HYBRID KLT-SVD IMAGE COMPRESSION

Patrick Waldemar[1] *Tor A. Ramstad*[1]

[1]Department of Telecommunications,
Norwegian University of Science and Technology (NTNU), Trondheim, Norway

ABSTRACT

This paper investigates a transform adaptation technique, applied to transform coding of images, as a way of exploiting the variation in local statistics within an image. The method makes use of the relationship between KLT and SVD, and their energy compaction properties. We compare this approach to a standard KLT coding system. Motivated by increased coding efficiency an analysis-by-synthesis approach using switching between the KLT coding system and the hybrid KLT-SVD system is proposed. The switching is implemented using a global rate-distortion criterion. The results are encouraging and the proposed techniques provide new insights on how to use SVD in an image compression system.

1. INTRODUCTION

The use of singular value decomposition (SVD) in image compression is motivated by its energy compaction property. The SVD is known to be the deterministically optimal separable transform for energy compaction [1]. This means that for a given image block $\mathbf{X}$ of size $N \times N$, the use of k_1, $k_1 < N$, singular values and $2k_1$ vectors will produce the optimal least squares approximation using separable basis functions in k_1 components of this block. For comparison, the use of k_2, $k_2 = k_1$ Karhunen-Loève coefficients and $2k_2$ vectors to approximate the same block will produce an optimal approximation in the mean square sense, assuming that the Karhunen-Loève vectors (basis functions) are obtained from the horizontal and vertical covariance matrices of the image source. Note that the k_1 singular values are chosen from the main diagonal of the coefficient matrix, where as the k_2 KLT coefficients are chosen from any position in the coefficient matrix, both in order of decreasing magnitude (see Section 2). For each block the least squares approximation will always be better or equal to the mean square approximation. However, in an SVD image compression system the singular vectors are part of the representation which must be quantized for every image block together with the singular values. This will change the energy compaction properties and part of the total bit-rate must be spent on the vectors. The best results for SVD image compression are obtained by combining SVD and vector quantization (VQ) of the singular vectors [2, 3, 4]. The effect of using VQ on the singular vectors results in a block adaptive transform coding system because the transform of each block will approximate the fully adaptive SVD transform. Previous work on adaptive transform coding [5] shows that finding the optimal transform for each block is a complex task. In this work we use the relationship between the Karhunen-Loève transform (KLT) and SVD to define a block adaptive transform with good energy compaction properties.

In Section 2 the fundamentals of SVD and KLT are briefly described and their relationship is discussed. Section 3 describes both the new hybrid KLT-SVD system and the analysis-by-synthesis system used to switch between the KLT and hybrid KLT-SVD coding methods. Specific simulation parameters are given i Section 4 and the corresponding results are presented and discussed in Section 5. Conclusions and proposals for further work are presented in Section 6.

2. FUNDAMENTALS

2.1. Singular Value Decomposition

For a given image block $\mathbf{X}$ of size $N \times N$ the singular vectors u_i and v_i are found as the eigenvectors of $\mathbf{X}\mathbf{X}^T$ and $\mathbf{X}^T\mathbf{X}$, respectively. The singular values s_i are equal to the square root of the nonzero eigenvalues of both $\mathbf{X}\mathbf{X}^T$ and $\mathbf{X}^T\mathbf{X}$. The block $\mathbf{X}$ can then be represented by

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^T, \tag{1}$$

and the result of the transform is given by

$$\mathbf{S} = \mathbf{U}^T\mathbf{X}\mathbf{V}, \tag{2}$$

where

$$\mathbf{U} = \begin{bmatrix} | & | & | & | \\ u_1 & . & . & u_N \\ | & | & | & | \\ | & | & | & | \end{bmatrix}, \mathbf{V} = \begin{bmatrix} | & | & | & | \\ v_1 & . & . & v_N \\ | & | & | & | \\ | & | & | & | \end{bmatrix}. \tag{3}$$

$\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices and $\mathbf{S}$ is a diagonal matrix with the singular values along the main diagonal.

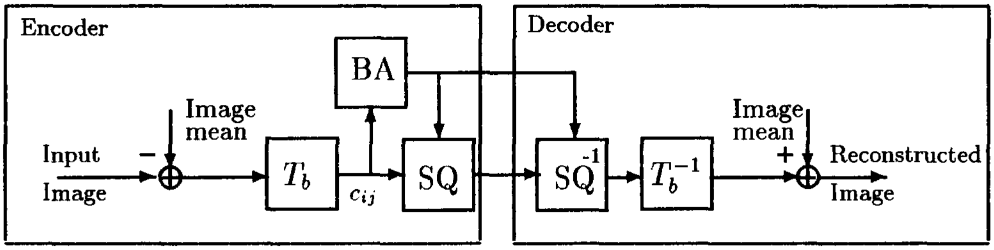

Figure 1. System description

2.2. Karhunen-Loève Transform

A thorough description of KLT is given in [6]. In this section we define the separable KLT for a zero mean image consisting of L image blocks of size $N \times N$, $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \ldots, \mathbf{X}^{(L)}$. The 2-D separable transform of one image block $\mathbf{X}^{(i)}$ is given by

$$\mathbf{C}^{(i)} = \mathbf{\Phi}_v^T \mathbf{X}^{(i)} \mathbf{\Phi}_h, \tag{4}$$

where $\mathbf{\Phi}_v$ and $\mathbf{\Phi}_h$ are given by

$$\mathbf{\Phi}_v = \begin{bmatrix} | & & | & | & | \\ \varphi_{v_1} & \cdot & \cdot & \varphi_{v_N} \\ | & & | & | & | \\ | & & | & | & | \end{bmatrix}, \mathbf{\Phi}_h = \begin{bmatrix} | & & | & | & | \\ \varphi_{h_1} & \cdot & \cdot & \varphi_{h_N} \\ | & & | & | & | \\ | & & | & | & | \end{bmatrix}. \tag{5}$$

φ_{v_i} and φ_{h_i} are found as the eigenvectors of the vertical and horizontal correlation matrices, $\mathbf{R}_{x_v}$ and $\mathbf{R}_{x_h}$, respectively. $\mathbf{R}_{x_v}$ and $\mathbf{R}_{x_h}$ can be estimated from the data matrices $\mathbf{X}_v$ and $\mathbf{X}_h$ by

$$\mathbf{R}_{x_v} = \frac{1}{L}\mathbf{X}_v^T \mathbf{X}_v, \quad \text{and} \quad \mathbf{R}_{x_h} = \frac{1}{L}\mathbf{X}_h^T \mathbf{X}_h. \tag{6}$$

The data matrices are given by

$$\mathbf{X}_h = \begin{bmatrix} \mathbf{X}^{(1)} \\ \cdot \\ \cdot \\ \mathbf{X}^{(L)} \end{bmatrix} \quad \text{and} \quad \mathbf{X}_v = \begin{bmatrix} \left(\mathbf{X}^{(1)}\right)^T \\ \cdot \\ \cdot \\ \left(\mathbf{X}^{(L)}\right)^T \end{bmatrix}. \tag{7}$$

The dimension of the data matrices is $LN \times N$. The coefficient matrices $\mathbf{C}^{(i)}$, $i \in [1, L]$ have in general elements $c_{ij} \neq 0$, and are not diagonal.

2.3. Relationship between KLT and SVD

We note that when the KLT is calculated over one image block $\mathbf{X}$ only, the correlation matrices can be estimated by

$$\mathbf{R}_{x_v} = \mathbf{X}\mathbf{X}^T, \quad \text{and} \quad \mathbf{R}_{x_h} = \mathbf{X}^T\mathbf{X}, \tag{8}$$

and the KLT is equal to the SVD transform for this block. The corresponding coefficient matrix $\mathbf{C}$ is diagonal and equal to $\mathbf{S}$.

From the above we conclude that an important difference between SVD and KLT coefficient matrices, $\mathbf{S}$ and $\mathbf{C}^{(i)}$, for a specific block is that $\mathbf{S}$ is a diagonal matrix and $\mathbf{C}^{(i)}$ is in general not a diagonal matrix. A second difference is the assumption of a global KLT for the whole picture, while the SVD needs a specific transform for each block. For successful compression we need to find efficient bit representations for the SVD transforms. Representing a matrix by its SVD, see Equation (1), keeps the number of representation values constant. The reason for this is the diagonal $\mathbf{S}$ and orthogonal $\mathbf{U}$ and $\mathbf{V}$ matrices. However, previous work [7, 8] show that finding an efficient bit representation of the $\mathbf{U}$ and $\mathbf{V}$ matrices is difficult. Therefore, in this work we propose a model where the first column vectors in the $\mathbf{U}$ and $\mathbf{V}$ matrices are represented using an ordinary quantization technique and the other column vectors are represented using an adaptation of the KLT basis vectors. A closer description of this adaptation is given in the next section. In a KLT compression system we need to find efficient bit representations of only two $N \times N$ correlation matrices for each image to be encoded.

3. SYSTEM DESCRIPTION

3.1. Hybrid KLT-SVD system

KLT is the optimal transform if only one transform is to be used for all blocks in the image, but SVD provides a better energy packing for a specific block. Therefore, if a block adaptive transform is to be used, it makes sense to create a system where approximations of the SVD transform is used. We choose to obtain the approximation of the SVD transform vectors using vector quantization (VQ). The orthogonal property of the transform matrices is preserved during the quantization by using a Gram-Schmidt orthogonalization procedure [9] on the codebooks with regard to previously found vectors. Separate codebooks need to be trained for each u_i and v_i, $i \in [1, N]$. However, to save the cost of using large codebooks for all vectors, it seems reasonable to use the i^{th} KLT vector as the only vector in the codebook for some of the vectors. Hence, the resulting transform matrix may be regarded as a hybrid KLT-SVD transform, where the directions of the KLT basis vectors are changed in order to approximate the direction of the SVD basis vectors.

The coding system is shown in Figure 1. The bit-allocation (BA) is performed using the greedy technique described in [10] (Chapter 8). Uniform scalar quantizers (SQ) assuming a Laplacian pdf are used on the transform coefficients c_{ij} created from the block adaptive transform T_b. The c_{11} coefficients are organized in a low-pass band image and an image adaptive third order DPCM system is used in the quantization of this band.

3.2. System with switching

It is clear that using the hybrid KLT-SVD transform described above has its cost in bit-rate, and for some blocks the KLT transform will perform very well. Therefore, it makes sense to consider a method which switches between pure KLT and hybrid KLT-SVD. Depending on the chosen quantization scheme this switching can be done for each $N \times N$ block or for a range of $N \times N$ blocks. We define the subband domain to contain $N \times N$ subbands such that each transform coefficients c_{ij} is located in subband i, j. The position within a subband is given by the position of the corresponding block in the image domain. Using bit-allocation with block size $B \times B$ in the subband domain makes it natural to perform the switching on $NB \times NB$ blocks in the image domain. In our coder both the KLT and the hybrid KLT-SVD coder must be run before the switching decision is taken, hence the hybrid KLT-SVD coder with switching becomes an analysis-by-synthesis system. The switching is done using a Lagrange technique which minimizes the total image reconstruction distortion under a rate constraint [11, 12]. For each $NB \times NB$ block i, $i \in [1..L]$, in the image domain we are able to find the rate and corresponding distortion of both the KLT and the hybrid KLT-SVD system. We define the rate and corresponding distortion of the two systems to be $R_{i,j}$ and $D_i(R_{i,j})$ for $i \in [1..L]$. $j = 1$ for the KLT system and $j = 2$ for the hybrid KLT-SVD system. The switching is then performed by choosing R_{ij} for each $i \in [1..L]$ so that the following Equation is minimized,

$$\min_{R_{ij}} \left[\sum_{i=1}^{L} D_i(R_{ij}) + \lambda R_{ij} \right], \tag{9}$$

under the constraint

$$\sum_{i=1}^{L} R_{ij} = R_T. \tag{10}$$

In Equations (9) and (10) $j = 1$ or $j = 2$ and R_T is a bit-rate threshold. Realizing that for a fixed λ Equation (9) may be written

$$\sum_{i=1}^{L} \min_{R_{ij}} \left[D_i(R_{ij}) + \lambda R_{ij} \right], \tag{11}$$

we note that a solution can be found by deciding which transform j to use for each $NB \times NB$ block i and then iterate over λ until the constraint in Equation (10) is met.

This method may easily be generalized by using more than two alternatives to switch between. Evidently, the cost of the generalization will be an increase in complexity and side information.

4. SIMULATIONS

The simulations were performed on the luminance component of the image Lenna of size 512×512. The transform block size was 8×8. The hybrid KLT-SVD transform was found by quantizing only the first singular vectors (u_1 and v_1) using a 6 bit vector quantizer, and then creating the other transform $N - 1$ basis vectors by orthogonalization of the KLT basis vector number k with respect to the previously found $k - 1$ vectors, $k \in [2..N]$. Here $N = 8$. A bit-allocation block size of 4×4 was used. Hence, the block size in the coder using switching was 32×32. The quantization in the low pass band was performed on the prediction error using a third order image adaptive DPCM configuration. Quantization in the other subbands were performed on normalized samples with normalization factor equal to the class standard deviation of the corresponding bit-allocation class. Eight classes were used. Uniform quantizers assuming a Laplacian pdf were used in all subbands. The KLT reference coder used image dependent transform matrices found as described in Section 2. The bit-allocation and quantization methods were identical for all three coding systems.

5. RESULTS

The coding results are given in Figure 2. The side information of the bit-allocation, approximately 0.10 bpp, is included in the bit-rate. Comparing the pure

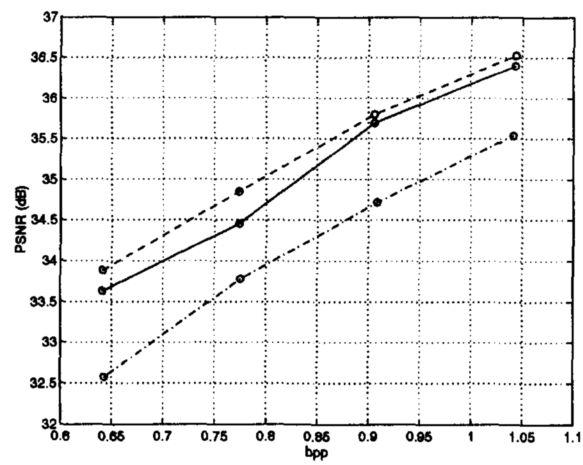

Figure 2. Coding results KLT (solid), Hybrid KLT-SVD (dashdot) and System with switching (dashed)

KLT to the hybrid KLT-SVD scheme we see that KLT gives the best results. However, the quantization of the u_1 and v_1 vectors in the hybrid KLT-SVD scheme has a cost of 0.1875 bpp. If we, for a moment, do not include this bit-rate in the total bit-rate we see that the hybrid KLT-SVD scheme gives an improved result. This confirms the assumption that using an approximate SVD transform may increase the energy compaction compared to using KLT. However, the cost on the total bit-rate from the vector quantization of the singular vectors is high, indicating that an adaptive switch between KLT and hybrid KLT-SVD is needed. Further on, we see that using the previously described system with switching gives an improvement of 0.1 - 0.4 dB compared to the pure KLT system.

When using the switching system it is interesting to find out how many blocks are used from the KLT codec and how many blocks are used from the hybrid KLT-SVD codec. In our simulations the percentage of used hybrid KLT-SVD blocks varied from 3.1 % to 9.8 %. In Figure 3 we show the reconstructed image using the switching system at a bit-rate of 0.77 bpp. Here 25 out of 256 blocks are coded using the hybrid KLT-SVD codec. The blocks using hybrid KLT-SVD are marked with a white border. It is difficult to find a pattern of which method that is used for the different blocks, but the knowledge of SVD indicates that transform blocks within the chosen 32×32 blocks have low rank.

Figure 3. System with switching: bit-rate 0.77 bpp PSNR 34.84 dB, hybrid KLT-SVD blocks are marked

6. CONCLUSIONS

The main contribution of this paper is the use of the relationship between KLT and SVD as a basis for creating a hybrid KLT-SVD transform which exploits the local statistics within an image. The presented results show that better reconstruction is achieved using the hybrid KLT-SVD transform instead of using KLT. However, the cost measured in terms of bits using this transform for every block is quite high. Therefore, further investigations finding a scheme using switching between the KLT and the hybrid KLT-SVD transform in order to enhance the performance and reduce the bit-rate was performed. The proposed scheme outperformed the KLT transform by 0.1 - 0.4 dB. Variable length coding will further enhance the rate-distortion performance.

REFERENCES

[1] A. Albert, "Regression and the moore-penrose pseudoinverse," *New York: Academic*, 1972.

[2] C. M. Goldrick, W. Dowling, and A. Bury, "Image coding using the singular value decomposition and vector quantization," in *Image Processing And Its Applications*, pp. 296–300, IEE, 1995.

[3] T. Saito and T. Komatsu, "Improvement on singular value decomposition vector quantization," *Electronics and Communications in Japan, Part 1*, vol. 73, no. 2, pp. 11–20, 1990.

[4] J.-F. Yang and C.-L. Lu, "Combined techniques of singular value decomposition and vector quantization for image coding," *IEEE Trans. Image Processing*, vol. 4, no. 8, pp. 1141–1145, 1995.

[5] R. D. Dony and S. Haykin, "Optimally adaptive transform coding," *IEEE Trans. Image Processing*, vol. 4, pp. 1358–1370, Oct. 1995.

[6] C. W. Therrin, *Discrete random signals and statistical signal processing.* Prentice-Hall, 1992.

[7] H. C. Andrews and C. L. Patterson, "Singular value decomposition (svd) image coding," *IEEE Trans. Commun.*, vol. 24, pp. 425–432, Apr. 1976.

[8] P. Waldemar and T. A. Ramstad, "Image compression using singular value decomposition with bit allocation and scalar quantization," in *Proc. NORSIG Conf.*, pp. 83–86, 1996.

[9] G. Strang, *Linear Algebra and its Applications.* Harcourt Brace Jovanovich, Inc., third ed., 1988.

[10] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* Boston, MA, USA: Kluwer Academic Publishers, 1992.

[11] Y. Shoham and A. Gersho, "Efficient bit allocation for arbitrary set of quantizers," *IEEE Trans. Acoust., Speech, Signal Processing*, 1988.

[12] P. H. Westerink, J. Biemond, and D. E. Boekee, "An optimal bit allocation algorithm for sub-band coding," in *Proc. Int. Conf. on Acoustics, Speech, and Signal Proc. (ICASSP)*, pp. 757–760, 1988.