

WAVELET BASED NONLINEAR SEPARATION OF IMAGES

Mariana S. C. Almeida and Luís B. Almeida

Instituto de Telecomunicações
Instituto Superior Técnico
Av. Rovisco Pais – 1, 1049-001 Lisboa, Portugal
mariana.almeida@lx.it.pt, luis.almeida@lx.it.pt

ABSTRACT

This work addresses a real-life problem corresponding to the separation of the nonlinear mixture of images which arises when we scan a paper document and the image from the back page shows through.

The proposed solution consists of a non-iterative procedure that is based on two simple observations: (1) the high frequency content of images is sparse, and (2) the image printed on each side of the paper appears more strongly in the mixture acquired from that side than in the mixture acquired from the opposite side.

These ideas had already been used in the context of nonlinear denoising source separation (DSS). However, in that method the degree of separation achieved by applying these ideas was relatively weak, and the separation had to be improved by iterating within the DSS scheme. In this paper the application of these ideas is improved by changing the competition function and the wavelet transform that is used. These improvements allow us to achieve a good separation in a single step, without the need to integrate the process into an iterative DSS scheme. The resulting separation process is both nonlinear and non-local.

We present experimental results that show that the method achieves a good separation quality.

1. INTRODUCTION

If we scan or photograph a paper document, the image from the back page often appears superimposed on the image from the front page, especially if the paper is thin or rather transparent. Here we deal with a difficult version of this problem, corresponding to the use of onion skin paper. This creates a strong nonlinear mixture. In our case the mixture images were obtained by printing photos and/or text on both sides of a sheet of onion skin, and then using a common desktop scanner to scan both sides of the onion skin. A more complete description of the way in which the mixture

images were obtained is given in [1]. The source and mixture images are available online at <http://www.lx.it.pt/~lbalmeida/ica/seethrough/index.html>, and are shown in Fig. 2.

Extracting the source images from the scanned mixtures is a nonlinear blind source separation (BSS) problem. This kind of problem is often approached by assuming that the sources are statistically independent from each other, which then justifies the use of an Independent Component Analysis (ICA) method. Linear ICA is a well understood problem. One of its main properties is the essential uniqueness of the solution [2]. Nonlinear ICA is still much less studied, however, and suffers, in general, from non-uniqueness of the solution. The problem studied in this paper involves real-life nonlinear, noisy mixtures of images, some of which are not independent from each other. This non-independence, together with the nonlinear character of the mixture, affects the quality of the results of ICA-based methods [1].

The image separation problem that we study here has been previously addressed by two other methods. The one described in [1] was based on the MISEP method of nonlinear ICA [3], using a multilayer perceptron as separating system. The other one [4] was based on the nonlinear denoising source separation (DSS) method, which uses some of the basic ideas that are used in this paper, albeit in a less efficient manner. The same image separation problem is also addressed in another paper in this conference [5]. That paper presents a solution based on a physical model of the mixture process, the model's parameters being trained by means of an ICA criterion through the MISEP method.

In contrast to ICA-based methods, the solution proposed in this paper does not assume independence of the sources. It uses, instead, other properties of images and of the mixture process. More specifically it uses, on the one hand, the sparsity of the coefficients of the wavelet decomposition of images and, on the other hand, the fact that, in this kind of mixture, each source image appears more strongly in the mixture acquired from the side of the paper in which that image is printed than in the mixture acquired from the opposite side. Since no independence assumption is made, this

This work was partially supported by FCT project POSC/EEA-CPS/61271/2004.

method is able to separate images that are not independent from each other.

The separation method proposed here is similar to the denoising step of the nonlinear DSS method described in [4]. However, while that denoising step yielded only a partial separation which had to be improved by iterating within the DSS scheme, the improvements proposed here allow a rather complete separation to be performed in a single step, without the need for any iterative DSS procedure. In contrast to nonlinear DSS, and also to the ICA-based methods mentioned above, which perform a pixel-by-pixel mapping, the method proposed here is non-pointwise, due to its use of wavelets.

This manuscript is organized as follows: Section 2 describes the separation method. Section 3 presents experimental results. Section 4 concludes and presents future research directions.

In the printed version of this paper, small details of the images are not visible. However, in the electronic version, images can be zoomed into, to view small details.

2. SEPARATION METHOD

2.1. High frequency competition

The separation procedure that we propose manipulates the images through a wavelet based representation, and uses only very basic information about the sources and the mixture process. More specifically, the information that is used consists of two observations:

1. The high frequency components of common images are sparse. In a wavelet based representation this is translated into the fact that wavelet coefficients have sparse distributions [6]. As a consequence, the wavelet coefficients from two different source images will seldom both have significant values in the same image location.
2. Each source is more strongly represented in the image acquired from the side of the paper in which that source is printed than in the image acquired from the opposite side.

The separation method is summarized in Fig. 1. After performing wavelet decomposition for a certain number of levels, we perform a competition between the corresponding wavelet coefficients from the two mixture images, according to

$$\sigma_i = \sqrt{\langle x_i^2 \rangle} \quad (1)$$

$$m_i = \frac{1}{1 + \exp\left(-a \frac{x_i^2 - x_{3-i}^2}{x_i^2 + x_{3-i}^2}\right)} \quad (2)$$

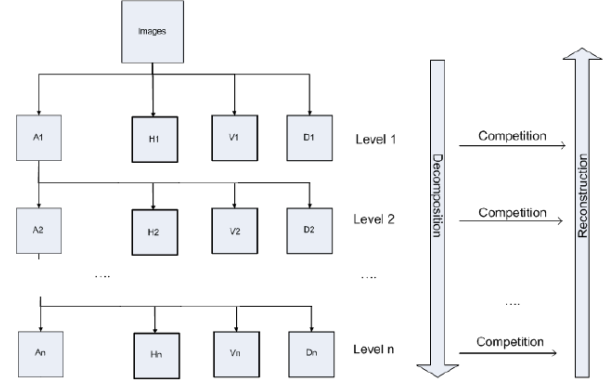


Fig. 1. Schematic representation of the wavelet-based separation method.

$$y_i = \sigma_i \frac{x_i m_i}{\sqrt{\langle x_i^2 m_i^2 \rangle}} \quad (3)$$

where $i \in \{1, 2\}$ indexes the two sides of the paper, x_i are the wavelet coefficients of a given type (for example vertical coefficients at the first decomposition level) of the decomposition of the i th mixture image, x_{3-i} are the corresponding coefficients from the other mixture image, and y_i are the corresponding coefficients that are used for synthesizing the i th separated image; $\langle \cdot \rangle$ denotes the mean computed across the coefficients of a given type, and a is a parameter that controls the strength of the competition.

This competition was applied to all horizontal, vertical and diagonal wavelet coefficients at all decomposition levels (represented, in Fig. 1, by blocks H_j – horizontal coefficients at level j , V_j – vertical coefficients at level j , and D_j – diagonal coefficients at level j). The competition computes mask m_i through a soft *winner take all* function given by Eq. (2), and then applies this mask to the wavelet coefficients of the mixture image, so that the coefficients are much intensified in the image in which they originally were strongest and much weakened in the image in which they were weakest.

The separated images were synthesized using, at all levels (except for the low frequency coefficients of the deepest level), the wavelet coefficients computed by this form of competition. For the low frequency coefficients at the deepest level (A_n in Fig. 1) we have used, in some cases, the coefficients from the corresponding mixture image and, in other cases, a value of zero, as detailed ahead.

2.2. Improvements in the separation method

One of the important aspects of the separation method has to do with the choice of the wavelet representation. The commonly used decimated wavelet transform showed not to

be very appropriate for the task at hand, leading to a rather incomplete separation. This was probably due to its shift-varying character, which makes it represent edges better or worse depending on their exact location. We tried two different wavelet transforms, to circumvent this limitation:

- The discrete complex wavelet transform [7], which is almost shift-invariant and is rather directionally selective. This transform has the disadvantage of having to use wavelets with a relatively large support, which make it less effective in handling the finer details of images.
- The stationary discrete wavelet transform [8], which is shift-invariant and can use short-support wavelets (e.g. Haar), at the cost of using a largely redundant representation, which translates into a somewhat higher computational cost.

An important improvement of the proposed separation method relative to the denoising step used in nonlinear DSS [4] consists of the use of a shift-invariant (or almost shift-invariant) wavelet transform instead of a decimated one. Another important improvement consists of the use of the competition procedure described in Eqs. (1-3). Together, these two improvements make it possible to perform the separation in a single step, just by means of the wavelet-based competition. The method avoids both the use of the multilayer perceptron and of the iteration that were needed in nonlinear DSS. Being non-iterative, the method is much more efficient than both nonlinear DSS and the ICA-based methods mentioned above.

2.3. Preprocessing

While the separation method that we've described can be directly applied to the mixture images, it makes sense to perform a partial separation by a linear procedure, as a preprocessing step. Many linear ICA methods use decorrelation for this purpose. In our case we were constrained by the fact that the mixture process was symmetrical (i.e. both sides of the paper were handled in virtually identical ways). Therefore the preprocessing procedure should also be symmetrical. Symmetry, in this context, means that the matrix that multiplies the mixture vector, Q , should obey $q_{12} = q_{21}$ and $q_{11} = q_{22}$. A decorrelation matrix doesn't normally obey the latter condition. We used a "quasi-decorrelating" matrix given by

$$Q = \begin{bmatrix} \frac{a_{11}+a_{22}}{2} & a_{12} \\ a_{21} & \frac{a_{11}+a_{22}}{2} \end{bmatrix}, \quad (4)$$

where a_{ij} are the elements of A , which is a decorrelating matrix given by the square root of the autocovariance matrix of the mixture data (A already obeys $a_{12} = a_{21}$).

3. EXPERIMENTAL RESULTS

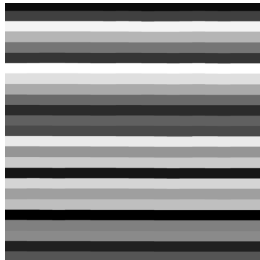
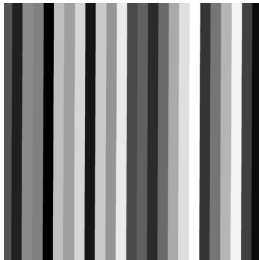
The separation method described in Section 2 was applied to the five pairs of mixtures shown in Fig. 2, both with and without the preprocessing described in Section 2.3. In the reconstruction of each separated image we used, for the low frequency coefficients (A_n in Fig. 1) the coefficients from the corresponding preprocessed image, in the cases in which preprocessing was used. In the cases in which no preprocessing was used, we have set $A_n = 0$ for the reconstruction. This yielded somewhat better results than keeping the coefficients from the corresponding mixture image. The value of parameter a used in Eq (2) was 510, which corresponds to a strong competition, the mask implementing almost a hard winner take all function.

We performed separations both with the complex wavelet transform and with the stationary wavelet transform with Haar wavelets. The depth of wavelet analysis was 7 for pairs 1, 2 and 3, and 8 for pairs 4 and 5. Using a 1.6 MHz Pentium-M (Centrino) processor, the separations of pairs 1 to 3 took approximately 3 seconds each with the complex wavelet transform, and 20 seconds each with the stationary wavelet transform. For pairs 4 and 5, which consist of larger images and which used a deeper wavelet analysis, the separations took approximately 14 seconds with the complex wavelet transform and 3.5 minutes with the stationary transform.

The best separation results were obtained with the stationary wavelet transform with preprocessing, and are shown in Fig. 3. Due to lack of space we only present all the separation results for this variant of the method. For other variants (without preprocessing and/or with complex wavelets) we only present a subset of the results. Results obtained with complex wavelets with preprocessing are shown in Fig. 4. Results obtained without preprocessing are shown in Figs. 5 and 6.

The quality of the results obtained with stationary wavelets and preprocessing (Fig. 3) is quite good. The main imperfection is a decrease in the contrast of each image in regions where the other image is darker. This is probably due to the fact that, in each mixture, the contrast of each source is strongly reduced in areas where the other source is dark. Since the separation method only performs a competition, without any attempt to compensate for this contrast reduction, the contrast variation appears in the separated images.

The results obtained with complex wavelets and preprocessing (Fig. 4) are still quite good, although they show somewhat stronger imperfections. The results obtained without preprocessing (Figs. 5 and 6) show significantly stronger degradations.



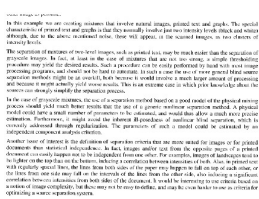
Pair #1



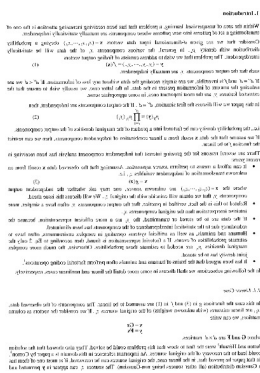
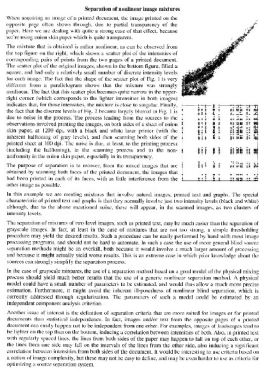
Pair #2



Pair #3



Pair #4



Pair #5

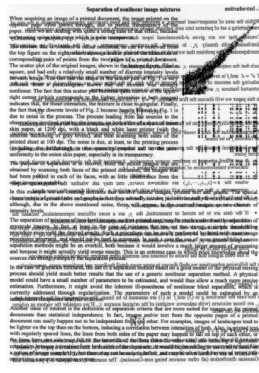
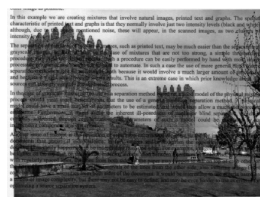
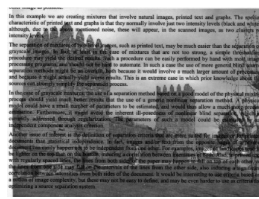
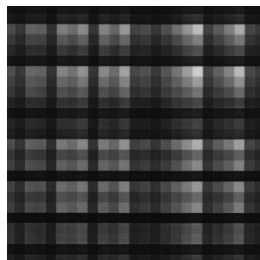
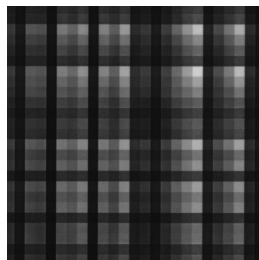


Fig. 2. Images used in the tests. Left: Source images. Right: Mixture images.



In this example we are using images that involve several images, printed text and graphs. The typical characteristics of printed text and graphs is that they normally consist of two primary levels: black and white. Although, due to the image acquisition noise, there will appear in the scanned images, as low contrast of the primary levels.

The acquisition of images of printed text, such as printed text, may be carried out using the operation of different images. In fact, at least in the case of images that are not too noisy, a simple thresholding procedure may yield the desired results. As a procedure that is easy to perform, the fast Fourier transform (FFT) is used to process the images. The images are then processed by the fast Fourier transform (FFT) to obtain the frequency components. The images are then processed by the fast Fourier transform (FFT) to obtain the frequency components. The images are then processed by the fast Fourier transform (FFT) to obtain the frequency components.

When applying an image of a printed document, the image appears as the upper part of the image. The image is then processed by the fast Fourier transform (FFT) to obtain the frequency components. The images are then processed by the fast Fourier transform (FFT) to obtain the frequency components. The images are then processed by the fast Fourier transform (FFT) to obtain the frequency components.



In this example we are using images that involve several images, printed text and graphs. The typical characteristics of printed text and graphs is that they normally consist of two primary levels: black and white. Although, due to the image acquisition noise, there will appear in the scanned images, as low contrast of the primary levels.

The acquisition of images of printed text, such as printed text, may be carried out using the operation of different images. In fact, at least in the case of images that are not too noisy, a simple thresholding procedure may yield the desired results. As a procedure that is easy to perform, the fast Fourier transform (FFT) is used to process the images. The images are then processed by the fast Fourier transform (FFT) to obtain the frequency components. The images are then processed by the fast Fourier transform (FFT) to obtain the frequency components.

When applying an image of a printed document, the image appears as the upper part of the image. The image is then processed by the fast Fourier transform (FFT) to obtain the frequency components. The images are then processed by the fast Fourier transform (FFT) to obtain the frequency components. The images are then processed by the fast Fourier transform (FFT) to obtain the frequency components.

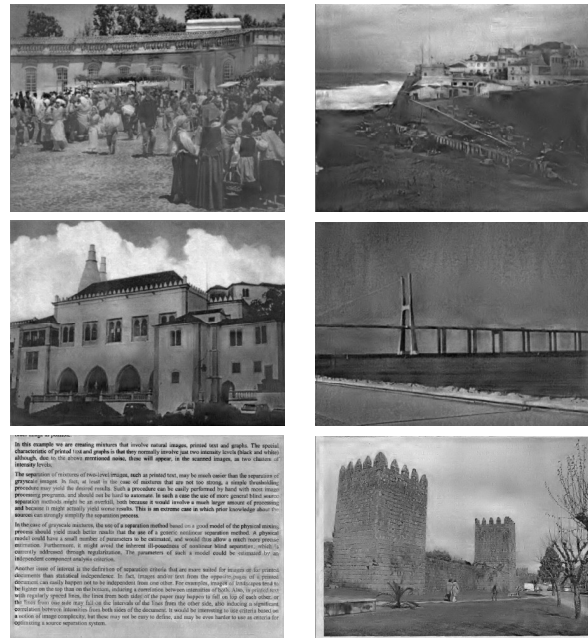


Fig. 4. Results obtained with the complex wavelet transform with preprocessing.

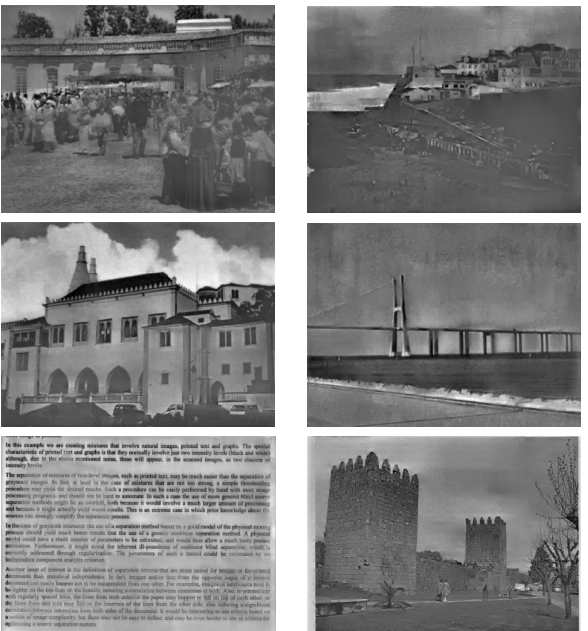


Fig. 5. Results obtained with the stationary wavelet transform without preprocessing.

Fig. 3. Results obtained with the stationary wavelet transform with preprocessing.

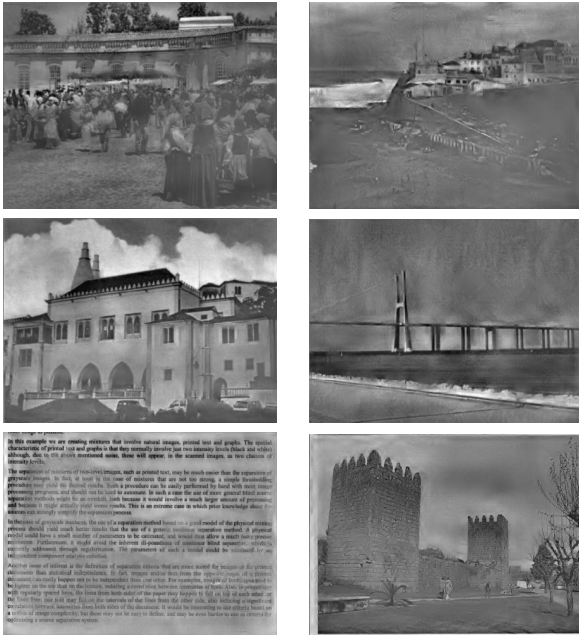


Fig. 6. Results obtained with the complex wavelet transform without preprocessing.

4. DISCUSSION AND FUTURE WORK

We presented a non-iterative method for separating a real-life nonlinear mixture of images. The method is fast and yields images that have a good perceptual separation quality. It does not assume independence of the sources, making use of other properties of images. Therefore the quality of the results is not affected by the possible non-independence of the source images.

The main imperfection of the separation results is a variation in the contrast of each image. This contrast depends on the other image's intensity in the same area. In the future we plan to incorporate a contrast compensation mechanism in the separation method, to eliminate this imperfection.

Contrary to other separation methods that have been presented in the literature, which are pointwise (in the sense that the intensity of each pixel of a separated image depends only on the intensities of the corresponding pixels of the mixture images), the method presented here is not pointwise. If we take into account that the largest Haar wavelets have 128×128 pixels in a 7 level decomposition and 256×256 pixels in an 8 level one, we see that the separation that is performed is strongly non-pointwise. Therefore, pointwise quality measures such as those that have been used with other separation methods [1, 4] are not appropriate for evaluating the perceptual quality of this method's results. We plan to assess the separation quality with other more perceptually oriented measures, such as those proposed in [9, 10].

5. ACKNOWLEDGMENT

We acknowledge the use of the free package available at <http://taco.poly.edu/WaveletSoftware/> for computing complex wavelet transforms.

6. REFERENCES

- [1] L.B. Almeida, "Separating a real-life nonlinear image mixture," *Journal of Machine Learning Research*, vol. 6, pp. 1199–1229, July 2005.
- [2] P. Comon, "Independent component analysis – a new concept?," *Signal Processing*, vol. 36, pp. 287–314, 1994.
- [3] L.B. Almeida, "MISEP – Linear and nonlinear ICA based on mutual information," *Journal of Machine Learning Research*, vol. 4, pp. 1297–1318, 2003.
- [4] M.S.C. Almeida, H. Valpola, and J. Särelä, "Separation of nonlinear image mixtures by denoising source separation," in *Independent Component Analysis and Blind Signal Separation*, J. Rosca, D. Erdogmus, J. Principe, and S. Haykin, Eds. 2006, vol. LNCS 3889, pp. 8–15, Springer-Verlag.
- [5] M.S.C. Almeida and L.B. Almeida, "Separating nonlinear image mixtures using a physical model trained with ica," in *IEEE Int. Worksh. on Machine Learning for Signal Processing*, Maynooth, Ireland, 2006.
- [6] P. Moulin and J. Liu, "Analysis of multiresolution image denoising schemes using generalized-gaussian and complexity priors," *IEEE Transactions on Information Theory*, vol. 14, no. 3, pp. 909–919, April 1999.
- [7] I.W. Selesnick, R.G. Baraniuk, and N.G. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Processing Magazine*, pp. 123–151, November 2005.
- [8] G.P. Nason and B.W. Silverman, "The stationary wavelet transform and some statistical applications," in *Wavelets and Statistics*. 1995, Lecture Notes in Statistics, pp. 281–299, Springer-Verlag.
- [9] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, April 2004.
- [10] H. Sheikh, A. Bovik, and G. Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," *IEEE Trans. Image Processing*, vol. 14, no. 12, pp. 2117–2128, December 2005.