Hyperspectral images repository using a hierarchical compression

Gashnikov M.V.
Samara State Aerospace University
mgash@smr.ru

Glumov N.I.
Samara State Aerospace University
nglu@smr.ru

ABSTRACT

The possibilities of hierarchical compression in hyperspectral images repository are investigated. The image analysis of «SpecTIR» and «AVIRIS» hyperspectrometers is carried out. In order to increase the compression ratio, the spectral bands approximation algorithms are proposed to provide fast access to individual bands. The effectiveness of the developed algorithms is investigated through computational experiments using real 16-bit hyperspectral images.

Keywords

Hierarchical compression, image approximation, hyperspectral image, image repository, maximum deviation, spectral band.

1. INTRODUCTION

An important problem of developing a hyperspectral [Che13a], [Bor04a] images repository is the large number of spectral bands that reaches several hundreds. For example, the results of «SpecTIR» hyperspectrometer shooting [Spe14a] contain more than 300 spectral bands, the results of «AVIRIS» shooting [Avi13a] contain more than 200 bands. The bit rate of each spectral band is 16 bits.

The result is a non-trivial task to store such images in the repository: an extremely large amount of hyperspectral data entails excessive demands on storage capacity and an unacceptably low rate of access to these data. Therefore, when creating a hyperspectral images repository, a compression should be applied.

In this paper, for the compression of hyperspectral images in the repository, we suggest using a modified method, based on a hierarchical grid [Gas10a] interpolation (HGI). This modification includes new algorithms for spectral bands approximation that reduce access time to the individual bands.

In addition, in this paper we present the results of computational experiments to investigate the effectiveness of the proposed algorithms on real 16-bit hyperspectral images.

2. RECENT SOLUTIONS

To store hyperspectral images one of these data formats is commonly used: hdf4, hdf5, lan, img+hdr, dat+hdr [Che13a], [Avi13a], [Spe14a]. These formats are not sufficiently effective for hyperspectral images repository, as do not support data compression.

An important feature of hyperspectral images is that common compression methods [Sal07a] for them are ineffective or even inapplicable. Transform coding techniques (primarily JPEG [Wal91a] and JPEG-2000 [Mar00a]), fractal methods [Sal07a] and video compression methods (including MPEG methods) [Has97a] are computationally too complex and do not allow a strict control of compression error (in particular, the maximum deviation, see below), which must be in the organization of a repository of unique hyperspectral data. Differential methods and 1-D approximation methods have not sufficiently high compression ratio even with compression of one-component images [Sal07a]. 2-D approximation methods based on segmentation [Pra07a] are overly complicated.

In [Gas14a] the HGI-method [Gas10a] was proposed for the compression of hyperspectral images and the results of numerical experiments demonstrating the benefits of HGI over JPEG method in this situation are given.

However, proposed in the above-mentioned article version of HGI method can not be used to store images in hyperspectral images repository, because it does not provide access to the individual spectral bands without decompressing a large part of the compressed image.

3. PROPOSED SOLUTION

In order to formulate the requirements for the compression method in the repository, analysis of the hyperspectral image characteristics was performed in this paper.

3.1. Analysis of hyperspectral images

Analysis of hyperspectral images features was carried out by the following data publicly available:

- five image of spectrometer «SpecTIR» [Spe14a] (360 bands of 16-bit pixels 320x600 size);
- five image of spectrometer «AVIRIS» [Avi13a] (224 bands of 16-bit pixels 614x1086 size).

Examples of spectral bands of these images are shown in Fig. 1.

Statistical characteristics of spectral bands were evaluated for analysis of the hyperspectral images features. Features of bands interdependencies were also evaluated. Some results are shown in the result section in Fig. 4-5. The following conclusions were formulated based on the analysis of these results.

- The difference between the maximum and minimum of the brightness gradations reaches thousands or tens of thousands. These images can not be converted to a "byte".
- Spectral bands are strongly interdependent, as the correlation [Pra07a] between the bands is extremely high.
- Most spectral bands have a high correlation within the band. It can be expected that the use of compression will have a significant effect.

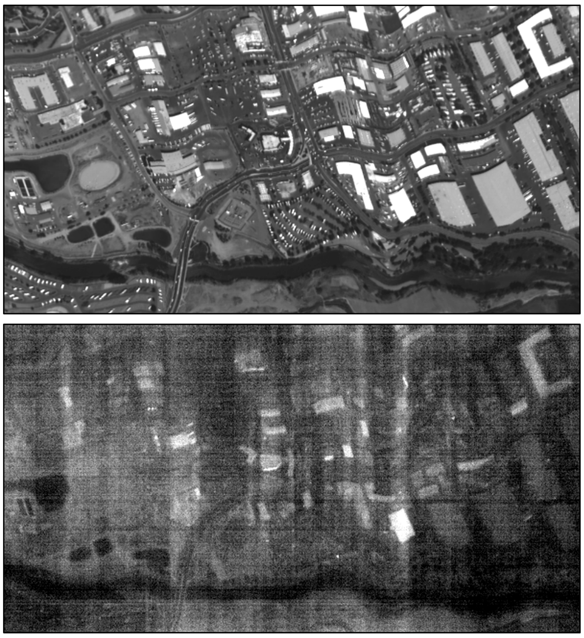

Figure 1. Spectral bands №27, №269 of image «Urban and Mixed Environment» of the spectrometer «SpecTIR»

3.2. Compression based on HGI

The method HGI [Gas10a] is proposed for hyperspectral image compression in [Gas14a]. The method uses the decimated image for interpolation of «less decimated» image. Interpolated pixels are subtracted from original pixels; the calculated differences are encoded to reduce data size.

The method provides control of the maximum deviation. Moreover, access time to an image fragment does not depend on the fragment scale, as the block-hierarchical [Gas10a] format of compressed data used in this method.

3.3. Independent bands compression based on HGI

Independent processing of spectral bands is the easiest way to use the HGI-method to compress hyperspectral images. This approach is very ineffective, since it does not use the correlation between the spectral bands.

3.4. Compression based on a «sliding approximation of bands»

To use the correlations between the spectral bands, the approximation of the spectral bands is proposed for use. Each spectral band is approximated on the basis of other bands that have already been compressed and restored. High correlation between the bands should provide good approximation accuracy. Compression of the difference between the original and approximated spectral bands instead of compression of the original spectral band should significantly improve the compression ratio.

Let X^s be a spectral band number s (matrix of numbers). Let S be a number of spectral bands X^s in the hyperspectral image $\{ X^s, 0 \le s < S \}$.

We will sequentially compress the spectral bands with the numbers 0, 1, 2 .. ($S-1$). When compressing each spectral band X^s we will first calculate approximating band:

$$\hat{X}^s = \sum_{i=0}^{N-1} k_i \overline{X}^{s-i-1}, \; 0 \le s < S \; ,$$

where $\overline{X}^i, i \ge 0$ are previous spectral bands that have already been compressed and decompressed,

N is the number of previous spectral bands that are used for approximation (algorithm parameter), $\{ k_i, 0 \le i < N \}$ are approximation coefficients, which are calculated on the basis of minimizing the mean square error of approximation. At the same time control of the maximum deviation is provided.

Thus, a set of base bands at the approximation is a «sliding window» in the spectral dimension.

 ISBN 978-80-86943-67-1

3.5. Compression based on the «non-overlapping sets of bands»

The above-described compression algorithm based on the «sliding approximation of bands» is not suitable for storage of hyperspectral images in the repository. The reason is the sequential compression of the spectral bands, wherein the previous bands are used for approximation of the next bands. As a result for decompression of any spectral band all previous bands should be decompressed. This is a serious obstruction in providing quick access to images.

In the repository decompression of any band should entail decompression as small as possible number of other bands, which may not be needed. To this end, we propose to use an approximation of the spectral bands within the «non-overlapping sets of bands».

While compressing the entire set of spectral bands is divided into «non-overlapping sets of bands». Each set contains N bands, see Fig. 2. Within each set the described algorithm of the «sliding approximation of bands» is used. Then for decompression of any spectral band we do not need to decompress all previous bands. Only previous bands of the corresponding set must be decompressed.

3.6. Compression based on the «shared base bands»

The compression ratio of the algorithm based on the «non-overlapping sets of bands» is significantly smaller than the compression ratio of the algorithm based on «sliding approximation of bands». The reason is that a much smaller number of base bands used in the algorithm based on the «non-overlapping sets of bands» for approximation.

To solve this problem, the approximation algorithm based on «shared base bands» (see Fig. 3) is proposed in this paper. «Shared base bands» are selected from the complete set of spectral bands with the step N. Also integer parameter C is specified.

Shared bands are compressed by the algorithm based on the «sliding approximation of bands». The number of base bands used for the approximation is equal to $N + C - 2$. Then, after the compression of shared bands, «other» bands are also compressed. Set of the «other» spectral bands is separated into subsets

$$\mathbf{X}_m = \left\{ X^s, mN < s < (m+1)N \right\}, 0 \le m \le S/N ,$$

located between the shared bands.

Each subset $\mathbf{X}_m$ is compressed by the algorithm based on the «sliding approximation of bands». As the base bands for approximation we use the bands of the same subset $\mathbf{X}_m$ and the nearest of the shared bands. The total number of bands used for the approximation is also equal to $N + C - 2$.

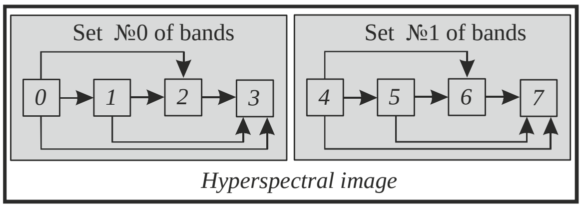

Figure 2. Bands approximation algorithm based on the «non-overlapping sets of bands» (N = 4)

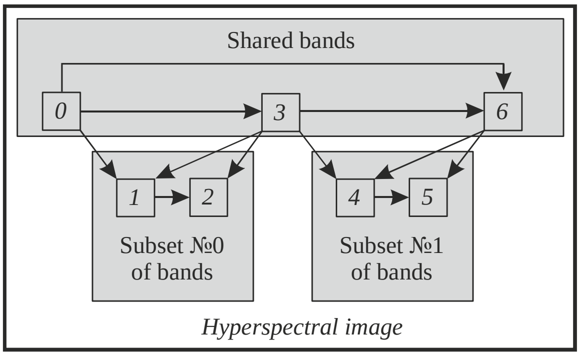

Figure 3. Bands approximation algorithm based on the «shared bands» (N = 3, C = 1)

4. EXPERIMENTAL RESULTS

4.1. Statistical characteristics of hyperspectral images

Statistical characteristics of spectral bands were evaluated for analysis of the hyperspectral images features. Some results are shown in Fig. 4-5.

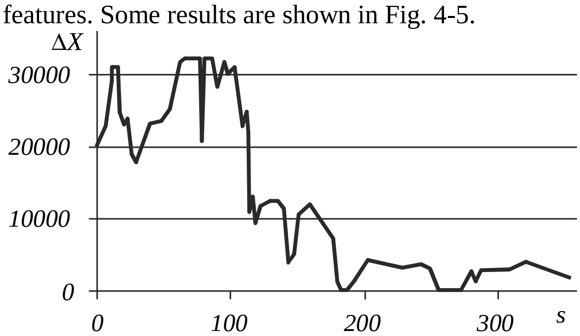

Figure 4. The difference ΔX between the maximum and the minimum depending on the band number s («SpecTIR»)

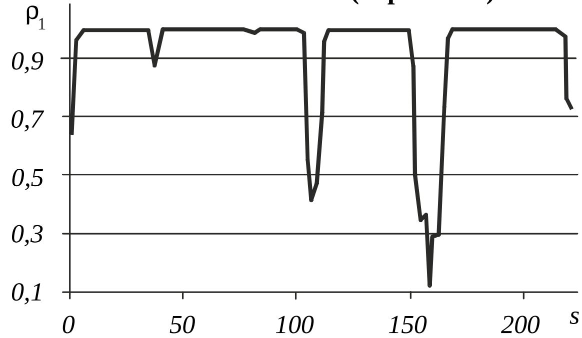

Figure 5. Evaluation of the correlation coefficient ρ_1 between the neighboring bands, depending on the number of band s («AVIRIS»)

 ISBN 978-80-86943-67-1

4.2. Efficiency of the HGI-compression in image repository

The dependency of compression ratio from error was calculated to evaluate the effectiveness of compression (see Fig. 6). Computational experiment was conducted on real 16-bit images of hyperspectrometers «SpecTIR» and «AVIRIS».

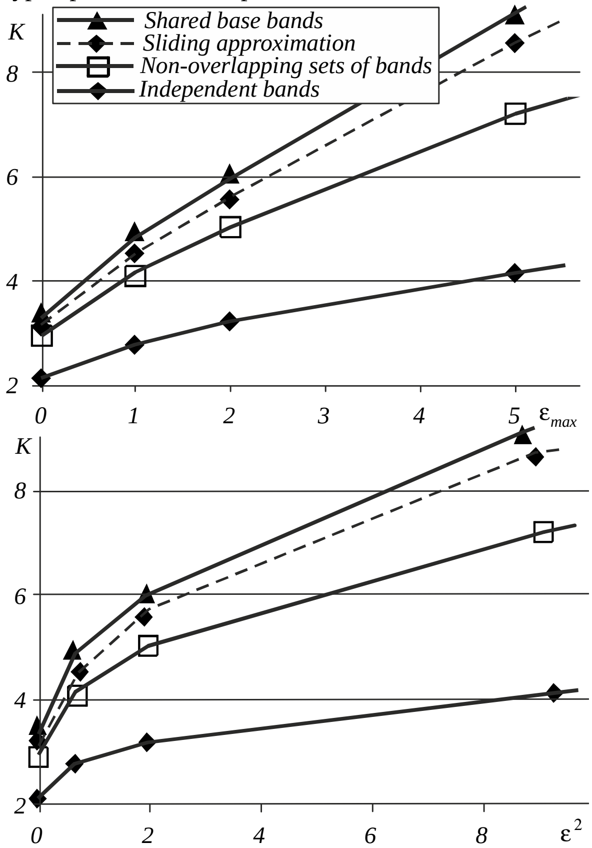

Figure 6. Averaged over five «SpecTIR» images compression ratio *K* depending on the maximum deviation $\varepsilon_{\max}$ and MSE ε^2 (L = 5, N = 7, C = 4, V_b x H_b = 112x614)

The following conclusions were formulated on the basis of the obtained results.

1. All the developed algorithms have a high compression ratio. Therefore, these algorithms can be effectively used in systems of hyperspectral images storage.
2. Algorithms based on the "non-overlapping sets of bands" and "shared bands" provide quick access to compressed data. These algorithms can be effectively used in hyperspectral image repository.
3. The use of any bands approximation can significantly improve the compression ratio.
4. The advantage of using the bands approximation increases with compression ratio increasing.
5. In a small error algorithms based on the "sliding approximation of bands " and "shared bands" demonstrate better results. The algorithm based on "shared bands" is preferred as it provides fast access to data.

5. CONCLUSION

New algorithms for approximate of the spectral bands developed for the compression method based on HGI. These algorithms are well adapted for use in the hyperspectral images repository. The effectiveness of the HGI-compression evaluated on real images of 16-bit hyperspectrometers. The prospects for using the HGI-compression in the hyperspectral images repository demonstrated.

6. ACKNOWLEDGEMENTS

This work was financially supported by the Russian Scientific Foundation (RSF), grant no. 14-31-00014 “Establishment of a Laboratory of Advanced Technology for Earth Remote Sensing”.

7. REFERENCES

[Avi13a] AVIRIS Data – Ordering Free AVIRIS Standard Data Products. Jet Propulsion Laboratory: http://aviris.jpl.nasa.gov/data/free_data.html

[Che13a] Chein-I, C. Hyperspectral Data Processing: Algorithm Design and Analysis. Wiley Press, 2013.

[Bor04a] Borengasser, M. Hyperspectral Remote Sensing – Principles and Applications. CRC Press, 128 p, 2004.

[Gas14a] Gashnikov, M.V., Glumov, N.I. Hierarchical grid interpolation for hyperspectral image compression. Computer optics, V. 38, No.1, pp. 87-93, 2014. (In Russian)

[Gas10a] Gashnikov, M.V., Glumov, N.I., Sergeev, V.V. A hierarchical compression method for space images. Automation and Remote Control, V. 71, No.3, pp. 501-513, 2010.

[Has97a] Haskel, B., Puri, A., Netravaly A., Digital Video An Introduction to MPEG-2. NY: Chapman&Hall, 1997.

[Mar00a] Marcellin, M., Gormish, M., Bilgin A., Boliek, M. An Overview of JPEG-2000. Proceedings of the 2000 IEEE Data Compression Conference, Snowbird, Utah, pp. 523-541, 2000.

[Pra07a] Pratt, W. Digital image processing. Wiley, 4ed, 2007.

[Sal07a] Salomon, D. Data Compression. The Complete Reference. Springer-Verlag, 4ed, 2007.

[Spe14a] SpecTIR Data – Advanced Hyperspectral and Geospatial Solutions. Corporate Headquarters SpecTIR Remote Sensing Division: http://www.spectir.com/free-data-samples.

[Wal91a] Wallace, G. The JPEG Still Picture Compression Standard. Communications of the ACM. V. 34, No. 4, pp. 30-44, 1991.