Signal Processing: *Image Communication* 25 (2010) 697–708

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

journal homepage: www.elsevier.com/locate/image

An efficient macroblock-based diverse and flexible prediction modes selection for hyperspectral images coding

Fan Zhao [a,b], Guizhong Liu [a,*], Xing Wang [a]

[a] *School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China*
[b] *Department of Information Science, Xi'an University of Technology, Xi'an 710048, China*

ARTICLE INFO

Article history:
Received 18 June 2008
Accepted 24 July 2010

Keywords:
Hyperspectral images
Compression coding
H.264/AVC
Prediction mode
Correlation coefficients

ABSTRACT

In this paper, an efficient macroblock-based diverse and flexible prediction modes selection algorithm is proposed for coding hyperspectral images, which is inspired by the prediction scheme of H264/AVC. Here, different modes are specified for the corresponding macroblocks (16×16 pixel regions of a band) of hyperspectral images other than the whole band image using only one reference band image for prediction. Only the 4×4 mode is employed for the intra-band prediction in view of the fact that correlation coefficients of pixels separated by not more than four pixels in the spatial domain are greater than 0.65 at most cases. The optimal reference band is determined by the fast reference band selection algorithm; thereafter, the best partition of the candidate macroblock in the optimal reference band is further selected for inter-band prediction of the current macroblock. Thus, the stronger correlation in the spectral direction or in the spatial domain is utilized for the prediction of the given macroblock. With a comparably low memory requirement, the prediction coding scheme is proposed to speed up the implemental process using the fast reference band selection algorithm, the integer DCT and the quantization, which just needs the multiplication and bit-shifts operations. Several AVIRIS images are used to evaluate the performance of the algorithm. The proposed scheme outperforms the state-of-the-art 3D-based compression algorithms at lower rates. Moreover, compared with the method by using all the prediction modes of H.264/AVC, about 80% encoding time can be saved by our method under the same experimental condition.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Hyperspectral images are represented by an image cube, whose bands are obtained from the same scene at hundreds of continuous spectral bands simultaneously. As the result of the development of remote sensing technology and the increased interest in hyperspectral images in diverse applications, the storage and transmission of the three-dimensional (3D) hyperspectral data sets have become significant, which makes efficient compression of hyperspectral images an active research area.

Both the spatial and spectral correlations exist in the hyperspectral images. How to make best use of these two correlations is the key to an efficient compression algorithm. For their high compression efficiency and excellent embedding feature, several promising lossy compression algorithms based on two- or three-dimensional wavelet transforms have been proposed recently, and the experimental results show that, at the lower compression bit ratios, the better methods are the ones which are based on the three-dimensional wavelet transform. The most classical 3D wavelet video coding algorithm is the 3DSPIHT (3D set partitioning in hierarchical trees) proposed in [1]. It is an

* Corresponding author. Tel./fax: +86 29 82667836.
E-mail addresses: vcu@xaut.edu.cn (F. Zhao), liugz@mail.xjtu.edu.cn (G. Liu), wangxing@stu.xjtu.edu.cn (X. Wang).

0923-5965/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.image.2010.07.003

extension of the original 2DSPIHT [2] and has a 3D symmetric tree structure. Lim et al. [3] adopted the 3DSPIHT algorithm for hyperspectral image compression. The 2DSPECK (2D Set Partitioning Embedded Block) coder was proposed in [4], which utilized the quadtree partitioning and octave band partitioning coding scheme. Tang, Pearlman and Modestino [5, 6] extended and modified it to 3DSPECK for hyperspectral image compression. This algorithm adopted the 3D wavelet transform for decorrelation and the octree partitioning to sort the significance of pixels. Experimental results show that 3DSPECK is not only comparable with 3DSPIHT in compression efficiency but also has lower computational complexity. Both the 3DSPIHT and the 3DSPECK algorithms use a conventional symmetric 3D DWT and they are close in the rate-distortion performance. Because the statistics of the hyperspectral images are not symmetric along the three directions (namely the spatial-horizontal, spatial-vertical, and spectral directions), the coding performance of the symmetric 3D DWT is not optimal. For more efficiently employing the higher correlation in the spectral direction than in the spatial domain for the hyperspectral images, some asymmetric 3D DWT based methods have been proposed recently. An algorithm AT-3DSPIHT (asymmetric tree based 3DSPIHT) was presented in [7] and [8], and another algorithm AT-3DSPECK (asymmetric transform 3DSPECK) was proposed in [9] for hyperspectral images compression. These asymmetric 3D DWT algorithms exhibited better performances than the corresponding symmetric 3D DWT ones.

Part II of JPEG2000 standard [10] makes provision for multi-component image compression, which is a good choice for hyperspectral image compression. Rucker et al. in [11] first presented three JPEG2000 coding strategies for hyperspectral images. A hybrid scheme was proposed in [12]. A principal component analysis (PCA) was deployed in JPEG2000 to provide spectral decorrelation as well as the reduction of the spectral dimensionality in [25]. A low complexity KLT has been integrated into a JPEG 2000 part 2 compliant scheme in [26, 27]. Although the superior decorrelation capabilities are embodied in the KLT- and PCA- based schemes, they have a few drawbacks. First the signal-adaptive transform matrix has to be computed for each of the input vectors, which would require the solution to numerically intensive eigenvector problems. Moreover, the transform matrix has to be transmitted along with the transform coefficients, thus causing a coding overhead. And in the case of high degree of nonstationarity, KLT would become far from optimal.

Rao and Bhargava [13] proposed a scheme with a simple block-based inter-band linear prediction followed by a block-based DCT (discrete cosine transform), which resembles the one used in the MPEG video coding standards. In their scheme, two fixed bands with a lower wavelength and a higher wavelength, respectively, are designated as the reference bands. After each of the reference bands is coded with high fidelity first, the previous adjacent band and the nearest one of the two selected reference bands are used for a bi-directional prediction to the current band. But in case of a large number of bands, the long distance between the current band and the reference band would result in distortion increase in coding the prediction error, especially when the correlation between the current band and its previous adjacent band is also comparably low.

In [22], a much more efficient scan based 3D implementation was proposed for video coding, which could easily employ the so called "sliding window" DWT technique along the spectral direction and thus could easily be deployed within the JPEG2000-MC framework. Although such a sliding-window JPEG2000-MC would be a light requirement on memory and not very computationally complex, a compromise has to be found between the filter length and the number of decomposition levels to determine a better frequency analysis. For an example, for a 3-level temporal wavelet transform, the memory requirement of the scan based DWT system is that of 28 frames for the 5-tap 5/3 filter banks and 48 frames for the 9-tap 9/7 filter banks respectively, and for a 4-level temporal wavelet transform, the memory requirement is that of 49 frames for the 5/3 filter banks and 87 frames for the 9/7 filter banks, respectively. Such a memory requirement like $87 \times 512 \times 614 \times 16$ bits for band-sequential format (BSQ) is nearly equivalent to the memory requirement $199 \times 614 \times 224 \times 16$ bits for band-interleaved-by-line format (BIL) in which the number of the lines is 199, and this would be a heavy constraint on the on-board compression. Note that, for fully making use of the spectral correlation, the number of the bands and the lines demanded for more than three decomposition levels 3D DWT are relatively large, which would undoubtedly constitute heavy burden of computation and storage.

Exploiting the strong correlation, a linear prediction between bands was proposed in [23] to greatly reduce the bit rate required to code images, in which several band orderings are compared and the best reverse ordering is used for all the simulations. However,in the complexity of the reordering there is $O(m^2)$, where m is the number of bands of the dataset. It is not so feasible for onboard compression due to both the memory and processing power limitations.

Considering the fact that the correlations between a current macroblock (a 16×16 pixels region of a band image) and the ones in the corresponding reference band images vary with the position of the macroblock, we advocate a macroblock-based rather than band-based multi-band prediction followed by a macroblock-based integer DCT and CABAC arithmetic coding embodied in [14, 15] for hyperspectral images. Each macroblock of the current band is predicted using the surrounding macroblocks in the current band or using the reference macroblocks at the identical position in the reference bands. One purpose of this paper is to present a fast algorithm for the prediction mode selection, in which the stronger correlation is chosen for prediction in the spectral direction or in the spatial domain. Since the H.264/AVC encoder considers 3 intra-frame modes and 7 inter-frame modes of predictions, the complexity of the corresponding mode selection is very high. To reduce the complexity and to fully utilize the characteristics of the hyperspectral images, we propose to use the one mode 4×4

Download English Version:

https://daneshyari.com/en/article/537310

Download Persian Version:

https://daneshyari.com/article/537310

Daneshyari.com