar{R} - \frac{1}{M} \sum_{i=1}^{M} R_i \right) \right]
$$

The solution of equation (10) is shown as follows

$$
R_i = \bar{R} + \frac{1}{2} \left[ \log_2 \sigma_i^2 - \frac{1}{M} \sum_{i=1}^{M} \log_2 \sigma_i^2 \right]
$$

$R_i$ is calculated as an integer (keep $\bar{R}$ fixed) to be the bit number allocated for $B_i$.

2.3.4. Reconstructor.
The data reconstructor is required both for data compression and decompression, which is used to calculate the raw data cyclically utilizing the quantized difference data $D_i^*$. The reconstructed data $\hat{F}_i$ is shown as follows

$$
\hat{F}_i = D_i^* + \hat{F}_{i-1}
$$

3. Experimental dataset
Before conducting the compression experiment of the proposed DCT-FDAQ algorithm, we first introduce the experimental datasets. The experimental datasets consist of raw data collected by three different GPR devices: the satellite-borne GPR, the aircraft-borne GPR and the vehicle-borne GPR.

The satellite-borne GPR is the SHARAD system\[6\] carried on the MRO(Mars Reconnaissance Orbiter) launched by NASA in 2006. The radar transmits signals have a center frequency of about 20MHz and a bandwidth of about 10MHz. The pulse repetition frequency is 700Hz, the duration of each pulse 85us, and the sampling rate about 26.67MHz. The raw data is stored on 8bpp(int 8bits).

The aircraft-borne GPR is the MCRDS-2 system\[7\] developed by the Kansas university, and the experimental dataset is collected in Antarctic Dome-A area. The radar transmits signals have a center frequency of about 195MHz and a bandwidth of about 30MHz. The pulse repetition frequency is 12KHz, the duration of each pulse 10us, and the sampling rate about 111.11MHz. The raw data is stored on 14bpp(int 14bits).

The vehicle-borne GPR is the Ice Detection Radar system\[8\] developed by the IECAS, and the experimental dataset is collected near the Kunlun station of Antarctica. The radar transmits signals have a center frequency of about 150MHz and a bandwidth of about 100MHz. The pulse repetition frequency is 8KHz, the duration of each pulse 8us, and the sampling rate about 500MHz. The raw data is stored on 16bpp(int 16bits).
4. Experimental result

In this section, the performance of the proposed DCT-FDBQ algorithm is evaluated with the different datasets mentioned above. The compression performance is measured in terms of the SQNR which is described by

\[
SQNR = \frac{\sum_{i=1}^{L} x_i^2}{\sum_{i=1}^{L} (x_i - \hat{x}_i)^2}
\]  

(13)

Where \( L \) denotes the number of GPR raw data. \( x_i \) denotes one data element of GPR raw data, and \( \hat{x}_i \) denotes one data element of reconstructed GPR data.

| Date       | Algorithm | Average quantization bit rate (bpp) |
|------------|-----------|-------------------------------------|
|            |           | 2bpp | 3bpp | 4bpp | 5bpp |
| Vehicle-borne GPR | FDQ       | 10.2881 | 16.0794 | 21.6214 | 26.2841 |
|            | FDAQ      | 13.7844 | 19.4623 | 24.9756 | 29.6991 |
|            | DCT-FDAQ  | **16.9187** | **22.7574** | **28.2688** | **32.8922** |
| Aircraft-borne GPR | FDQ       | 9.5338 | 15.0791 | 20.1226 | 23.8525 |
|            | FDAQ      | 12.4723 | 17.8920 | 23.0722 | 26.8473 |
|            | DCT-FDAQ  | **15.4222** | **20.8469** | **25.9947** | **29.6736** |
| Satellite-borne GPR | FDQ       | 8.5547 | 13.8792 | 18.7272 | 22.3061 |
|            | FDAQ      | 10.2289 | 15.5478 | 20.4966 | 24.0230 |
|            | DCT-FDAQ  | **12.0871** | **17.3858** | **22.2033** | **25.7497** |

Figure 3. Imagings of the compressed MCRDS-2 data with proposed algorithms (5bpp): (a) raw data, (b) FDQ algorithm, (c) FDAQ algorithm, (d) DCT-FDAQ algorithm
In order to study the influence of DCT transform and adaptive bit allocation mechanism on the performance of the proposed algorithm, the following three algorithms are tested simultaneously. The FDQ algorithm does not include DCT transform and adaptive bit allocation mechanism, and the FDAQ algorithm does not include DCT transform compared to the DCT-FDAQ algorithm. Table 1 shows the performance of the three algorithms with different quantization bits.

As shown in Table 1, the overall performance of the proposed DCT-FDAQ algorithm is much better than the compared FDQ and FDAQ algorithm, which proves the DCT transform and adaptive bit allocation mechanism would obtain high compression gains. Taking the aircraft-borne GPR raw data as an example, the SQNRs of the DCT-FDAQ algorithm with different quantization bits are improved by an average of nearly 5.9dB and 3.0dB compared with the FDQ and FDAQ algorithm.

Figure 3 shows the imaging comparison of the aircraft-borne MCRDS-2 raw data and compressed data with the proposed algorithms. Imagings of the compressed data with FDQ and FDAQ algorithm produce noise spots, and some areas with massive spots will destroy the continuity of the image. The imaging of the compressed data with 5bpp DCT-FDAQ algorithm is almost the same as that of the raw data, and there is little visual difference can be distinguished.

5. Conclusion
This paper proposes a novel DCT-FDAQ algorithm for GPR raw data compression, which obtains high compression gains through the DCT transform and a kind of adaptive bit allocation mechanism. Experimental results with different GPR datasets show the effectiveness of the proposed algorithm. And engineering implementation of the proposed algorithm is planned as part of future work.

Acknowledgments
This work was supported by National Natural Science Foundation of China (No. u1531141) and the Key Laboratory of Electromagnetic Radiation and Detection Technology, Chinese Academic of Science.

References
[1] Ho, K. C., L. Carin, P. D. Gader, and J. N. Wilson, "An investigation of using the spectral characteristics from ground penetrating radar for landmine/clutter discrimination," IEEE Transactions on Geoscience and Remote Sensing, 46, pp.1177–1191,(2008).
[2] W. Zhao, E. Forte, R. R. Colucci and M. Pipan, "High-resolution glacier imaging and characterization by means of GPR attribute analysis," Geophysical Journal International, 206, pp.1366-1374, (2016).
[3] M. S. El-Mahallawy and M. Hashim, "Material Classification of Underground Utilities From GPR Images Using DCT-Based SVM Approach," IEEE Geoscience and Remote Sensing Letters, 10, pp.1542-1546, (2013).
[4] G. Roe. "Quantizing for Minimum Distortion". Ire Transactions on Information Theory, 6, pp.7-12, (2003).
[5] S. H. Yang and P. B. Vo, "Adaptive Bit Allocation for Consistent Video Quality in Scalable High Efficiency Video Coding," IEEE Transactions on Circuits and Systems for Video Technology, 27, pp.1555-1567, (2017).
[6] Croci R, Seu R, Flamini E, Russo E, "The SHAallow RADar (SHARAD) Onboard the NASA MRO Mission". Proceedings of the IEEE, 99, pp.794-807, (2011).
[7] S. Lang, B. Zhao, X. Liu and G. Fang, "Two-Dimensional Imaging of Ice Sheets of Airborne Radar Sounder via a Combined Modified Range Migration Algorithm Based on ISFT and Beamforming Using Curvelets," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8, pp.76-89, (2015).
[8] X. Liu, S. Lang, B. Zhao "A high-resolution polar ice penetrating radar and experiments in the 28th Chinese National Antarctic Research Expedition," Proceedings of the 15th International Conference on Ground Penetrating Radar, Brussels, pp.719-722, (2014).