4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003), April 2003, Nara, Japan

APPLICATION OF ICA TO LOSSLESS IMAGE CODING

Michel Barret and Michel Narozny[⋆]

Supélec
Équipe Signaux et Systèmes Électroniques
2 rue É. Belin 57070 Metz (France)
Michel.Barret@supelec.fr and Michel.Narozny@supelec.fr

[*]*This work was supported in part by the Lorraine Region.*

ABSTRACT

In this paper we give a lifting factorization of any matrix of order 3 or 4 whose determinant is equal to one. In the case of matrices of order 4, the factorization has been obtained by solving a system of algebraic equations, thanks to a Gröbner basis computation. The lifting factorization associated with rounding allows to construct a transformation that maps integers to integers and that is close to the separation matrix of an instantaneous mixture of independent components, given by any blind source separation algorithm. We have applied such transforms to images in order to reduce mutual information between inter-band pixels in a multiresolution decomposition. Results of our simulations are presented.

1. INTRODUCTION

In many applications, the transmission or the storage of digital images involve such an amount of information that current lossless coding do not satisfy users. Moreover, the lossless codecs do not assure a fixed bit rate: whatever the codec may be, it will exist an image that will be badly compress by the lossless codec. Therefore current lossless coding schemes are incompatible with data transmission via band limited channels. However, for some images classes, like seismic, medical or satellite images, the users fear artifacts due to lossy codecs and they are used to lossless image coding. For these users, embedded progressive coding from lossy to lossless may be a good compromise. Multiresolution decompositions with fixed filter coefficients, like the $S+P$ transform [13] and the wavelet decompositions [3], or with filter coefficients that adapt to the encoded image by linear mean square estimation [2], are well suited to embedded progressive coding, both in quality and in resolution. In the transformed domain, the correlation between pixels, in both intra-band and inter-band subimages, has significantly decreased. Therefore, it might be relevant to go beyond second-order moments, by means of independent component analysis (ICA) in order to further reduce redundancy after the multiresolution decomposition. In this paper we investigate the idea of adding a stage, at each level of decomposition, made of a separation matrix with the aim of reducing mutual information between pixels of inter-bands. The idea of using ICA techniques in lossless image coding has already been studied by S. M. Marusic and G. Deng [12] and others. Our approach is quite different : the introduction of an ICA stage in a multiresolution decomposition allows progressive coding. The lifting scheme introduced by W. Sweldens is well suited to construction of multiresolution decompositions that map integers to integers [3], the same holding the M-band nonlinear subband decompositions introduced by F. J. Hampson and J-C. Pesquet [10].

In Section 2 we recall well known results of the theory of information [6] and of ICA applied to blind source separation (BSS). In Section 3 we present a factorization of any matrix $\mathbf{A}$ of order 3 or 4, whose determinant is equal to one, into the product of three triangular matrices with ones on the principal diagonal. This factorization, associated with rounding to the nearest integer, permits the construction of a transformation that maps integers to integers and which is close to the $\mathbf{A}$ transform. Indeed, when we neglect the rounding, the transforms are the same. The factorization has been obtained by computing first a Gröbner basis and then by solving a system of algebraic equations, this method can be extended to matrices of any order M. In Section 4, we apply the separation matrix, obtained with a BSS algorithm, to four different multiresolution decompositions, which are the $S+P$ transform [13], the $5/3$ wavelet [1], the LAE and the GAE methods described in [2], in order to reduce the mutual information between inter-band pixels.

2. REVERSIBLE TRANSFORM AND CODING

Let $\underline{X} = (X_1, \ldots, X_M)^t$ be a discret random vector which takes its values in a finite set $\mathcal{V}$. Let $H(\underline{X})$ and $H(X_i)$ denote the entropies of $\underline{X}$ and X_i $(1 \leq i \leq M)$

respectively. These entropies are related to the mutual information $I(\underline{X})$ between the components of $\underline{X}$ according to the relation

$$H(\underline{X}) = \sum_{i=1}^{M} H(X_i) - I(\underline{X}). \tag{1}$$

The mutual information $I(\underline{X})$ is always non negative. Let F be a bijective application between the set $\mathcal{V}$ and another set $\mathcal{V}'$ and let us introduce the random vector

$$F(\underline{X}) = \underline{X}' = (X'_1, \ldots, X'_m)^t, \tag{2}$$

where m is not necessarily equal to M. Since the transformation F is reversible (bijective) the entropy is preserved : $H(\underline{X}') = H(\underline{X})$ and the relation (1) still holds for the random vector $\underline{X}'$. Moreover, on the one hand, when the components of $\underline{X}$ are coded independently, the size of the bit stream is, on average, equal to $\frac{1}{M}\sum_{i=1}^{M} H(X_i)$ bits per component of $\underline{X}$. On the other hand, when the redundancy between components is entirely exploited by the coder, the mean size of the bit stream is reduced to $\frac{1}{M}H(\underline{X})$ bits per component of $\underline{X}$. Therefore, the quantity $\frac{1}{M}I(\underline{X})$ corresponds to the optimal coding gain (given in bits per component of $\underline{X}$) that can be expected in applying first a reversible transformation F to $\underline{X}$ and then in coding independently the new components X'_i.

Let us suppose now and until the end of this section that the random vector $\underline{X} = (X_1, \ldots, X_M)^t$ is centered and takes real values. BSS algorithms applied to $\underline{X}$ are broken into two stages [4]. First whitening: $\underline{Z} = \mathbf{W}\underline{X}$ (we assume that the number of independent sources is equal to M), then reduction of mutual information via an orthogonal transformation $\mathbf{Q}$: $\underline{Y} = \mathbf{Q}\underline{Z}$. By applying a singular value decomposition to the correlation matrix of $\underline{X}$, we can express $\mathbf{W} = \mathbf{DP}$, where $\mathbf{P}$ is an orthogonal matrix and $\mathbf{D}$ an invertible diagonal matrix. The indeterminate values of the sources' variances are arbitrarily fixed to 1 in BSS, but this is not the best choice for coding applications as we shall see in the next sections. So we introduce an arbitrarily invertible diagonal matrix $\mathbf{\Delta}$ and consider the vector

$$\underline{X}' = \mathbf{\Delta QDP}\underline{X} \quad \text{with} \quad \mathbf{PP}^t = \mathbf{QQ}^t = \mathbf{I}, \tag{3}$$

where $\mathbf{I}$ denotes the identity matrix. The BSS algorithms return each of the matrices $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{D}$.

Let $S = \underline{X}(1), \underline{X}(2), \ldots, \underline{X}(N)$ be a population of independent samples of $\underline{X}$, and let $(\mathbf{P}, \mathbf{Q}, \mathbf{D})$ be the returned matrices of a BSS algorithm applied to the population S. Any orthogonal matrix is invertible. However, this does not imply that the transformation (3) is reversible when it is implemented with a floating point arithmetic processor. In the following we show that the matrix $\mathbf{\Delta QDP}$ admits a lifting factorization (see Fig. 1), provided the matrix $\mathbf{\Delta}$ is well chosen (we just process the cases $M = 3$ and $M = 4$). Then, by rounding intermediate values (as in [3]) it will be possible to construct a reversible decomposition $\mathcal{F}$ that maps integers to integers and such that $\underline{X}'' = \mathcal{F}(\underline{X}) \simeq \underline{X}'$. Therefore, we shall reasonably expect that $I(\underline{X}'') \ll I(\underline{X})$, (i.e., a significant coding gain).

3. LIFTING FACTORIZATION OF MATRICES

Proposition 1 *Let* $\mathbf{A} = [a_{ij}]$ $(1 \le i, j \le 3)$ *be a matrix of order 3 and* $\delta = a_{12}a_{23} - a_{22}a_{13}$. *If* $\det(\mathbf{A}) = 1$ *and* $\delta \neq 0$, *then* $\mathbf{A}$ *can be factored into*

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{bmatrix} \begin{bmatrix} 1 & d & e \\ 0 & 1 & f \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ a_1 & 1 & 0 \\ b_1 & c_1 & 1 \end{bmatrix}, \tag{4}$$

with $e = a_{13}$, $f = a_{23} - aa_{13}$ *(*a *is chosen in order to satisfy* $f \neq 0$*),* $c_1 = (a_{22} - aa_{12} - 1)/f$, $d = [a_{13}(1 - a_{22}) + a_{12}a_{23}]/f$, $b = -[a_{33}(fc_1 + 1) - fa_{32} - 1]/\delta$, $c = (-d + a_{12}a_{33} - a_{13}a_{32})/\delta$, $b_1 = a_{31} - ba_{11} - c(a_{21} - aa_{11})$, $a_1 = a_{21} - aa_{11} - fb_1$.

Fig. 1 shows lifting steps of a matrix when $M = 3$. We give now a lifting factorization for any square matrix $\mathbf{A}$ of order four, whose determinant is equal to one. We wish to find two lower triangular matrices

$$\mathbf{L}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ b & c & 1 & 0 \\ d & e & f & 1 \end{bmatrix}, \mathbf{L}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a_2 & 1 & 0 & 0 \\ b_2 & c_2 & 1 & 0 \\ d_2 & e_2 & f_2 & 1 \end{bmatrix} \tag{5}$$

and one upper triangular matrix such that

$$\mathbf{A} = \mathbf{L}_1 \mathbf{U} \mathbf{L}_2. \tag{6}$$

If we put

$$\mathbf{U} = \begin{bmatrix} 1 & a_1 & b_1 & c_1 \\ 0 & 1 & d_1 & e_1 \\ 0 & 0 & 1 & f_1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{A} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ z_1 & z_2 & z_3 & z_4 \\ t_1 & t_2 & t_3 & t_4 \end{bmatrix}, \tag{7}$$

the equation (6) can then be recast as a system $\mathcal{E}$ of 16 algebraic equations with 18 unknowns, which are

$$\mathcal{F} = \{a, b, c, d, e, f, a_i, b_i, c_i, d_i, e_i, f_i \,(1 \le i \le 2)\} \tag{8}$$

and 16 parameters: $\mathcal{P} = \{x_i, y_i, z_i, t_i \,(1 \le i \le 4)\}$. Each algebraic equation can be expressed as $P = 0$, where P is a polynomial of the variables $\mathcal{F}$ with parameters in $\mathcal{P}$. The solvability of such a system of multivariate polynomial equations as well as the computation of all its solutions are two fundamental problems of algebraic geometry, which have been shown to be reducible, by relatively easy algorithms, to the problem of constructing Gröbner bases [7]. The main result of Gröbner bases theory is that the problem of constructing such a basis can be solved algorithmically. As it is relatively involved to present and understand the underlying theory of Gröbner bases, we content ourselves with giving a very intuitive definition of such a base, which, in the first place, will hopefully help the non initiate reader gain a good insight into the reason for computing Gröbner bases. More details about Gröbner bases theory can be found in [7]. Let's denote $\mathcal{E}$ the set of multivariate polynomials in the indeterminates $\mathcal{F}$ involved in the above set of algebraic equations. A Gröbner basis is a finite set $\mathcal{G}$ of polynomials such that $\mathcal{E}$ and $\mathcal{G}$ have exactly the same zeros and $\mathcal{G}$ is in a certain canonical form that, roughly, is a generalization of the triangular form well known for linear multivariate polynomials. In practice, the computation of a Gröbner basis makes it possible to turn our original system of algebraic equations into a new one which is simpler to handle. The following results can be obtained by using any symbolic computation software systems providing an implementation of the Gröbner bases method.

$$\left.\begin{array}{rcl}
\det \mathbf{A} - 1 & = & 0 \\
eq_1(d_2, e_2, f_2) & = & 0 \\
eq_2(b_2, c_2, d_2, e_2, f_2) & = & 0 \\
eq_3(a_2, b_2, c_2, d_2, e_2, f_2) & = & 0 \\
eq_4(f_1, b_2, c_2) & = & 0 \\
eq_5(e_1, a_2, b_2, c_2) & = & 0 \\
eq_6(d_1, a_2, d_2, e_2, f_2) & = & 0 \\
eq_7(c_1) & = & 0 \\
eq_8(b_1, f_2) & = & 0 \\
eq_9(a_1, c_2, e_2, f_2) & = & 0 \\
eq_{10}(f, d_2, e_2, f_2) & = & 0 \\
eq_{11}(e, b_2, c_2, d_2, e_2, f_2) & = & 0 \\
eq_{12}(d, a_2, b_2, c_2, d_2, e_2, f_2) & = & 0 \\
eq_{13}(c, b_2, c_2, d_2, e_2, f_2) & = & 0 \\
eq_{14}(b, a_2, b_2, c_2, d_2, e_2, f_2) & = & 0 \\
eq_{15}(a, a_2, b_2, c_2, d_2, e_2, f_2) & = & 0
\end{array}\right\} \quad (9)$$

Note that the result of the computation of a Gröbner basis is not unique and depends upon the ordering of the power products of the indeterminates. Many different orderings can be used. However, not every admissible ordering allow to simplify the original system of algebraic equations. In our case, we chose to use the lexicographic ordering which can be defined as follows. The lexicographic ordering ($\succ$) is analogous to the ordering of words used in dictionaries. For instance, `arrow` $\succ$ `arson` since the third letter of `arson` comes after the third letter of `arrow` in alphabetical order, whereas the first two letters are the same in both. In our problem, the ordering of the variables is chosen as follows: $a \succ b \succ c \succ d \succ e \succ f \succ a_1 \succ b_1 \succ c_1 \succ d_1 \succ e_1 \succ f_1 \succ a_2 \succ b_2 \succ c_2 \succ d_2 \succ e_2 \succ f_2$.

Using $\mathcal{G}$, we can turn our original system of algebraic equations into a new one which is much simpler to handle. It is worth noting that the new set $\mathcal{G}$ of polynomials contains the constant polynomial $\det \mathbf{A} - 1$, simply bringing to light the underlying working hypothesis on the determinant of $\mathbf{A}$. Using the system (9) of algebraic equations, it is now straightforward to get all the solutions by solving each equation line of the system one after the other, just as we would proceed with a triangular system in the case of linear multivariate polynomials. The inspection of the new system of algebraic equations reveals the existence of three degrees of freedom. By setting $c_2 = d_2 = f_2 = 0$, we obtain the following solutions

$$\left.\begin{array}{rcl}
e_2 & = & (1-\delta_{4,4})/\delta_{4,2};\, f=\delta_{3,4}+e_2\delta_{4,2} \\
b_2 & = & (1+e_2|x_1,y_4|-|x_1,y_2|)/(|x_2,y_3|+e_2|x_3,y_4|) \\
a_2 & = & (x_1-1-b_2x_3)/(x_2-e_2x_4) \\
f_1 & = & \delta_{4,3}+b_2\delta_{4,1};\, d=t_1-b_2t_3-a_2(t_2-e_2t_4) \\
e_1 & = & |x_1,y_4|-a_2|x_2,y_4|-b_2|x_3,y_4| \\
d_1 & = & |x_1,y_3|-a_2|x_2,y_3|-a_2e_2|x_3,y_4| \\
c_1 & = & x_4;\, b_1=x_3;\, a_1=x_2-e_2x_4 \\
e & = & |x_1,t_2|-e_2|x_1,t_4|+b_2(|x_2,t_3|+e_2|x_3,t_4|) \\
c & = & |x_1,z_2|-e_2|x_1,z_4|+b_2(|x_2,z_3|+e_2|x_3,z_4|) \\
b & = & z_1-b_2z_3+a_2(e_2z_4-z_2) \\
a & = & y_1-b_2y_3+a_2(e_2y_4-y_2)
\end{array}\right\} \quad (10)$$

where $\delta_{i,j}$ is the minor of order 3 obtained by eliminating the i^{th} row and the j^{th} column of $\mathbf{A}$ and $|u_i, v_j|$ is the minor of order 2 obtained by keeping the rows u, v and the columns i, j; for instance

$$\delta_{4,2} = \begin{vmatrix} x_1 & x_3 & x_4 \\ y_1 & y_3 & y_4 \\ z_1 & z_3 & z_4 \end{vmatrix}, \; |x_3, z_4| = \begin{vmatrix} x_3 & x_4 \\ z_3 & z_4 \end{vmatrix}. \quad (11)$$

With the factorization (6), it becomes possible to construct a nonlinear transformation that maps integers to integers and that is close to the linear transform $\mathbf{A}$, by introducing rounding to the nearest integer, as in [3].

4. REDUCTION OF MUTUAL INFORMATION BETWEEN INTER-BAND PIXELS

The well known formula [6] for measuring mutual information between the M components of a discrete random vector $\underline{X} = (X_1, \ldots, X_M)^t$ taking its values in $\mathcal{V}$ is given by

$$I(\underline{X}) = \sum_{\underline{x} \in \mathcal{V}} p(\underline{x}) \log_2 \frac{p(\underline{x})}{p(x_1) \cdots p(x_M)} \quad (12)$$

where $\underline{x} = (x_1, \ldots, x_M)$, $p(\underline{x}) = P(\underline{X} = \underline{x})$ and $p(x_i) = P(X_i = x_i)$. Let us introduce the ranges $I_1, \ldots I_M$ of x_1, $\ldots x_M$ respectively, and the set $[\![1, K]\!]$ of natural integers in $[1, K]$. Let $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_K$ be K réalisations of the random vector $\underline{X}$. The joint probabilities and the marginal probabilities involved in (12) are defined, for any $\underline{x} \in \mathcal{V}$, $x_i \in I_i$ $(1 \le i \le M)$, by $p(\underline{x}) = \text{card}(\{k \in [\![1, K]\!] : \underline{x}_k = \underline{x}\})/K$ and $p(x_i) = \text{card}(\{k \in [\![1, K]\!] : x_{ki} = x_i\})/K$, where $\underline{x}_k = (x_{k1}, \ldots, x_{kM})^t$ $(1 \le k \le K)$.

Using the factorization given in Proposition 1, we add to the lifting schemes of transforms that map integers to integers the separation matrices (see Fig. 1) obtained by the blind source separation algorithm called `jadeR` [5]. Let us call "separation" step, this additional nonlinear transformation in the decomposition. The multiresolution decompositions tested are the $S+P$ transform [13], the bi-orthogonal wavelet called $5/3$ in [1] and the adapted decompositions GAE and LAE [2]. All the images are initially coded with 8 bits per pixel. The GAE decomposition uses filters of orders $(2,2)$ and the LAE uses the parameters $r_1 = 16$ and $r_2 = 8$. After each level of decomposition, we compute the separation matrix returned by `jadeR` when the input data are the three sub-images of detail. Let D_{11}, D_{21} and D_{31} be the three detail subimages obtained after the first level of decomposition and localized respectively in the top-right, bottom-right and bottom-left hand corner of the transformed image. More generally, let D_{1i}, D_{2i} and D_{3i} denote the detail subimages of the i^{th} level of decomposition. For an image of size $M \times N$, decomposed into L levels, the weighted mutual information between detail subimages, expressed in bpp, is given by the relation

$$I_1 = \left(1 - \frac{1}{4^L}\right)^{-1} \sum_{i=1}^{L} \frac{I(D_{1i}, D_{2i}, D_{3i})}{3 \times 4^i}. \qquad (13)$$

Let us introduce the random vector $\underline{X} = (X_1, X_2, X_3)^t$, whose components are the pixel values of the three detail sub-images (at the same localization) obtained after the decomposition and before the "separation" step. Let $\overline{H}(X)$ be the average of the components' entropies ($\sum_{i=1}^{3} H(X_i) = 3\overline{H}(X)$). After the "separation" step (including rounding), the vector whose components are the pixel values of the detail sub-images becomes $\underline{X}' = (X_1', X_2', X_3')^t$. Since the transformation is reversible, the entropy does not vary and it results from the equation (1) that

$$I(\underline{X}) - I(\underline{X}') = 3(\overline{H}(X) - \overline{H}(X')). \qquad (14)$$

In other words the loss of mutual information is equal to three times the coding gain expected (in bpp).

The separation matrix $\mathbf{B} = \mathbf{QDP}$ returned by `jadeR` does not satisfy the condition $\det \mathbf{B} = 1$, which is required for lifting factorization. However, the condition of unit variance of the estimated sources (that is generally assumed in BSS) is not useful for reducing mutual information between the estimated sources. Indeed, if $I(\underline{X})$ denotes the mutual information between the components of a random vector $\underline{X}$ and if $\boldsymbol{\Delta}$ is a diagonal matrix, it can be easily verified that $I(\boldsymbol{\Delta}\underline{X}) = I(\underline{X})$. Hence, we have left multiplied the returned matrix $\mathbf{B}$ with a diagonal matrix $\boldsymbol{\Delta}$: $\mathbf{A} = \boldsymbol{\Delta}\mathbf{B}$, in order to satisfy the condition $\det \mathbf{A} = 1$.

Because of the rounding introduced in the lifting scheme, the result of compression depends on the diagonal matrix $\boldsymbol{\Delta}$. Surprisingly, we observed that it depends greatly on $\boldsymbol{\Delta}$. The problem is then to choose a good diagonal matrix $\boldsymbol{\Delta}$, i.e., a matrix such that the nonlinear transform (mapping integers to integers) associated with the lifting factorization of $\mathbf{A}$, reduces the average first order entropy of the detail sub-images. We have tested five different possibilities for $\boldsymbol{\Delta}$, which are

$$\left.\begin{array}{l} \boldsymbol{\Delta}_1 = \mathrm{diag}(\frac{1}{\det \mathbf{B}}, 1, 1),\ \boldsymbol{\Delta}_2 = \mathrm{diag}(1, \frac{1}{\det \mathbf{B}}, 1) \\ \boldsymbol{\Delta}_3 = \mathrm{diag}(1, 1, \frac{1}{\det \mathbf{B}}),\ \boldsymbol{\Delta}_4 = \frac{\mathbf{I}}{(\det \mathbf{B})^{1/3}} \\ \boldsymbol{\Delta}_5 = \mathbf{D}^{-1} \end{array}\right\} \qquad (15)$$

where $\mathbf{D}$ is the diagonal matrix computed during the whitening step of `jadeR`.

In all our simulations we have observed that the multiplication by the diagonal matrices $\boldsymbol{\Delta}_4$ or $\boldsymbol{\Delta}_5$ do not improve the decomposition for lossless coding. Indeed, the average entropy of the detail sub-images in highest resolutions do not decrease after the "separation" step. Moreover, we have estimated the mutual information between detail sub-images after multiplication by the matrix $\mathbf{B}$ and without rounding (in this case, the pixels take real values and we have used the algorithm described in [8] for estimating the mutual information, like in [11]). We have noted that the mutual information estimated is generally not smaller than the mutual information before the multiplication by the matrix $\mathbf{B}$ (see Table 2). However, we have observed that with rounding, the "separation" step associated with a diagonal matrix of the form $\boldsymbol{\Delta}_1$, $\boldsymbol{\Delta}_2$ or $\boldsymbol{\Delta}_3$ can reduce, sometimes significantly (for some medical MRI images) the mutual information. It can be used for reducing the weighted first order entropy of decomposed image (see Table 1) in following this procedure: for the two first levels of decomposition (the others do not modify significantly the weighted first order entropy of the decomposed image) a "separation" step is added if and only if it reduces the inter-band mutual information (compared with no "separation" step). And when a "separation" step is added, it is the best among the three possible $\boldsymbol{\Delta} = \boldsymbol{\Delta}_1$, $\boldsymbol{\Delta}_2$ or $\boldsymbol{\Delta}_3$. We emphasize that the size of the bit stream required for coding the separation matrices is insignificant.

Av. Natural Images	Av. MRI Images	Av. Sat. Images
4.62	2.61	4.86
4.57	2.45	4.82

Table 1. *Weighted first order entropy in* bpp. *First row: without "separation" step, second row: after "separation" step following the procedure given in the text. For each image, we have used the decomposition giving the smallest weighted first order entropy among* $S+P$, $5/3$, *GAE and LAE. The averages have been obtained on 17 (for MRI), 10 (for satellite) and 12 (for natural) images.*

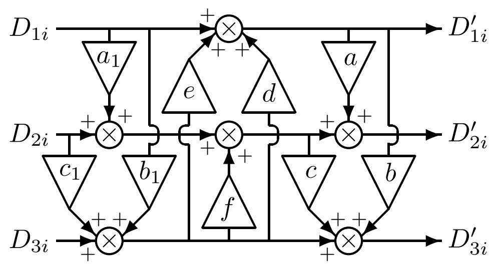

Fig. 1. *Lifting scheme of a separation matrix (according to Proposition 1) between the three detail subimages.*

In Table 2, we present the weighted mutual information given by the relation (13) between inter-band pixels of the decomposed image, for three different cases: first without "separation" step, second with multiplication by $\mathbf{B}$ and without rounding, and third with "separation" step . The tests were carried out on different kinds of images (medical MRI, satellite[1] and natural images), all coded with 8 bits per pixel and of dimension 512×512. At each level of decomposition, the mutual information between the three sub-images of detail is computed assuming the pixels in a same sub-image are independent.

5. CONCLUSION

In this paper, we investigated ICA applied to lossless image coding through multiresolution decompositions. We first gave a lifting factorization to any matrix of order three or four, whose determinant is equal to one, by using the mathematical framework of Gröbner bases. We then applied this method to separation matrices of order three with the aim to reduce the mutual information between inter-band pixels, in the process of multiresolution decomposition.

6. REFERENCES

[1] M. D. Adams and F. Kossentini, "Reversible Integer-to-Integer Wavelet Transforms for Image Compression: Performance Evaluation and Analysis", IEEE Trans. Image Processing, vol. 9, no. 6, pp. 1010-1024, Jun. 2000.

[2] H. Bekkouche and M. Barret, " Comparison of lossless codecs for satellite and MRI medical images", Vol. II, pp. 475–478, EUSIPCO'02, Toulouse France, Sept. 2002

[3] A. Calderbank, I. Daubechies, W. Sweldens and B. Yeo, "Wavelet transforms that map integers to integers", Appl. Comput. Harmon. Anal., vol. 5, pp. 332–369, Jul. 1998.

[4] J. F. Cardoso, "Blind signal separation: statistical principles", Proc. of IEEE, vol. 9, no. 10, pp. 2009–25, Oct. 1998.

[1] The satellite images have been given as a favor by the French National Center of Spatial Studies (CNES) and by the society named SPOT Image.

[5] http://www.tsi.enst.fr/icacentral/Algos/cardoso/

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

[7] D. Cox, J. Little and D. O'Shea, *Ideals, varieties and algorithms*, New-York, Springer-Verlag, 1997.

[8] G. A. Darbellay and I. V. Darbellay, "Estimation of the Information by an Adaptive Partitioning of the Observation Space", IEEE Transactions on Information Theory, Vol. 45, No. 4, pp. 1315–1321, May 1999.

[9] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps", J. Fourier Anal. Appl., vol. 4, no. 3, pp. 247–269, 1998.

[10] F. J. Hampson and J. C. Pesquet, "M-band nonlinear subband decompositions with perfect reconstruction", IEEE Trans. Image Processing, vol. 7, no. 11, Nov. 1998.

[11] J. Liu and P. Moulin, "Information-theoretic analysis of interscale and intrascale dependencies between image wavelet coefficients", IEEE Trans. Image Processing, vol. 10, no. 11, pp. 1647–1658, Nov. 2001.

[12] S. M. Marusic and G. Deng, "ICA-FIR based image redundancy reduction", IEEE ICA'99 Proc., pp. 191–195, Aussois, France, Jan. 1999.

[13] A. Said and W. A. Pearlman, "An image multiresolution representation for lossless and lossy compression", IEEE Trans. Image Processing, vol. 5, pp. 1303–1310, Sep. 1996.

Image	$S+P$	$5/3$	GAE	LAE
Lena	0.11	0.14	0.14	0.13
	1.35	1.26	1.18	1.31
	0.17	0.19	0.21	0.15
Mandrill	0.40	0.41	0.41	0.40
	0.11	0.14	0.07	0.19
	0.43	0.46	0.32	0.44
Peppers	0.13	0.14	0.16	0.15
	1.13	1.16	0.91	0.86
	0.13	0.17	0.18	0.35
Av. Nat.	0.19	0.20	0.23	0.20
	0.83	0.82	0.80	0.87
	0.27	0.28	0.21	0.27
Av. MRI	0.20	0.21	0.24	0.18
	0.85	0.43	0.73	0.93
	0.19	0.17	0.20	0.21
Av. Sat.	0.19	0.20	0.20	0.19
	0.63	0.66	0.56	0.60
	0.28	0.30	0.24	0.30

Table 2. *Estimation of the weighted mutual information (in* bpp*) of the decomposed image. First row: without "separation" step, second row: with multiplication by* $\mathbf{B}$ *with no scaling nor rounding (using the algorithm described in [8]), third row: with "separation" steps associated with the diagonal matrix* $\mathbf{\Delta}_i$ *(*$1 \le i \le 3$*) giving the smallest result and with rounding.*