Theoretical Foundation: Our work is developed based on the two following theoretical foundations:

- An enhanced Lifting Scheme based on second generation of wavelet transform.
- Non Negative Matrix Factorization (NMF), which is intended to provide a low rank approximation to a given matrix $X$ that involve two nonnegative factors $M$ and $S$ such as $X = MS$. The possible pairs of factor are selected with regard to an entropy-based objective function.

This two theories are combined and used to solve the convolutive version of the blind source separation (BSS) problem.

Contribution: Our contribution is two-fold. First, we propose an adaptive quincunx lifting scheme to preprocess and extract features from multidimensional data in order to subsequently achieve a successful blind separation. Second, we use our own entropy-based minimization criterion to perform a nonnegative factorization which is well suited to the data produced by the lifting scheme. This criterion will be used as the objective cost function of the nonnegative matrix factorization process. It is worth noting that the combination of the two techniques mentioned above has, to our knowledge, not been studied elsewhere.

closest work: To cope with convolutive blind source separation, many algorithms where tailored for the frequency domain but all of these works are limited to fourier transform. In our work, we develop a suitable wavelet transform for the frequency domain. We mention also that unlike all existing NMF algorithms, the objective function here is based on entropy.

Confirmation: We confirm that this manuscript is our original work.
Non-Negative Matrix Factorization for Blind Source Separation in Wavelet Transform Domain

Jamel Hattay¹, Samir Belaid² and Wady Naanaa³
¹LR-99-ES21 Systemes de communications SYS'COM, Ecole Nationale d’ingénieurs de TUNIS ENIT, TUNISIA
E-mail : Jamel.hattay@tunisietelecom.tn;
²Faculty of Sciences, University of Monastir, TUNISIA
E-mail : Samir.Belaid@fsm.rnu.tn;
³Faculty of Sciences, University of Monastir, TUNISIA
E-mail : wady.naanaa@fsm.rnu.tn

July 13, 2014

Abstract

This paper describes a new multilevel decomposition method for the separation of convolutive image mixtures. The proposed method uses an Adaptive Quincunx Lifting Scheme (AQLS) based on wavelet decomposition to preprocess the input data, followed by a Non-Negative Matrix Factorization whose role is to unmix the decomposed images. The unmixed images are, thereafter, reconstructed using the inverse of AQLS transform.

Experiments carried out on images from various origins showed that the proposed method yields better results than many widely used blind source separation algorithms.

keywords: Convolutive blind source separation, nonnegative matrix factorization, wavelet transform, multiscale analysis, adaptive lifting scheme, quincunx sampling.

1 INTRODUCTION

In the field of signal processing, one of the most studied research areas is the separation of blindly mixed signals. The problem, which is commonly known as blind source separation, involves mixtures of some unknown source signals obtained by an unknown mixing process, and the goal is to retrieve the source signals. The simplest mixing model is the one referred to as the linear and instantaneous model. This model assumes that each observation is a weighted sum of coefficients from the original source signals, which results in the following equation [1]:

\begin{equation}
\mathbf{x} = \mathbf{A} \mathbf{s} + \mathbf{n}
\end{equation}
where $X$ is a rectangular matrix holding in its rows the detected signals, $M$ is the mixing matrix whose entries are the unknown mixing coefficients, $S$ is a matrix holding in its rows the unknown source signals and $\theta$ is a matrix modelling the sensor noise.

Independent Component Analysis (ICA) [2] is one of the more powerful technique used in solving the blind source separation problem, giving rise to many ICA-based algorithms which have especially been proposed for the linear and instantaneous settings [3–6]. The problem as it is described by (1) has, in general, many solutions. To narrow the solution space, ICA-algorithms rely on the assumption that the source signals are as statistically independent as possible given the observed data. The latter hypothesis, however, does not always hold true, in particular when the source signals are images taken from the same context.

Yet another track explored for the purpose of unmixing linear and instantaneous mixtures is the one relying on convex analysis. In the latter approach, the unmixing process does not assume the “as statistically independent as possible” hypothesis to apply correctly. Nonetheless, source nonnegativity is used instead. CAMNS-LP and DEDS are two unmixing algorithms which best represent the convex analysis approach [7,8]. Moreover, each of these two algorithms uses an additional working hypotheses, namely, local orthogonality for CAMNS-LP and weak local orthogonality for DEDS. The convex analysis-based algorithms have proved to be more effective in separating highly correlated images, even though they are, in general, more time consuming than those based on ICA.

However, all the unmixing algorithms described above were developed for mixtures obtained by an instantaneous process which is hard to realize in practice. Indeed, in real situations, such as in image processing, another difficulty is added to the extent that all source pixels are weighted and delayed. The observations recorded by the sensors are, therefore, seen as a sum with multiple delays of propagation. This mixing process is known as the convolutive mixing process.

Several methods have been proposed to unmix convolutive mixtures in time domain [4,9], but these methods were limited and computationally expensive [10]. Other methods were dedicated to solving the problem in the frequency domain [11–13]. Motivated by these works, Ciaramella et al proposed a novel approach in which the fixed-point ICA algorithm in complex domain is combined with Short-Time Fourier Transform (STFT) [14]. Relying also on Fourier transform, Sawada proposed a multichannel extensions of nonnegative matrix factorization (NMF) to solve the convolutive separation problem in the frequency domain [15]. Always in the frequency domain, Daniel et al proposed a separation approach based on second order statistics [16]. The latter approach uses the decorrelation technique in frequency domain of convolutive mixtures and can be extended in order to provide an algorithm that is capable of perform-
ing joint acoustic echo cancellation and blind source separation. By examining
the working condition described in [17], the author demonstrated why the sep-
aration performance in the frequency domain using STFT is poor when there is
long reverberation. To overcome this drawback the author used the condition
$T \geq P$, where $T$ is the frame size of the STFT and $P$ is the length of a room
impulse response.

In the present work, we keep the unmixing process in the frequency domain,
but we make use of an enhanced wavelet packet transforms instead of STFT.
This work is motivated by the one presented in [18], where it was demonstrated
that, by focusing on the sparseness of sources and their mixtures once they are
projected onto a proper space of sparse representation, one can improves the
quality of separation. Our goal is to use the properties of multiscale transforms,
namely wavelet packets, to decompose signals into sets of local features with
various degrees of sparseness. To this end, we propose as a first step to apply a
lifting scheme. The latter scheme is a simple and general construction of second
generation wavelets where the original signal is recursively decomposed into ap-
proximation and detail subsets. In the context of images, i.e., two-dimensional
signals, the use of separable wavelets results in decomposing an input image into
an approximation subband and three details subbands which are respectively
oriented horizontally, vertically and diagonally. The extension of this decompo-
sition to an adaptive non separable version is called Adaptive Quincunx Lifting
Scheme (AQLS). A similar scheme was applied, in a preview work, to image
coding and compression [19]. In the present work, we aim to adapt and apply
this efficient technique to the blind source separation of convolutive image mix-
tures.

The paper is organized as follow: in Section 2, we remind the convolutive
mixing model for images. The proposed blind deconvolution method is detailed
in Section 3. Section 4 presents AQLS, our wavelet-based decomposition scheme.
A description of the NMF-based unmixing algorithm is described in Section 5.
Various simulation results are reported in Section 6. Finally, Section 7 is a brief
conclusion.

2 Convolutive Mixing Model for Images

A more general mixing model is required in situations where the linear and in-
stantaneous mixing model does not hold true. Such a model is the one referred
to as the convolutive model. Although the latter model presents some similarities
with the linear and instantaneous one from the experimental point of view,
(see Figure 1 which describes both models), they are theoretically very distinct.

To better understand the convolutive model, we first describe it for one-
dimensional signals.

Consider $P$ source signals $s_p(n), p : 1, \ldots, P$ and $Q$ observed signals $x_q(n), \ldots$
Figure 1: Mixing and Unmixing System.

$q : 1, \ldots, Q$ generated with a mixing system $M$. The goal is to take into account the influence of the delaying factor which is due to the contribution of source signal samples of previous moments. This can be done by means of an $L$-vector which can be viewed as a $M_{q,p}$ filter. The role of this filter is to allow adding the contribution of the $L$ neighboring samples of a given position. The convolutive model can therefore be described by the following equation.

$$x_q(n) = \sum_{p=1}^{P} \sum_{k=0}^{L-1} M_{q,p}(k) s_p(n-k) + \theta(n)$$

(2)

where $M_{q,p}$ is a vector containing the $L$ coefficients which stand for the delaying factors.

For image data, things are slightly different. The convolutive mixing system can be caused by defocus blur as it is the case, for instance, in hyperspectral, tomographic and microscopic imaging [20]. Convolutive mixtures are also involved in simple physical example of dynamic reflections, such as video recorded through the windshield of a car or the canopy on an airborne platform where the reflected (virtual) image is superimposed on a dynamic visual scene. In this example, each scene is at a different distance from the camera, thus differently defocus blurred in the mixtures [21]. Convolutive mixtures can also be obtained when the images are acquired using airborne cameras outfitted on a remotely piloted vehicle (RPV). Such images are usually contaminated by the interference of semi-transparent clouds [22]. The goal if, therefore, to extract a clear landscape view by separating the landscape from the thin semi-transparent layers of clouds. This problem of separation of clouds from landscape is more complex than the separation of reflections, described above. The presence of clouds involves a weighting and delay for each pixel of recorded images. They are rather
obtained by a more complex process involving multiplicative and convolutive masks. In such a setting, each mask $M_{q,p}$ becomes an $(2L + 1) \times (2L + 1)$ matrix. Here, we consider $P$ two-dimensional signals which are denoted by $[s_1, \ldots, s_P]^T$. In the noiseless case, the convolutive mixing process using the matrix $M_{q,p}$ produces $Q \in \mathbb{N}^*$ observed images $[x_1, \ldots, x_Q]^T$ according to the following convolutive model.

$$x_q(n_1, n_2) = \sum_{p=1}^{P} \sum_{k_1=-L}^{L} \sum_{k_2=-L}^{L} M_{q,p}(k_1, k_2)s_p(n_1 - k_1, n_2 - k_2)$$  \hspace{1cm} (3)$$

where the double-index notation $(n_1, n_2)$ designates a pixel position in the image. The unmixing process consists in inverting the operation described above as indicated in Figure 1. It involves $((2L + 1) \times (2L + 1))$-matrices denoted by $W_{q,p}$ which operate according to the following equation.

$$y_p(n_1, n_2) = \sum_{q=1}^{Q} \sum_{k_1=-L}^{L} \sum_{k_2=-L}^{L} W_{q,p}(k_1, k_2)x_q(n_1 - k_1, n_2 - k_2)$$  \hspace{1cm} (4)$$

where $y_p$ denotes an estimated image sources. According to (4), the process is inverted by computing estimates of $W_{q,p}$, $q : 1, \ldots, Q$, $p : 1, \ldots, P$ since $x_q$, $q = 1, \ldots, Q$ are known.

## 3 Wavelet-Based Separation Algorithm

The proposed deconvolution algorithm is composed of two stages. In the first stage, the observed images are transformed into the frequency domain using a second generation wavelet transform, namely, the adaptive quincunx lifting scheme (AQLS). This scheme is applied adaptively according to the spatial local activity of the input image as it will be detailed in Section 4. The second stage consists in applying a Nonnegative Matrix Factorization (NMF) algorithm, which plays the role of the unmixing algorithm (see Section 5). An entropy-based objective function is used to guide the transform scheme as well as the NMF algorithm. Accordingly, the adaptive quincunx lifting scheme transforms the input data into the frequency domain. Next, the transformed data are used as input to the NMF algorithm. The latter algorithm is guided by an entropy-based cost function which is well suited to sparse signals. The modifications introduced on the wavelet transform and on the nonnegative matrix factorization, that is the use of entropy, provide a perfect complementarity between the two stages which significantly outperform the effect of a simple combination. Figure 2 depicts the overall block diagram of our method.

The proposed method proceeds by decomposing the mixed signals using the Adaptive Quincunx Lifting Scheme (AQLS). Applying this tool till a given level results in a wavelet packet tree whose nodes correspond to signals decomposed at various scales. The next step consists in locating the node in the wavelet
packet tree that corresponds to the sparsest component. The more sparsest node is chosen based on an entropy criterion. Once the mixtures are denoised, the decomposed data or subimages are used as input for the entropy-based NMF algorithm that will be presented in Section 5. Finally, an inverse AQLS transform is applied to retrieve an estimate of the source images in real spatial domain.

Algorithm 1 summarizes the overall unmixing process whose steps are detailed in the following sections.

**Algorithm 1: AQLS-NMF**

Data: $X$
Result: $M, S$

// Wavelet Transform Decomposition Step
$(a_J, d_1, \ldots, d_J) := \text{AQLS}(X)$

// Denoising Step
$\sigma := 0$
for $j = 1$ to $J$
  $\sigma_j := \text{STEIN_ESTIMATOR}(d_j)$
end for
$\sigma := \text{mean}(\sigma_1, \ldots, \sigma_J)$
$\tilde{a}_J := \text{DENOISE}(a_J, d_1, \ldots, d_J, \sigma)$

// Separation Step
$\tilde{S} := \text{ENTROPY_NMF}(\tilde{a}_J)$

// Inverse Wavelet Transform Step
$S := \text{INVERSE_AQLS}(\tilde{S}, d_1, \ldots, d_J)$
$M := XS^{\dagger}$ // $S^{\dagger}$ denotes the pseudo-inverse of $S$
4 Adaptive QLS for convolutive BSS

The Lifting Scheme (LS) is a valuable tool dedicated to reversible multiresolution representations of images [23, 24]. Using this tool, we describe how the input images are transformed into the frequency domain. In order to get a better extraction of correlation within a mixture image, we propose to extract the spatial and frequency redundancies contained in each mixture. This is done by adapting a recently proposed lifting scheme called Quincunx LS (QLS) [19]. The latter scheme takes into account the bi-dimensional features of the input image mixtures in a straightforward way. As shown in Figure 3, QLS take into account the quincunx sampling lattice of the 2D pixels neighborhood.

![Figure 3: Separable (left) and quincunx (right) sampling lattices.](image)

In what follows, we begin by explaining the concept of nonseparable wavelet decomposition based on Quincunx Lifting Scheme (QLS). Then, we describe the predictive coding of images with optimal prediction coefficients. Finally, we describe an adaptive lifting scheme based on a quadtree decomposition together with a wavelet-based denoising procedure.

### 4.1 The Lifting Concept

The past decades have witnessed significant advances in the field of image processing using the transform based approach, especially wavelet-based methods. Indeed, the first wavelet generation can be viewed as a valuable tool to image processing, especially as a sparse transform [25]. For instance, in [26], the authors has demonstrated that the signal sparseness can be better highlighted in transform domain such as the one obtained by wavelet transform. The advantage of this transform for our context is in its ability to provide a good estimate of the mixing process.

The second wavelet generation has been developed in the context of sparse image representation. It allows a perfect decomposition with reversible reconstruction thanks to a specific structure called lifting scheme (LS). The LS, which involves a nonlinear subband decomposition aims at removing the redundancies
that exist between adjacent spatial samples [27]. The basic idea is to split a
1D signal into even and odd samples. Then, the odd samples are predicted
using the even ones. The resulting prediction error is used to update the even
samples. The extension of the described scheme to 2D signals is possible in the
following two manners.

- **Separable scheme**: the decomposition is applied to all the lines of the
  input image then to the columns of the resulting components. As a conse-
  quence, the image is split into an approximation subband and three detail
  subbands oriented vertically, horizontally and diagonally.

- **Nonseparable scheme**: the lifting scheme is applied to the whole image
  by considering the quincunx polyphase components of the image. Two
  subbands consisting of an approximation subband and a detail subband
  are obtained. The nonseparable scheme is called Quincunx Lifting Scheme
  (QLS).

### 4.2 Quincunx Lifting Scheme

A lot of interest was devoted to QLS because it provides an unified decompo-
sition framework for both images sampled on a quincunx or on a rectangular
grid [28,29]. Besides, the spatial contents of the image can be better taken into
account by nonseparable scheme. These are the reasons why we chosen to use
QLS.

Let us denote by \( x(n_1,n_2) \) the pixels of a image mixture (rectangularly
sampled image), the quincunx sampling provides the two polyphase components
according to the following two equations:

\[
\begin{align*}
  x_{1/2}(n_1,n_2) &= x(n_1 - n_2, n_1 + n_2) \\
  \tilde{x}_{1/2}(n_1,n_2) &= x(n_1 - n_2 + 1, n_1 + n_2)
\end{align*}
\]  

(5)

The pixels in \( x_{1/2} \) and \( \tilde{x}_{1/2} \) correspond, respectively, to the white points and
black points in Figure 3. Then, the prediction step consists in estimating the
\( \tilde{x}_{1/2} \) samples from those of \( x_{1/2} \). The prediction residual coefficients forming
vector \( d_1 \) are used to update the \( x_{1/2} \) samples into an approximation \( a_1 \). Passing
from the resolution 1/2 to the next resolution can, therefore, be expressed as follows:

\[
\begin{align*}
  d_1(n_1,n_2) &= \tilde{x}_{1/2}(n_1,n_2) - x_{1/2}(n_1,n_2)p_{1/2} \\
  a_1(n_1,n_2) &= x_{1/2}(n_1,n_2) + [d_1(n_1,n_2)u_{1/2}]
\end{align*}
\]  

(6)

where

- \( x_{1/2}(n_1,n_2) \) is the vector containing the four neighboring samples of \( x_{1/2}(n_1,n_2) \);
- \( d_1(n_1,n_2) \) is the reference detail vector and
- \( p_{1/2} \) and \( u_{1/2} \) are respectively the prediction and the update coefficient
  vectors.
More generally, for a multiresolution analysis, \( a_{j/2} \) will represent the approximation of the observed mixture at resolution \( j/2, j \in \mathbb{N}^* \). The next step consists in splitting the input signal in order to generate two polyphase components \( x_{j/2} \) and \( \tilde{x}_{j/2} \) as follows:

\[
\begin{align*}
  x_{j/2}(n_1, n_2) &= a_{j/2}(n_1 - n_2, n_1 + n_2) \\
  \tilde{x}_{j/2}(n_1, n_2) &= a_{j/2}(n_1 - n_2 + 1, n_1 + n_2)
\end{align*}
\]

Similarly, the prediction step consists in approximating the \( \tilde{x}_{j/2} \) samples from those of \( x_{j/2} \) ones. Finally, the residual error prediction coefficients forming vector \( d_{j/2} \) are used to update the \( x_{j/2} \) samples into the approximation \( a_{(j+1)/2} \) at the next stage. The described scheme is iterated \( J \) times in order to provide a multiresolution representation. Passing from the resolution \( j/2 \) to the next resolution, i.e., \( (j + 1)/2 \), can be expressed as follows:

\[
\begin{align*}
  d_{(j+1)/2}(n_1, n_2) &= \tilde{x}_{j/2}(n_1, n_2) - \lfloor x_{j/2}(n_1, n_2) p_{j/2} \rfloor \\
  a_{(j+1)/2}(n_1, n_2) &= x_{j/2}(n_1, n_2) + \lfloor d_{j/2}(n_1, n_2) u_{j/2} \rfloor
\end{align*}
\]

where

- \( x_{j/2}(n_1, n_2) \) is the vector containing the four neighboring samples of \( x_{j/2}(n_1, n_2) \);
- \( d_{j/2}(n_1, n_2) \) is the residual prediction vector and
- \( p_{j/2} \) and \( u_{j/2} \) are respectively the prediction and the update coefficient vectors.

The way following which \( p_{j/2}, d_{j/2} \) and \( u_{j/2} \) are computed will be detailed below. Accordingly, the final multiresolution representation of a given image mixture after a recursive decomposition over \( J \) stages is equivalent to the sum of the last approximation \( a_{J/2} \) and all details \( d_{j/2}, 1 \leq j \leq J \).

### 4.3 Predictors Optimization

The performance of the QLS to solve the blind source separation problem is very sensitive to the choice of the predictors used for extracting the dependency and correlation between the samples. In this work, we propose to use entropy-based predictors. Our choice is motivated by the mutual information theory where the entropy has been intensively used as a quantitative measurement for the dispersion of joint distribution and the degree of independence between random variables [30].

Accordingly, in view of having an efficient extraction of dependency and correlation, the predictors should minimize an entropy-based objective cost function. More precisely, we have chosen to design predictors that minimize the weighted entropy \( \mathcal{H}_J \) of the input images. Recall that the weighted entropy is a simple and efficient measure of divergence [31]. For each observation, the weighted entropy \( \mathcal{H}_J \) is the weighted sum of the entropy \( \mathcal{H}(a_{J/2}) \) of the coarsest approximation and the entropies \( \mathcal{H}(d_{j/2}) \) of the detail sub-images at every resolution levels \( j = 1, \ldots, J \):
\[ H_J = \frac{1}{2^J} H(a_{j/2}) + \sum_{j=1}^{J} \frac{1}{2^j} H(d_{j/2}) \]  

Recall that the entropy of a discrete random variable \( X \) taking values in \( \omega = \{x_1, x_2, \ldots, x_n\} \) with probability distribution \( p(X) \) is given by:

\[ H(X) = - \sum_{x \in \omega} p(x) \log p(x) \]

Because \( H_J \) is a nonlinear function of the predictor, its optimization is a delicate task for which a method, based on statistical modelling of the detail subbands, has already been proposed in [19]. In the sequel, we briefly outline this method.

For each mixture, the approximation and details subbands corresponding to the wavelet coefficients are used as input to the rest of the unmixing process. Here, we focus on the optimization of the operator of the prediction step. The details subbands are seen as a continuous random variable \( \xi \) whose realizations are the coefficients of each subband. The distribution of \( \xi \), denoted by \( f \), is modelled by a zero-mean Generalized Gaussian Density (GGD) function expressed as follows:

\[ f(\xi) = \frac{\beta_{j/2}}{2\alpha_{j/2}^{\beta_{j/2} - 1}} \Gamma(\beta_{j/2}) \exp\left(-\frac{\|\xi\|^2}{\alpha_{j/2}}\right) \]  

where \( \Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t} dt \). It is possible to estimate the scale parameter \( \alpha_{j/2} \) and the shape parameter \( \beta_{j/2} \) by a method of the moments [33] or by a maximum likelihood method [34]. The authors of [35] have proposed an explicit approximation where the coefficient entropy is equivalent to the differential entropy given by the following equation.

\[ \tilde{H}(\alpha_{j/2}, \beta_{j/2}) = -E[\log f(\xi)] \]  

Using Equation 10, the approximated differential entropy becomes:

\[ \tilde{H}(\alpha_{j/2}, \beta_{j/2}) = \log \left( \frac{2\alpha_{j/2}^{\beta_{j/2} - 1}}{\beta_{j/2}} \right) + \beta_{j/2}^{-1} \]  

As a consequence, the global entropy of one mixture \( H_J \) given in Equation (9) could be approximated as follows:

\[ H_J \simeq \frac{1}{2^J} H(a_{j/2}) + \sum_{j=1}^{J} \frac{1}{2^j} \tilde{H}(d_{j/2})(\alpha_{j/2}, \beta_{j/2}) \]  

From Equation (13), we deduce that the vector \( p_{j/2} \) that minimize the mixture entropy \( H_J \) also minimizes \( \tilde{H}(d_{j/2})(\alpha_{j/2}, \beta_{j/2}) \) and then maximizes the maximum likelihood.
\( L(p_{j/2}; \alpha_{j/2}, \beta_{j/2}) = \sum_{k=1}^{K_{j/2}} \log \left( f(\tilde{\xi}_{j/2}(k) - \xi_{j/2}(k)p_{j/2}) \right) \) \( (14) \)

where \( \tilde{\xi}_{j/2}(1), \ldots, \tilde{\xi}_{j/2}(K_{j/2}) \) and \( \xi_{j/2}(1), \ldots, \xi_{j/2}(K_{j/2}) \) are realizations of \( \tilde{\xi}_{j/2} \) and \( \xi_{j/2} \). So, the ML estimator of \( p_{j/2} \) can be expressed as:

\[
L(p_{j/2}; \alpha_{j/2}, \beta_{j/2}) = \sum_{k=1}^{K_{j/2}} \log(\beta_{j/2}) - \log(2\alpha_{j/2}\Gamma(\beta_{j/2} - 1)) - \exp \left[ \beta_{j/2} \log \left( \frac{|\tilde{\xi}_{j/2}(k) - \xi_{j/2}(k)p_{j/2}|}{\alpha_{j/2}} \right) \right] \quad (15)
\]

This also amounts to minimizing the following \( \ell^{\beta_{j/2}} \) criterion:

\[
\ell^{\beta_{j/2}}(p_{j/2}) = \sum_{k=1}^{K_{j/2}} |\tilde{\xi}_{j/2}(k) - \xi_{j/2}(k)p_{j/2}|^{\beta_{j/2}} \quad (16)
\]

In general, there is no explicit expression of \( p_{j/2} \) except in the case \( \beta_{j/2} = 2 \). So, we can use an iterative algorithm (like the Gauss-Newton one) in order to minimize the \( \ell^{\beta_{j/2}} \) criterion. Such algorithm is guaranteed to converge to the global minimum of (16) as it is a convex function of \( p_{j/2} \) when \( \beta_{j/2} \geq 1 \).

### 4.4 Adaptive QLS process

The QLS performances can be enhanced by the use of adaptive predictors. As described in [19], the Adaptive QLS (AQLS) considers a block-based adaptation procedure which is coupled with a classified prediction approach. In the context of blind image separation, the adaptive prediction approach uses several pairs of predictor and update operators. The proposed adaptation must depend on the local activity of the observations in order to extract the maximum correlation between them. The adaptation procedure is achieved through a quadtree construction using an entropy-based image splitting rule. If the splitting rule is satisfied, the whole image is subdivided into four blocks. Each block is in turn subdivided into four subblocks if the splitting rule is satisfied again. The quadtree splitting procedure is recursively repeated until a minimum block size is reached. The simulations indicated that the more suitable minimum block size is \( 32 \times 32 \) pixels. In the following, we give an overview of the quadtree construction together with the computation of the QLS predictors giving rise to the following three steps:

- Initially, QLS is applied to every mixed signal \( x_q \) which is represented by the father node \( f_q \). As results, we get four child quadrants \( c_{q,1}, \ldots, c_{q,4} \) for every mixed signal.
Next, the predictors $p_{j/2}(f_q)$ and $p_{j/2}(c_{q,i})$ of the $J$th stage of QLS are computed for each mixture $x_q$, by minimizing the associated $\ell^{2/2}$ function, for $i = 1, \ldots, 4$, $j = 1, \ldots, J$. Let $\mathcal{H}_J(f_q)$ and, $\mathcal{H}_J(c_{q,i})$ denote respectively the entropy of father node $f_q$ and one of its children $c_{q,i}$.

Verify whether the splitting of the current block into its four children results in a more decorrelated data. This is done by checking the splitting rule which resumes to the following condition

$$\frac{1}{4} \sum_{i=1}^{4} \left( \mathcal{H}_J(c_{q,i}) \right) < \mathcal{H}_J(f_q)$$

(17)

If the average entropy of the four children is smaller than the the entropy of father block then the splitting of the current block into its four children blocks is retained. In that case, the same steps are recursively and separately applied to the resulting four children. Otherwise, the splitting of the father block is not advantageous and the father block is viewed as a leaf-node. The procedure is repeated until blocks of size $32 \times 32$ are reached.

The adaptive QLS is achieved by applying the splitting procedure using the condition given by (17). In contrast of the static QLS where the same predictors are used for the input images, adaptive QLS (AQLS) uses, for each correlated region, a specific predictor to detect the maximum amount of correlation between data.

4.5 Wavelet Denoising

Next to the use of the AQLS to achieve the decomposition step, the resulting subimages are as follow: an approximation subimage at the level $J$, which corresponds to the more sparsest node of the wavelet tree and the various details subimages at levels $j$, $1 \leq j \leq J$.

In the noisy case, the latter wavelet decomposition of a noisy mixture results a noisy approximation and noisy details subimages. These subimages are encoded by the noisy values of wavelet coefficients. We assume that the wavelet coefficients of both the approximation and details subimages are corrupted by a random noise distribution. One of the most efficient techniques of wavelet denoising is the one based on a thresholding procedure [36]. In practice, the noisy wavelet coefficients smaller than a given threshold are set to zero and the coefficients above the threshold are either left unchanged (hard thresholding) or reduced to the value of the threshold (soft thresholding). Afterwards, the noise-free image estimate is obtained through an inverse wavelet transform. The reason for effectiveness of such a procedure lies in the fact that in wavelet domain, a small number of high-valued detail coefficients represent areas around sharp transitions (edges) while the bulk of small and close to zero detail coefficients represents the smooth image regions. In others words, this thresholding
principle exploits the fact that the most of wavelet coefficients are zero or very close to zero (details subimages) and only few of these wavelet coefficients are significant (approximation subimage).

Since the noise is uniformly distributed on the whole mixture, it will dominate especially low amplitude coefficients and the denoising task is perfectly achieved by thresholding these coefficients. In this context, two approaches are used: VISUShrink and SUREShrink. As described in [37], VISUShrink uses an universal threshold $T$ whereas SUREShrink calculates the threshold based on Stein estimator. This threshold is obtained by a risk minimization calculated by the Stein formula explained in [38]. Many experimental studies showed that the SUREShrink provides a better results for image denoising, which motivates our choice to use SUREShrink. The denoising process consists in estimating the noise deviation from all detail subimages. The value of the noise deviation is subsequently subtracted from all approximation and details coefficients in order to obtain the denoised mixture.

5 NMF-based unmixing procedure

Matrix factorization is a unifying theme in numerical linear algebra. A wide variety of matrix factorization algorithms have been developed over many decades, providing a numerical platform for matrix operations such as solving linear systems, spectral decomposition, and subspace identification. Recent work in blind source separation has focused on matrix factorizations that directly target some of the special features of statistical data analysis. Since its first proposition in 1999, the non-negative matrix factorization (NMF) method has demonstrated a high degree of efficiency to solving various problems in many areas. Among these areas is the blind source separation (BSS) [39–41].

The NMF decomposition assumed that the observed mixtures, the mixing matrix and the source matrix are nonnegative. Since only the observation matrix $X$ is known, the goal of the factorization is to decompose $X$ into two matrices which correspond to the original source matrix and the mixing matrix [42, 43]. More formally, given a non-negative observation matrix $X$, the aim is to find nonnegative estimated mixing matrix $\hat{M}$ and an estimated source matrix $\hat{S}$ such that:

$$X \simeq \hat{M} \hat{S}$$

In the present work, the solution of the Equation (18) is optimized according to the entropy-based objective function. Using the the generalized gaussian distribution modelling, the input observation $X$ and the estimation $\hat{M} \hat{S}$ are regarded as normalized probability distributions. Once the AQLS is applied to the observation matrix, the unmixing task is achieved using a NMF algorithm enhanced by an entropy-based divergence measurement between $X$ and $\hat{M} \hat{S}$. The entropy $\mathcal{H}(X||\hat{M} \hat{S})$, introduced in [43] measures the divergence between
the observed mixtures and the product of the estimates of $M$ and $S$. This entropy can be calculated as follow:

$$
H(X||\tilde{M}\tilde{S}) = \sum_{i,j} (X_{i,j} \log \left( \frac{X_{i,j}}{\tilde{M}_{i,j}\tilde{S}_{i,j}} \right) - X_{i,j} + \tilde{M}_{i,j}\tilde{S}_{i,j})
$$

(19)

The divergence decreases by minimizing the value of the entropy. An outline of the entropy-based NMF algorithm is given in Algorithm 2.

**Algorithm 2: entropy_NMF**

**Data:** $X$

**Result:** $\tilde{M}, \tilde{S}$

$(p,m) := \text{size}(X)$

// Initializing step

$\tilde{M} := \text{rand}(P,P)$ // $P$ is equal to the number of sources

$\tilde{S} := \text{rand}(P,C)$ // $C$ is equal to the number of columns in $X$

$h := H(X||\tilde{M}\tilde{S})$ // $h$ is initial entropy divergence value between $X$ and $\tilde{M}\tilde{S}$ calculated according to Equation 19

repeat

// Update step using the formula of entropy divergence

for $i=1$ to $P$

    for $j=1$ to $P$

        $\tilde{m}_{i,j} := \tilde{m}_{i,j} \frac{\tilde{s}_{i,k}x_{j,k}}{\sum_k \tilde{s}_{j,k}}$

    end for

    for $i' = 1$ to $C$

        $\tilde{s}_{i,i'} := \tilde{s}_{i,i'} \frac{\sum_k \tilde{m}_{k,i}x_{k,i'}}{\sum_k \tilde{m}_{i,k}}$

    end for

end for

$h := H(X||\tilde{M}\tilde{S})$

// $h$ is the entropy divergence value between $X$ and $\tilde{M}\tilde{S}$ calculated according to Equation 19

until $h > \epsilon$

Return $\tilde{M}$ and $\tilde{S}$

**6 Experimental Results**

In this section, we consider two types of mixed images: noiseless and noisy mixed images. In a first experiment, we used four natural test images extracted from a standard image database. A second experiment is carried out in order to test our algorithm on images from a differing area. We therefore used four MRI images which are rather sparse, that is, contain many dark pixels. These
test images are convolutedly mixed using a set of mixing matrices generated at random. This set contains various types of matrices provided by the NMFLAB platform [3]. The elements of these matrices are normally distributed with zero mean and unit variance. As described by Equation (3), before mixing the source images, each source pixel is weighted and mixed by its \(L\)-pixels distance neighbors. This requires using a \((2L + 1) \times (2L + 1)\) matrix as mask filter. Next, all delayed images are mixed using the mixing matrices. To obtain hard to separate mixtures, we used ill-conditional random matrices, Hilbert, Toeplitz and Hankel mixing matrices. In the noisy case, we assumed that the mixed images are corrupted by an Additive White Gaussian Noise (AWGN) whose power is set to 40 dB.

Figure 4 shows one of the four mixed images of each experiment where the subfigures a, b, c and d depict respectively one mixture of the noise-free mixed natural images, noisy mixed natural images, noise-free mixed MRI images and noisy mixed MRI images.

The separation performance of the proposed algorithm, that is AQLS-NMF, was compared with the performance of the fixed-point ICA algorithm (FPICA), which operates in the complex domain. We emphasize that FPICA is dedicated to the unmixing of convolutive mixtures [14]. We also compared AQLS-NMF against both the decorrelation-based algorithm proposed in [16] and the multichannel extension of nonnegative matrix factorization (MCHNMF) also dedicated to unmixing convolutedly mixed [15]. We used the following index as criterion to measure the performances of the various algorithms:

- The separability index (SI) and Amari index (AI) which quantify the separating accuracy [3].
- The PSNR as a quality measurement of the output image compared with original one.

Table 1 contains the average values using four mixing matrices provided by NMFLAB indicated above. In this table, we reported a comparative study between AQLS-NMF, FPICA, MCHNMF and FDBSS. For the two types of images, the results indicate that AQLS-NMF provides very competitive performances in noise free and noisy mixtures. On natural images, for example, AQLS-NMF outperformed the other two algorithms by a difference of \(\pm 0.12\) in terms of SI from noise free and noisy mixtures, a difference of 0.18 and 0.13 in terms of AI respectively on noise free and noisy mixtures and provided a gain of \(\pm 12\) dB in term of reconstructed image quality (PSNR). The results obtained on MRI images are also in favor of AQLS-NMF, which has given the best values for both separability and image quality index.

In addition, to illustrate the visual aspect of the various results, we also report the separated images. Figure 5 and Figure 6 show the estimated source images obtained by AQLS-NMF. These two figures reveal a very successful separation since the PSNR is always superior to 50 dB in the two types of images, i.e., natural and MRI images. The latter value is considered as an indicator of a perfect reconstruction because conventionally, beyond 50 dB, the image is
Figure 4: All Mixed Images
(a) BARBARA PSNR = 58.19 dB
(b) BOAT PSNR = 54.58 dB
(c) LENA PSNR = 56.98 dB
(d) CAMERAMAN PSNR = 55.33 dB

Figure 5: Estimated Natural Images by AQLS-NMF in Noise-Free Case

assumed to have ideal quality. Figure 7 and Figure 8 shows the results of the separation process for both natural and MRI images from noisy mixtures where the PSNR is also superior to 50 dB.

7 Conclusion

This paper presented a convolutive blind source separation method for image mixtures. The method is based on a wavelet packet transform and a NMF unmixing algorithm. The wavelet packet transform is enhanced by a non separable and adaptive version, resulting in a technique called Adaptive Quincunx Lifting Scheme (AQLS). The AQLS decomposes the mixed images, and the most relevant component is used as input of the unmixing NMF algorithm. Simulation results performed on images from various origins clearly demonstrated the effectiveness of the proposed method even on noisy mixtures.
Figure 6: Estimated MRI Images by AQLS-NMF in Noise-Free Case
Figure 7: Estimated Natural Images by AQLS-NMF in Noisy Case
Figure 8: Estimated MRI Images by AQLS-NMF in Noisy Case

- (a) CSF PSNR = 58.97 dB
- (b) Gray Matter PSNR = 57.91 dB
- (c) Muscle and Skin PSNR = 56.89 dB
- (d) White Matter PSNR = 57.91 dB
Table 1: Overall Comparison of performance separation.

<table>
<thead>
<tr>
<th></th>
<th>Noise Free Natural Mixtures</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Index</td>
<td>FPICA</td>
<td>MCHNMF</td>
<td>FDBSS</td>
</tr>
<tr>
<td></td>
<td>Separability Index</td>
<td>0.3871</td>
<td>0.1783</td>
<td>0.1031</td>
</tr>
<tr>
<td></td>
<td>Amari Index</td>
<td>0.7841</td>
<td>0.3831</td>
<td>0.2782</td>
</tr>
<tr>
<td></td>
<td>Average PSNR (dB)</td>
<td>42.22</td>
<td>43.98</td>
<td>51.44</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Noise Free MRI Mixtures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Index</td>
<td>FPICA</td>
<td>MCHNMF</td>
<td>FDBSS</td>
</tr>
<tr>
<td></td>
<td>Separability Index</td>
<td>0.1745</td>
<td>0.1179</td>
<td>0.0651</td>
</tr>
<tr>
<td></td>
<td>Amari Index</td>
<td>0.2695</td>
<td>0.1622</td>
<td>0.1057</td>
</tr>
<tr>
<td></td>
<td>Average PSNR (dB)</td>
<td>44.61</td>
<td>47.21</td>
<td>51.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Noisy Natural Mixtures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Index</td>
<td>FPICA</td>
<td>MCHNMF</td>
<td>FDBSS</td>
</tr>
<tr>
<td></td>
<td>Separability Index</td>
<td>0.4169</td>
<td>0.1852</td>
<td>0.1161</td>
</tr>
<tr>
<td></td>
<td>Amari Index</td>
<td>0.8382</td>
<td>0.4177</td>
<td>0.3188</td>
</tr>
<tr>
<td></td>
<td>Average PSNR (dB)</td>
<td>41.16</td>
<td>42.88</td>
<td>46.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Noisy MRI Mixtures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Index</td>
<td>FPICA</td>
<td>MCHNMF</td>
<td>FDBSS</td>
</tr>
<tr>
<td></td>
<td>Separability Index</td>
<td>0.2001</td>
<td>0.1213</td>
<td>0.0947</td>
</tr>
<tr>
<td></td>
<td>Amari Index</td>
<td>0.2871</td>
<td>0.1719</td>
<td>0.1163</td>
</tr>
<tr>
<td></td>
<td>Average PSNR (dB)</td>
<td>43.69</td>
<td>46.55</td>
<td>49.91</td>
</tr>
</tbody>
</table>
References


23


