ICA-BASED ALGORITHMS APPLIED TO IMAGE CODING

Michel Narozny and Michel Barret

SUPELEC
Signal Processing Systems Team
2, rue E. Belin 57070 Metz - France
E-mail: Firstname.Lastname@supelec.fr

ABSTRACT

Recently, Narozny *et al* [1] proposed a new viewpoint in variable high-rate transform coding. They showed that the problem of finding the optimal 1-D linear block transform for a coding system employing entropy-constrained uniform quantization may be viewed as a modified independent component analysis (ICA) problem. By adopting this new viewpoint, two new ICA-based algorithms, called `GCGsup` and `ICAorth`, were then derived for computing respectively the optimal linear transform and the optimal orthogonal transform. In this paper, we show that the transforms returned by `GCGsup` and `ICAorth` can achieve better visual image quality (better preservation of lines and contours) than the KLT and 2-D Discrete Cosine Transform (DCT) when applied to the compression of well-known grayscale images.

Index Terms— Transform coding, independent component analysis, DCT, KLT, image coding

1. INTRODUCTION

Classically, the best-known results on optimal 1D linear block transforms for transform coding [2] apply only for Gaussian sources. High resolution theory shows that the Karhunen-Loève transform (KLT) is optimal for Gaussian sources [3], and the asymptotic low resolution analysis does likewise [4].

Recently, Narozny *et al* [1] proposed a new viewpoint in variable high-rate transform coding. They showed that the problem of finding the optimal 1-D linear block transform for a high-rate transform coding system employing entropy-constrained uniform quantization may be viewed as a modified independent component analysis (ICA) problem. This result applies without the presumption of Gaussianity or orthogonality. By adopting this new viewpoint, two new modified ICA algorithms, called `GCGsup` and `ICAorth`, were then derived for computing respectively the optimal 1-D linear transform and the optimal 1-D orthogonal transform. Unlike other more intuitive attempts of applying ICA to transform coding [5, 6], the work of Narozny *et al* [1], which relies on a theoretical analysis, revealed one underlying link between ICA and transform coding.

This paper aims at assessing the performances of algorithms `GCGsup` and `ICAorth` when applied to the compression of well-known grayscale images (*Lenna*, *Mandrill*, *Peppers* and *Boat*) each of size 512×512 pixels and coded using 8 bits per pixel (bpp). In order to present a self-contained paper, we begin with a brief review of the main results presented in [1]. The modified ICA bases learned from the set of test images are presented in section 3. In section 4, we show that the new transforms can achieve better visual image quality (better preservation of lines and contours) than the 2-D DCT in medium-to-low bit rate compression. Perspectives of future evolution are presented in section 5.

2. LINK BETWEEN ICA AND TRANSFORM CODING

The transform coding scheme we are interested in applies a linear invertible transform $\mathbf{T}$ to an input vector $\mathbf{X} = (X_1, \dots, X_N)^T$ (where the exponent T denotes transposition) in order to obtain a vector $\mathbf{Y} = (Y_1, \dots, Y_N)^T$ better suited to coding than $\mathbf{X}$. To construct a finite code, each coefficient Y_i is first approximated by a quantized variable $\widehat{Y}_i$. We concentrate on uniform scalar quantizers. The quantized coefficients are then entropy coded. The coded representation is stored or communicated over an error-corrected (lossless) channel. The receiver (decoder) provides an approximation $\widehat{\mathbf{X}} = (\widehat{X}_1, \dots, \widehat{X}_N)^T$ of the original signal $\mathbf{X}$ by applying a linear transform $\mathbf{U}$ to the quantized signal $\widehat{\mathbf{Y}}$. In this paper, we assume $\mathbf{U} = \mathbf{T}^{-1}$. The end-to-end distorsion is measured by the Mean Square Error (MSE): $D = 1/N \sum_{i=1}^{N} \mathrm{E}[(X_i - \widehat{X}_i)^2]$, where E denotes the expected value.

In [1], Narozny *et al* showed that the optimal linear transform for a high-rate transform coding system using entropy-constrained uniform quantization is the transform $\mathbf{T}$ that maximizes the generalized coding gain

$$G^\star = \frac{\left(\prod_{i=1}^{N} c_i^\star\right)^{1/N} \left(\prod_{i=1}^{N} \sigma_i^{\star 2}\right)^{1/N}}{\left(\prod_{i=1}^{N} w_i c_i\right)^{1/N} \left(\prod_{i=1}^{N} \sigma_i^{2}\right)^{1/N}}. \tag{1}$$

where the weight w_i is equal to the square euclidean norm of the ith column of $\mathbf{T}^{-1}$, $\sigma_i^{\star 2}$ (resp. σ_i^2) is the variance of X_i (resp. Y_i), and $c_i^\star$ (resp. c_i) is the constant associated with the variable $\widetilde{X}_i = (X_i - \mathrm{E}[X_i])/\sigma_i^\star$ (resp. $\widetilde{Y}_i = (Y_i - \mathrm{E}[Y_i])/\sigma_i$) according to the relation $c_i^\star = \frac{2^{2h(\widetilde{X}_i)}}{12}$ (resp. $c_i = \frac{2^{2h(\widetilde{Y}_i)}}{12}$), where $h(\widetilde{X}_i)$ (resp. $h(\widetilde{Y}_i)$) is the differential entropy of $\widetilde{X}_i$ (resp. $\widetilde{Y}_i$). The generalized coding gain is the factor by which the distorsion is reduced because of the linear transform $\mathbf{T}$, assuming high rate and optimal bit allocation.

In [1], the authors also showed that maximizing the generalized coding gain is equivalent to minimizing the criterion:

$$C(\mathbf{T}) = I(Y_1; \ldots; Y_N) + \frac{1}{2}\log_2 \frac{\det \mathrm{Diag}[\mathbf{T}^{-T}\mathbf{T}^{-1}]}{\det \mathbf{T}^{-T}\mathbf{T}^{-1}}, \quad (2)$$

where $I(Y_1; \ldots; Y_N)$ is the mutual information between the random variables $Y_1, \ldots, Y_N$, and for any square matrix $\mathbf{C}$, $\mathrm{Diag}(\mathbf{C})$ denotes the diagonal matrix having the same main diagonal as $\mathbf{C}$. The criterion (2) may be decomposed into the sum of two terms: $C(\mathbf{T}) = C_{\mathrm{ICA}}(\mathbf{T}) + C_O(\mathbf{T})$. The first term $C_{\mathrm{ICA}}(\mathbf{T}) = I(Y_1; \ldots; Y_N)$ corresponds to the mutual information, which is a classical criterion in ICA. The second term $C_O(\mathbf{T})$ measures a pseudo-distance to orthogonality of the transform $\mathbf{T}$; it is non negative, and zero if and only if the column vectors of $\mathbf{T}^{-1}$ are orthogonal. Two algorithms for the minimization of the criterion (2) were proposed in [1]. The first one, called `GCGsup` for Generalized Coding Gain Supremum, consists of a modified version of the mutual information based ICA algorithm by Pham [7] called `ICAinf`. In the second new algorithm, called `ICAorth` for Independent Component Analysis Orthogonal, the algorithm `ICAinf` was modified in order to compute the optimal orthogonal matrix that minimizes $C(\mathbf{T})$.

3. BASES ESTIMATION

3.1. Learning schemes

The modified ICA bases (i.e., the columns of $\mathbf{T}$) were estimated according to two learning schemes. The first one yields 8 different bases (2 per image): for each test image, the algorithms `GCGsup` and `ICAorth` were applied to a training set consisting of 4 096 non overlapping image blocks each of size 8×8 pixels extracted from the test image. Moreover, the KLT is estimated on the same training set. The first column of Fig. 1 displays the estimated modified ICA bases as well as the practically achieved KLT bases obtained from the test image *Peppers*. As for the second scheme, it yields only 2 different bases. The modified ICA bases were learned from one training set consisting of 12 288 non overlapping image blocks each of size 8×8 pixels extracted from three test images (*Lenna*, *Peppers* and *Boat*), again the KLT basis was learned on the same training set. The second row of Fig. 1 displays the estimated modified ICA bases (denoted $\mathbf{T}^\star_{\mathrm{orth}}$ and $\mathbf{T}^\star_{\mathrm{opt}}$) as well as the KLT basis (denoted KLT*).

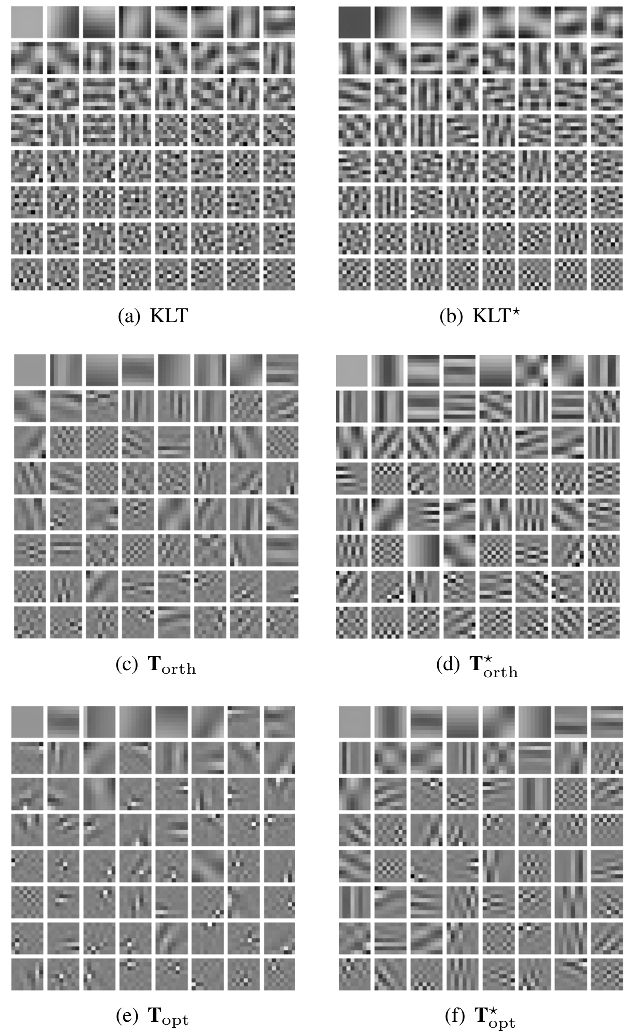


Fig. 1. KLT, $\mathbf{T}_{\mathrm{orth}}$ and $\mathbf{T}_{\mathrm{opt}}$ basis vectors obtained from *Peppers* (on first column) and KLT*, $\mathbf{T}^\star_{\mathrm{orth}}$ and $\mathbf{T}^\star_{\mathrm{opt}}$ basis vectors obtained from the image training set (on second column).

3.2. Bases properties

Examining Fig. 1 closely reveals the features found with ICA modified algorithms are much more localized in space than the checkerboardlike basis vectors obtained with the KLT. Notice also the more pronounced edge-like nature of the modified ICA bases, regardless of the learning scheme employed.

The computation of $C_O(\mathbf{T})$, i.e., , the pseudo-distance to orthogonality, reveals that the bases estimated with `GCGsup` are quasi-orthogonal as can be seen in Tab. 1.

	Lenna	*Mandrill*	*Peppers*	*Boat*
$\mathbf{T}_{\mathrm{opt}}$	0.009	0.008	0.016	0.012
$\mathbf{T}^\star_{\mathrm{opt}}$	0.006	0.006	0.006	0.006

Table 1. Pseudo-distance to orthogonality (in bpp) of $\mathbf{T}_{\mathrm{opt}}$ and $\mathbf{T}^\star_{\mathrm{opt}}$ for each test image.

Tab. 2 shows estimations of the generalized coding gain for each tested transform and each test image. The average generalized coding gain of each transform computed over the set of test images is also given. The estimation method used is the same as that described in [1]. Looking at the average values of the generalized coding gain reveals that, whatever the learning scheme, the modified ICA transforms perform best followed respectively by the 2-D DCT and the KLT. The coding gain of any of the modified ICA transforms relative to the 2-D DCT is about 0.33 dB (resp. 0.12 dB) when the first (resp. second) learning scheme is used suggesting that a transform-based image coder could benefit from using any of the modified ICA transforms.

	Lenna	*Mandrill*	*Peppers*	*Boat*	Average
KLT	18.13	6.99	17.23	15.35	14.42
$\mathbf{T}_{\mathrm{orth}}$	18.58	7.42	17.89	15.96	14.96
$\mathbf{T}_{\mathrm{opt}}$	18.60	7.49	17.84	15.94	14.96
KLT*	17.76	6.98	17.08	14.79	14.15
$\mathbf{T}^{\star}_{\mathrm{orth}}$	18.34	7.27	17.63	15.74	14.74
$\mathbf{T}^{\star}_{\mathrm{opt}}$	18.31	7.28	17.70	15.72	14.75
2-D DCT	18.25	7.17	17.50	15.61	14,63

Table 2. Generalized coding gain (in dB) of the KLT, $\mathbf{T}_{\mathrm{orth}}$, $\mathbf{T}_{\mathrm{opt}}$, KLT*, $\mathbf{T}^{\star}_{\mathrm{orth}}$, $\mathbf{T}^{\star}_{\mathrm{opt}}$, and 2-D DCT for each test image. Last column yields the average generalized coding gain of each transform computed over the set of test images.

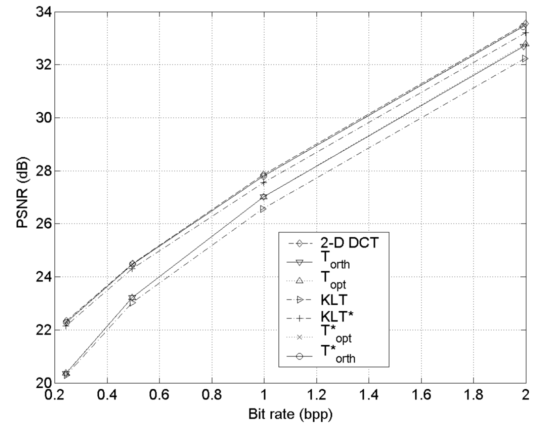
4. MEDIUM-TO-LOW BIT RATES COMPRESSION

4.1. Context

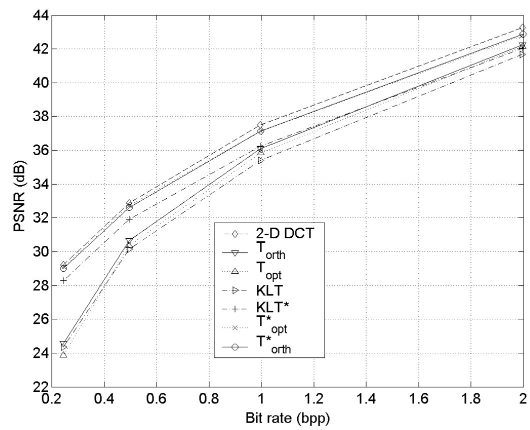
The image coder used in our experiment is a transform coder originally developed by Davis[1]. It has been designed for experimentation and is not intended to outperform current state-of-the-art image coders. It is very modular and allows for simple replacements of individual components (quantizer, entropy coder, transform). It was modified so that it ressembles a JPEG-like coder. The bases obtained using the first learning scheme are transmitted with the image since they are data-dependent bases. As for the bases estimated using the second learning scheme, they are *not* transmitted with the image. Quantization steps are chosen to minimize the end-to-end distorsion subject to bit rate constraint. The bit allocation procedure is based on integer programming algorithms described in [8] which provide optimal or near-optimal allocations for the scalar uniform quantizers included here. Entropy coding of the quantizer output is carried out by an adaptive arithmetic coder.

4.2. Results

We now compare the compression performances of the following transforms: KLT, $\mathbf{T}_{\mathrm{opt}}$, $\mathbf{T}_{\mathrm{orth}}$, KLT*, $\mathbf{T}^{\star}_{\mathrm{opt}}$, $\mathbf{T}^{\star}_{\mathrm{orth}}$ and 2-D DCT. The results for *Mandrill* and *Boat* are displayed in Fig. 2, in which we present the peak signal to noise ratio (PSNR) as a function of bit-rate. Whatever the image, the plots have these common characteristics.



(a) *Mandrill*



(b) *Boat*

Fig. 2. PSNR vs bit-rate for the images *Mandrill* and *Boat*.

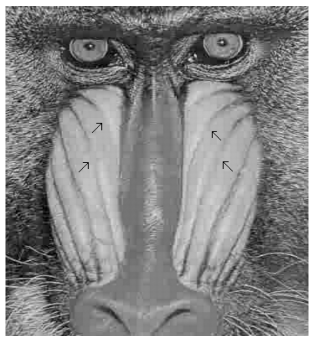
1) The transform codes based on KLT*, $\mathbf{T}^{\star}_{\mathrm{opt}}$, $\mathbf{T}^{\star}_{\mathrm{orth}}$ and the 2-D DCT perform better than those based on the KLT, $\mathbf{T}_{\mathrm{opt}}$ and $\mathbf{T}_{\mathrm{orth}}$, regardless of the bit-rate. The poor coding performances of the KLT, $\mathbf{T}_{\mathrm{opt}}$ and $\mathbf{T}_{\mathrm{orth}}$ are mainly due to the coding penalty resulting from coding the basis vectors (11 bits were allocated on average to each matrix coefficient resulting in a coding precision of 10^{-3}).

2) We observed that the high-resolution hypothesis is verified for bit-rates greater than about 1.6 bpp (in this case, the slope of the curve is about 6 dB per bits).

3) No meaningful performance difference can be observed between the 2-D DCT and the class-adapted transform codes based on $\mathbf{T}^{\star}_{\mathrm{opt}}$ and $\mathbf{T}^{\star}_{\mathrm{orth}}$. Thus, our approach has made it possible to learn two bases which are competitive with the well-known 2-D DCT basis according to the standard PSNR measure.

[1] http://www.geoffdavis.net/dartmouth/wavelet/wavelet.html

(a) 2-D DCT (0.495 bpp; 32.88 dB) (b) $\mathbf{T}^{\star}_{\mathtt{orth}}$ (0.495 bpp; 32.61 dB)



(c) 2-D DCT (0.496 bpp; 24.48 dB) (d) $\mathbf{T}^{\star}_{\mathtt{orth}}$ (0.496 bpp; 24.51 dB)

Fig. 3. Images *Boat* and *Mandrill* coded at about 0.5 bpp. Black arrows point towards details which are not present or blurred on the corresponding image coded with the 2-D DCT.

Visual inspection of the image quality was also carried out. Fig. 3(d) and 3(b) show respectively the images *Mandrill* and *Boat* coded with $\mathbf{T}^{\star}_{\mathtt{orth}}$. Black arrows point towards details which are not present or blurred on the corresponding images coded with the 2-D DCT (best seen enlarged on a computer screen). These details represent lines (e.g., some ropes in the case of the image *Boat*) which are well preserved with $\mathbf{T}^{\star}_{\mathtt{orth}}$. These results suggest that the class-adapted modified ICA bases are better suited to coding fine details such as lines and edges compared to the 2-D DCT. This is not quite surprising given the more pronounced edge-like nature of the the modified ICA bases (see Fig. 1).

5. CONCLUSION

Recently, two new ICA-based algorithms, called `GCGsup` and `ICAorth`, were proposed to compute optimal transforms in transform coding. In this paper, we were interested in the performances of the transforms returned by `GCGsup` and `ICAorth` when applied to the compression of well-known grayscale images. Experimental results showed that the new transforms are 1) comparable to the classical 2-D DCT according to the PSNR measure and 2) yield better visual image quality (better preservation of lines and contours) than the 2-D DCT.

A further work under consideration is to design a low complexity image coder based on the class-adapted transforms returned by `GCGsup` and `ICAorth` and compare its performance to JPEG and JPEG2000. Furthermore, some work have already begun on the application of `GCGsup` and `ICAorth` to the compression of satellite images [9].

6. ACKNOWLEDGMENT

This work has been partially supported by the Lorraine Region. Moreover, the authors are grateful to Jacques Weidig who took part in the simulation work.

7. REFERENCES

[1] M. Narozny, M. Barret, D. T. Pham, I. P. Akam Bita, “Modified ICA algorithms for finding optimal transforms in transform coding”, *4th Int. Symposium on Image and Signal Processing and Analysis*, Zagreb (Croatia), 15–17 Sept. 2005.

[2] V. K. Goyal, “Theoretical foundations of transform coding,” *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 9–21, Sept. 2001.

[3] J.-Y. Huang and P. M. Schultheiss, “Block quantization of correlated Gaussian random variables,” *IEEE Trans. Commun.*, vol. COM-11, pp. 289–296, Sept. 1963.

[4] V. K. Goyal, J. Zhuang, and M. Vetterli, “Transform coding with backward adaptive updates,” *IEEE Trans. Inform. Theory*, vol. 46, pp. 1623–1633, July 2000.

[5] A. Ferreira and M. Figueiredo, “Class-adapted image compression using independent component analysis,” *in IEEE International Conference on Image Processing - ICIP’2003*, Barcelona, Spain, Sept. 2003.

[6] A. Ferreira and M. Figueiredo, “Image compression using orthogonalised independent components bases,” *in Proceedings of the IEEE Neural Networks for Signal Processing Workshop*, Toulouse, France, Sept. 2003.

[7] D. T. Pham, “Fast algorithms for mutual information based independent component analysis,” *IEEE Transactions on Signal Processing*, vol. 52, no.10, pp. 2690–2700, Oct. 2004.

[8] Y. Shoham and A. Gersho, “Efficient bit allocation for an arbitrary set of quantizers,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 9, pp. 1445–1453, Sept. 1988.

[9] I. P. Akam Bita, M. Barret et D. T. Pham, “Compression of multicomponent satellite images using independent component analysis,” *ICA 2006*, Charleston (USA), 5-8 March 2006.