

Pairwise KLT-Based Compression for Multispectral Images

Yongjian Nian¹ · Yu Liu¹ · Zhen Ye²

Received: 7 September 2015 / Revised: 25 December 2015 / Published online: 21 January 2016
© Springer Science+Business Media New York 2016

Abstract This paper presents a pairwise KLT-based compression algorithm for multispectral images. Although the KLT has been widely employed for spectral decorrelation, its complexity is high if it is performed on the global multispectral images. To solve this problem, this paper presented a pairwise KLT for spectral decorrelation, where KLT is only performed on two bands every time. First, KLT is performed on the first two adjacent bands and two principle components are obtained. Secondly, one remaining band and the principal component (PC) with the larger eigenvalue is selected to perform a KLT on this new couple. This procedure is repeated until the last band is reached. Finally, the optimal truncation technique of post-compression rate-distortion optimization is employed for the rate allocation of all the PCs, followed by embedded block coding with optimized truncation to generate the final bit-stream. Experimental results show that the proposed algorithm outperforms the algorithm based on global KLT. Moreover, the pairwise KLT structure can significantly reduce the complexity compared with a global KLT.

Keywords Multispectral images · Lossy compression · Spectral decorrelation · Pairwise KLT

✉ Yongjian Nian
yjnian@126.com

¹ School of Biomedical Engineering, Third Military Medical University, Chongqing 400038, China

² School of Electronics and Control Engineering, Chang'an University, Xi'an 710064, China

1 Introduction

Multispectral images are of interest for a large number of applications including material identification, environmental monitoring, geological research, and other purposes. Currently, a large number of satellites have carried multispectral imaging sensors to observe the earth, such as IKONOS, GeoEye, SPOT, WorldView and so on. Such multispectral imaging sensors usually capture several bands in the wavelength interval from visible light to near-infrared. With the increasing performances of imaging systems and the wide diffusion of the images acquired, the problem of compactly and faithfully representing such images for efficient transmission and storage becomes ever more urgent. In this situation, the use of an efficient compression techniques is therefore very critical in reducing the volumes of data to be transmitted.

Unlike natural images, multispectral images have both spatial correlation and spectral correlation. The popular transforms for spectral decorrelation of multispectral images mainly includes discrete wavelet transform (DWT) and Karhunen–Loève transform (KLT). However, KLT provides much higher performance than that of DWT, because KLT can more closely match the original images while the wavelet filters are fixed [1]. For this reason, KLT-based algorithm has been more prominently used for lossy compression of multispectral images. Gelli [2] presented an algorithm which segments the multispectral images into several regions and applies KLT to remove the spectral correlation of each region and DCT to remove spatial correlation, experimental results show that this technique outperforms the conventional transform coding technique by 1–2 dB at all rates of interest. Du [3] investigated the impact of using KLT for dimensionality reduction and the performance of JPEG2000 with KLT-based spectral decorrelation in terms of information preservation, experiments indicate that principal component analysis (PCA) outperforms the DWT for spectral decorrelation, the best performance occurs when only a small PCs is coded instead of the entire set. Moreover, the simple linear heuristic that developed in this paper can estimate the optimal number of PCs from both the compression performance as well as information-preservation perspectives. Wang [4] proposed a new transform scheme of multiplierless reversible time-domain lapped transform (RTDLT) and KLT for lossy-to-lossless compression, experiments show that the proposed transform scheme outperforms existed DWT or KLT-based algorithms. Lee [5] proposed a hybrid compression method with pre-encoding discriminant information which employs a feature extraction method to generate the discriminant information, experiments indicate that the proposed method provides better compression efficiency with improved classification accuracy than conventional compression methods. Rucker [6] evaluated three different rate allocation schemes for KLT-based spectral decorrelation followed by JPEG2000 compression, results show that the scheme of simultaneously truncating all codeblock bitstreams from all codeblocks from all principal components significantly outperforms the other schemes considered in terms of not only the compression performance but also the classification accuracy.

Since KLT is optimal only for Gaussian sources, Bitz [7] presented the criterion satisfied by an optimal transform of a JPEG2000 compatible compression

scheme by using KLT, under high resolution quantization hypothesis and without the Gaussianity assumption, tested results show that it obtains a little but significant gain at medium and high bit-rates compared to the conventional KLT; however, the actual drawback of the optimal transforms is their heavy computational complexity. For the parallel compression in which the data set is partitioned to multiple independent processing nodes, Du [8] proposed a segmented approach to KLT to mitigate the detrimental effects of transform-matrix overhead, experimental results show that it outperform the wavelet-based decorrelation that entails no such overhead. Cheng [9] proposed a lossless to lossy compression scheme based on dual-tree Binary Embedded Zerotree Wavelet algorithm, which employs KLT and DWT to achieve 3-D integer reversible hybrid transform and remove spectral and spatial correlation, results show that the achieved lossless compression performance is comparable with the predictive algorithms, while its computational cost is comparable with those of transform-based algorithms. Cagnazzo [10] proposed a region-based transform which employs a reliable segmentation algorithm, the region adaptive KLT to remove spectral decorrelation and shape-adaptive wavelet-based transform to remove spatial correlation, results show that the proposed algorithm outperforms state-of-the-art techniques, such as JPEG-2000 multicomponent or set partitioning in hierarchical trees (SPIHT)-based encoders; however, the complexity of this algorithm is extremely high due to the segmentation processing.

Penna [11] proposed a low-complexity KLT for spectral decorrelation with virtually no performance loss, however, the proposed low-complexity KLT only reflects the estimation of covariance matrix, which carries out on a subset of vectors selected at random. According to the above description, it can be seen that KLT has been widely used in the multispectral images compression, because KLT is much more effective for spectral decorrelation than the other transforms. However, it also has several disadvantages which have been also pointed out in [4, 7, 10, 11]: high computational cost due to a high number of arithmetic operations, large memory requirements due to the requirement of the whole image for transform, and higher implementation difficulty due to the difficulties posed by the iterative numerical algorithm commonly used in the calculation of eigenvalues and eigenvectors. Although the algorithm proposed in [11] can reduce the complexity of KLT for spectral decorrelation, however, this reduction is very limited. In [12], Blanes proposed a novel transform based on the KLT which achieves a better coding performance than the wavelets. In this paper, we investigate in depth the KLT-based approach, and we propose an efficient algorithm based on pairwise KLT, which employs pairwise-KLT for spectral decorrelation and DWT for spatial decorrelation with proper rate allocation, followed by embedded block coding with optimized truncation (EBCOT) coding scheme. Experimental results show that the proposed algorithm obtains an excellent compression performance, moreover, the proposed algorithm can effectively solve the above disadvantages.

The paper is organized as follows: Sect. 2 discusses the disadvantages of KLT for spectral decorrelation; Sect. 3 describes the proposed compression algorithm, including pairwise KLT-based structure and the encoding scheme; Sect. 4 demonstrates the experimental results for several test images and Sect. 5 draws conclusions.

2 Disadvantages of KLT

When using KLT to remove the spectral correlation of multispectral images, it is typically performed on all bands simultaneously. Let $\mathbf{X} = [\mathbf{x}_{1,1}, \mathbf{x}_{1,2}, \dots, \mathbf{x}_{MN}]$ be the multispectral images with height M , width N and total band number L , where

$$\mathbf{x}_{i,j} = [x_{i,j,1}, x_{i,j,2}, \dots, x_{i,j,L}], \quad i = 1, 2, \dots, M; \quad j = 1, 2, \dots, N \quad (1)$$

is the spectral vector in the spatial position of i th row and j th column, and $x_{i,j,l}$ ($l = 1, 2, \dots, L$) is pixel value in the spatial position of i th row, j th column and the l th band. The average estimated covariance matrix can be computed as

$$\mathbf{C}_X = \frac{1}{MN} E[(\mathbf{x}_{i,j} - \mathbf{m}_X)(\mathbf{x}_{i,j} - \mathbf{m}_X)^T] \quad (2)$$

where \mathbf{m}_X is the spectral mean vector and m_i is the mean value of the i th band, which can be written as

$$\mathbf{m}_X = [m_1, m_2, \dots, m_L] \quad (3)$$

The eigenvalues λ_l ($l = 1, 2, \dots, L$) and eigenvectors \mathbf{u}_l ($l = 1, 2, \dots, L$) can be obtained by decomposing the symmetric matrix \mathbf{C}_X , that is $\mathbf{C}_X \mathbf{u}_l = \lambda_l \mathbf{u}_l$. The KLT kernel is a orthogonal matrix $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_L]$, whose columns are the eigenvectors arranged in descending order of eigenvalue magnitude. This matrix is used to transform the zero-mean spectral vector as

$$\mathbf{y}_{i,j} = \mathbf{U}^T (\mathbf{x}_{i,j} - \mathbf{m}_X) \quad (4)$$

Then the outcome after applying the transform is given as $\mathbf{Y} = [\mathbf{y}_{1,1}, \mathbf{y}_{1,2}, \dots, \mathbf{y}_{MN}]$.

According to the above description, the orthogonal matrix \mathbf{U} is obtained by using eigenvalue decomposition of the covariance matrix \mathbf{C}_X . Note that the size of the covariance matrix \mathbf{C}_X is $L \times L$. Jacobi iterative algorithm is widely used for eigenvalue decomposition (ED). The main idea of Jacobi method is to obtain a number of symmetrical squared matrices \mathbf{A}_k by using a number of rotation on \mathbf{C}_X . For each rotation, the two elements a_{ij} and a_{ji} of \mathbf{A}_k may become zero. After several iterations, \mathbf{A}_k will become a diagonal matrix \mathbf{D} , where the elements in the main diagonal is just the eigenvalues of \mathbf{C}_X . The rotation transform equation from \mathbf{A}_k to \mathbf{A}_{k+1} can be written as

$$\begin{cases} \mathbf{A}_{k+1} = \mathbf{T}^T \mathbf{A}_k \mathbf{T} & k = 0, 1, \dots, n \\ \mathbf{A}_0 = \mathbf{C}_X \end{cases} \quad (5)$$

where \mathbf{T} is the orthogonal rotation squared matrix with size of $L \times L$. After n times of iterations, we can obtain the final symmetrical squared matrix as follows

$$\mathbf{D} = \mathbf{T}_1^T \mathbf{T}_2^T \dots \mathbf{T}_n^T \mathbf{A}_0 \mathbf{T}_1 \mathbf{T}_2 \dots \mathbf{T}_n \quad (6)$$

where $\mathbf{T}_1 \mathbf{T}_2 \dots \mathbf{T}_n$ is the orthogonal matrix \mathbf{U} .

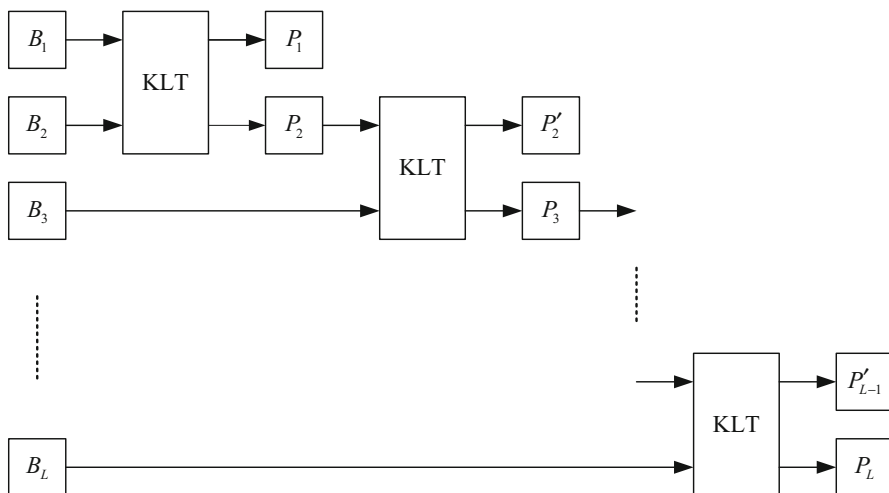
According to the above description, we can see that with the increase of L value, the complexity of ED will significantly increase due to the iterative numerical

algorithm commonly used for eigenvalue decomposition. Another disadvantage is that, because the global images are required simultaneously, the memory requirement is extremely high. Therefore, although KLT can provide higher compression performance than the other transforms because it is trained for each image to provide optimal decorrelation, it also has significant disadvantages, such as extremely high computational cost and memory requirement mentioned above.

3 The Proposed Algorithm

In order to reduce the complexity of KLT, a pairwise KLT structure is proposed for spectral decorrelation of multispectral images. The pairwise KLT perform on only two bands of multispectral images every time, and these operations are organized in a multi-level structure. In each level, bands are decorrelated in pairs and only the resulting component with the larger energy of each pair is further decorrelated in a next level. The pairwise KLT is depicted in Fig. 1. where B_l ($l = 1, 2, \dots, L$) is the l th band of multispectral images.

The first two bands (B_1 and B_2) are transformed by KLT to remove the correlation between them, and then the PC with larger energy will be further decorrelated with the next original band B_3 . This procedure is repeated until the last original band B_L is the final level. It is obviously that the total level for KLT operation is $L - 1$. The final results obtained by the pairwise KLT with $L - 1$ levels are $P_1, P'_2, P'_3, \dots, P_L$. With the proposed structure, most of the energy is accumulated in the P_L since KLT for each two-band is regarded as a classic KLT, grouping most of the energy in one of the two resulting PCs, the PC with larger energy will be decorrelated allows most of the energy of multispectral images to accumulated to the resulting PCs of the last levels.



The pairwise KLT can easily solve the disadvantages of global KLT: firstly, with the use of only the two-component KLTs, the ED procedure is greatly simplified because it does not require any iterative process to obtain the orthogonal matrix U . Suppose the matrix C_X and U are expressed as follows [12]

$$C_X = \begin{bmatrix} a & b \\ b & c \end{bmatrix}, \quad U = \begin{bmatrix} r & s \\ u & v \end{bmatrix}$$

The elements of the matrix U can be simply computed as follows

$$s = \sqrt{(a - c)^2 + 4b^2}$$

$$u = -s = \frac{b}{|b|} \sqrt{0.5 - \frac{a - c}{2s}}$$

$$r = v = \sqrt{0.5 + \frac{a - c}{2s}} = \sqrt{1 - u^2}$$

The eigenvalue value can be computed as follows

$$\lambda_1 = \frac{a + c + s}{2}, \quad \lambda_2 = \frac{a + c - s}{2}$$

The other advantages is that, the pairwise KLT structure significantly reduces the transformed memory requirement because only two bands are required every time. According to the flowchart of the proposed algorithm given by Figs. 1, 2 gives the corresponding hardware implementation of the proposed algorithm.

An efficient spectral transform is usually the first stage of the coding process. Hence, to assess the compression gains, a spectral transform has to be combined with an optimal encoding scheme. Once all the PCs are obtained by the pairwise KLT, the post-compression rate-distortion optimization (PCRD-opt) is employed to realize rate allocation. PCRD-opt is a critical part of the EBCOT algorithm [13], providing excellent coding performance. For the proposed algorithm, spatial DWT

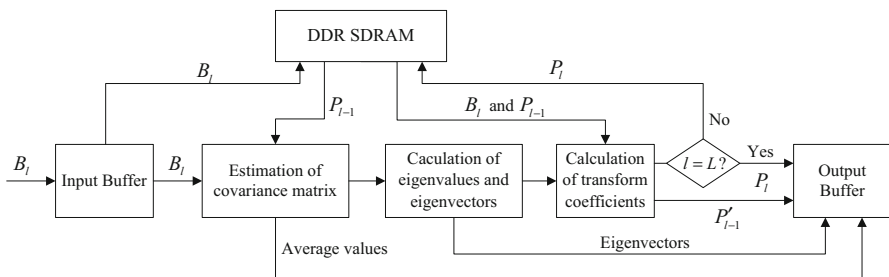


Fig. 2 The hardware implementation of the proposed algorithm

is performed on each PC, and then each PC is partitioned into small blocks (code-blocks), and the PCRD-opt is performed simultaneously across all the code-blocks of all PCs. PCRD-opt first computes the rate-distortion (RD) curve of each code-block by encoding it at high rate, and then deallocate resources progressively, acting each time on the object with the lowest RD slope to keep an approximate equi-slope condition while reaching the desired overall rate. If an optimal truncation point of each code-block is obtained, the truncated code-blocks are concatenated together to generate a single bit-stream.

4 Experimental Results

Several multispectral images acquired by the SPOT-*Vegetation* detector with 1 km spatial resolution are employed for performance evaluation. In particular, four raw images have been used, which are shown in Fig. 3. Note that all the images are represented on two bytes; each image has 1024 lines, 4 bands, and 1024 pixels per line.

4.1 Compression Performance

In this paper, bpp (bit per pixel) and signal to noise ratio (SNR) are used to evaluate the performance of various compression algorithms. We compare the results of the proposed algorithm with JPEG2000(Part 1), DWT-JPEG2000, KLT-JPEG2000 and low-complexity KLT-JPEG2000 [11] under a large range of bit-rates. JPEG2000 is a well-known compression standard that is primarily used for still image compression; its lossy function is used in this paper. DWT-JPEG2000 is also a popular lossy compression algorithm that removes spectral correlation using wavelet transform followed by a spatial wavelet transform with a full post-compression rate-distortion optimization (PCRD-opt). Since multispectral images have a few number of bands, the wavelet basis of haar filter is used for spectral transform, while the wavelet basis of 9/7 tap biorthogonal filter is employed for spatial transform. KLT-JPEG2000 is similar to DWT-JPEG2000, with the only difference lying in spectral decorrelation; the former removes spectral correlation through the use of a KLT transform. At present, KLT-JPEG2000 has become the most popular lossy compression for multispectral images because it provides

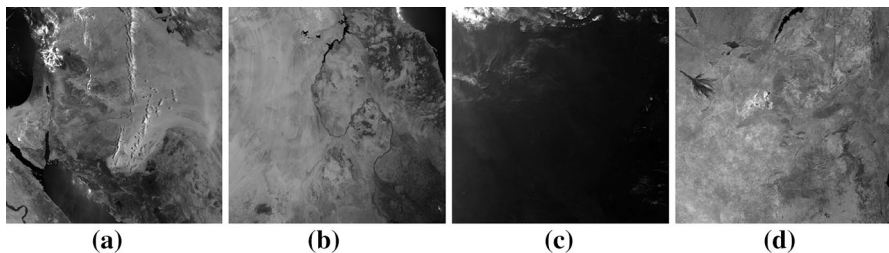


Fig. 3 Tested multispectral images for performance evaluation. **a** Vegetation 1. **b** Vegetation 2. **c** Vegetation 3. **d** Vegetation 4

satisfied compression performance. Low-complexity KLT-JPEG2000 that proposed in [11] is similar with KLT-JPEG2000, which use a low-complexity version of the KLT with virtually no performance loss with respect to the full-complexity transform, where the only different between the low version of KLT and the common KLT is the evaluation of the average correlation matrix. In the following, we use the global DWT, the global KLT and the pairwise KLT to represent the DWT-JPEG2000 and KLT-JPEG2000 and the proposed algorithm respectively. The compression results on Vegetation 1–4 obtained by various algorithms are shown in Fig. 4, where the level of spatial wavelet decomposition is 4 for all algorithms.

For the multispectral images compression, to obtain high rate-distortion performance, it is very important to remove the spectral correlation. As seen from Fig. 4, although JPEG2000 has the perfect rate-distortion performance for still images, its RD performance for multispectral images is the worst because it does not take the spectral correlation into account. The global KLT provides the better RD performance than that of the global DWT, because the global KLT matrix is

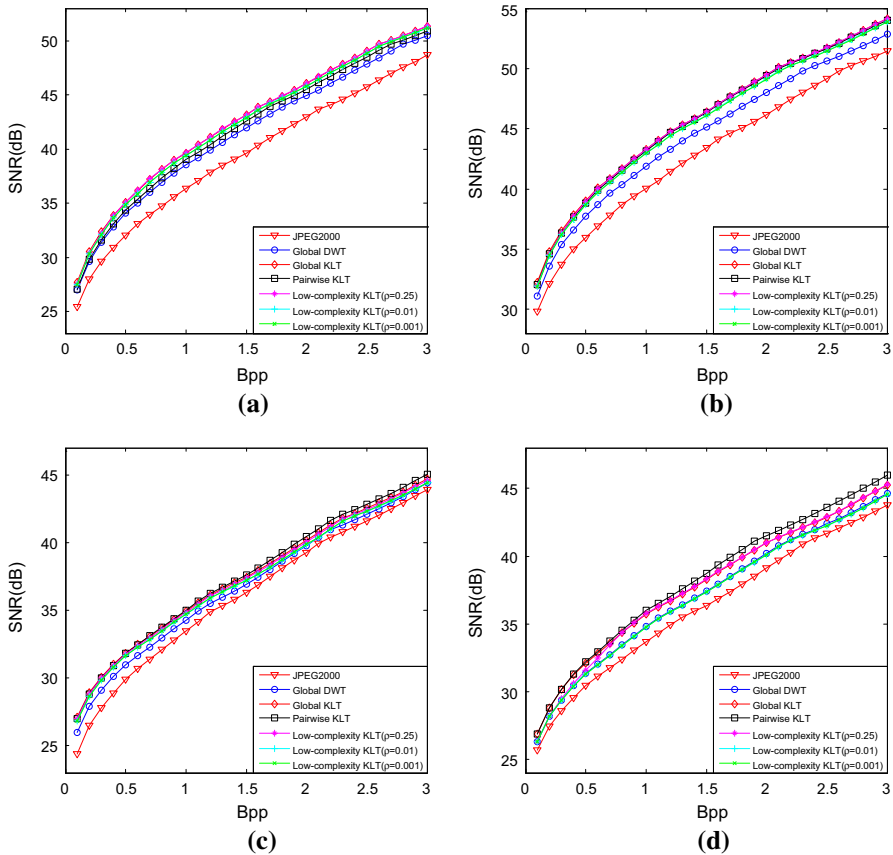


Fig. 4 Comparison of the RD performance of various algorithms. **a** Vegetation 1. **b** Vegetation 2. **c** Vegetation 3. **d** Vegetation 4

computed according to the statistics of the original images while the basic function of DWT is fixed and cannot adapt to the original images. As a consequence, a spectral DWT still leaves a significant degree of correlation in the spectral orientation. In summary, KLT is the optimal linear orthogonal transform and provides higher decorrelation and energy compaction than DWT; thus, the encoder using KLT as a spectral transform provides a performance gain compared with DWT. As for the pairwise KLT, its performance is better than that of the global DWT. For Vegetation 1, the performance achieved by the pairwise KLT is lower than that achieved by the global KLT, however, as can be seen from Fig. 4b–d, the pairwise KLT is competitive with the global KLT, sometimes the pairwise KLT can achieve a performance gain compared with the global KLT. In general, the spectral

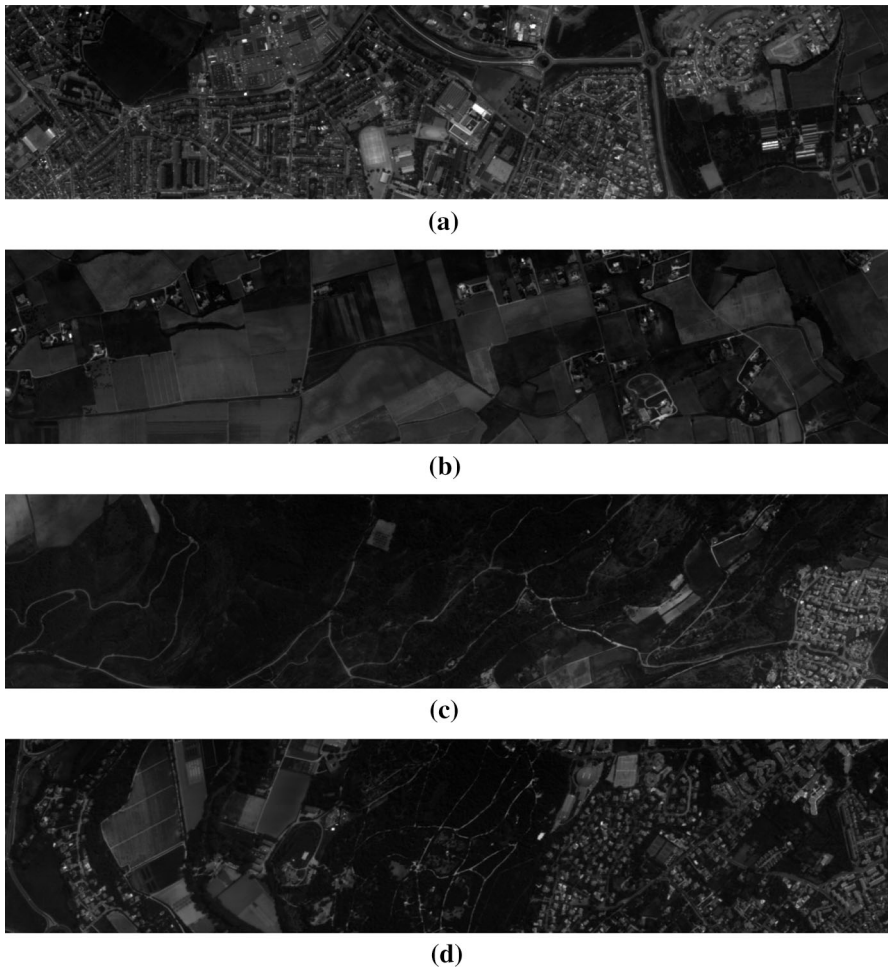


Fig. 5 The tested multispectral images for performance evaluation. **a** Pleiades_perpignan 1. **b** Pleiades_perpignan 2. **c** Pleiades_montpellier 1. **d** Pleiades_montpellier 2

correlation of multispectral images is weak, sometimes KLT is unable to achieve perfect performance when it is performed on the global multispectral images if the spectral correlation is very weak. As for the pairwise KLT, it is performed on two bands every time, which can take advantages of local spectral correlation, this is the reason why the pairwise KLT outperforms the global KLT. However, if the correlation between bands is strong enough, the performance achieved by the pairwise KLT may be lower than that achieved by global KLT. In general, the KLT evaluates the average covariance matrix over all spectral vectors not only for the global KLT but also for the pairwise KLT. As for the low-complexity KLT, KLT evaluates the average covariance matrix on a subset of vectors in order to reduce the complexity of KLT operation. Let $\rho = M'N'/MN$ be the percentage of spectral vectors employed to evaluate the covariance matrix. Obviously, the smaller is ρ , the lower is the complexity of this KLT. Figure 4 gives the compression performance obtained by the low-complexity KLT with $\rho = 0.25, 0.01$ and 0.001 . In fact, when $\rho = 1$, the low-complexity KLT becomes the global KLT. As can be seen from

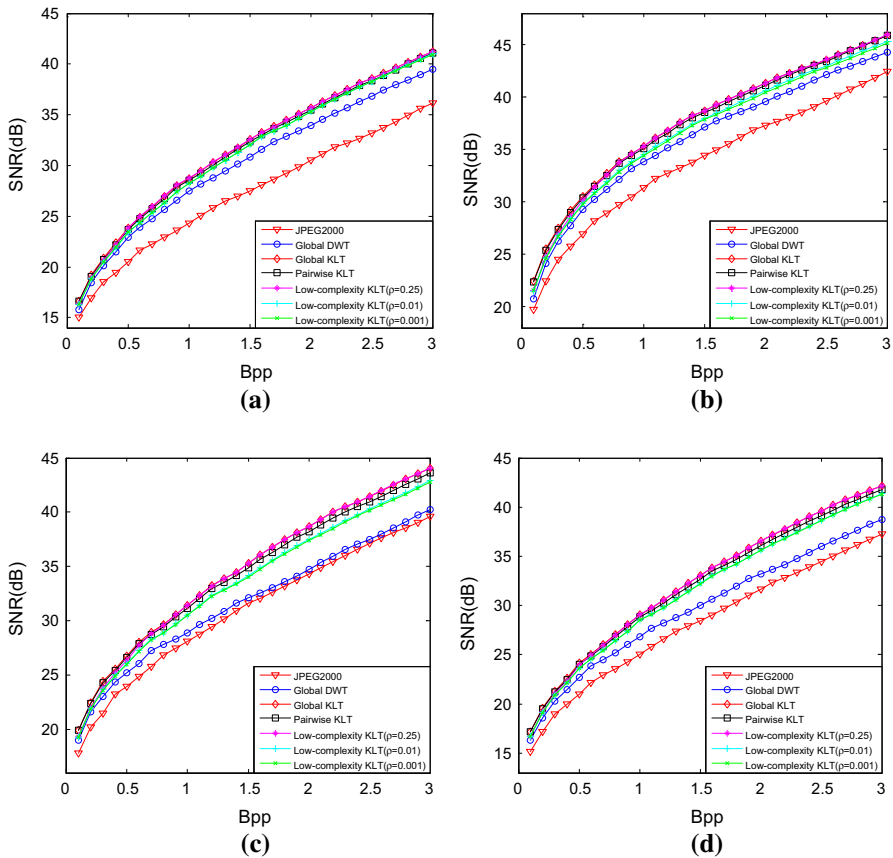
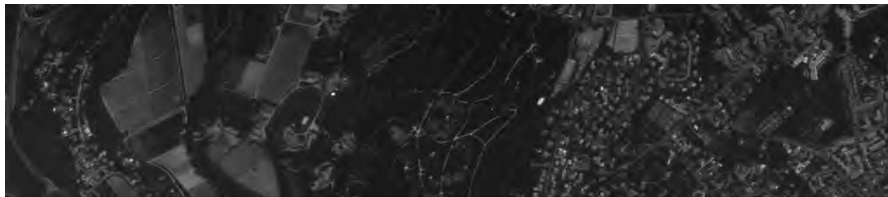
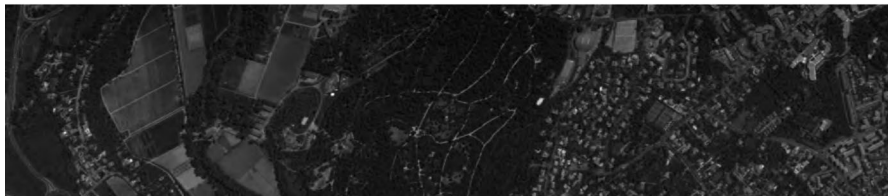


Fig. 6 Comparison of RD performance of various algorithms. **a** Pleiades_perpignan 1. **b** Pleiades_perpignan 2. **c** Pleiades_montpellier 1. **d** Pleiades_montpellier 2

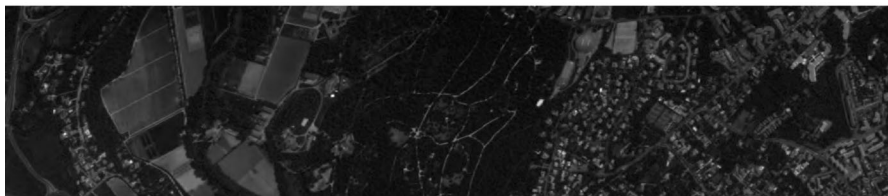
Fig. 4, with de decrease of the ρ value, the precision of the evaluation of KLT correlation matrix decreases, which also decrease the final compression performance. In summary, the pairwise KLT has the competitive performance with the global KLT and the low-complexity KLT.



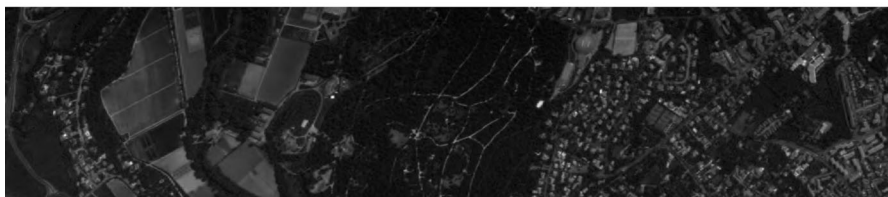
(a)



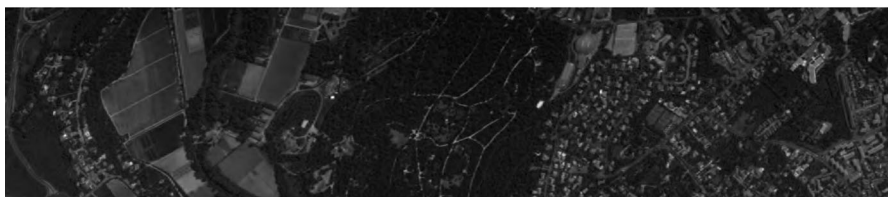
(b)



(c)



(d)



(e)

Fig. 7 Comparison of reconstructed images. **a** JPEG2000. **b** Global DWT. **c** Global KLT. **d** Pairwise KLT. **e** Low-complexity KLT ($\rho = 0.001$)

Table 1 The comparison of complexity of spectral KLT

	The first contribution		The second contribution		The third contribution	
	Addition	Multiplication	Addition	Multiplication	Addition	Multiplication
Global KLT	$MN(L+2)$	MNL^2	$2L^3 - L^2 - L$	$4L^3 - L^2 - 3L$	MNL^2	MNL^2
Low-complexity KLT	$\rho MN(L+2)$	ρMNL^2	$2L^3 - L^2 - L$	$4L^3 - L^2 - 3L$	MNL^2	MNL^2
Pairwise KLT	$MN(L+2)$	$4MN(L-1)$	$10(L-1)$	$22(L-1)$	$4MN(L-1)$	$4MN(L-1)$

We also evaluate the performance of the proposed algorithm on the other four additional multispectral images acquired by Pleiades Satellite. They are named Pleiades_perpignan 1, Pleiades_perpignan 2, Pleiades_montpellier 1 and Pleiades_montpellier 2 respectively, each with 2.8 m spatial resolution, 224 lines, 4 bands, and 1024 pixels per line. Figure 5 gives the one band of the above four images.

Figure 6 depicts the comparison of RD performance obtained by various algorithms. From this figure, we can see that the performance of JPEG2000 is still the worst. As for Pleiades_perpignan 1, 2 and Pleiades_montpellier 2, the global DWT can significantly improve the RD performance while slightly improve the RD performance for Pleiades_montpellier 1. The performance of the pairwise KLT is much higher than that of the global DWT while slightly lower than that of the global KLT, which means that the pairwise KLT has the competitive compression performance with the global KLT. Note that when the ρ value is notably low, the performance achieved by the low-complexity KLT will significantly decreased. That is to say, the low-complexity KLT reduce the complexity of KLT operation by expanding the cost of reducing compression performance.

The comparison of the reconstructed images obtained by the above algorithms under the $\text{bpp} = 0.5$ is shown in Fig. 7. As can be seen, the visual quality of reconstructed image by JPEG2000 is the worst. Global DWT can provide better visual quality of reconstructed image, however, the visual quality obtained by the global DWT is still unsatisfied. There is no obvious visual difference between the reconstructed images obtained by the proposed algorithm and the global KLT. However, with a low ρ value, the visual quality of the reconstructed image achieved by the low-complexity KLT may slightly worse than that achieved by the global KLT and the pairwise KLT. According to the above description, we can see that the comparison of visual quality is coincident with the comparison of RD performance.

4.2 Complexity

Complexity is also an important factor for the evaluation of compression performance. The total times of addition and multiplication are used to measure the complexity. There are three contributions to perform spectral KLT on multispectral images, where the first contribution is the evaluation of the covariance matrix, the second contribution is the solution of the eigenvector problem, and the third contribution is the computation of transform coefficients. Table 1 reports the comparison of complexity of each contribution for global KLT, the low-complexity KLT and the pairwise KLT. As can be seen, the difference between the global KLT and the low-complexity KLT only lies in the first contribution. In fact, the complexity of the second contribution is the largest, which means that the low-complexity cannot significantly reduce the complexity of KLT operation. As for the pairwise KLT, it mainly change the second contribution, which can significantly reduce the overall complexity compared with global KLT when the band number L is larger than 2, especially for the second contribiton. Moreover, the pairwise KLT is easily to be realized for hardware implementation because it dose not require iteration for ED.

Table 2 Comparison of running time of encoder

Algorithm	Encoding time (ms)
JPEG2000	114
Global DWT	120
Global KLT	139
Low-complexity KLT ($\rho = 0.25$)	137
Low-complexity KLT ($\rho = 0.01$)	136
Low-complexity KLT ($\rho = 0.001$)	134
Pairwise KLT	123

In order to quantitatively measure the encoder complexity, we have taken the running time of each band of a software implementation of the encoders, in C language, on windows 7 system. The complexity of the pairwise KLT has mainly been compared with that of JPEG2000, the global DWT, the global KLT and the low-complexity KLT. The results are reported in Table 2. As can be seen, the complexity of JPEG 2000 is the lowest, because it does not perform the spectral decorrelation. The complexity of the global DWT is slightly higher than that of JPEG2000. Note that the global KLT has the highest complexity while providing a significant performance gain. The low-complexity KLT can only slightly reduce the complexity compared with the global KLT even when ρ is low because it only reduces the complexity of the first contribution of KLT operation. The complexity of the pairwise KLT is comparative with that of the global DWT, which is significantly lower than that of the global KLT. Moreover, the pairwise KLT only performs on two bands every time, which can significantly reduce the memory requirements compared with the other algorithms.

5 Conclusions

Multispectral images compression has attracted more and more attention. The KLT is the most common spectral transform, however, it has several disadvantages for spectral decorrelation of multispectral images. In this paper, a pairwise KLT is proposed for spectral decorrelation which is based on the composition of multiple two-band instances of the KLT, which allows benefits in computational cost, memory requirements, and implementation difficulty. Experimental results on several multispectral images show that the proposed algorithm has competitive compression performance with the global KLT-based algorithm. Moreover, the proposed algorithm can significantly reduce the complexity compared with the global KLT-based algorithm, which is easy for hardware implementation.

Acknowledgments This work has been sponsored by a grant from the National Natural Science Foundation of China (No. 41201363). The authors would like to thank Dr Fanqiang Kong for his assist for the revision of this article and those anonymous reviewers for their insightful comments, in improving the quality of this paper. Moreover, The images are CNES Copyright 2007 that are provided by CNES, distributed by SpotImage and produced by VITO.

References

1. Lee, H. S., Younan, N. H., & King, R. L. (2002). Hyperspectral image cube compression combining JPEG-2000 and spectral decorrelation. In *Proceedings of the international geoscience and remote sensing symposium* (Vol. 6, pp. 3317–3319), Toronto, Canada.
2. Gelli, G., & Poggi, G. (1999). Compression of multispectral images by spectral classification and transform coding. *IEEE Transactions on Image Processing*, 8(4), 476–489.
3. Du, Q., & Fowler, J. E. (2007). Hyperspectral image compression using JPEG2000 and principal component analysis. *IEEE Geoscience and Remote Sensing Letters*, 4(2), 201–205.
4. Wang, L., Wu, J., Jiao, L., & Shi, G. (2009). Lossy-to-lossless hyperspectral image compression based on multiplierless reversible integer TDLT/KLT. *IEEE Geoscience and Remote Sensing Letters*, 6(3), 587–591.
5. Lee, C., Youn, S., Jeong, T., Lee, E., & Serra-Sagristà, J. (2015). Hybrid compression of hyperspectral images based on PCA with pre-encoding discriminant information. *IEEE Geoscience and Remote Sensing Letters*, 12(7), 1491–1495.
6. Rucker, J. T., Fowler, J. E., & Younan, N. H. (2005). JPEG2000 coding strategies for hyperspectral data. In *Proceedings of the international geoscience and remote sensing symposium* (pp. 128–131), Seoul, South Korea.
7. Bitá, I. P. A., Barre, M., & Pham, D.-T. (2010). On optimal transforms in lossy compression of multicomponent images with JPEG2000. *Signal Processing*, 90, 759–773.
8. Du, Q., Zhu, W., He, Y., & Fowler, J. E. (2009). Segmented principal component analysis for parallel compression of hyperspectral imagery. *IEEE Geoscience and Remote Sensing Letters*, 6(4), 713–717.
9. Cheng, K.-J., & Dill, J. (2014). Lossless to lossy dual-tree BEZW compression for hyperspectral images. *IEEE Transaction on Geoscience and Remote Sensing*, 52(9), 5765–5770.
10. Cagnazzo, M., Poggi, G., & Verdoliva, L. (2007). Region-based transform coding of multispectral images. *IEEE Transactions on Image Processing*, 16(12), 2916–2926.
11. Penna, B., Tillo, T., Magli, E., et al. (2007). Transform coding techniques for lossy hyperspectral data compression. *IEEE Transactions on Geoscience and Remote Sensing*, 45(5), 1408–1421.
12. Blanes, I., & Serra-Sagrista, J. (2011). Pairwise orthogonal transform for spectral image coding. *IEEE Transactions on Geoscience and Remote Sensing*, 49(3), 961–972.
13. Taubman, D. (2000). High performance scalable image compression with EBCOT. *IEEE Transactions on Image Processing*, 9(7), 1158–1170.