

Lossless Compression for Hyperspectral Images based on Adaptive Band Selection and Adaptive Predictor Selection

Fuquan Zhu^{1,2}, Huajun Wang^{1*}, Liping Yang³, Changguo Li⁴, and Sen Wang⁵

¹ College of Geophysics, Chengdu University of Technology, Chengdu 610059, China

² Academic Affairs Office, Sichuan Police College, Luzhou 646000, China

³ General Education Department, Sichuan Police College, Luzhou 646000, China

⁴ College of Fundamental Education, Sichuan Normal University, Chengdu 610068, China

⁵ Technical Center of Shanghai Shentong Metro Group Co., Ltd, Shanghai 201102, China

*Corresponding author: Huajun Wang [e-mail: hjwang@cdut.edu.cn]

*Received February 21, 2020; revised April 29, 2020; revised June 14, 2020; accepted July 18, 2020;
published August 31, 2020*

Abstract

With the wide application of hyperspectral images, it becomes more and more important to compress hyperspectral images. Conventional recursive least squares (CRLS) algorithm has great potentiality in lossless compression for hyperspectral images. The prediction accuracy of CRLS is closely related to the correlations between the reference bands and the current band, and the similarity between pixels in prediction context. According to this characteristic, we present an improved CRLS with adaptive band selection and adaptive predictor selection (CRLS-ABS-APS). Firstly, a spectral vector correlation coefficient-based k-means clustering algorithm is employed to generate clustering map. Afterwards, an adaptive band selection strategy based on inter-spectral correlation coefficient is adopted to select the reference bands for each band. Then, an adaptive predictor selection strategy based on clustering map is adopted to select the optimal CRLS predictor for each pixel. In addition, a double snake scan mode is used to further improve the similarity of prediction context, and a recursive average estimation method is used to accelerate the local average calculation. Finally, the prediction residuals are entropy encoded by arithmetic encoder. Experiments on the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) 2006 data set show that the CRLS-ABS-APS achieves average bit rates of 3.28 bpp, 5.55 bpp and 2.39 bpp on the three subsets, respectively. The results indicate that the CRLS-ABS-APS effectively improves the compression effect with lower computation complexity, and outperforms to the current state-of-the-art methods.

Keywords: hyperspectral image lossless compression, conventional recursive least squares, k-means clustering, adaptive band selection, adaptive predictor selection

This work was supported by the National Natural Science Foundation of China (Grant No. 61373162), the Research Foundation of the Sichuan Department of Education (Grant No. 15ZB0044), and the Research Foundation of Sichuan Normal University (Grant No. 2015KYQD312).

1. Introduction

Hyperspectral image is a radiance imagery acquired by imaging spectrometer in the ultraviolet, visible, near infrared, shortwave, medium wave and long wave infrared regions of the electromagnetic spectrum. Because of containing abundant spatial and spectral information, hyperspectral images provide a great capability to classify and identify ground objects [1,2]. Therefore, hyperspectral images have been widely used in resource management, environmental monitoring, military reconnaissance, geological exploration, agricultural production and so on. With the improvement of spectral and spatial resolution, the data volume of hyperspectral image increases rapidly. For instance, the hyperspectral image acquired in Lunar lake regions of the United States by the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) in 2009, consists of 6955 rows, 781 columns and 224 bands, and needs more than 2.3 GB of storage space [3]. Moreover, with the wider application of hyperspectral remote sensing technology, more and more hyperspectral images are produced. However, the transmission bandwidth and storage capacity of spaceborne or airborne remote sensing system are usually very limited. Therefore, it is necessary to compress hyperspectral images effectively before transmission and storage.

Hyperspectral image compression techniques are classified into lossy compression and lossless compression. Lossy compression can achieve a higher compression ratio but will cause distortion [4,5]. Due to the distortion, the original images cannot be recovered. That has an unpredictable effect on subsequent research and application of hyperspectral images. Thus, lossless compression is the preferred scheme of hyperspectral image compression. Lossless compression techniques can be classified into three types, namely, transform-based methods [6], vector-quantization-based methods [7], and prediction-based methods. The latter has attracted much attention for low computation complexity and good compression effect.

Prediction-based compression methods are implemented by removing correlations in data. Hyperspectral images have two correlations, namely, the spatial correlation between pixels and the inter-spectral correlation between bands. However, due to the low spatial resolution, spatial correlation is often relatively weak. Therefore, the compression methods of natural images mainly based on spatial correlation are not suitable for hyperspectral images. For example, because the inter-spectral correlation is not utilized, the context-based adaptive lossless image coding (CALIC) [8] method and the low complexity lossless compression for images (LOCO-I) [9] method perform well on natural images, but fail to achieve a good compression effect on hyperspectral images. Thanks to the high spectral resolution, the inter-spectral correlation is very strong. That provides the possibility of high-performance compression. Lookup table (LUT) [10] is a typical compression method utilizing inter-spectral correlation. In the method, the value of the pixel in the previous band at the same spatial position as the current pixel is used as an index of lookup table, and the corresponding value of the lookup table is used as prediction value. Locally averaged inter-band scaling (LAIS)-LUT [11] improves LUT by choosing the closest value to the local average inter-band scaling estimate as the prediction value from two lookup tables. In [12], LAIS-LUT is improved by quantizing the indices with adaptive quantization factors. Because the time-consuming search process is replaced by the lookup table, the LUT series methods have very low complexity. But only the regularities in calibrated images and few reference bands are utilized, their prediction accuracy is not high, especially on uncalibrated images. Clustered difference pulse code modulation (C-DPCM) [13] adopts classified prediction strategy and multiple linear regression models to improve compression effect.

C-DPCM with adaptive prediction length (APL) [14] adopts exhaustive search strategy to search the best number of reference bands for each band, and achieves a better compression effect. But the exhaustive search is a very time-consuming process. C-DPCM with removal of local spectral outliers (RLSO) [15] improves C-DPCM by removing the spectral outliers in each cluster and employing the high-order predictors for each band. In [16], recurrent neural network (RNN) is applied to the lossless compression for hyperspectral images. In [17], a one-pass framework performing lossless data hiding and compression of remote sensing data is introduced.

In recent years, adaptive filtering algorithms have been widely applied in hyperspectral image compression. Fast lossless (FL) [18] adopts least mean square (LMS) filter to calculate prediction coefficients. Because of its low computation complexity, FL has been recommended as the standard of on-board real-time hyperspectral imagery compression for the Consultative Committee for Space Data Systems (CCSDS) [19]. In [20], FL is used to compress multitemporal hyperspectral image data. Third-order inter-band prediction (IP3) [21] employs third-order wiener filter to improve prediction accuracy. But the wiener filter needs to solve the inverse of the correlation coefficient matrix. That increases computational complexity significantly. To avoid solving inverse matrix, recursive least square (RLS) filter is adopted in [22]. For obtaining a better compression effect, fast RLS with adaptive length prediction (Fast-RLS-ALP) [23] adopts adaptive length prediction strategy. CRLS with adaptive prediction bands (A-CRLS) [24] uses a forgetting factor to reduce the influence of the old data and selects the optimal number of reference bands by exhaustive searching, and achieves a very good compression results. Bimodal CRLS (B-CRLS) [25] defines fully-spectral mode and spatio-spectral mode to generate input vector of CRLS filter and selects the best mode to obtain minimum residual for each band. Although CRLS has strong tracking ability, the lack of stationarity reduces its prediction accuracy. To solve the problem, Clustered CRLS (C-CRLS) [26] adopts the classified prediction strategy and uses different CRLS predictors to predict pixels of different clusters. But C-CRLS does not use clustering information to calculate local average. B-SuperCRLS [27] uses superpixel segmentation to partition the hyperspectral image to multiple small regions and uses the pixels in the same regions to calculate local average, and achieves the best compression effect. However, due to the low spatial resolution, the classification effect of superpixel is not better than that of k-means clustering.

In addition, for further improving the compression performance, researchers propose some effective preprocessing methods, such as band reordering, pixel normalization and scanning mode [28-30]. To reduce the compression time, researchers introduce the graphics processing units (GPUs) based parallel computation to hyperspectral image compression [31,32].

Due to the strong tracking performance, CRLS shows great potential in lossless compression for hyperspectral images. However, as an online prediction algorithm, CRLS uses the predicted pixels in the reference bands and the current band to calculate the linear prediction coefficient. Therefore, the correlation between the reference bands and the current band, and the similarity between the pixels in prediction context and the current pixel, both directly affect the prediction accuracy of CRLS predictor. In order to utilize the characteristic to improve compression effect, an effective lossless compression method using CRLS with adaptive band selection and adaptive predictor selection (CRLS-ABS-APS) is proposed. As an extension from CRLS, CRLS-ABS-APS is mainly improved as follows. Firstly, clustering map based adaptive predictor selection strategy is employed to select the optimal CRLS predictor for each pixel. Secondly, in order to improve the prediction effect

obtained from clustering, the spectral vector correlation coefficient is used as the measure of similarity in k-means clustering algorithm. Thirdly, an adaptive band selection strategy is used to improve the correlation between the reference bands and the current band. Fourthly, a double snake scan mode is adopted to further improve the stationarity of prediction process. Fifthly, a recursive local average estimation method is employed to accelerate calculating local average.

The rest of the paper is organized as follows. In section 2, the proposed method is described in detail. Section 3 presents experimental results. Finally, some conclusions are given in section 4.

2. Proposed Method: CRLS-ABS-APS

CRLS-ABS-APS consists of clustering module, JPEG-LS predictor, adaptive band selection module, adaptive predictor selection module and arithmetic encoder. As shown in Fig. 1, the spectral vector correlation coefficient-based k-means clustering module is used to generate the clustering map used to select predictor. Afterwards, the first band image is predicted by JPEG-LS predictor [25], and the rest band images are predicted by adaptive predictor selection module. The adaptive predictor selection module contains multiple CRLS predictors, and selects the best one for each pixel according to its cluster number. The adaptive band selection module is used to search the best reference bands for each band. Finally, all prediction residuals are entropy coded by arithmetic encoder [25]. Due to the demand of decompression, the clustering map and the results of adaptive band selection are stored or transmitted as side information. The former is removed correlation by JPEG-LS predictor before being encoded. The latter, some arrays composed of band numbers, is directly converted to bit stream.

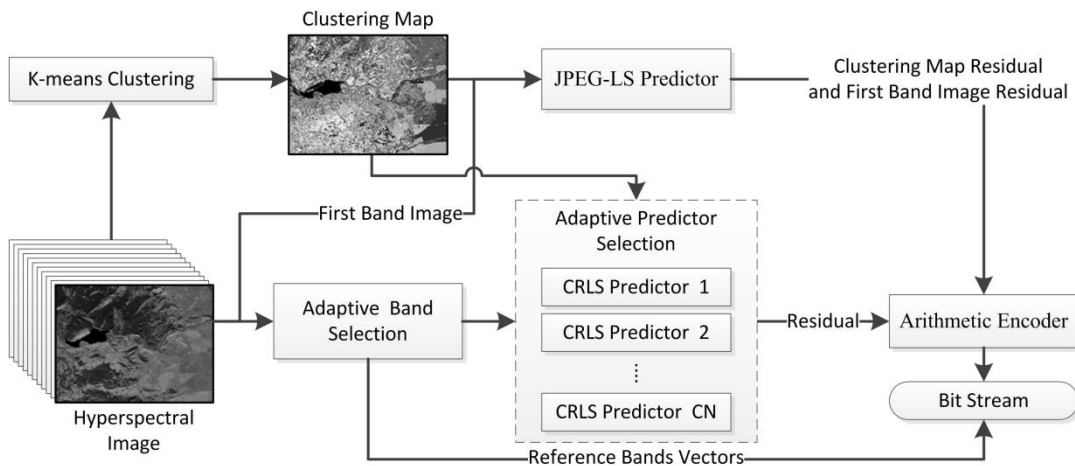


Fig. 1. Block diagram of CRLS-ABS-APS

Next, the k-means clustering module, the adaptive band selection module and the adaptive predictor selection module are introduced in detail.

2.1 K-means Clustering based on Spectral Vector Correlation Coefficient

Because of low spatial resolution, adjacent pixels often represent different ground objects in hyperspectral images. This means that the values of adjacent pixels in the same band may

vary greatly. Therefore, using adjacent pixels to calculate the local average and prediction coefficient of the current pixel will reduce prediction accuracy. In [26], to solve the problem, the Euclidean distance is adopted to clustering spectral vectors. However, Euclidean distance is susceptible to outliers, especially for hyperspectral images because of the large scale variation between their bands. That easily results in erroneous clustering of spectral vectors. More importantly, Euclidean distance is not the best expression of linear correlation which is the basis for CRLS predictor to be efficient. Therefore, we choose spectral vector correlation coefficient to replace Euclidean distance as the measure of similarity in k-means clustering algorithm, because the former can represent the linear correlation between bands better.

Let $\mathbf{p}_i = [p_i(1) p_i(2) \dots p_i(N)]$ and $\mathbf{p}_j = [p_j(1) p_j(2) \dots p_j(N)]$ are two spectral vectors of hyperspectral image, where N denotes the number of bands in hyperspectral image. The spectral vector correlation coefficient between \mathbf{p}_i and \mathbf{p}_j is defined as follows

$$\rho(\mathbf{p}_i, \mathbf{p}_j) = \frac{\sum_z [p_i(z) - \bar{p}_i][p_j(z) - \bar{p}_j]}{\sqrt{\sum_z [p_i(z) - \bar{p}_i]^2 \sum_z [p_j(z) - \bar{p}_j]^2}}, \quad (1)$$

where \bar{p}_i and \bar{p}_j denote the average of the corresponding spectral vectors, z denotes the band number.

The details of the k-means clustering based on spectral vector correlation coefficient are given in [Algorithm 1](#).

Algorithm 1: The k-means clustering algorithm based on spectral vector correlation coefficient

Input: CN // CN is the number of clusters.

Output: CM // CM is the clustering map.

- (1) **initialization:** $\mu_i \leftarrow \mathbf{p}_j, i = 1, 2, \dots, CN$, $cm = 0$, $iter = 0$ // μ_i denotes the cluster center.
// cm is the convergence mark, $iter$ is the number of iterations.
 - (2) **repeat:**
 - (3) $iter \leftarrow iter + 1$
 - (4) **repeat:**
 - (5) $j \leftarrow 1$
 - (6) assign the spectral vector \mathbf{p}_j to the most similar cluster by $\arg \max_i \rho(\mathbf{p}_j, \mu_i)$
 - (7) **until** $j = W \times H$ // W and H are the width and height of hyperspectral image, respectively.
 - (8) update the cluster centers $\mu_i \leftarrow \frac{1}{N_i} \sum_{j=1}^{N_i} \mathbf{p}_j, i = 1, 2, \dots, CN$ // \mathbf{p}_j denotes the vector belong to
// the i -th cluster. N_i is the number of corresponding spectral vectors.
 - (9) if all the cluster centers are unchanged
 - (10) $cm \leftarrow 1$
 - (11) **until** $cm = 1$ or $iter = 100$
-

2.2 Adaptive Band Selection

The prediction accuracy of the CRLS predictor depends on the correlation between the reference bands and the current band. However, it is prevalent in hyperspectral images that the correlations between closer bands may be weaker than those between distant bands. As shown in Fig. 2, the inter-spectral correlation coefficients between the 155 to 165 band and the 170 band are lower than those between the 113 to 154 band and the 170 band. Therefore, it is difficult to achieve the best prediction effect by using adjacent bands as reference bands.

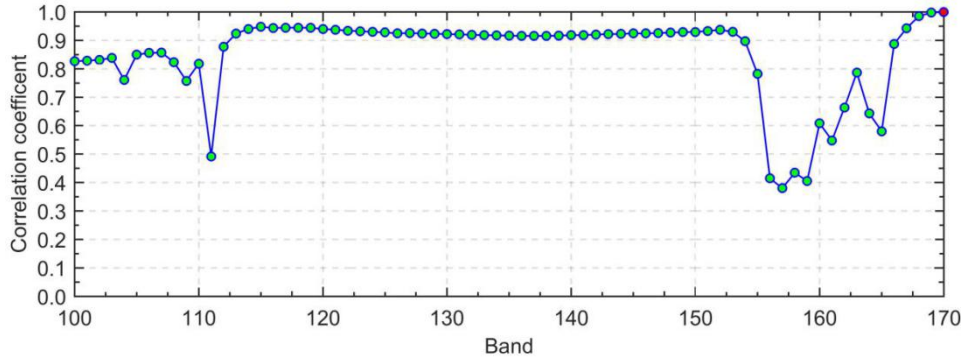


Fig. 2. Inter-spectral correlation coefficients between the 100 to 169 band and the 170 band in the calibrated Yellowstone scene 0 of the AVIRIS 2006 data set.

To solve this problem, we adopt the adaptive band selection strategy, that is, to select reference bands based on maximum correlation principle. The effective way to evaluate the inter-band correlation is to calculate the inter-spectral correlation coefficient, which is defined as follows

$$c_{i,j} = \frac{\sum_{x=1}^H \sum_{y=1}^W [p_i(x,y) - \bar{p}_i][p_j(x,y) - \bar{p}_j]}{\sqrt{\sum_{x=1}^H \sum_{y=1}^W [p_i(x,y) - \bar{p}_i]^2 \sum_{x=1}^H \sum_{y=1}^W [p_j(x,y) - \bar{p}_j]^2}}, \quad (2)$$

where i and j denote the i -th band and j -th band, respectively. $p_i(x,y)$ denotes the value of pixel at the x -th row and y -th column in the i -th band, and \bar{p}_i denotes the average value of pixels in the i -th band. W and H denote the width and height of hyperspectral image, respectively.

The details of adaptive band selection module are given in Algorithm 2. It is to simplify the search process that the correlation coefficient searched in step (8) is set to -1 in step (9). That does not affect the correctness of the results, because the correlation coefficient is always between -1 and 1. The output \mathbf{rb}_z is an array of band numbers, which will be used to generate image fragments in the prediction stage.

Algorithm 2: Implementation of the adaptive band selection module

Input: RN // RN is the default number of reference bands.

Output: \mathbf{rb}_z // \mathbf{rb}_z denotes the reference bands of the z -th band.

- (1) **if** $z \leq RN + 1$
 - (2) **return** $\mathbf{rb}_z = [1, 2, \dots, z - 1]$
 - (3) **else**
 - (4) **for** $i = 1$ to $z - 1$ **do**
 - (5) calculate inter-spectral correlation coefficients $c_{i,z}$ by Eq. 2
 - (6) **end for**
 - (7) **for** $j = 1$ to RN **do**
 - (8) $rb_z(j) = \arg \max_i (c_{i,z}, 1 \leq i \leq z - 1)$
 - (9) $c_{rb_z(j),z} = -1$
 - (10) **end for**
 - (11) **return** $\mathbf{rb}_z = [rb_z(1)rb_z(2)\dots rb_z(RN)]$
 - (12) **end if**
-

2.3 Predicting using Adaptive Predictor Selection

Scan order affects the stationarity of the prediction process, hence CRLS predictor is extremely sensitive to scan order. In order to ensure the stationarity of the prediction process as possible, we design a double snake scan mode. As shown in Fig. 3, the clustering map used to search for similar pixels, is divided into belts with fixed width. The snake scan mode is used both inside and between belts. The belt width is called snake scanning width (SW). When $SW=1$, it is the normal transverse snake scan mode; when SW is equal to the image width, it is the normal vertical snake scan mode. Compared with the traditional raster scan order mode, the double snake scan mode can ensure that the previous pixel in the same cluster is spatially closer to the current pixel when SW takes an appropriate value.

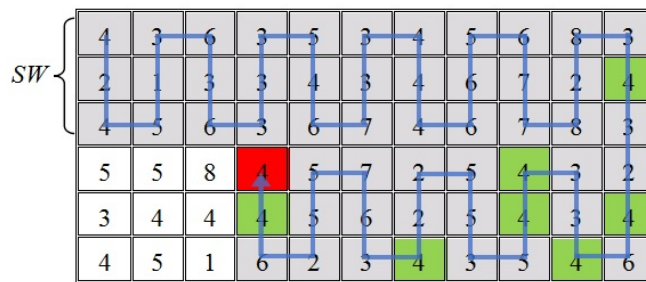


Fig. 3. Double snake scan mode. The numbers in the boxes denote the cluster number. The blue broken line with arrow denotes the scan order. The red box denotes the current pixel. The gray boxes denote the predicted pixels. The green boxes denote the similar neighborhood of the current pixel.

For preferable prediction performance, CRLS predictor needs to calculate local average to generate input vector [26]. There are two ways to calculate local average. One is to calculate the average of pixels in local neighborhood. The other is to calculate the average of pixels in similar neighborhood, which is a set of pixels with the same cluster as the current pixel, as shown in Fig. 3. The latter can obtain better prediction effect, but needs to generate similar

neighborhood by searching for the same cluster pixels from the predicted pixels in reverse scan order. To avoid the time-consuming searching process, we adopt a simple and efficient recursive local average estimation method. The recursive local average estimation only needs reading and updating in predicting. Before predicting, the local average vector of each predictor is initialized to a zero vector and cached. In predicting, the local average vector is read according to the cluster number of the current pixel to generate the input vector and then updated by

$$\boldsymbol{\mu}_i = \boldsymbol{\mu}_i + \frac{\mathbf{p} - \boldsymbol{\mu}_i}{pn}, 1 \leq pn \leq PN \quad (3)$$

where i is the predictor index, pn is the number of pixels in similar neighborhood, PN denotes the default size of similar neighborhood, $\mathbf{p} = [p(1)p(2)\dots p(rn)p(rn+1)]$ is a vector consisting of the current pixel and the corresponding pixels of reference bands, and rn denotes the actual number of reference bands for the current band.

Adaptive predictor selection module selects the optimal predictor for each pixel from predictor set. Predictor set consists of CN CRLS predictors with different prediction coefficient, where CN is the number of clusters. Predictor is selected according to the cluster number of the current pixel. The details of the adaptive predictor selection method are shown in [Algorithm 3](#).

Each band except the first is predicted by Algorithm 3. The reference bands as the part of the input are generated according to the reference bands array produced by adaptive band selection module. Step (2) implements the double snake scan mode. Adaptive predictor selection method is performed from step (3) to step (9), and CRLS predictor is performed from step (6) to step (9). Recursive local average estimation is implemented in step (4) and (10).

Algorithm 3: Prediction process of the z -th band using adaptive predictor selection method

Input: $\mathbf{HS}, \mathbf{CM}, PN, SW$ // \mathbf{HS} denotes the hyperspectral image segment consisting of the reference bands and the current band.

Output: \mathbf{e}_z // $\mathbf{e}_z = [e_z(1)e_z(2)\dots e_z(W \times H)]$ denotes prediction residuals of the z -th band.

(1) **initialization:** $\delta = 0.001$, $\lambda = 0.9995$, $\Phi_{ci} = \delta \mathbf{I}$, $\mathbf{w}_{ci} = [\mathbf{0}_{1 \times m} \ \mathbf{0}_{1 \times rn}]^T$, $\boldsymbol{\mu}_{ci} = [\mathbf{0}_{1 \times m} \ \mathbf{0}_{1 \times rn}]^T$

// δ denotes the regularization parameter, λ denotes the forgetting factor,

// Φ_{ci} denotes inverse correlation matrix, \mathbf{I} is a $rn \times rn$ identity matrix,

// \mathbf{w}_{ci} denotes the weight coefficient, $\boldsymbol{\mu}_{ci}$ denotes the local average,

// $ci = 1, 2, \dots, CN$.

(2) **for** $i = 1$ **to** $W \times H$ **do** // i is calculated by double snake scan mode.

(3) get the cluster number ci

(4) get the local average $\boldsymbol{\mu}_{ci} = [\mu_{ci}(1)\mu_{ci}(2)\dots \mu_{ci}(rn)\mu_{ci}(rn+1)]^T$

(5) calculate the input vector

$$\mathbf{u} = [p_i(1) - \mu_{ci}(1), p_i(2) - \mu_{ci}(2), \dots, p_i(rn) - \mu_{ci}(rn)]^T$$

(6) calculate the prediction residual $e_z(i) = p_i(rn+1) - \mathbf{u}^T \mathbf{w}_{ci} - \mu_{ci}(rn+1)$

-
- (7) calculate the gain vector $\mathbf{k} = \frac{\Phi_{ci} \mathbf{u}}{\lambda + \mathbf{u}^T \Phi_{ci} \mathbf{u}}$
- (8) update the inverse correlation matrix $\Phi_{ci} \leftarrow \frac{1}{\lambda} (\Phi_{ci} - \mathbf{k} \mathbf{u}^T \Phi_{ci})$
- (9) update the weight coefficient vector $\mathbf{w}_{ci} \leftarrow \mathbf{w}_{ci} + e_z(i) \mathbf{k}$
- (10) update the local average vector μ_{ci} by Eq. 3
- (11) **end for**
-

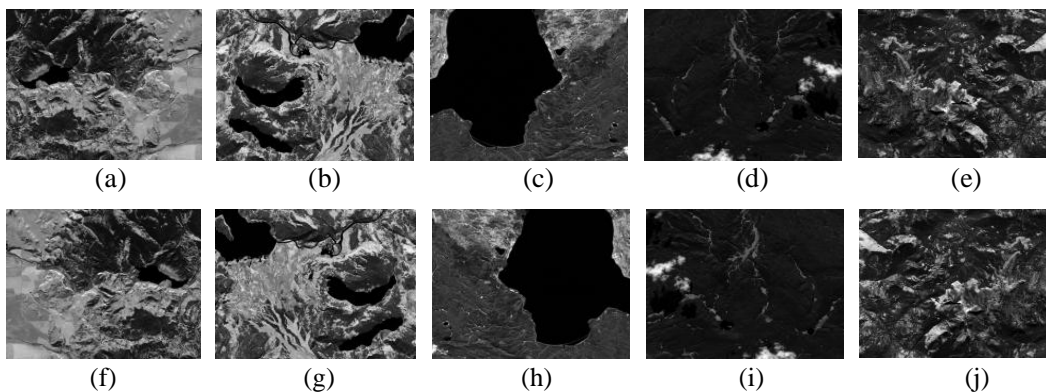
For reducing the memory occupied by the algorithm, the prediction starts from the last band to the first band, and each band image is overwritten with its residuals. The decompression is performed in the opposite order.

3. Experimental Results

In order to verify the validity of CRLS-ABS-APS, we carried out the simulation experiments using MATLAB 2018a on a personal computer configured with 3.40 GHz Intel Core i7 processor and 16 GB RAM. The test data is the AVIRIS 2006 data set [33], which is the latest data set dedicated to compression testing and evaluation provided by the CCSDS multispectral and hyperspectral data compression working group. AVIRIS 2006 data set consists of 12 hyperspectral images, and can be divided into three subsets, namely, a 16-bit calibrated image subset, a 16-bit uncalibrated image subset and a 12-bit uncalibrated image subset. The specifications of AVIRIS 2006 data set are shown in Table 1. The 128th band images of the data set are shown in Fig. 4. Bit rate is adopted to estimate the compression effect in the experiments.

Table 1. Specifications of AVIRIS 2006 data set

Name	Size	Calibrated	Formation	Scene Number
Yellowstone	512×677×224	YES	Signed 16-bit	0,3,10,11,18
Yellowstone	512×680×224	NO	Unsigned 16-bit	0,3,10,11,18
Hawaii	512×614×224	NO	Unsigned 12-bit	1
Maine	512×680×224	NO	Unsigned 12-bit	10



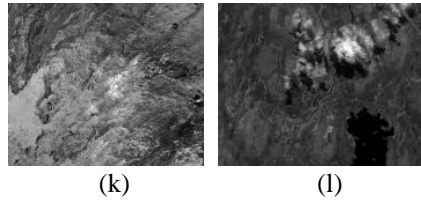


Fig. 4. The 128th band images of AVIRIS 2006 data set. (a) - (e) 16-bit calibrated Yellowstone 0, 3, 10, 11, 18. (f) - (j) 16-bit uncalibrated Yellowstone 0, 3, 10, 11, 18. (k) - (l) 12-bit uncalibrated Hawaii 1 and Maine 10.

3.1 Determination of Parameters

There are six adjustable parameters in the CRLS-ABS-APS method, namely, regularization parameter δ , forgetting factor λ , clustering number CN , number of pixels in similar neighborhood PN , snake scan width SW and number of reference bands RN . As in [26], δ and λ are set to 0.001 and 0.9995, respectively. The values of the rest parameters are obtained by experiments. The determination of RN is presented in next section.

Because CN , PN and SW affect each other, for simplicity, the three parameters are paired for experiments to obtain the optimal values. Average bit rates obtained by CRLS-ABS-APS on the three image subsets are shown in Fig. 5. It can be observed that the parameter pairs (PN, SW) have the same effect on the three image subsets. When $PN=20$ and $SW=20$, the lowest average bit rates are achieved on the all three image subsets. In theory, the more clusters, the better the compression performance. However, because the clustering map needs to be stored as side information, increasing the clustering number will increase the bit rate. Therefore, the optimal value of CN needs to be selected. Because of the different data specifications, the clustering numbers for the best compression effect on the three image subsets may be different. For instance, the optimal values of CN for the three image subsets are 20, 16 and 8, respectively. In the following experiments, we select the above optimal values.

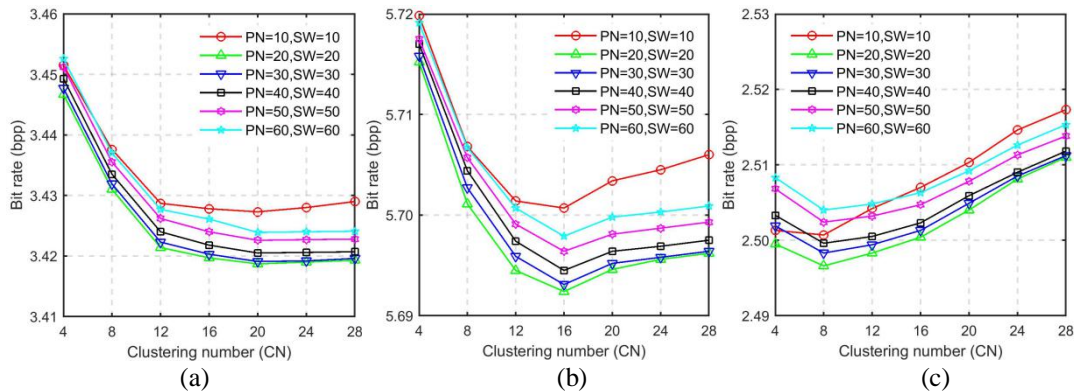


Fig. 5. Average bit rates obtained by CRLS-ABS-APS on the three image subsets for various clustering number. (a) 16-bit calibrated image subset, (b) 16-bit uncalibrated image subset, and (c) 12-bit uncalibrated image subset. The number of reference bands $RN=8$.

3.2 Improvement of Compression Effect

In order to show the effect of using spectral vector correlation coefficient instead of Euclidean distance in clustering preprocessing, the differences of the bit rates obtained by the two methods are shown in Fig. 6. It can be seen that spectral vector correlation coefficient based clustering obtains lower bit rate in most bands. Further observations

indicate that the improvements obtained on calibrated images are more pronounced than those on uncalibrated images.

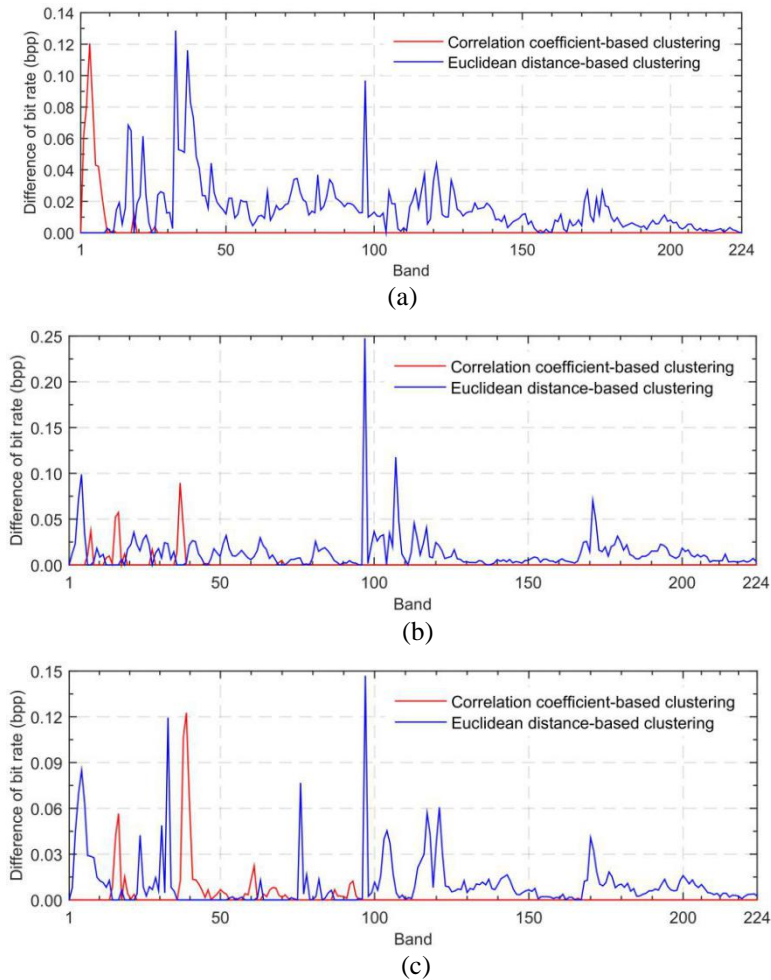


Fig. 6. Difference of the bit rates obtained by Euclidean distance based clustering and spectral vector correlation coefficient based clustering in preprocessing. (a) 16-bit calibrated Yellowstone 18, (b) 16-bit uncalibrated Yellowstone 3, and (c) 12-bit uncalibrated Maine 10. The number of reference bands $RN=8$.

In order to illustrate the effectiveness of the other improvement strategies, the average bit rates obtained by CRLS, C-CRLS, CRLS-APS, and CRLS-ABS-APS are compared in **Fig. 7**. Note that CRLS-APS indicates the proposed algorithm without adaptive band selection module. As shown in **Fig. 7**, CRLS-APS obtains a better compression effect than CRLS and C-CRLS, while CRLS-ABS-APS achieves the lowest average bit rate. Even on the 12-bit uncalibrated image subset where C-CRLS does not get any improvements, both CRLS-APS and CRLS-ABS-APS achieve obvious improvements. However, the comparison between CRLS-ABS-APS and CRLS-APS also shows that the compression effect obtained by adaptive band selection is very obvious when RN is small, but decreases gradually as RN increases. Moreover, it can be seen that the optimal value of RN is 56, because a further increase in RN cannot bring meaningful improvements. Therefore, parameter RN is set to 56 in this paper.

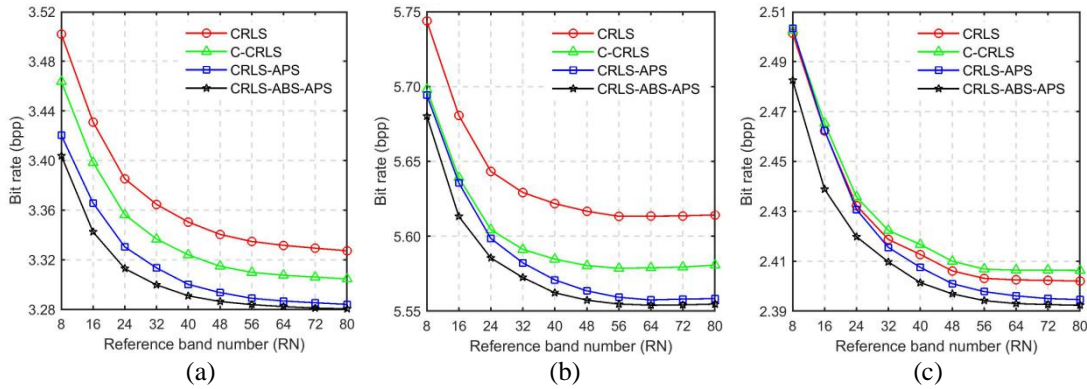


Fig. 7. Average bit rates obtained by different methods on the three image subsets for various number of reference bands. (a) 16-bit calibrated image subset, (b) 16-bit uncalibrated image subset, and (c) 12-bit uncalibrated image subset.

3.3 Comparison with Other Methods

In this section, CRLS-ABS-APS is compared with the typical lossless compression methods, such as FL, IP3, RLS, C-DPCM-APL, CRLS, C-CRLS, A-CRLS, B-CRLS and B-SuperRLS. The results of these methods are given in **Table 2**. For simplicity, each image is represented by the initials of name and calibration status and the scene number. For example, the calibrated Yellowstone scene 0 is represented as YC0, and the uncalibrated Yellowstone scene 0 is represented as YU0. The results of FL, IP3, RLS, C-DPCM-APL, A-CRLS and B-CRLS are obtained from [25], and that of B-SuperRLS are obtained from [27]. CRLS[#] and C-CRLS[#] represent CRLS and C-CRLS with the same number of reference bands (i.e. $RN=56$) as the proposed method, respectively. The average bit rates for each image subset and the lowest bit rates for each image are marked in boldface.

Table 2. Compression results of the AVIRIS 2006 data set (in bits per pixel)

Image	FL	IP3	RLS	C-DP CM- APL	CR LS [#]	C-C RLS [#]	A-C RLS	B-C RLS	B-S upe rRL	CRLS -ABS- APS
YC0	3.94	3.84	3.77	3.52	3.52	3.51	3.50	3.50	3.51	3.48
YC3	3.83	3.74	3.63	3.36	3.40	3.39	3.39	3.39	3.39	3.35
YC10	3.34	3.21	3.24	2.93	3.06	3.01	3.01	3.01	2.95	2.97
YC11	3.61	3.51	3.48	3.25	3.25	3.22	3.24	3.23	3.22	3.20
YC18	3.92	3.75	3.67	3.42	3.43	3.41	3.41	3.41	3.41	3.38
Average	3.73	3.61	3.56	3.30	3.33	3.31	3.31	3.31	3.29	3.28
YU0	6.23	6.08	6.01	5.81	5.78	5.78	5.79	5.75	5.77	5.76
YU3	6.10	5.96	5.86	5.65	5.66	5.64	5.67	5.63	5.65	5.63
YU10	5.57	5.44	5.46	5.17	5.30	5.27	5.28	5.25	5.19	5.23
YU11	5.84	5.74	5.71	5.47	5.55	5.47	5.50	5.46	5.46	5.45
YU18	6.29	6.00	5.90	5.71	5.77	5.73	5.70	5.66	5.67	5.67
Average	6.00	5.84	5.79	5.56	5.61	5.58	5.59	5.55	5.55	5.55
HU1	2.57	2.58	2.50	2.35	2.32	2.33	2.32	2.31	2.33	2.30
MU10	2.68	2.68	2.63	2.51	2.49	2.48	2.49	2.44	2.48	2.48
Average	2.63	2.63	2.57	2.43	2.41	2.41	2.41	2.38	2.41	2.39

As can be seen from **Table 2** and **Fig. 4**, like other methods, the compression performance of the CRLS-ABS-APS is also affected by the image specifications and content. However, as shown in **Table 2**, CRLS-ABS-APS obtains the best compression effect in all methods. Compared with C-CRLS[#], CRLS-ABS-APS reduces the average bit rates by 0.03 bpp, 0.03 bpp and 0.02 bpp on the three image subsets, respectively. Even compared with the current state-of-the-art methods, e.g., B-CRLS and B-SuperRLS, CRLS-ABS-APS also has obvious superiority on 16-bit calibrated image subset. In addition, it should be emphasized that CRLS-ABS-APS also obtains an improvement in compression effect on the 12-bit image subset, on which other similar methods based on classified compression such as C-CRLS and B-SuperRLS fail to further reduce the bit rate of these images.

3.4 Computational Complexity Analysis

CRLS-ABS-APS mainly contains four steps, namely clustering, reference band selection, prediction and encoding. For one hyperspectral image, clustering module requires $4 \times 16 \times 100 \times WHN$ addition and $3 \times 16 \times 100 \times WHN$ multiplication, where W , H and N represent the width, height and number of band for hyperspectral image, respectively. Adaptive band selection module needs about $4 \times 112 \times WHN$ addition and $3 \times 112 \times WHN$ multiplication. The prediction using adaptive predictor selection demands $56^2 \times WHN$ addition and $3 \times 56^2 \times WHN$ multiplication. Thanks to using the algorithm provided by MATLAB, the computation times of clustering and reference bands selecting are greatly reduced. By statistics, the average computation time for compressing one hyperspectral image is about 2188 s. The average time spent on the four steps is 23 s, 1 s, 2080 s and 84 s, respectively.

Table 3 shows the average computation times of the mentioned methods. FL method needs the least time due to the use of low complexity LMS predictor and the fewest reference bands. Because of using less reference bands, RLS method takes less time than the proposed method. C-DPCM-APL and A-CRLS need multiple predictions for each band, so they take the longest time. Without fast mode selection or parallel computation, B-CRLS and B-SuperRLS need twice predictions to choose the best prediction mode, hence they require long time too. Due to the need to calculate inverse of correlation matrix, IP3 takes more time than CRLS-ABS-APS. Thanks to adopting the recursive local average estimation, CRLS-ABS-APS needs only two vector addition and one vector multiplication to calculate local average for each pixel. For this reason, CRLS-ABS-APS requires less time than CRLS[#] and C-CRLS[#].

Table 3. Average computation times of the AVIRIS 2006 data set (in seconds)

FL	IP3	RLS	C-DPCM-APL	CRLS [#]	C-CRLS [#]	A-CRLS	B-CRLS	B-SuperRLS	CRLS-ABS-APS
113	2382	284	29861	2313	2332	25635	3387	3936	2188

In the terms of space complexity, CRLS-ABS-APS requires about 392 kB memories to cache process variables such as inverse correlation matrix, prediction coefficient and local average for each predictor. That is about 16 times than other similar methods. However, thanks to using residuals image to overwritten original image, CRLS-ABS-APS requires less total memory than other methods.

4. Conclusion

With the wide application of hyperspectral images, the compression for hyperspectral images is becoming more and more important. According to the characteristics of CRLS algorithm, we present an effective lossless compression method, namely, CRLS-ABS-APS. In this method, the adaptive band selection strategy is used to improve the correlation between the reference bands and the prediction band, and the adaptive predictor selection strategy is used improving the similarity of prediction context. In order to further improve the prediction effect, the spectral vector correlation coefficient is adopted as the measure of similarity in k-means clustering algorithm. Additionally, the double snake scan mode is adopted to further improve the stationarity of prediction process, and the recursive local average estimation method is employed to accelerate calculating local average. Experimental results indicate that the CRLS-ABS-APS improves the compression effect with a lower computation complexity effectively, and outperforms to the current state-of-the-art methods.

In future work, we plan to study the parallel computation of the proposed method using graphics processing units (GPUs) to further improve the computation speed.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 61373162), the Research Foundation of the Sichuan Department of Education (Grant No. 15ZB0044), and the Research Foundation of Sichuan Normal University (Grant No. 2015KYQD312).

References

- [1] X. Yu, R. Wang, B. Liu and A. Yu, "Salient feature extraction for hyperspectral image classification," *Remote Sensing Letters*, vol. 10, no. 6, pp. 553-562, 2019. [Article \(CrossRef Link\)](#)
- [2] Y. Liu, G. Gao and Y. Gu, "Tensor matched subspace detector for hyperspectral target detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 4, pp. 1967-1974, April, 2017. [Article \(CrossRef Link\)](#)
- [3] Jet Propulsion Laboratory, NASA Airborne visible infrared imaging spectrometer website. [Online]. Available: <http://aviris.jpl.nasa.gov>
- [4] W. Fu, S. Li, L. Fang and J. A. Benediktsson, "Adaptive spectral-spatial compression of hyperspectral image with sparse representation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 2, pp. 671-682, February, 2017. [Article \(CrossRef Link\)](#)
- [5] R. Nagendran and A. Vasuki, "Hyperspectral image compression using hybrid transform with different wavelet-based transform coding," *International Journal of Wavelets Multiresolution & Information Processing*, vol. 18, no. 1, pp. 481-484, 2020. [Article \(CrossRef Link\)](#)
- [6] S. Álvarez-Cortés, N. Amrani and J. Serra-Sagrìstà, "Low complexity regression wavelet analysis variants for hyperspectral data lossless compression," *International Journal of Remote Sensing*, vol. 39, no. 7, pp. 1971-2000, July, 2018. [Article \(CrossRef Link\)](#)
- [7] B. Huang, A. Ahuja and M.D. Goldberg, "Fast precomputed VQ with optimal bit allocation for lossless compression of ultraspectral sounder data," in *Proc. of Data Compression Conference, 2005. Proceedings. DCC 2005 IEEE*, 2006. [Article \(CrossRef Link\)](#)
- [8] X. Wu and N. Memon, "Context-based adaptive, lossless image coding," *IEEE Transactions on Communications*, vol. 45, no. 4, pp. 437-444, April, 1997. [Article \(CrossRef Link\)](#)

- [9] M. Weinberger, G. Seroussi and G. Sapiro, "The LOCO-I lossless image compression algorithm: Principles and standardization into JPEG-LS," *IEEE Transactions on Image Processing*, vol. 9, no. 8, pp. 1309–1324, August, 2000. [Article \(CrossRef Link\)](#)
- [10] J. Mielikainen, "Lossless compression of hyperspectral images using lookup tables," *IEEE Signal Processing Letters*, vol. 13, no. 3, pp. 157–160, March, 2006. [Article \(CrossRef Link\)](#)
- [11] B. Huang and Y. Sriraja, "Lossless compression of hyperspectral imagery via lookup tables with predictor selection," in *Proc. of SPIE*, vol. 6365, pp. 63650L-1–63650L-8, September 2006. [Article \(CrossRef Link\)](#)
- [12] J. Mielikainen and P. Toivanen, "Lossless compression of hyperspectral images using a quantized index to lookup tables," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 3, pp. 474–478 July, 2008. [Article \(CrossRef Link\)](#)
- [13] J. Mielikainen and P. Toivanen, "Clustered DPCM for the lossless compression of hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 12, pp. 2943–2946, December, 2003. [Article \(CrossRef Link\)](#)
- [14] J. Mielikainen and B. Huang, "Lossless compression of hyperspectral images using clustered linear prediction with adaptive prediction length," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 6, pp. 1118–1121, November, 2012. [Article \(CrossRef Link\)](#)
- [15] J. Wu, W. Kong, J. Mielikainen and B. Huang, "Lossless compression of hyperspectral imagery via clustered differential pulse code modulation with removal of local spectral outliers," *IEEE Signal Processing Letters*, vol. 22, no. 12, pp. 2194–2198, December, 2015. [Article \(CrossRef Link\)](#)
- [16] J. Luo, J. Wu, S. Zhao, et al, "Lossless compression for hyperspectral image using deep recurrent neural networks," *International Journal of Machine Learning & Cybernetics*, vol. 10, pp. 2619–2629, 2019. [Article \(CrossRef Link\)](#)
- [17] B. Carpentieri, A. Castiglione, A. De Santis, et al, "One-pass lossless data hiding and compression of remote sensing data," *Future generation computer systems*, vol. 90, pp. 222-239, Jan, 2019. [Article \(CrossRef Link\)](#)
- [18] M. Klimesh, "Low-complexity adaptive lossless compression of hyperspectral imagery," in *Proc. of SPIE*, vol. 6300, pp. 63000N-1–63000N-9, September, 2006. [Article \(CrossRef Link\)](#)
- [19] D. Bascones, C. Gonzalez and D. Mozos, "FPGA implementation of the CCSDS 1.2.3 standard for real-time hyperspectral lossless compression," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 4, pp. 1158–1165, April, 2018. [Article \(CrossRef Link\)](#)
- [20] H. Shen, Z. Jiang and W. D. Pan, "Efficient lossless compression of multitemporal hyperspectral image data," *Journal of Imaging*, vol. 4, no. 12, pp. 1–15, December, 2018. [Article \(CrossRef Link\)](#)
- [21] C. C. Lin and Y. T. Hwang, "An efficient lossless compression scheme for hyperspectral images using two-stage prediction," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 3, pp. 558–562, July, 2010. [Article \(CrossRef Link\)](#)
- [22] J. Song, Z. Zhang and X. Chen, "Lossless compression of hyperspectral imagery via RLS filter," *Electronics Letters*, vol. 49, no. 16, pp. 992–993, August, 2013. [Article \(CrossRef Link\)](#)
- [23] J. Song, L. Zhou, C. Deng and J. An, "Lossless compression of hyperspectral imagery using a fast adaptive length prediction RLS filter," *Remote Sensing Letters*, vol. 10, no. 4, pp. 401–410, 2019. [Article \(CrossRef Link\)](#)
- [24] F. Gao and S. Guo. "Lossless compression of hyperspectral images using conventional recursive least squares predictor with adaptive prediction bands," *Journal of Applied Remote Sensing*, vol. 10, no. 1, pp. 1–9, February, 2016. [Article \(CrossRef Link\)](#)
- [25] A. C. Karaca and M. K. Güllü, "Lossless hyperspectral image compression using bimodal conventional recursive least-squares," *Remote Sensing Letters*, vol. 9, no. 1, pp. 31–40, 2018. [Article \(CrossRef Link\)](#)

- [26] F. Gao, C. Sun, Q. Shao and S. Guo, “Lossless compression of hyperspectral images using k-means clustering and conventional recursive least-squares predictor,” *Journal of Electronics & Information Technology*, vol.38, no. 11, pp. 2709–2714, 2016. [Article \(CrossRef Link\)](#)
- [27] A. C. Karaca and M. K. Güllü, “Superpixel based conventional recursive least-squares method for lossless compression of hyperspectral images,” *Multidimensional Systems and Signal Processing*, vol. 30, pp. 903–919, 2019. [Article \(CrossRef Link\)](#)
- [28] R. Pizzolante and B. Carpentieri, “Visualization, band ordering and compression of hyperspectral Images,” *Algorithms*, vol. 5, no. 1, pp. 76-97, 2012. [Article \(CrossRef Link\)](#)
- [29] A. S. Mamatha, Vipula Singh and Rajath Kumar M P, “Significance of preprocessing and its impact on lossless hyperspectral image compression,” *The Imaging Science Journal*, vol. 65, no. 5, pp. 270 –281, May, 2017. [Article \(CrossRef Link\)](#)
- [30] M. I. Afjal, M. A. Mamun and M. P. Uddin, “Band reordering heuristics for lossless satellite image compression with 3D-CALIC and CCSDS,” *Journal of Visual Communication and Image Representation*, vol. 59, pp. 514-526, 2019. [Article \(CrossRef Link\)](#)
- [31] C. Li, “Parallel Implementation of the Recursive Least Square for Hyperspectral Image Compression on GPUs,” *KSII Transactions on Internet and Information Systems*, vol. 11, no. 7, pp. 3543- 3557, July. 2017. [Article \(CrossRef Link\)](#)
- [32] J. Li, J. Wu and G. Jeon, “GPU acceleration of clustered DPCM for lossless compression of hyperspectral images,” *IEEE Transactions on Industrial Informatics*, vol. 16, no. 5, pp. 2906–2916, May, 2019. [Article \(CrossRef Link\)](#)
- [33] CCSDS CWE, CCSDS Collaborative work environment website. [Online]. Available: <https://cwe.ccsds.org>



Fuquan Zhu is currently a lecturer in Sichuan Police College, and pursuing a doctoral degree in Chengdu University of Technology, China. He received the M. S. degree from Sichuan Normal University, China, in 2007. His research interests include image processing and data mining.



Huajun Wang is currently a professor at Chengdu University of Technology, China. He received the Ph.D. degree from University of Electronic Science and Technology, China, in 2000. His research interests mainly include sensor networks, network security, image processing, embedded and so on.



Liping Yang is currently a lecturer at Sichuan Police College, China. She received the M. S. degree from Sichuan Normal University, China, in 2007. Her research interests include image processing and data mining.



Changguo Li is currently an associate professor at Sichuan Normal University, China. He received the Ph.D. degree from Chengdu University of Technology, China, in 2015. His research interests mainly include hyperspectral image processing, parallel computing and pattern recognition.



Sen Wang is currently an engineer at Shanghai Shentong Metro Group Co., Ltd. He received the Master Degree of Engineering from China Academy of Telecommunications Technology in 2014. His research interests include signal processing, multimedia security and forensic.