Clustered Reversible-KLT for Progressive Lossy-to-Lossless 3d Image Coding *

Ian Blanes and Joan Serra-Sagristà

Department of Information and Communications Engineering
Universitat Autònoma Barcelona
Cerdanyola del Vallès (Barcelona), Spain
Ian.Blanes@uab.es, Joan.Serra@uab.es

Abstract

The RKLT is a lossless approximation to the KLT, and has been recently employed for progressive lossy-to-lossless coding of hyperspectral images. Both yield very good coding performance results, but at a high computational price. In this paper we investigate two RKLT clustering approaches to lessen the computational complexity problem: a normal clustering approach, which still yields good performance; and a multi-level clustering approach, which has almost no quality penalty as compared to the original RKLT. Analysis of rate-distortion evolution and of lossless compression ratio is provided. The proposed approaches supply additional benefits, such as spectral scalability, and a decrease of the side information needed to invert the transform. Furthermore, since with a clustering approach, SERM factorization coefficients are bounded to a finite range, the proposed methods allow coding of large three dimensional images within JPEG2000.

1 Introduction

This paper deals with three dimensional image coding as commonly employed in two scenarios: Geographic Information Systems (GIS) and medical imaging. Whereas in GIS the third dimension contains spectral information, in medical imaging the nature of the third dimension may be either "spatial" or also spectral. Both of them may require the use of progressive lossy-to-lossless techniques, where the decoded image quality is to be progressively refined up to lossless, for instance, because of legal issues.

A common setup in this situation is to use a spectral decorrelating transform followed by a state-of-the-art encoder. Among the state-of-the-art coders, JPEG2000 [1] is a suitable candidate due to its good performance, progressive lossy-to-lossless capabilities, and standardization. The 3d-TCE technique could also be an adequate encoder [2, 3].

Among the possible decorrelating transforms, the Karhunen-Loêve Transform (KLT) is the one known to produce the best results for most situations [4], since the KLT

*This work was supported in part by the Spanish and Catalan Governments, and by FEDER, under grants TSI2006-14005-C02-01, SGR2005-00319, and FPU2008.

is optimal for centered Gaussian sources. Nevertheless, when the image contains discontinuities or noise, this assumption may not hold true, and another approach should be taken. On the other hand, perfect reversibility is compulsory in progressive lossy-to-lossless coding, but the KLT is not fully reversible because of rounding errors when finite-precision arithmetic is used. To address this shortcoming, approximations of the KLT that map integers to integers have emerged. A common and general approach is the Reversible Karhunen-Loêve Transform (RKLT), based on the reversible integer mapping first suggested in [5], and later developed in [6, 7].

As spectral decorrelating transform, the KLT presents two further problems: the high computational complexity of the transform, and the lack of spectral scalability, since individual spectral components may only be accessed after a complete decoding. In addition, as currently defined within JPEG2000 standard Part 2, the RKLT cannot be used for images with a large amount of components, because of floating point overflow.

These problems have been previously investigated. A parallelization of the full RKLT is presented in [8]. A statistical subsampling to reduce the cost of the covariance matrix creation is used in [9]. And a clustered Multi-Level (ML) approach to improve lossless Compression-Ratio (CR) of MRI images is employed in [10, 11].

Here, we examine the use of clustered RKLT, which helps to decrease the computational complexity, and brings some degree of spectral scalability. Figure 1 reports the coding performance of a clustered RKLT as compared to the performance of a lossy KLT and to an original RKLT for image Jasper Ridge. On very low to moderate bitrates, the clustered RKLT suffers a severe quality penalty. This poor performance might be largely improved with the use of multiple-level clustering, which even yields the best coding performance at very low bitrates because ML decreases the ancillary side information that the transform requires.

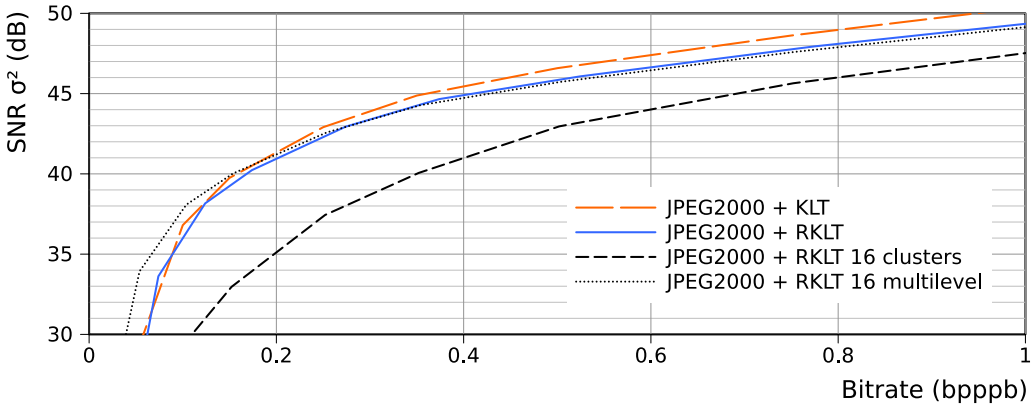


Figure 1. Quality penalty of a clustered RKLT, compared to a RKLT and a multi-level clustered RKLT. Image is Jasper Ridge from the AVIRIS corpus.

The outline of this paper is as follows: the clustered RKLT and the multi-level clustered RKLT are described in Section 2, experimental results are provided in Section 3, and conclusions are drawn in 4.

2 Clustered and Multi-Level Clustered RKLT

2.1 The KLT and RKLT Transforms

The KLT guarantees stochastic independence and optimality for centered Gaussian sources, since, for each set of sources, the optimal transform is built. For simplicity, let X be a matrix that has N rows, one for each source, and one or more samples as columns. The transform is defined as

$$Y = \mathcal{KLT}_{\Phi_X}(X) = Q^T X, \tag{1}$$

where $\Phi_X = \frac{1}{N} X X^T$ is the covariance matrix of X, and Q is the orthonormal matrix obtained from the Singular Value Decomposition (SVD) of $\Phi_X = Q \Lambda Q^{-1}$ ($\Lambda = diag(\lambda_1, \ldots, \lambda_N)$, $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N \geq 0$).

The lossless approach to the KLT is based on the ladder structure [5], which consists of transforming Q^T in a sequence of simple reversible operations. The two main strategies to factor a matrix into reversible operations are the Triangular Elementary Reversible Matrix (TERM) factorization, and the Single-row Elementary Reversible Matrix (SERM) factorization. They are described in detail in [6]. Both are based on decomposing Q^T into a product of Elementary Reversible Matrices (ERMs) and rounding intermediate results. A factorization based in SERMs is one where

$$Q^T = P S_N S_{N-1} \ldots S_0, \tag{2}$$

with P a permutation matrix, and

$$S_i = \begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ s_{j,1} & s_{j,2} & 1 & \ldots & s_{j,N} & \\ & & & 1 & & \\ & & & & 1 & \end{pmatrix}, j = ((i-1) \bmod N) + 1. \tag{3}$$

With the use of the rounding operator ($[\cdot]$), each of the steps can be losslessly applied and reversed in the following way:

$$Y = [S_i X] = (x_1, \ldots, x_i + [s_{i,1} x_1 + \cdots + s_{i,N} x_N], \ldots, x_N)^T \tag{4}$$

$$X = \left[S_i^{-1} Y\right] = (y_1, \ldots, y_i - [s_{i,1} y_1 + \cdots + s_{i,N} y_N], \ldots, y_N)^T. \tag{5}$$

The TERM factorization has this form:

$$Q^T = P L U S_0, \tag{6}$$

where P is again a permutation matrix, L and U are unitary triangular matrices, and S_0 is a SERM step. In this case, triangular matrices are applied and reversed as

follows:

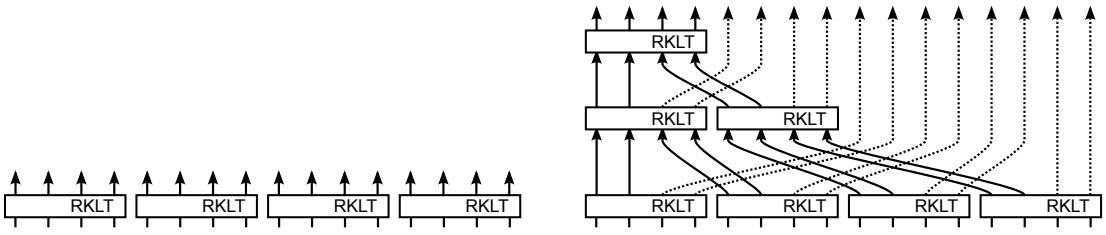
$$Y = [LX] = \begin{pmatrix} x_1 \\ \vdots \\ x_i + \left[\sum_{1\le j<i}^{N} l_{(i,j)} x_j\right] \\ \vdots \\ x_N + \left[\sum_{1\le j<N}^{N} l_{(N,j)} x_j\right] \end{pmatrix}, X = \begin{pmatrix} y_1 \\ \vdots \\ y_i - \left[\sum_{1\le j<i}^{N} l_{(i,j)} x_j\right] \\ \vdots \\ y_N - \left[\sum_{1\le j<N}^{N} l_{(N,j)} x_j\right] \end{pmatrix}. \quad (7)$$

While both factorizations can be applied without each other, the TERM factorization is required in order to create the SERM factorization. The SERM factorization is a bit cheaper to apply, and is standardized in JPEG2000. The problem is that if large matrices are factored, their coefficients soon become large enough to produce sporadic floating point overflows. We have seen this occurring on over 200 components decompositions performed with IEEE 754 single-precision floats. Details on this phenomenon are presented at the end of this section.

In order to bring the performance of the RKLT approximation closer to that of the KLT, several authors have examined distinct pivoting approaches [12, 13, 14, 15]. We have found [12] to be a good trade-off between error reduction and complexity. It is based on the introduction of an additional degree of freedom by turning the S_0 SERM into another TERM. This approach is then based on 3 TERM applications, and will be referenced as 3TERM. This improvement is incompatible with the original SERM factorization.

2.2 Clustering

Figure 2 depicts the two proposed clustering strategies, which are now described.



(a) Normal clustering (b) ML clustering

Figure 2. Clustering strategies

Normal Clustering This strategy consists of virtually splitting an image into multiple images, each with just some of the spectral portion of the original. We define the parameter c as the number of partitions or clusters. Each of the c clusters is spectrally transformed, and all are concatenated again to perform the additional encoding stages. Volumetric Rate-Distortion [16] is used.

Multi-Level Clustering With the aim of improving the energy compaction capability of the previous strategy, multi-level clustering is devised. The scheme we propose is structured hierarchically, applying multiple levels of normal clustering. It iteratively performs the transform on the high energy components of the

distinct clusters. After one pass of normal clustering, half of the components with more energy of each cluster proceed to the next level. The remaining components will be coded without further modification. In the next level, another normal clustering pass is applied, but now with half the number of clusters, until, in the last stage, only one cluster is applied. In this case, we define c as the number of clusters in the first level. Notice that with this strategy, the cluster size remains constrained to the size of the clusters in the first level.

More sophisticated schemes to the multilevel structure may be proposed. Our scheme is inspired by similar schemes in analogous transforms (*e.g.* the tree-structured wavelet transform, or the Hadamard transform in H.264). It remains to be seen if other schemes can still significantly reduce the number of operations performed. As will be seen next, this scheme is, in the worst case, twice as expensive as a Normal Clustering.

2.3 Complexity, Orthonormality, and Coefficient Magnitude Bounds

Each of the transforms described in this paper has been analyzed from the computational complexity point of view, reporting in Table 1 the number of flops for each stage. Costs are parameterized as a function of: n – the number of components, m – the number of spatial locations (width $\times$ height), and c – the number of clusters. Figure 3 shows operation costs for typical values of n, m and c. The ML clustering costs are, in all but the shifting stages, $(2-\frac{1}{c})$ times the cost of a Normal Clustering.

	Method			
Stage	Lossy KLT	RKLT TERM	RKLT SERM	RKLT 3TERM
Shift -$\overline{x}$ (force zero mean)	$2mn$	$2mn$	$2mn$	$2mn$
Covariance Matrix	$m(n^2+n)$	$m(\frac{n^2}{c}+n)$	$m(\frac{n^2}{c}+n)$	$m(\frac{n^2}{c}+n)$
Singular Value Decomposition	$\simeq 9n^3$	$\simeq \frac{9n^3}{c^2}$	$\simeq \frac{9n^3}{c^2}$	$\simeq \frac{9n^3}{c^2}$
TERM Factorization	–	$\frac{n^3}{c^2}-\frac{n^2}{2c}+\frac{n}{2}-c$	$\frac{n^3}{c^2}-\frac{n^2}{2c}+\frac{n}{2}-c$	$\frac{5n^3}{3c^2}+\frac{n^2}{2c}-\frac{37n}{6}+5c$
SERM Factorization	–	–	$\frac{8n^3}{3c^2}-\frac{9n^2}{2c}+\frac{11}{6}n$	–
Apply / Remove (done twice)	$m(2n^2-n)$	$m(\frac{2n^2}{c}+2n-3c)$	$m(\frac{2n^2}{c}+n-c)$	$m(\frac{3n^2}{c}-3c)$
Remove shift	mn	mn	mn	mn

	Total ($\simeq$)
Lossy KLT	$9n^3+5mn^2+2mn$
RKLT TERM	$\frac{10}{c^2}n^3+\frac{10m-1}{2c}n^2+\frac{10m+1}{2}n-c(6m+1)+3mn$
RKLT SERM	$\frac{38}{3c^2}n^3+\frac{(5m-5)}{c}n^2+\frac{(9m+7)}{3}n-c(2m+1)+3mn$
RKLT 3TERM	$\frac{32}{3c^2}n^3+\frac{(14m+1)}{2c}n^2+\frac{(6m-37)}{6}n-c(6m-5)+3mn$
RKLT TERM ML	$(2-\frac{1}{c})(\frac{10}{c^2}n^3+\frac{10m-1}{2c}n^2+\frac{10m+1}{2}n-c(6m+1))+3mn$
RKLT SERM ML	$(2-\frac{1}{c})(\frac{38}{3c^2}n^3+\frac{(5m-5)}{c}n^2+\frac{(9m+7)}{3}n-c(2m+1))+3mn$
RKLT 3TERM ML	$(2-\frac{1}{c})(\frac{32}{3c^2}n^3+\frac{(14m+1)}{2c}n^2+\frac{(6m-37)}{6}n-c(6m-5))+3mn$

Table 1. KLT and RKLT operation costs (in flops). SVD cost from [17], TERM/SERM application cost from [6].

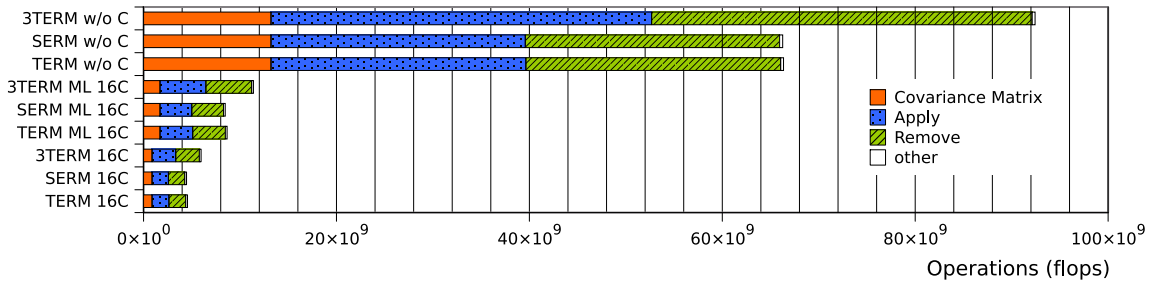


Figure 3. Operation costs for a typical AVIRIS image (z=224; y,x=512)

Energy conservation is also a desired feature, because transforms that distort or do not conserve energy are problematic for Rate-Distortion (RD). The KLT is an orthonormal linear transform, and, according to Parseval's relation, energy conservation can be established. The RKLT is an approximation of the KLT and some degree of energy conservation should be expected. However, several rounding operations are applied before the final result, which may lead to imprecision. Also, in a ML strategy, rounding operations are not "evenly" distributed, therefore, some components are rounded more times than others.

The energy conservation of the KLT approximation is reported in Figure 4 for two remote sensing images and a medical image. Box-plots show the energy gain distribution over all the spatial locations of an image. For most of the considered clustering strategies, energy is conserved. Nonetheless, there are very rare spatial location where up to 3% of energy is gained or lost. It seems that, despite those peak values, energy gain is stable at 1, therefore, no global energy correction needs to be applied. Similar results are obtained for other images in the corpus.

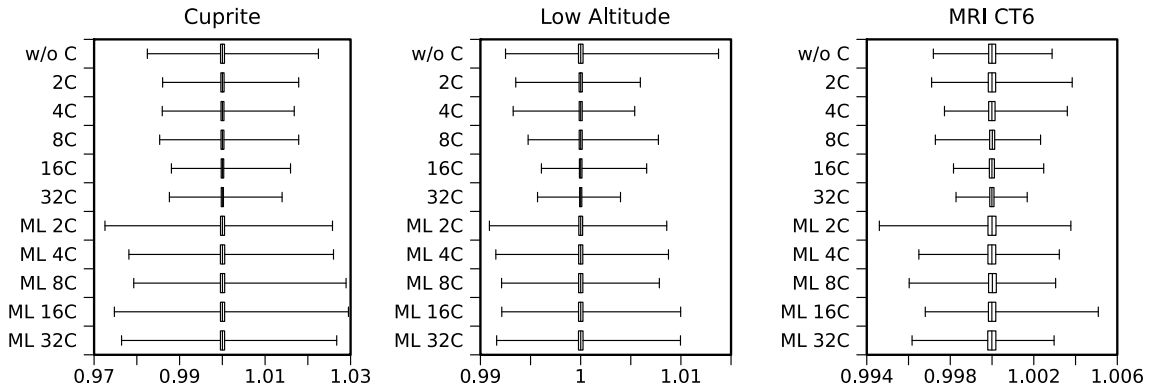


Figure 4. Energy gain for 3TERM factorization

With regard to the possible overflow of the SERM decomposition, the largest coefficient in a SERM factorization seems to grow as the number of transform steps gets larger, as reported in Figure 5. This growth seems exponential and renders the SERM decomposition unusable. From our experience, the SERM decomposition is not feasible for images with $N = 256$ components or larger. For each matrix size N, 2048 random orthonormal matrices are generated. The peak coefficient of the SERM factorization of each matrix is used to build the box plot. The plot depicts the range of the peak coefficients for decomposition of size N. Our findings indicate that the

partial pivoting suggested to reduce the approximation error in [6], actually reduces the peak values of the decomposition.

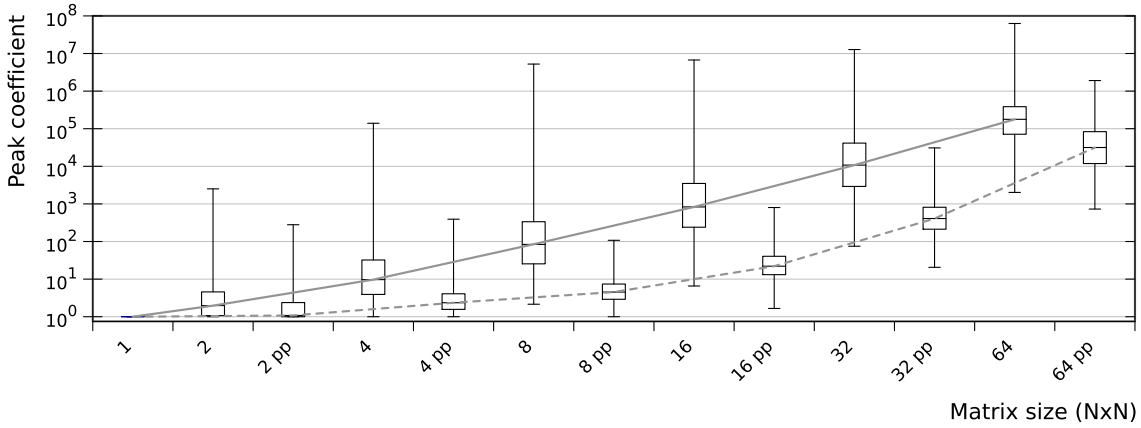
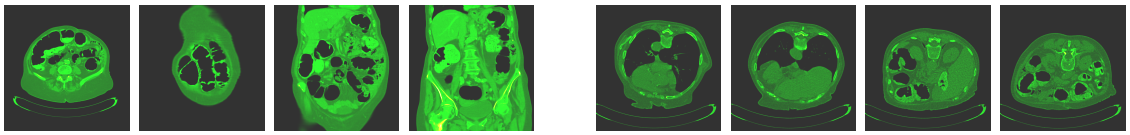


Figure 5. Peak coefficient for SERM factorizations box plot (*pp.* – partial pivoting)

3 Experimental Results

Experiments are performed on the NASA AVIRIS hyperspectral image corpus and on two Computed Tomography (CT) images[1]. All images have signed 16 bpppb samples and a spatial size of 512 × 512. AVIRIS images have 224 spectral components; CT images have 32 and 256 components respectively. Slices of CT images are presented in Figure 6. Due to the nature of CT images, the third dimension is also a "spatial" dimension, and discontinuities are noticeable; see, for instance, CT4.



(a) CT4 (b) CT6

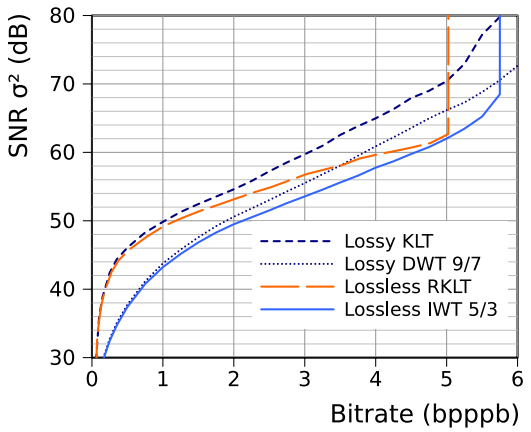
Figure 6. Medical image corpus

In a first experiment, Figure 7a reports the RD performance[2] of two different spectral transforms, namely the wavelet transform and the KLT, for both the lossy and the lossless case, for the image Low Altitude. As expected, the lossy KLT provides the best results. The approximation capability of the plain RKLT can be assessed: for very low to moderate bitrates, RKLT provides very competitive performance; for medium to high bitrates, its performance decreases, although it is the first spectral decorrelating transform to reach the lossless regime. Similar results are obtained for other images in the AVIRIS corpus. However, these results cannot be extended to the CT images. As seen in Figure 7b, the RKLT yields lower coding performance than the Integer Wavelet Transform (IWT) for bitrates higher than 2.5 bpppb. Acording to [20]

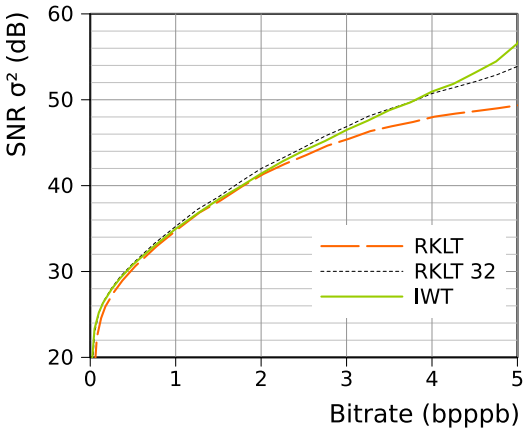
[1]Courtesy of Center for Medical Digital Imaging, UDIAT, Parc Tauli Health Corporation (Spain).

[2]JPEG2000 encodings have been performed with Kakadu Software [18]. For RKLT, clustering, and ML clustering, our own open-source implementation has been used [19].

the IWT already provides better performance than the lossy KLT. The 3TERM version of the RKLT is the one used for these results.



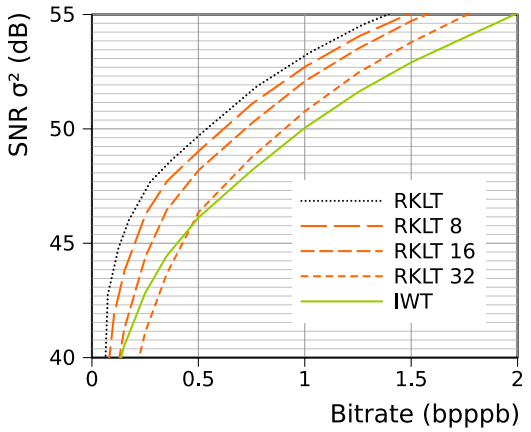
(a) Lossless RKLT and IWT with their lossy equivalents as reference (Low Altitude).



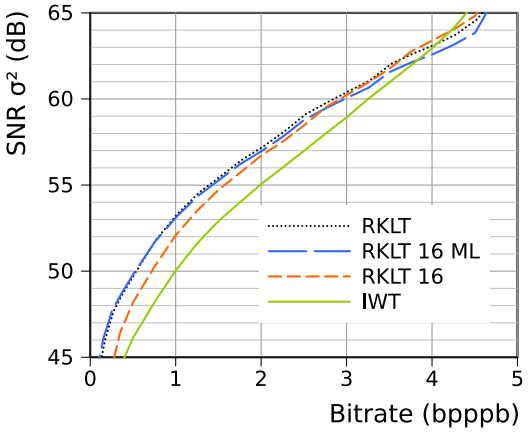
(b) RKLT on the image CT6

Figure 7. RKLT rate-distortion properties.

In progressive lossy-to-lossless coding, it is important to provide good performance while in the lossy range. In a second experiment, in Figure 8a we evaluate the quality penalty introduced by the use of clusters. As was expected, normal clustering trades quality for coding speed, and a moderate quality loss it produced. As the number of clusters increases, the amount of decorrelation is lower, and so is the quality.



(a) Quality degradation produced in a normal clustering



(b) Coding performance of ML RKLT and RKLT

Figure 8. Clustered RKLT. Image is Cuprite from the AVIRIS corpus.

With the introduction of ML clustering, very competitive results are achieved for the AVIRIS corpus, yielding a performance similar to that of RKLT; see Figure 8b. As compared with the IWT (with 5 decomposition levels), both RKLT clustering schemes provide a competitive advantage over IWT (although, above 4 bpppb, IWT outperforms RKLT). ML clustering results for medical imaging are not reported; the IWT provides better results than ML clustering.

	AVIRIS				CT	
	Cuprite	Jasper Ridge	Low Altitude	Lunar Lake	CT4	CT6
none	**7.01**	**7.66**	**7.83**	**6.91**	**4.87**	**5.76**
3d IWT	5.28	5.54	5.95	5.30	5.13	**5.39**
RKLT	**4.85**	**4.84**	5.19	4.96	**5.40**	5.52
RKLT 2	**4.85**	4.85	5.19	**4.95**	5.24	5.48
RKLT 4	4.87	4.89	5.23	4.96	5.12	5.43
RKLT 8	4.95	4.98	5.31	5.02	5.01	5.41
RKLT 16	5.04	5.12	5.45	5.11	4.93	5.40
RKLT 32	5.22	5.35	5.66	5.28	**4.87**	5.42
RKLT ML 2	**4.85**	**4.84**	5.19	**4.95**	5.33	5.51
RKLT ML 4	**4.85**	**4.84**	**5.18**	**4.95**	5.23	5.48
RKLT ML 8	**4.85**	4.85	5.19	**4.95**	5.12	5.45
RKLT ML 16	4.86	4.87	5.21	4.96	5.03	5.44
RKLT ML 32	4.87	4.92	5.23	4.97	**4.87**	5.42

Table 2. Lossless compression ratios achieved with different spectral decorrelating transforms.

Finally, to properly assess the benefits of the proposed clustered and multi-level clustered RKLT, in a third experiment we report their lossless achievement. Table 2 presents results for several spectral decorrelating transform approaches. CR results for AVIRIS images are positive. With normal clustering, competitive results are achieved, always better than those obtained with IWT. ML clustering achieves similar CR performances as an unclustered RKLT. For the CT images, as expected, we have found the RKLT not to be an adequate transform. We suspect that this is due to the discontinuous nature of the image. On a side remark, we note that results with 32 clusters on CT4 are exactly the same as without spectral transform, since the image has only 32 components.

4 Conclusions

Progressive lossy-to-lossless three dimensional image coding is needed in several scenarios, such as remote sensing and medical imaging. The KLT commonly provides the best performance as spectral decorrelating transform for lossy coding, but it cannot be employed for lossless compression. In addition, its use is limited because of two drawbacks: its computational complexity and its lack of spectral scalability. In order to address these issues, we present a clustered Reversible KLT approximation.

Two clustering RKLT strategies have been investigated: normal clustering and multi-level clustering. The two clustering strategies reduce the computational complexity, provide some degree of spectral scalability, and achieve good coding performance, both in the rate-distortion sense and in the lossless compression regime. For lossy coding, the normal clustering provides worst qualities than the classical RKLT, although still superior to the IWT; the ML strategy provides the same quality as the usual RKLT, and, at very low bit rates, its yields even higher coding performance.

Regarding the SERM decomposition, the floating point overflow problem has been studied, suggesting a feasible solution with the ML strategy and a constrained cluster size. With this strategy, the RKLT as defined in JPEG2000 Part 2 can be applied to

images with a large number of components. JPEG2000 Part 2 is flexible enough to allow multiple stages of transform, and transforms of subsets of components; therefore, a multilevel scheme composed of SERMs is a viable solution to code three dimensional images.

References

[1] D. Taubman and M. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards, and Practice*. Kluwer International Series in Engineering and Computer Science, 2002, vol. 642.

[2] J. Yang, Y. Wang, W. Xu, and Q. Dai, "Image coding using dual-tree discrete wavelet transform," *IEEE Transactions on Image Processing*, vol. 17, no. 9, pp. 1555–1569, September 2008.

[3] J. Zhang, J. Fowler, and G. Liu, "Lossy-to-lossless compression of hyperspectral imagery using three-dimensional TCE and an integer KLT," *IEEE Geoscience and Remote Sensing Letters*, to appear.

[4] B. Penna, T. Tillo, E. Magli, and G. Olmo, "Transform coding techniques for lossy hyperspectral data compression," *IEEE Trans. Geoscience Remote Sensing*, vol. 45, no. 5, pp. 1408–1421, May 2007.

[5] F. Bruekers and A. van den Enden, "New networks for perfect inversion and perfect reconstruction," *IEEE Journal on Selected Areas in Communications*, vol. 10, no. 1, pp. 130–137, 1992.

[6] P. Hao and Q. Shi, "Matrix factorizations for reversible integer mapping," *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2314–2324, 2001.

[7] ——, "Reversible integer KLT for progressive-to-lossless compression of multiple component images," *ICIP 2003. Proceedings of 2003 International Conference on Image Processing*, vol. 1, pp. 633–636, 2003.

[8] Y. She, P. Hao, and Y. Paker, "Matrix factorizations for parallel integer transforms," *Parallel Architectures, Algorithms and Networks, 2004. Proceedings. 7th International Symposium on*, pp. 254–259, May 2004.

[9] B. Penna, T. Tillo, E. Magli, and G. Olmo, "A new low complexity KLT for lossy hyperspectral data compression," *IGARSS 2006. IEEE International Conference on Geoscience and Remote Sensing Symposium, 2006.*, pp. 3525–3528, 2006.

[10] Y. Wongsawat, S. Oraintara, and K. R. Rao, "Integer sub-optimal karhunen-loeve transform for multi-channel lossless eeg compression," *European Signal Processing Conference*, 2006.

[11] Y. Wongsawat, "Lossless compression for 3-d mri data using reversible klt," *Audio, Language and Image Processing, 2008. ICALIP 2008. International Conference on*, pp. 1560–1564, July 2008.

[12] L. Galli and S. Salzo, "Lossless hyperspectral compression using KLT," *IGARSS 2004: IEEE International Geoscience and Remote Sensing Symposium Proceedings*, vol. 1-7, pp. 313–316, 2004.

[13] P. van Vugt and M. Adams, "Reversible integer-to-integer wavelet transforms with improved approximation properties," *39th Asilomar Conference on Signals, Systems and Computers*, vol. 1-2, pp. 247–251, 2005.

[14] S. Pei and J. Ding, "Improved reversible integer transform," *Proceedings of 2006 IEEE International Symposium on Circuits and Systems*, vol. 1-11, pp. 1091–1094, 2006.

[15] X. Luo, L. Guo, and Z. Liu, "Lossless compression of hyperspectral imagery using integer principal component transform and 3-D Tarp coder," *SNPD 2007. Proceedings of Eighth Acis International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing*, vol. 1, pp. 553–558, 2007.

[16] J. Serra-Sagristà and F. Aulí-Llinàs, *Computational Intelligence for Remote Sensing*, ser. Studies in Computational Intelligence. Springer, 2008, vol. 133, ch. Remote Sensing Data Compression, pp. 27–61.

[17] G. Golub and C. van Loan, *Matrix Computations (Johns Hopkins Studies in Mathematical Sciences)*. The Johns Hopkins University Press, October 1996.

[18] D. Taubman, "Kakadu software," http://www.kakadusoftware.com/, 2000.

[19] Group on Interactive Coding of Images, "RKLT software," http://gici.uab.cat/, 2008.

[20] R. Leung and D. S. Taubman, "Transform and embedded coding techniques for maximum efficiency and random accessibility in 3-d scalable compression," *IEEE Transactions on Image Processing*, vol. 14, no. 10, pp. 1632–1646, October 2005.