Modified ICA algorithms for finding optimal transforms in transform coding

Michel Narozny*, Michel Barret
Équipe Systèmes de Traitement du Signal ,
Supélec, 2, rue É. Belin, 57070, Metz, France
firstname.lastname@supelec.fr

*This work was partially supported by the Lorraine Region.

Dinh-Tuan Pham, Isidore Paul Akam-Bita
Laboratoire de Modélisation et Calcul, IMAG
B.P. 53X, 38041 Grenoble Cedex, France
Dinh-Tuan.Pham@imag.fr
Isidore-Paul.Akambita@imag.fr

Abstract

In this paper we present two new algorithms that compute the linear optimal transform in high-rate transform coding, for non Gaussian data. One algorithm computes the optimal orthogonal transform, and the other the optimal linear transform. Comparison of the performances in high-rate transform coding between the classical Karhunen-Loève Transform (KLT) and the transforms returned by the new algorithms are given. On synthetic data, the transforms given by the new algorithms perform significantly better that the KLT, however on real data all the transforms, included KLT, give roughly the same coding gain.

1. Introduction

Common sources such as speech and images have considerable "redundancy" that scalar quantization cannot exploit. Strictly speaking the term "redundancy" refers to the statistical correlation or dependence between the samples of such sources and is usually referred to as *memory* in the information theory literature. It is well known that "removing the redundancy" in the data before scalar quantizing it leads to much improved codes. In transform coding, the redundancy is reduced by transforming the data before a scalar quantization. Generally the transform is linear. In this type of transform coding, an input vector X of size N is transformed into another vector S of the same dimension; the components S_i, $i = 1, \ldots, N$, of that vector are then independently quantified and fed to the encoder. High resolution theory shows that the Karhunen-Loève transform (KLT) is optimal for Gaussian sources [1], and the asymptotic low resolution analysis does likewise [2]. Transform coding has been extensively developed for coding images and video (for example, H.261, H.263, JPEG, and MPEG), where the discrete cosine transform (DCT) is most commonly used because of its computational simplicity and its good performance. Special transform codes are subband codes which decompose an image into separate images by using a set of linear filters. The resulting subbands can then be quantized, e.g., by scalar quantizers. The discrete wavelet transform (DWT) is a particular subband code, which is used in the new image compression standard JPEG 2000.

The basic idea that leads to the use of linear pre-processing before coding lies in the de-correlation effect between the pixel values that allows the use of simple source encoders. Though appealing, this idea is hampered by the fact that linear processing alone may not achieve total independence in the case of non-Gaussian sources. This explains why most of the compression methods (JPEG, JPEG 2000) that perform well use linear pre-processing and some form of context modeling. Through context modeling it is possible to extract the dependencies remaining in the data after linear pre-processing in order to improve the compression performance.

For non Gaussian data, the linear transform that performs best in high rate transform coding does not decorrelate the data in general, and this result remains valid when the transform is constrained to be orthogonal. Hence, the KLT is not the best linear transform in high rate transform coding for non Gaussian data. Moreover in image coding it is well known (see e.g., [3]) that after a DWT the wavelets coefficients obtained from an image are not Gaussian and hence, pixels are not Gaussian, even if we neglect the fact that they are quantized data. Independent component analysis (ICA) is a recently developed technique which aims at finding a linear transform that minimizes the statistical dependence between the transform coefficients. The mutual information is a natural measure of the dependence between random variables. Therefore, finding a transform which minimizes the mutual information between the transform coefficients is a very natural way of performing ICA. One may expect that this ICA transform is optimal for a linear transform coding system, since it reduces the redundancy between the components as far as it can. But it is not quite so, since the distortion is measured in term of the mean squared error, which favors orthogonal transforms and the ICA transform is not orthogonal in general.

In section 2, we show that the optimal linear transform for a high-rate linear transform coding system employing entropy-constrained uniform quantization is the one that minimizes a contrast $\mathcal{C}$ which is equal to the sum of the mutual information between the transform coefficients and another term which may be interpreted as a kind of distance to orthogonality of the transform. A presentation of ICA is

given in section 3 and its link to transform coding is elaborated on section 4. In section 5, we propose two new algorithms for the minimization of $\mathcal{C}$. Both algorithms are derived from the mutual information based ICA algorithm by Pham called `ICAinf` [4]. A comparison between the performances of the transforms returned by the new algorithms and that of the KLT, using both synthetic and real data, is given in section 6.

2. Optimal transform in transform coding

The class of signals to be encoded is represented by a random vector $X = (X_1, \ldots, X_N)^T$ of size N. Let $(\Omega, \mathcal{E}, P)$ be the probability space associated to $X: \Omega \rightarrow \mathbb{R}^N$. A transform coder applies a transform $\mathbf{A}: \mathbb{R}^N \rightarrow \mathbb{R}^N$ to X in order to obtain a random vector S better suited to coding than X. To construct a finite code, each coefficient S_i is approximated by a quantized variable $\widehat{S}_i$. We concentrate on scalar quantizations, which are most often used for transform coding. The decoder then applies a transformation $\mathbf{B}$ to $\widehat{S} = (\widehat{S}_1, \ldots, \widehat{S}_N)^T$ in order to obtain an approximation $\widehat{X}$ of X. In this paper we assume $\mathbf{B} = \mathbf{A}^{-1}$ and the transform $\mathbf{A}$ is linear.

2.1. Entropy-constrained scalar quantization

A scalar quantizer Q approximates a random variable Z by a quantized variable $\widehat{Z}$. It is a mapping from the source alphabet $\mathbb{R}$ to a reproduction codebook $\{\widehat{z}_k\}_{k\in\mathcal{K}} \subset \mathbb{R}$, where $\mathcal{K}$ is an arbitrary countable index set. Quantization can be decomposed into two operations: the lossy encoder $\alpha: \mathbb{R} \rightarrow \mathcal{K}$ is specified by a partition of $\mathbb{R}$ into partition cells $S_k = \{z \in \mathbb{R} \,|\, \alpha(z) = k\}$, $k \in \mathcal{K}$, and the reproduction decoder $\beta: \mathcal{K} \rightarrow \mathbb{R}$ is specified by the codebook $\{\widehat{z}_k\}_{k\in\mathcal{K}}$. We denote $p_k = \Pr\{Z \in S_k\} = \Pr\{\widehat{Z} = \widehat{z}_k\}$. The Shannon theorem [5] proves that the entropy $H(\widehat{Z}) = -\sum_k p_k \log_2 p_k$ is a lower bound of the average number of bits per symbol used to encode the values of $\widehat{Z}$. Arithmetic entropy coding achieves an average bit rate that can be arbitrarily close to the entropy lower bound (see e.g., [9]); therefore, we shall consider that this lower bound is reached. An *entropy constrained scalar quantizer* is designed to minimize $H(\widehat{Z})$ for a fixed mean square distortion $D = \mathrm{E}[(Z - \widehat{Z})^2]$, where $\mathrm{E}[Z]$ denotes the expectation of Z. Consider the variance σ^2 of Z and $\tilde{Z} = (Z - \mathrm{E}[Z])/\sigma$ the standardized random variable associated to Z; let $h(\tilde{Z})$ be the differential entropy of $\tilde{Z}$, $h(\tilde{Z}) = -\int_{-\infty}^{+\infty} p(\tilde{z}) \log_2 p(\tilde{z}) d\tilde{z}$, where $p(\tilde{z})$ denotes the probability density function (pdf) of $\tilde{Z}$. A result from high resolution quantization theory (see e.g., [5]) is that the quantizer performance is described by

$$D \simeq c\,\sigma^2\, 2^{-2R}, \qquad \text{with} \quad c = \frac{2^{2h(\tilde{Z})}}{12}, \tag{1}$$

where $R = H(\widehat{Z})$ is the minimum average bit rate, and the constant c depends only on the pdf shape.

2.2. Generalized coding gain

Coding (quantizing and entropy coding) each transform coefficient S_i separately splits the total number of bits among the transform coefficients in some manner. This bit allocation problem can be stated this way: one is given a set of quantizers described by their high rate-distortion performances—see (1)—as $D_i \simeq c_i\,\sigma_i^2\,2^{-2R_i}$ ($i = 1, \ldots, N$), where σ_i^2 is the variance of S_i and the constant c_i is associated with the standardized variable $\tilde{S}_i$ of S_i. The problem is to minimize the end-to-end distortion $D = N^{-1}\sum_{i=1}^N \mathrm{E}[(X_i - \widehat{X}_i)^2]$ given a maximum average rate $R = N^{-1}\sum_{i=1}^N R_i$. Let us introduce the elements $b_{m,n}$ of the matrix $\mathbf{A}^{-1} = [b_{m,n}]$. If we assume that the quantizer error signals, $S_i - \widehat{S}_i$, $i = 1, \ldots, N$, are white and mutually uncorrelated, the end-to-end distortion can then be directly computed by a weighting sum of the distortion of each transform coefficient as

$$\begin{aligned} D &= \frac{1}{N}\sum_{j=1}^N \mathrm{E}\Big[\Big(X_j - \widehat{X}_j\Big)^2\Big] \qquad (2)\\ &= \frac{1}{N}\sum_{j=1}^N \mathrm{E}\Big[\Big|\sum_{i=1}^N b_{j,i}(S_i - \widehat{S}_i)\Big|^2\Big] = \frac{1}{N}\sum_{i=1}^N w_i D_i, \end{aligned}$$

where the weight w_i corresponds to the square euclidean norm of the ith column of $\mathbf{A}^{-1}$.

The arithmetic mean of the $w_i D_i$s is equal to or greater than their geometric mean, with equality if and only if all the terms are equal. Therefore, under the constraint of a given average bit rate R, the distortion D is minimum if and only if all the $w_i D_i$s are equal, in which case the minimum value of the end-to-end distortion can be approximated as follows

$$D_{\mathbf{A}}(R) \simeq \left[\prod_{i=1}^N w_i\, c_i\, \sigma_i^2\right]^{\frac{1}{N}} 2^{-2R}. \tag{3}$$

Let $\mathbf{I}$, $\sigma_i^{\star 2}$ and $c_i^\star$ be respectively the identity transform, the variance of X_i and the constant associated with the standardized random variable of X_i according to (1). The distortion rate (3) may then be used to define a figure of merit that we call the *generalized coding gain*

$$G^\star = \frac{D_{\mathbf{I}}(R)}{D_{\mathbf{A}}(R)} = \left[\prod_{i=1}^N \frac{c_i^\star\,\sigma_i^{\star 2}}{w_i c_i \sigma_i^2}\right]^{\frac{1}{N}}. \tag{4}$$

The generalized coding gain is the factor by which the distortion is reduced because of the linear transform $\mathbf{A}$, assuming high rate and optimal bit allocation.

2.3. Optimal transform for coding

Finding the matrix $\mathbf{A}$ which maximizes $G^\star$ is the same problem as finding the matrix $\mathbf{A}$ which maximizes the generalized maximum reducible bits $R^\star_{\max} = \frac{1}{2}\log_2 G^\star$, or

equivalently, finding the linear transform which minimizes the contrast

$$\mathcal{C}(\mathbf{A}) = I(S_1; \ldots; S_N) + \frac{1}{2}\log_2 \frac{\det \mathrm{Diag}[\mathbf{A}^{-T}\mathbf{A}^{-1}]}{\det \mathbf{A}^{-T}\mathbf{A}^{-1}}, \quad (5)$$

where the first term of (5) is the mutual information $\int_{\mathbb{R}^N} p(s) \log_2 \frac{p(s)}{p(s_1)\cdots p(s_N)}\, ds$ between the random variables S_1, ..., S_N, and for any square matrix $\mathbf{C}$, $\mathrm{Diag}(\mathbf{C})$ denotes the diagonal matrix having the same main diagonal as $\mathbf{C}$. Indeed, using the relation (1) and the following relations $h(X) = \sum_{i=1}^{N} h(X_i) - I(X_1; \cdots; X_N)$, a similar one with S and $h(S) = h(X) + \log_2 |\det \mathbf{A}|$ (see e.g. [9] for notions of information theory), some calculus give $R^{\star}_{\max} = \frac{1}{N} I(X_1; \cdots; X_N) - \frac{1}{N} I(S_1; \cdots; S_N) - \frac{1}{N}\log_2 |\det \mathbf{A}| - \frac{1}{2N}\log_2 \prod_{i=1}^{N} w_i$ and the last two terms are equal to the opposite of the second term of (5).

The mutual information of S_1, ..., S_N is a measure of the statistical dependence between the transform coefficients S_i: it is always non-negative, and zero if and only if the variables are statistically independent. As for the second term in (5), it is always non-negative, and zero if and only if $\mathbf{A}^{-1}$ is a transform with orthogonal columns. The columns may not be of unit Euclidean norm. In other words, the second term of the contrast $\mathcal{C}(\mathbf{A})$ can be interpreted as a kind of distance to orthogonality for the transform $\mathbf{A}$. Furthermore, if $\mathbf{D}$ is a diagonal matrix, one can verify that $\mathcal{C}(\mathbf{DA}) = \mathcal{C}(\mathbf{A})$, i.e., the contrast is scale invariant. As a consequence, one can normalize the rows of $\mathbf{A}^{-1}$ to have unit norm and using equal quantizer step sizes.

3. Independent component analysis

A common problem encountered in a variety of disciplines, including data analysis, signal processing, and compression, is finding a suitable representation of multivariate data. For computational and conceptual simplicity, such a representation is often sought as a linear transformation of the original data. Well-known linear transformation methods include, for example, principal component analysis (PCA). A recently developed linear transformation method is the independent component analysis, in which the desired representation is the one that minimizes the statistical dependence of the components of the representation. Although non-linear forms of ICA also exist, we shall only consider the linear case here.

Hyvärinen [7] gives the following definition for the noise-free ICA model, which is of primary interest in our study.

Definition 1 (Noise-free ICA model) *ICA of a random vector X of size N consists of estimating the following generative model for the data:*

$$X = \mathbf{B}S \quad (6)$$

where $\mathbf{B}$ is a constant $N \times M$ "mixing" matrix, the latent variables (components) S_i in the vector $S = (S_1, \ldots, S_M)^T$ are assumed independent.

In the following, we assume that the dimension of the observed data equals the number of the independent components, i.e., $N = M$, and that the matrix $\mathbf{B}$ is invertible. In this situation, the identifiability of the noise-free ICA model can be assured under the following fundamental restrictions (in addition to the basic assumption of statistical independence) that all the independent components S_i, with the possible exception of one component, must be non-Gaussian [6].

Note that identifiability here means only that the independent components and the columns of $\mathbf{B}$ can be estimated *up to a multiplicative constant and a permutation.* Indeed, any multiplication of an independent component in (6) by a constant could be canceled by a division of the corresponding column of the mixing matrix $\mathbf{B}$ by the same constant. Further, the definition of the noise-free ICA model implies no ordering of the independent components, which is in contrast to, e.g., PCA.

The estimation of the data model of ICA is usually performed by formulating an objective function and then minimizing or maximizing it. The mutual information is a natural measure of the dependence between random variables. Finding a transform that minimizes the mutual information between the components S_i is a very natural way of estimating the ICA model [6]. The problem with mutual information is that it is difficult to estimate. One needs a good estimate of the density. This problem has severely restricted the use of mutual information in ICA estimation. Some authors have used approximations of mutual information based on polynomial density expansions [6], which lead to the use of higher-order cumulants. More recently, in [4], Pham has proposed fast algorithms to perform ICA based on the use of mutual information.

4. Link between transform coding and ICA

The criterion (5) may be decomposed into

$$\mathcal{C}(\mathbf{A}) = \mathcal{C}_{\mathrm{ICA}}(\mathbf{A}) + \mathcal{C}_{\mathrm{O}}(\mathbf{A}), \quad (7)$$

where $\mathcal{C}_{\mathrm{ICA}}(\mathbf{A}) = I(S_1; \ldots; S_N)$, and

$$\mathcal{C}_{\mathrm{O}}(\mathbf{A}) = \frac{1}{2}\log_2 \left[\frac{\det \mathrm{Diag}(\mathbf{A}^{-T}\mathbf{A}^{-1})}{\det \mathbf{A}^{-T}\mathbf{A}^{-1}}\right]. \quad (8)$$

The first term $\mathcal{C}_{\mathrm{ICA}}(\mathbf{A})$ corresponds to the mutual information criterion in ICA. The second term $\mathcal{C}_{\mathrm{O}}(\mathbf{A})$ measures a pseudo-distance to orthogonality of the transform $\mathbf{A}$: it is non negative and can be zero if and only if the columns of $\mathbf{A}^{-1}$ are orthogonal. In general, the optimal transform $\mathbf{A}_{\mathrm{opt}}$ in transform coding, i.e., the transform which minimizes the contrast defined in the relation (5), will be different from that $\mathbf{A}_{\mathrm{ICA}}$ which minimizes the first term of (5), i.e., the solution of the ICA problem. Note that the contrast $\mathcal{C}(\mathbf{A})$ is always non-negative, and that it is equal to zero if and only if $\mathbf{A}$ is a transform with orthogonal columns which produces independent components. Therefore, when such a transform exists, it is both the solution of the compression

problem and that of the ICA problem. Unfortunately, for most sources, it is very unlikely to find orthogonal transforms that produce independent components.

It is important to notice here that the classical assumption made in blind source separation problems, that is the observations are obtained from a linear mixing of independent sources, is not really required in the problem of finding the transform that maximizes the generalized coding gain.

The expression of the contrast (5) depends on the definition of the distortion. In this work, we measure the distortion as mean squared error (MSE). Therefore, it is not surprising that orthogonal transforms are favored over other linear transforms since they are energy-preserving.

5. Modified ICA algorithms for coding

In this section, we propose two algorithms for the minimization of the contrast (5). The first algorithm, called `GCGsup` for *Generalized Coding Gain Supremum*, consists of a modified version of the mutual information based ICA algorithm by Pham [4] called `ICAinf`. The second term of (5) has been incorporated in `ICAinf` in order to find the optimal linear transform $\mathbf{A}_{\rm opt}$ which minimizes the contrast (5). In the second new algorithm, called `ICAorth` for *Independent Composent Analysis Orthogonal*, the algorithm `ICAinf` has been modified in order to find the optimal orthogonal matrix $\mathbf{A}_{\rm orth}$ that minimizes the contrast $\mathcal{C}(\mathbf{A})$.

5.1. Algorithm `GCGsup`

The minimization of the criterion (5) can be done through a gradient descent algorithm, but a much faster method is the Newton algorithm (which amounts to using the natural gradient [8]). As in [4], because of the multiplicative structure of our optimization problem, we use multiplicative increment of the parameter $\mathbf{A}$ rather than additive increment. Starting with a current estimator $\widehat{\mathbf{A}}$, it consists of expanding $\mathcal{C}(\widehat{\mathbf{A}}+\mathcal{E}\widehat{\mathbf{A}})$ with respect to the matrix $\mathcal{E}$ up to second order and then minimizing the resulting quadratic form in $\mathcal{E}$ to obtain a new estimate. Note that the parameter $\mathcal{E}$ is a matrix of order N. This method requires the computation of the Hessian[1] of $\mathcal{C}(\widehat{\mathbf{A}}+\mathcal{E}\widehat{\mathbf{A}})$ with respect to $\mathcal{E}$, which is quite involved. For this reason, we will approximate it by the Hessian of $\mathcal{C}(\widehat{\mathbf{A}}+\mathcal{E}\widehat{\mathbf{A}})$, computed under the assumption that the transform coefficients $\widehat{S}_i$ are independent. The method is then referred to as quasi-Newton.

Although those simplifications result in a slower convergence speed about the solution, they cause the robustness of the algorithm to be improved by reducing the risk of divergence when the initial estimator $\widehat{\mathbf{A}}_0$ is far from the final solution. Note that the final solution is the same as that obtained without simplification since the algorithm consists of cancelling the first order terms in the expansion of $\mathcal{C}(\mathbf{A}+\mathcal{E}\mathbf{A})$.

[1]The Hessian of a function of several variables is the matrix of its second partial derivatives.

Given that (see e.g., [9]) $h(X) = \sum_{i=1}^N h(X_i) - I(X_1;\ldots;X_N)$, $h(S) = \sum_{i=1}^N h(S_i) - I(S_1;\ldots;S_N)$, $h(S) = h(X) + \log_2|\det\mathbf{A}|$, and the term $h(X)$ does not depend on $\mathbf{A}$, minimizing the contrast (5) is the same as minimizing $\widetilde{\mathcal{C}}(\mathbf{A}) = \mathcal{C}_{\rm O}(\mathbf{A}) + \widetilde{\mathcal{C}}_{\rm ICA}(\mathbf{A})$ where

$$\widetilde{\mathcal{C}}_{\rm ICA}(\mathbf{A}) = \sum_{i=1}^N h(S_i) - \log_2|\det\mathbf{A}|. \tag{9}$$

Using the results of [10] it can be seen that the Taylor expansion of $\widetilde{\mathcal{C}}_{\rm ICA}(\mathbf{A}+\mathcal{E}\mathbf{A})$ up to second order may be approximated as follows

$$\widetilde{\mathcal{C}}_{\rm ICA}(\mathbf{A}+\mathcal{E}\mathbf{A}) = \widetilde{\mathcal{C}}_{\rm ICA}(\mathbf{A}) + \sum_{1\le i\ne j\le N}\mathrm{E}[\psi_{S_i}(S_i)S_j]\mathcal{E}_{ij} + \\ + \frac{1}{2}\sum_{1\le i\ne j\le N}\{\mathrm{E}[\psi^2_{S_i}(S_i)]\,\mathrm{E}[S_j^2]\mathcal{E}_{ij}^2 + \mathcal{E}_{ij}\mathcal{E}_{ji}\} + \cdots, \tag{10}$$

where the function ψ_{S_i} is equal to the derivative of $-\log_2 p(s_i)$ and is known as the score function, which can be viewed as the gradient of the entropy functional. This approximation concerns only the second order terms in the expansion, but *not the first order terms*. It relies essentially on the assumption of independent transform coefficients, which may not be valid if the solution of the ICA problem is far from the solution that minimizes the contrast (5). But it is quite useful since it leads to a decoupling in the quadratic form of the expansion.

Let $\mathbf{M} = \mathbf{A}^{-T}\mathbf{A}^{-1}$. One may verify that the Taylor expansion of $\mathcal{C}_{\rm O}(\mathbf{A}+\mathcal{E}\mathbf{A})$ with respect to $\mathcal{E}$ and around $\mathcal{E}=\mathbf{0}$, up to second order, is given by

$$\mathcal{C}_{\rm O}(\mathbf{A}+\mathcal{E}\mathbf{A}) = \mathcal{C}_{\rm O}(\mathbf{A}) - \sum_{1\le i\ne j\le N}\frac{M_{ji}}{M_{ii}}\mathcal{E}_{ji} - \frac{1}{2}\sum_{1\le i\ne j\le N}\mathcal{E}_{ij}\mathcal{E}_{ji} \\ + \sum_{\substack{1\le i,j,k\le N\\ j\ne i \text{ and } k\ne i}}\left[\left(\frac{M_{jk}}{2M_{ii}} - \frac{M_{ij}M_{ik}}{M_{ii}^2}\right)\mathcal{E}_{ji}\mathcal{E}_{ki} + \frac{M_{kj}}{M_{kk}}\mathcal{E}_{ji}\mathcal{E}_{ik}\right] \\ + \sum_{1\le i\ne j\le N}\frac{M_{ji}}{M_{jj}}\mathcal{E}_{ii}\mathcal{E}_{ji} + \cdots \tag{11}$$

The quadratic form associated with the above expansion is quite involved and is not positive. One possible approximation consists in neglecting the non diagonal elements of $\mathbf{M}$, which amounts to assuming that the optimal linear transform is close to an orthogonal transform. Under this hypothesis, one may verify that

$$\mathcal{C}_{\rm O}(\mathbf{A}+\mathcal{E}\mathbf{A}) \approx \mathcal{C}_{\rm O}(\mathbf{A}) - \sum_{1\le i\ne j\le N}\frac{M_{ji}}{M_{ii}}\mathcal{E}_{ji} + \\ + \frac{1}{2}\sum_{1\le i\ne j\le N}\left[\frac{M_{jj}}{M_{ii}}\mathcal{E}_{ji}^2 + \mathcal{E}_{ji}\mathcal{E}_{ij}\right] + \cdots \tag{12}$$

The quadratic form associated with the above expansion is now positive, but not positive definite. However, this is sufficient for the matrix associated with the quadratic form of

the Taylor expansion of $\widetilde{\mathcal{C}}(\mathbf{A})$ to be positive definite, which ensures the stability of the iterative algorithm. Finally we have

$$\widetilde{\mathcal{C}}(\mathbf{A}+\mathcal{E}\mathbf{A}) \approx \widetilde{\mathcal{C}}(\mathbf{A}) + \sum_{1\le i\ne j\le N} \mathcal{E}_{ij}\Big[\mathrm{E}[\psi_{S_i}(S_i)S_j] - \frac{M_{ij}}{M_{jj}}\Big] + \\ + \frac{1}{2}\sum_{1\le i\ne j\le N}\Big\{\Big[\mathrm{E}[\psi^2_{S_i}(S_i)]\mathrm{E}[S_j^2] + \frac{M_{ii}}{M_{jj}}\Big]\mathcal{E}_{ij}^2 + 2\mathcal{E}_{ij}\mathcal{E}_{ji}\Big\} \\ + \cdots \qquad (13)$$

Explicitly, the iteration consists of solving the linear equations

$$\begin{bmatrix} \mathrm{E}[\psi^2_{S_i}(S_i)]\mathrm{E}[S_j^2] + \frac{M_{ii}}{M_{jj}} & 2 \\ 2 & \mathrm{E}[\psi^2_{S_j}(S_j)]\mathrm{E}[S_i^2] + \frac{M_{jj}}{M_{ii}} \end{bmatrix}\begin{bmatrix}\mathcal{E}_{ij}\\ \mathcal{E}_{ji}\end{bmatrix} \\ = \begin{bmatrix} \frac{M_{ij}}{M_{jj}} - \mathrm{E}[\psi_{S_i}(S_i)S_j] \\ \frac{M_{ji}}{M_{ii}} - \mathrm{E}[\psi_{S_j}(S_j)S_i]\end{bmatrix}. \qquad (14)$$

The indeterminate diagonal terms $\mathcal{E}_{ii}$ are arbitrarily fixed to zero. Then the estimator $\widehat{\mathbf{A}}$ is left multiplied by $\mathbf{I}+\mathcal{E}$ in order to update it. In this expression, the probability density functions being unknown, the score function $\psi_{S_i}(s_i)$ is replaced by an estimation (see [4]) and the expectations are estimated by empirical means.

5.2. Algorithm `ICAorth`

In this section, we propose a modified version of the mutual information based ICA algorithm by Pham [4] in order to find the orthogonal transform that minimizes the contrast (5). Since the second term of (5) vanishes for any othogonal matrix $\mathbf{A}$, this amounts to finding the orthogonal transform which minimizes the first term of (5), or equivalently, which minimizes $\widetilde{\mathcal{C}}_{\mathrm{ICA}}(\mathbf{A})$. If the matrix $\mathbf{A}$ is orthogonal, so is $\mathbf{A}+\mathcal{E}\mathbf{A}$, providing that $\mathbf{I}+\mathcal{E}$ be orthogonal. This last condition will be satisfied up to second order if $\mathcal{E}$ is anti-symmetric, since $(\mathbf{I}+\mathcal{E})^T(\mathbf{I}+\mathcal{E}) = \mathbf{I}+\mathcal{E}^T\mathcal{E}$ differs from the identity only by second order terms. Let $\mathcal{E}$ be anti-symmetric. The Taylor expansion of $\widetilde{\mathcal{C}}_{\mathrm{ICA}}(\mathbf{A}+\mathcal{E}\mathbf{A})$ becomes

$$\widetilde{\mathcal{C}}_{\mathrm{ICA}}(\mathbf{A}+\mathcal{E}\mathbf{A}) = \widetilde{\mathcal{C}}_{\mathrm{ICA}}(\mathbf{A}) + \\ + \sum_{1\le i<j\le N}\big\{\mathrm{E}[\psi_{S_i}(S_i)S_j] - \mathrm{E}[\psi_{S_j}(S_j)S_i]\big\}\,\mathcal{E}_{ij} + \\ + \frac{1}{2}\sum_{1\le i<j\le N}\mathcal{E}_{ij}^2\Big[\mathrm{E}[\psi^2_{S_i}(S_i)]\mathrm{E}[S_j^2] + \mathrm{E}[\psi^2_{S_j}(S_j)]\mathrm{E}[S_i^2] - 2\Big] \\ + \cdots, \qquad (15)$$

and the minimization of the second term in the above expansion yields

$$\mathcal{E}_{ij} = \frac{\mathrm{E}[\psi_{S_j}(S_j)S_i] - \mathrm{E}[\psi_{S_i}(S_i)S_j]}{\mathrm{E}[\psi^2_{S_i}(S_i)]\,\mathrm{E}[S_j^2] + \mathrm{E}[\psi^2_{S_j}(S_j)]\,\mathrm{E}[S_i^2] - 2}. \qquad (16)$$

Actually, $\mathbf{A}+\mathcal{E}\mathbf{A}$ is not a true orthogonal transform. This may be overcome by replacing $\mathbf{A}+\mathcal{E}\mathbf{A}$ with $e^{\mathcal{E}}\mathbf{A} = (\mathbf{I} + \mathcal{E} + \mathcal{E}^2/2! + \cdots)\mathbf{A}$, which is an orthogonal matrix differing from $\mathbf{A}+\mathcal{E}\mathbf{A}$ only by second order terms.

6. Simulation results

In this section, we are interested in comparing the performances of $\mathbf{A}_{\mathrm{opt}}$, $\mathbf{A}_{\mathrm{orth}}$, $\mathbf{A}_{\mathrm{ICA}}$, and the KLT for the independent coding of N information sources which are both statistically dependent and non-Gaussian. Our objective metric to measure the performance of a transform is the generalized coding gain (4). Given both N sources and a transform, the estimation of (4) requires good estimates of the pdfs of each source as well as each transformed component, which may be very difficult to obtain. In the next section, we shall elaborate on a more "practical" way of evaluating (4) which consists in actually coding the transformed components and measuring both the actual bit-rate (i.e., the first order entropy of quantized data) and the actual end-to-end distortion.

6.1. Evaluation methodology

The sources under investigation are either 1D signals or multidimensional-like images. In the latter, each image is first converted into a 1D-signal by scanning vertically column by column. The multidimensional signal $X = (X_1, \ldots, X_N)^T$ resulting from this pre-processing is then linearly transformed using an invertible linear transform $\mathbf{A}$ to produce the vector $S = (S_1, \ldots, S_N)^T$ whose components are coded separately. The i-th component S_i is first high-rate quantized with a uniform scalar quantizer of quantization step q_i. This gives $\widehat{S}_i$. The bit-rate R_i is then estimated by computing the empirical first order entropy of $\widehat{S}_i$ and the inverse transform $\mathbf{A}^{-1}$ is applied to $\widehat{S} = (\widehat{S}_1, \ldots, \widehat{S}_N)^T$ in order to reconstruct an approximation $\widehat{X} = (\widehat{X}_1, \ldots, \widehat{X}_N)^T$ of X. The distortion is the end-to-end one, given in the relation (2) and the total average rate is the empirical mean of the N rates R_i $(1 \le i \le N)$. It is well known that under the high-resolution hypotheses the optimal allocation of rates between the transform coefficients results in equal weighted distortions $w_i D_i$ (see section 2). Moreover, using uniform scalar quantizers, bit allocation amounts to choosing a quantization step q_i for each of the N components, and for small q_i (i.e., satisfying the hypothesis of high resolution quantization for S_i, [5]) the distortion D_i may be well approximated by $q_i^2/12$. Therefore, the equal-weighted distortion property of the analytical bit allocation solution gives a simple rule (for high bit-rate): let c be a constant, then make all the quantization steps $q_1, \ldots, q_N$ such that $w_i D_i = c$. This gives $q_i = \sqrt{12c/w_i}$, for $i = 1, \ldots, N$. When the constant c varies (under the assumption of high resolution quantization for each component) we obtain the classical asymptotic curve (distortion versus bit-rate, or equivalently bit-rate versus distortion). In our tests, we consider that the hypothesis of high resolution quantization is valid when for each component S_i, the relative deviation between the actual distortion $\mathrm{E}[(S_i - \widehat{S}_i)^2]$ (where the expectation is estimated by empirical mean) and $q_i^2/12$ is not greater than 1%. For a given high bit-rate, the ratio between the end-to-end distor-

tion read on the asymptotic curve obtained using the identity transform and that read on the asymptotic curve obtained using $\mathbf{A}$ yields the generalized coding gain of $\mathbf{A}$. We have tested two kinds of data, one consists in synthetic memoryless sources and the sources of the second are wavelet coefficients in the same sub-band (HH sub-band) of a multispectral satellite image Landsat.

In the first test, we have $N = 6$ and the data size is 2^{16}. First we produce a white vector $(\widetilde{S}'_1, \ldots, \widetilde{S}'_N)^T = \widetilde{S}'$ whose the i-th component is the standardized random variable associated to $S'_i = \mathrm{Sign}(Y_i).|Y_i|^\alpha$, where $(Y_1, \ldots, Y_N)^T$ is a standardized white Gaussian random vector. The exponent α is an arbitrary positive real number. When $\alpha > 1$ (resp. $\alpha < 1$), $\widetilde{S}'_i$ is super- (resp. sub-) Gaussian. Then, the vector X is obtained by the operation $X = \mathbf{M}\widetilde{S}'$, where $\mathbf{M}$ is an arbitrary orthogonal matrix. The random vector X being white, the KLT does nothing on it and the generalized coding gain $G^\star$ of the KLT is equal to $0\,\mathrm{dB}$. However, the components of X are not independent, and any algorithm among `GCGsup`, `ICAorth` and `ICAinf` gives the same result: the components S_i are independent, and the generalized coding gain $G^\star$ is the same for $\mathbf{A}_{\mathrm{ICA}}$, $\mathbf{A}_{\mathrm{orth}}$ and $\mathbf{A}_{\mathrm{opt}}$. In Tab. 1 we present $G^\star$ for different values of α. We remark that when α increases the hypothesis of

α	0.5	1	1.5	2	2.5
$G^\star$ (dB)	1.4	0.0	1.65	3.2	4.8

Table 1. Generalized coding gain of $\mathbf{A}_{\mathrm{orth}}$.

high-resolution quantization is satisfied for an increasing rate (e.g., when $\alpha = 2.5$, the rate must be greater than about 7 bits per coefficient).

The second test deals with real data, obtained from a satellite multispectral image Landsat of dimension $512 \times 512 \times 6$ and coded with 8 bits per pixel. Each component is decomposed in wavelet coefficients using the Daubechies 9-7 DWT. In our test, $N = 6$ and the six information sources we use are the six HH-subbands obtained by this way. In

	KLT	$\mathbf{A}_{\mathrm{orth}}$	$\mathbf{A}_{\mathrm{opt}}$	$\mathbf{A}_{\mathrm{ICA}}$
$G^\star$ (dB)	3.05	3.05	3.1	1.55

Table 2. Generalized coding gain when the sources are sub-band signals.

table 2, it can be seen that $\mathbf{A}_{\mathrm{opt}}$ performs slightly better than the KLT. As for $\mathbf{A}_{\mathrm{orth}}$, it yields practically the same result as the KLT. Therefore, for this type of data, when one consider orthogonal transform, achieving decorrelation is equivalent to minimizing the mutual information in terms of coding performances. Those results need further investigations.

7. Conclusion

In this paper we have presented two new algorithms that compute the linear optimal transform for a high-rate linear transform coding system employing entropy-constrained uniform scalar quantization. One algorithm computes the optimal orthogonal transform, and the other the optimal linear transform. These algorithms are both derived from an algorithm by D. T. Pham that minimizes the mutual information of the transformed components. Comparison of the performances in high-rate transform coding between the classical Karhunen-Loève Transform (KLT) and the transforms returned by our algorithms are given. On synthetic data, the transforms given by the new algorithms perform significantly better that the KLT, however on real data all the transforms, included KLT, give roughly the same coding gain.

References

[1] 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.

[2] 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.

[3] M. Antonini, M. Barlaud, P. Mathieu and I. Daubechies, "Image coding using wavelet transform", *IEEE Trans. Image Processing*, vol 1, pp. 205–220, 1992.

[4] D. T. Pham, "Fast algorithms for mutual information based independent component analysis", *IEEE Transaction on Signal Processing*, vol. 52, no. 10, pp. 2690–2700, Jan. 2004.

[5] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Boston, Kluwer, 1992.

[6] P. Comon, "Independent component analysis – a new concept?," *Signal Processing*, vol. 36, pp. 287–314, 1994.

[7] A. Hyvärinen, "Survey on independent component analysis," *Neural Computing Surveys*, vol. 2, pp. 94–128, 1999.

[8] S. Amari, "Natural gradient works eficiently in learning", *Neural Computation*, vol. 10, no. 2, pp. 251–276, 1998.

[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, New York, Wiley & Sons, 1991.

[10] D. T. Pham, "Entropy of a variable slightly contaminated with another" *IEEE Signal Processing Letter*, To appear.