Neurocomputing 72 (2008) 57–70

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Wavelet-based separation of nonlinear show-through and bleed-through image mixtures

Mariana S.C. Almeida *, Luís B. Almeida

Instituto de Telecomunicações, Instituto Superior Técnico, Av. Rovisco Pais, 1, 1049-00 Lisboa, Portugal

ARTICLE INFO

Available online 1 October 2008

Keywords:
Nonlinear separation
Image mixture
Show-through
Bleed-through
Image enhancement

ABSTRACT

This work addresses the separation of the nonlinear real-life mixture of images that occurs when a page of a document is scanned or photographed and the back page shows through. This effect can be due to partial paper transparency (*show-through*) and/or to bleeding of the ink through the paper (*bleed-through*). These two causes usually lead to mixtures with different characteristics.

We propose a separation method based on the fact that the high-frequency components of the images are sparse and are stronger on one side of the paper than on the other one. The same properties were already used in nonlinear denoising source separation (DSS). However, we developed significant improvements that allow us to achieve a competitive separation quality by means of a one-shot processing, with no iteration. The method does not require the sources to be independent or the mixture to be invariant, and is suitable for separating mixtures such as those produced by bleed-through, for which we do not have an adequate physical model.

1. Introduction

This paper focuses on the separation of two-image mixtures that occur in a well known practical situation: when we scan or photograph a document and the back page shows through. This effect is often due to partial transparency of the paper (which we designate by *show-through*). Another possible cause is bleeding of ink through the paper, a phenomenon that is more common in old documents, in which the ink has had more time to bleed. The latter phenomenon is commonly designated by *bleed-through*. The two phenomena may be simultaneously present in the same document.

In this work we use, as test examples, three different kinds of mixtures. The first kind essentially only contains the show-through effect: five pairs of images were printed on the two sides of five sheets of tracing paper[1] which, due to its high transparency, creates very strong mixtures. The second type corresponds to an old manuscript letter written in very thin "air-mail" paper (also called *onion skin paper*, which is rather transparent, causing show-through to occur). This document also has some areas in which bleed-through appears to have occurred. The third kind of mixture corresponds to images of old manual transcripts of music (partitures), which mostly contain the bleed-through effect. For each document, scanning or photographing both sides allowed us to obtain two different mixtures of the contents of the two pages. In this paper we address the source separation problem whose aim is to recover, form the two acquired images of each document, the original page images.

Show-through is known to lead to nonlinear mixtures [2,5,3]. A physical model of the show-through mixture of gray-level images printed with halftoning has been presented in [3]. Bleed-through probably is a much more complex phenomenon, which is much harder to model.

Source separation is often performed by assuming that the sources are statistically independent from each other, an assumption which leads to the use of independent component analysis (ICA) techniques. While linear ICA is a well studied problem for which several efficient solutions exist [6,7,27], nonlinear ICA is a much less studied problem [1,13,10,8]. Nonlinear ICA has the additional difficulty of being ill-posed, having an infinite number of solutions without any simple relationship with one another [14,11]. The mixtures addressed in this work are nonlinear and noisy, and the letter and partiture mixtures are spatially variant. Besides these challenging properties, most of the sources studied in this work do not completely obey the independence assumption, a fact which affects the quality of the results obtained through ICA-based methods [2,5].

Instead of assuming independence of the source images, we propose a solution that uses other properties of images and of the

* Corresponding author. Tel.: +351 218418387; fax: +351 218418472.
E-mail addresses: mariana.almeida@lx.it.pt (M.S.C. Almeida), luis.almeida@lx.it.pt (L.B. Almeida).

[1] In previous publications, this tracing paper has been improperly called "onion skin paper". The latter is, for example, the very thin paper commonly used, some decades ago, for air-mail letters, and actually appears in the "air-mail letter" mixture that we used in this work. Tracing paper is the semi-transparent pater often used in professional drawing.

0925-2312/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2007.12.048

mixture process. We use the well known fact that high-frequency components of images are sparse (and that high-frequency wavelet coefficients are also sparse), and we formulate a competition based on the observation that each source is more strongly represented in one of the mixture components than in the other one. Making assumptions that are suited to the present problem, our method achieves a good perceptual separation quality even when the sources are non-independent and the mixture is spatially variant. The separation method that we propose is similar to the denoising step used by nonlinear denoising source separation (DSS) [5]. However, we use an improved form of competition, and also a wavelet transform that is more suited to the problem at hand. These improvements lead to a method that performs the separation in a single step, without the iterative procedure required by nonlinear DSS.

Both the old manuscript letter and the old partitures with the bleed-through effect are addressed for the first time in this paper. On the other hand, the tracing paper mixtures have already been studied in other works [2,5,3]. Contrasting with the method proposed here, which, due to its use of wavelets, performs a non-point-wise transformation, all the other mentioned methods performed point-wise separation. One of them [2] used the MISEP method of nonlinear ICA [1] to train a regularized MLP which performed the separation. In another one [3], MISEP was used to train a nonlinear physical model of the mixture process. Nonlinear DSS has also been applied to some of the tracing paper mixtures [5]. Nonlinear DSS does not assume independence of the sources, but assumes spatial invariance of the mixture. It uses the same basic ideas that are used in this paper, albeit in a less efficient manner. Show-through and/or bleed-through mixtures have also been addressed in [19,22–24,9], but in different settings from the one considered here. In [19,24] separation is archived through linear models, which were shown to be too restrictive to separate tracing paper mixtures [2]. Refs. [22,9] focus only on the restoration of text documents, for which linear separation yields relatively good results (see [2]). In [23], the contents of both pages are assumed to consist of text, and separation is linear and is based on a single color image from one side of the document.

This manuscript is structured as follows: Section 2 describes the three kinds of mixtures that were studied, as well as the processes of image acquisition and alignment. Section 3 describes the proposed separation method. Section 4 presents experimental results, and Section 5 concludes.

The Matlab separation routines and the images used in this work are available at http://www.lx.it.pt/~mscla/. The routines for performing image alignment are available at http://www.lx.it.pt/~lbalmeida/ica/seethrough/.

2. Experimental setup

2.1. Mixtures and acquisition

The method proposed in this work was applied to three kinds of image mixtures. Although they have completely different origins, all the three mixture processes are from real life (i.e., they are not synthetic), and are noisy and significantly nonlinear. Some of them also are spatially variant.

- *Tracing paper images*: Five different pairs of images (including synthetic bars, photos and text) were used as sources of five pairs of mixtures. These pairs of sources, shown in Figs. 1 and 2, were printed on opposite pages of tracing paper, which was chosen for its high transparency. Image printing and scanning were performed as symmetrically as possible, regarding the two images of each pair. Printing was performed with a 1200 dpi laser printer, using the printer's default halftoning system. Both pages were then scanned with a desktop scanner at a low resolution of 100 dpi, which was chosen so that the printer's halftoning grid would not be apparent in the scanned images. Figs. 3 and 4 show the acquired images after the alignment procedure that is described below. The mixtures involved here are highly nonlinear and, except for the pair of text images, linear approaches do not archive reasonable separation results (see [2] for a detailed analysis).
- *Air-mail letter*: The document used in this case was an old, handwritten air-mail letter. Old air-mail paper was very thin, having a high transparency that, as can be seen in the acquired pages shown in Fig. 5, leads to strong mixtures. Since the letter was very thin and old, it was hard to place in the scanner in a good position, without any wrinkles. The letter had three folds and, although the acquisition was preformed as carefully as possible, these folds and some wrinkles could not be eliminated (the wrinkles are especially visible in the bottom corners of the acquired images). This led to a mixture that is not space-invariant, but instead has characteristics that vary from one place to another. This mixture is not compatible with a global mixture/separation model. Also, differently from the tracing paper images, whose middle-tone levels were generated by a halftoning process, the middle-tones in these sources are due to variable transparency of the ink. The acquired images show that the ink was not homogeneously distributed in different words: its density depended on the pressure applied by the writer. Furthermore, in a few areas, the ink seems to have bled through the paper.
  Many of the lines of text of both pages are aligned with one another, which makes the original images non-independent. This is a disadvantage for ICA-based methods, and also a challenge for our method due to the increase in the number of superimposed edges from both pages. The images were acquired using the same desktop scanner that was used for the tracing paper mixtures. However, since the writing ink was blue, the images were acquired in color. They were then converted to grayscale, specifically for this work. Contrary to the tracing paper mixtures, which were specially produced for studies of this kind, we could not access the actual sources corresponding to the air-mail letter.
- *Old partitures*: These mixtures were obtained from very old handwritten partitures, in which bleed-through is the main effect. We used two partiture sheets (see Fig. 6) which were chosen, among those that we had available, for presenting the strongest bleed-through effects. Bleed-through is present, to a significant level, in a few areas of these documents. The images were acquired by a photographic process, and only the photographs were available to us, not the original documents. As with the air-mail letter, they were originally acquired in color and then converted to grayscale. They also show some evidence of wrinkles. And, as in the case of the air-mail letter, we did not have access to the original source images.

2.2. Alignment

Image separation methods usually require that the components of the mixture be precisely registered with one another. This alignment is essential both in ICA-based methods and in our method, which explicitly assumes edges to be in the same spatial position in the two mixture components. To achieve this correspondence, one of the images of each pair had first to be horizontally flipped. After that, an alignment procedure was applied, to correct misalignments due to the different positions of

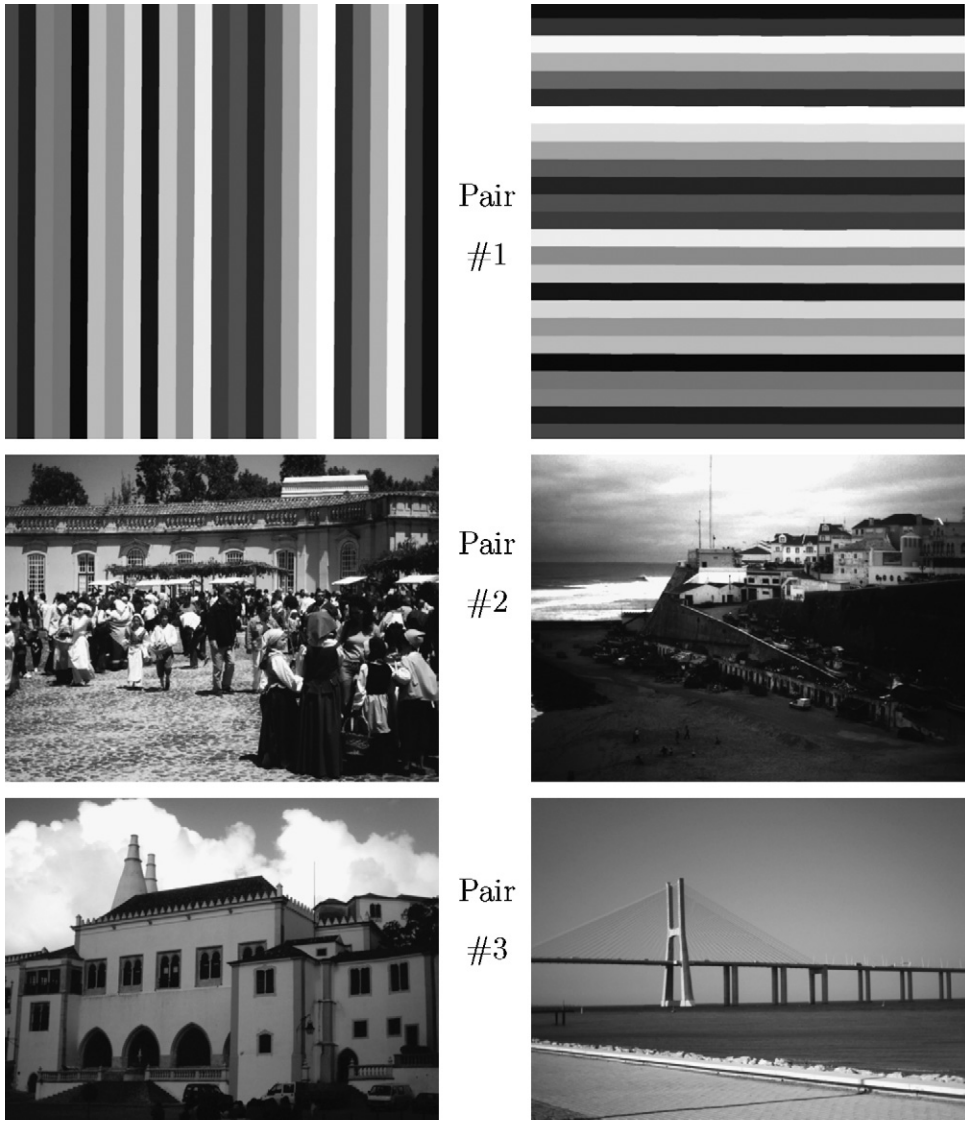


Fig. 1. The first three first pairs of tracing paper source images. In this and all subsequent figures containing images, one of the images of each pair has been horizontally flipped.

the paper during the two scanning acquisitions. In order for it to be precise, the alignment had to be performed locally. This local alignment was needed even for documents that were not wrinkled, probably due to some geometrical imperfections of the scanner. In the case of the air-mail letter, which was significantly wrinkled, the local alignment was even more important.

All mixture pairs were subject to an initial crude alignment, which was performed by hand, with the aid of an image editing program. After this manual alignment, the tracing paper and air-mail letter mixtures were subject to a local alignment procedure, which was specially developed for this problem. Mixture images were first expanded in resolution, by a factor of four in each direction, using bicubic interpolation. Then, one of the images of the pair was divided into 100×100 pixel squares (corresponding to 25×25 pixels in the original image) and, for each square, the best displacement was found, based on the maximum of the cross-correlation with the other image. Aligned images were then reduced to the original resolution. Therefore this alignment method performed a local alignment with a resolution of $\frac{1}{4}$ pixel. More information about the tracing paper images and the alignment process is available in [2]. Since the partitures only presented significant bleed-through in a few areas of the images, showing a weak mixture on most other areas, the alignment routine based on small-block correlations did not work satisfactorily for this case. Therefore, these mixtures were only subject to a careful manual alignment.

3. Separation method

Instead of assuming independence of the source images, the method that we propose uses a property of common images and a property of the mixture process to perform the separation. These properties are:

(1) High-frequency components of common images are sparse. This translates into the fact that high-frequency wavelet coefficients have sparse distributions [16]. Consequently, the high-frequency wavelet coefficients from two different source images will seldom both have significant values in the same image location.

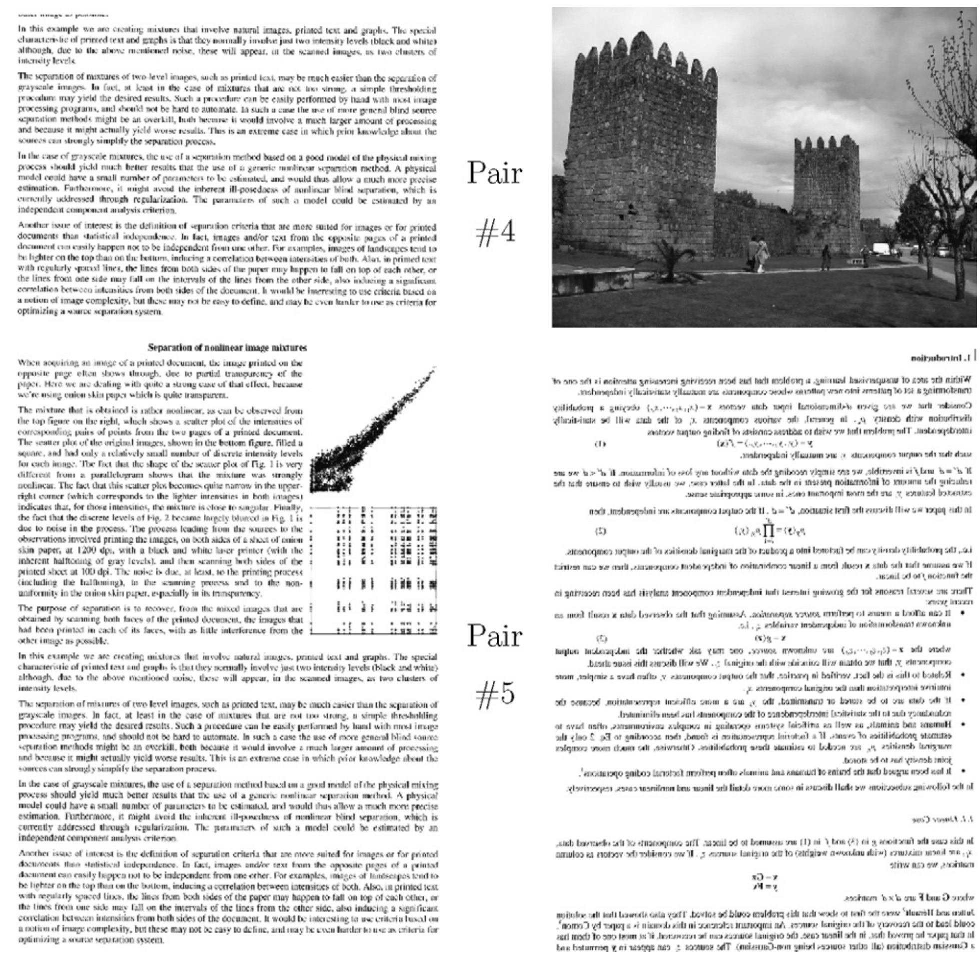


Fig. 2. The fourth and fifth pairs of tracing paper source images.

(2) In the mixture processes considered here, each source is represented more strongly in one of the mixture components (the one acquired from the side where that source is printed or drawn) than in the other component.

3.1. High-frequency competition

The separation method, which is summarized in Fig. 8, manipulated the images through a wavelet-based representation. First, both images were subject to a wavelet decomposition down to a certain depth. Then, the corresponding high-frequency wavelet coefficients of both mixture images were subject to the following competition process:

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

$$y_i = x_i m_i \qquad (2)$$

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) from the ith mixture image, x_{3-i} are the corresponding coefficients from the other image of the same mixture, and y_i are the coefficients that are used for synthesizing the ith separated image; a is a parameter that controls the strength of the competition. The first equation computes a soft winner-take-all mask m_i which is then used, in the second equation, to control the strength of the corresponding high-frequency component in the separated image. This procedure preserves the coefficients that are stronger in the mixture component under consideration than in the opposite component, and weakens the coefficients that are weaker than in the opposite component. Fig. 7 illustrates how the soft winner-take-all mask (1) rules the competition process described in (1)–(2). We should note that the exact definition of the competition mask is not important, as long as it has the general behavior illustrated in Fig. 7.

The competition process described in (1)–(2) was applied to all horizontal, vertical and diagonal wavelet coefficients at all decomposition levels (represented, in Fig. 8, 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 separated images were obtained by wavelet reconstruction using the high-frequency coefficients (H_j, V_j, D_j) after competition. For the low-frequency coefficients (A_n in Fig. 8) we used the coefficients obtained from the decomposition of the corresponding mixture image, with no change.

3.2. Wavelet representation

An important aspect of the proposed 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

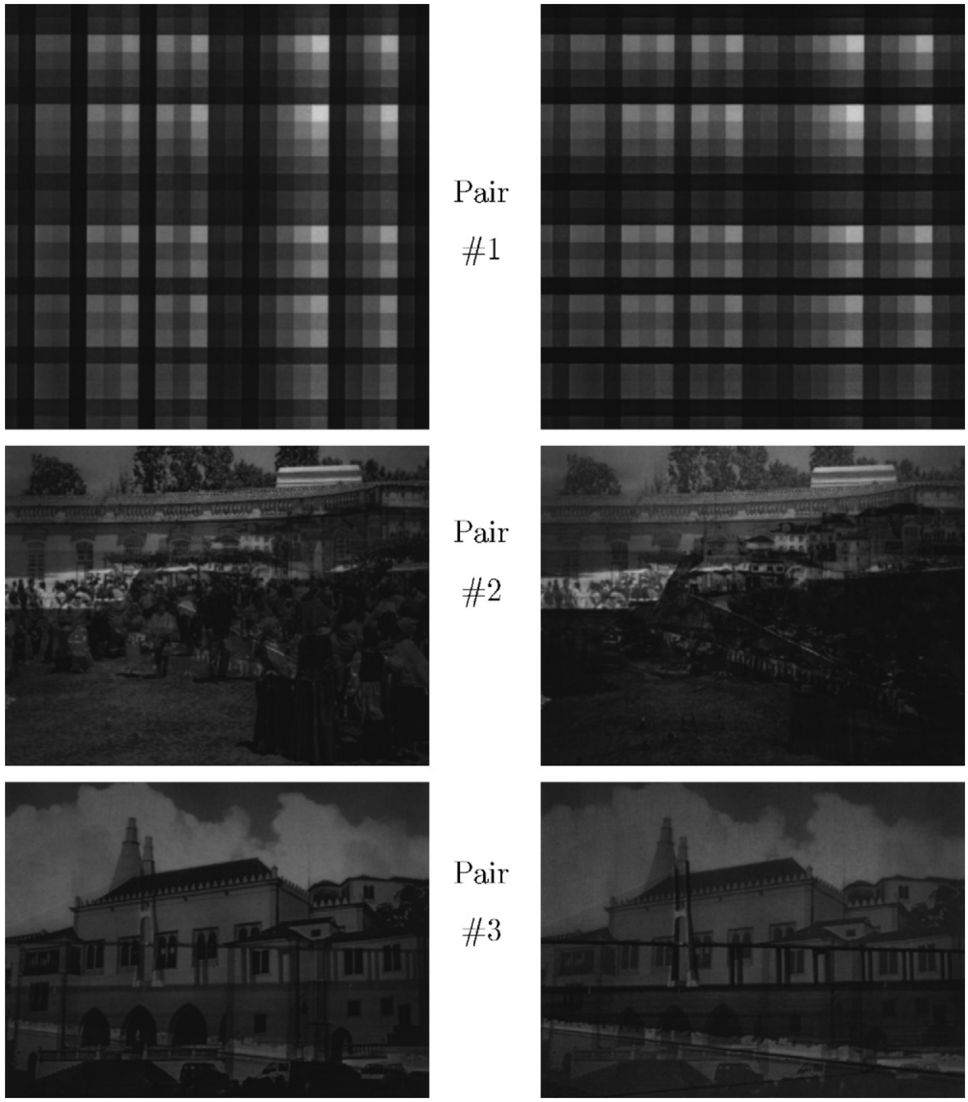


Fig. 3. First three pairs of tracing paper mixtures.

it represent edges better or worse depending on their exact locations. Since this representation did not yield satisfactory results, we tried two different wavelet transforms which circumvent this limitation:

- The stationary discrete wavelet transform [17], which is shift-invariant and can use short-support wavelets (e.g. Haar). This transform uses a very redundant representation, which translates into a somewhat higher computational cost.
- The discrete complex wavelet transform [18], 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, making it less effective in handling the finer details of the images.

3.3. Comparison with nonlinear DSS

The method that we propose is similar to the denoising step used in nonlinear DSS [5], but incorporates two important improvements. One corresponds to the use of a more suitable wavelet transform, which is shift-invariant or almost shift-invariant. The other improvement has to do with the use of a better form of competition. Together, these two improvements led to a one-step procedure that is, by itself, sufficient to separate the images. The proposed method avoids the use of both the multilayer perceptron and the iteration that were required in nonlinear DSS. Being non-iterative, the method is much more efficient than both nonlinear DSS and the ICA-based methods that have previously been proposed.

3.4. Linear preprocessing

Although the separation method that we have described can be directly applied to the mixture images, it can make sense to use a preprocessing step that performs an initial linear decorrelation. This is a form of preprocessing that is used by many linear ICA methods. We used an approximate decorrelation that was constrained to be symmetrical (i.e., it processed both mixture components identically). This constraint is translated into the fact that the matrix Q that multiplies the mixture vector must obey $q_{12} = q_{21}$ and $q_{11} = q_{22}$. It is not possible, in general, to perform an exact decorrelation that is symmetrical in this sense. Designating by A the square root of the autocovariance matrix of the mixture

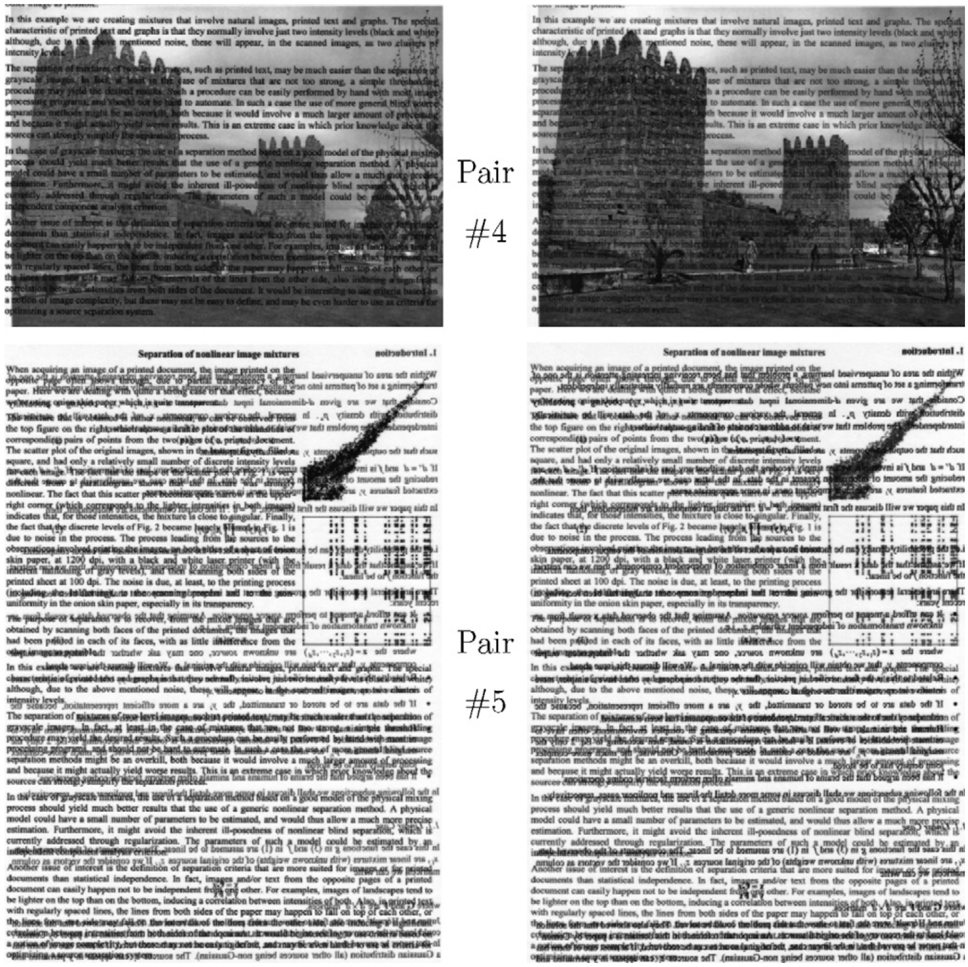


Fig. 4. Fourth and fifth pairs of tracing paper mixtures.

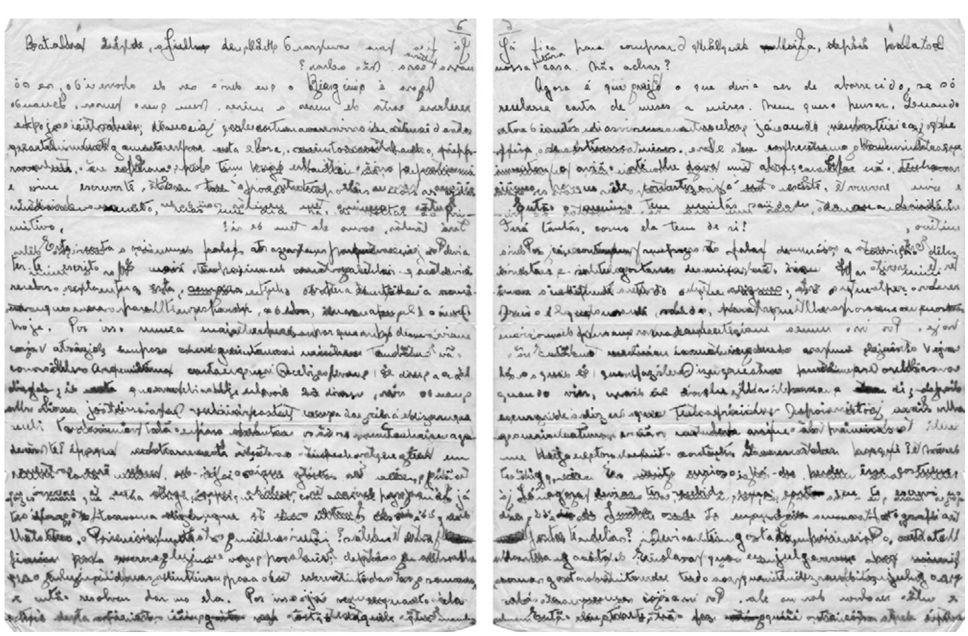


Fig. 5. Acquired images of the old air-mail letter.

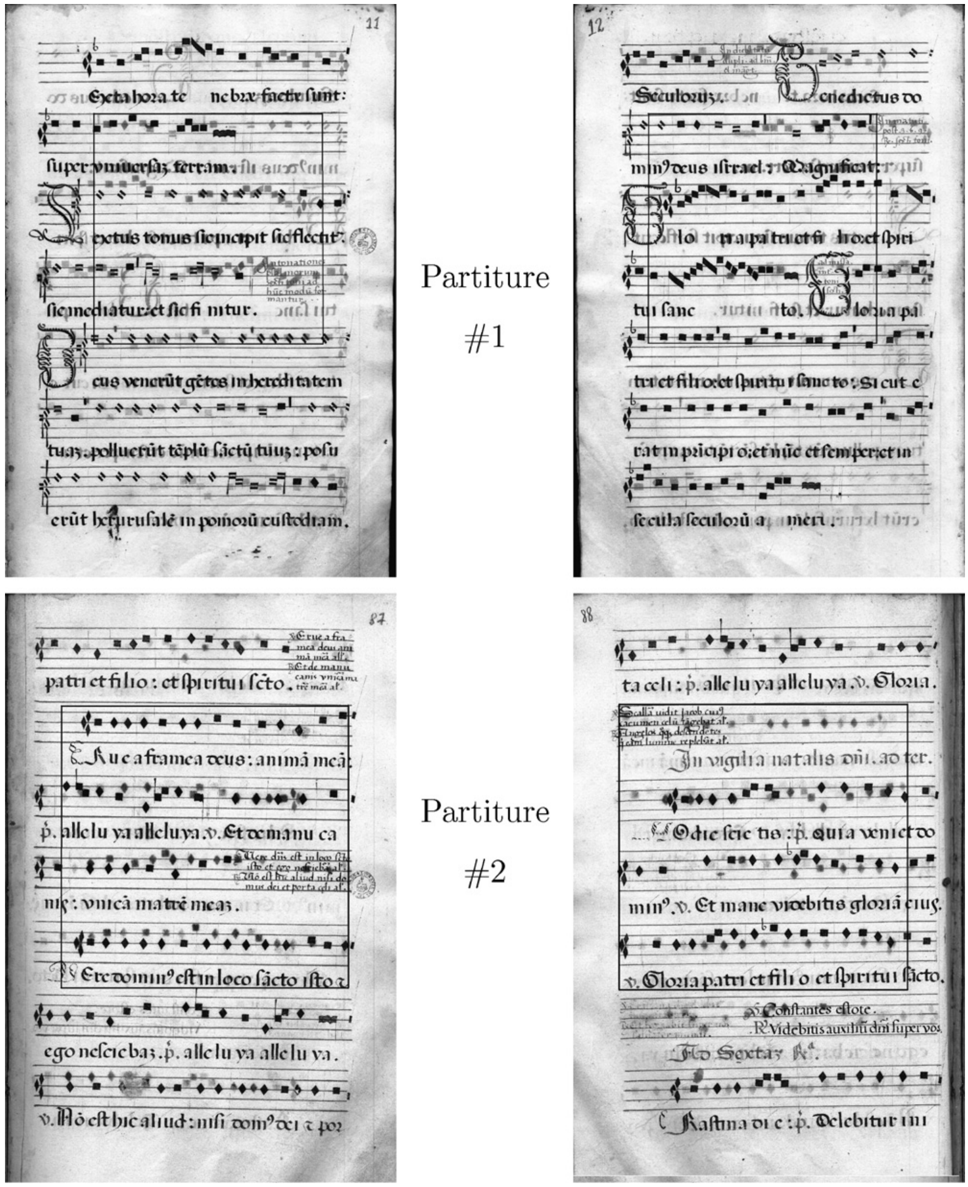


Fig. 6. Acquired images of the old partitures. The squares indicate the areas that were selected for separation.

vector, A satisfies $a_{12} = a_{21}$ but, in general, $a_{11} \neq a_{22}$. The linear preprocessing that we performed used a matrix Q defined by

$$Q = \begin{bmatrix} \dfrac{a_{11}+a_{22}}{2} & a_{12} \\ a_{21} & \dfrac{a_{11}+a_{22}}{2} \end{bmatrix} \tag{3}$$

This preprocessing was only applied to the tracing paper mixtures. Both the manuscript letter and the old manuscript documents contain space-variant mixtures, and this linear preprocessing, which treats the whole mixture equally, was found not to be appropriate for them.

3.5. Contrast compensation

The previously described method tries to identify, through the competition detailed in Section 3.1, which edges correspond to each source. Since edges contain the main information that a human being extracts from an image [21,12], edge identification leads to a good perceptual separation of images. However, the intensity of the reconstructed edges may not be correct. When the separation method described above was applied to the tracing paper mixtures, the contrast of each recovered source image was found to be reduced in the areas where the other source image was darker. This effect is quite visible in Fig. 9, which shows the separation obtained for the fourth pair of tracing paper mixtures without contrast compensation. One can observe a reduction of the contrast of the separated text in the areas where the opposite component is darker (castle and ground areas). This imperfection was to be expected since, in each mixture component, the contrast of each source is lower where the other source is darker, and the competition that we described above does not compensate for this effect: it simply assigns the high-frequency components to the sources, but does not correct their intensities.

We incorporated a contrast compensation mechanism in the separation method, to reduce the imperfections resulting from the interference from the opposite source's intensity. This mechanism compensates that interference by applying a gain that is a function of the other source's estimated intensity.

The contrast compensation uses a parameter ($g_{max} \geqslant 1$) that controls how strong the maximal compensation will be. The gain applied to each estimated source is an affine function of the other source's estimated intensity, varying between 1 (when the other source is white) and g_{max} (when the other source is black). For this contrast compensation we need to use an estimate of the other source's intensity. At each wavelet reconstruction level j, we have access to the low-frequency image of that level (A_j), obtained from the wavelet reconstruction of the lower levels, which were already subject to the competition process. In the proposed method, this low-frequency image is used for estimating the other source's intensity, to control the gain to be applied to each of the high-frequency wavelet coefficients in the reconstruction at level j (referring to Fig. 8, the reconstructed image A_j from one component is used to control the gain applied to H_j, V_j and D_j of the other component). This contrast compensation mechanism is thus applied while the sources are being reconstructed, using the information of lower levels that have already been reconstructed. It leads to a moderate increase of the computational complexity of the separation method, as will be detailed ahead.

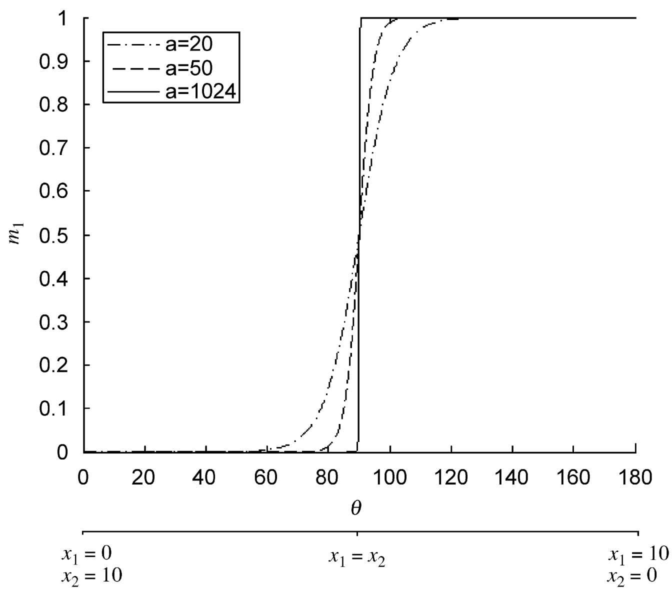


Fig. 7. Behavior of the mask m_1 for different values of parameter a. Focusing on the case where $x_1^2 + x_2^2 = 100$, which is of the order of magnitude found in our images, mask m_1 is plotted against θ, with $x_1 = 10 \sin\theta$ and $x_2 = 10 \cos\theta$.

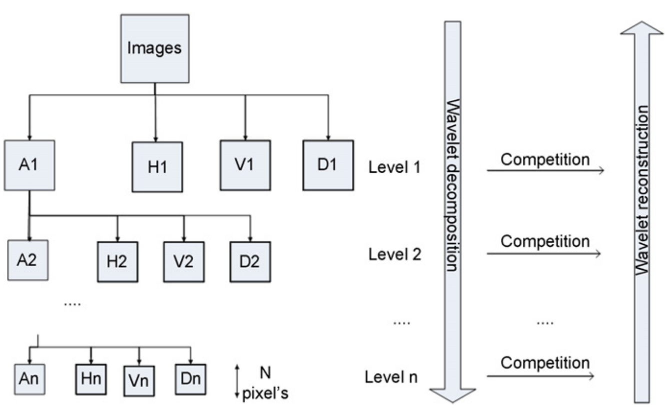


Fig. 8. Schematic representation of the wavelet-based separation method.

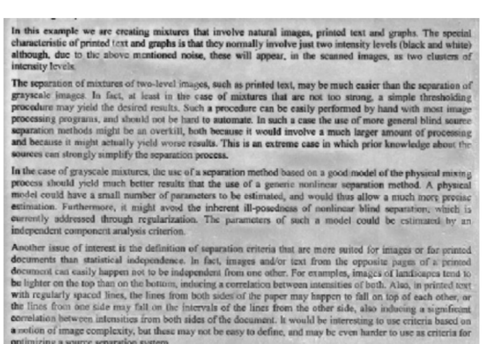
4. Experimental results

In this section we present the experimental results obtained with the proposed method, and a brief comparison with results from other methods. Due to space limitations, the images are shown much smaller than real size. In the electronic version of this paper it is possible to zoom in on the images to better examine their details.

The separation method presented in Section 3 was applied to the three mixture sets that were described in Section 2. For all experiments, the value of the competition parameter a (see (1)) was set to 1024, which yields a mask that is almost a hard winner-take-all function.

The tracing paper mixture pairs were separated using the proposed method with and without decorrelation preprocessing, with and without contrast compensation, and using both the complex wavelet transform and the stationary discrete wavelet transform with the Haar wavelet. All separations were performed with a 7-level wavelet analysis. The results that we considered to be best are shown in Figs. 10 and 11, and correspond to using the stationary discrete wavelet transform with the Haar wavelet, with decorrelation preprocessing and with contrast compensation using $g_{max} = 3$. The separation results of other variants of the method, for the bars mixture, are shown in Fig. 12. The impact of the various options (preprocessing, contrast compensation and choice of wavelet transform) is clearly visible. The results of Figs. 10 and 11 have a perceptual separation quality that we consider better than the one of the results obtained with the other existing methods (see [2,5,3]). In Section 4.1 we present a formal evaluation of some of these results.

Regarding the choice of wavelet transform, we subjectively considered all separations obtained with the shift-invariant transform with Haar wavelet better than those obtained with the complex wavelet transform, which typically presented some


In this example we are creating mixtures that involve natural images, printed text and graphs. The special characteristic of printed text and graphs is that they normally involve just two intensity levels (black and white) although, due to the above mentioned noise, these will appear, in the scanned images, as two clusters of intensity levels.

The separation of mixtures of two-level images, such as printed text, may be much easier than the separation of grayscale images. In fact, at least in the case of mixtures that are not too strong, a simple thresholding procedure may yield the desired results. Such a procedure can be easily performed by hand with most image processing programs, and should not be hard to automate. In such a case the use of more general blind source separation methods might be an overkill, both because it would involve a much larger amount of processing and because it might actually yield worse results. This is an extreme case in which prior knowledge about the sources can strongly simplify the separation process.

In the case of grayscale mixtures, the use of a separation method based on a good model of the physical mixing process should yield much better results that the use of a generic nonlinear separation method. A physical model could have a small number of parameters to be estimated, and would thus allow a much more precise estimation. Furthermore, it might avoid the inherent ill-posedness of nonlinear blind separation, which is currently addressed through regularization. The parameters of such a model could be estimated by an independent component analysis criterion.

Another issue of interest is the definition of separation criteria that are more suited for images or for printed documents than statistical independence. In fact, images and/or text from the opposite pages of a printed document can easily happen not to be independent from one other. For examples, images of landscapes tend to be lighter on the top than on the bottom, inducing a correlation between intensities of both. Also, in printed text with regularly spaced lines, the lines from both sides of the paper may happen to fall on top of each other, or the lines from one side may fall on the intervals of the lines from the other side, also inducing a significant correlation between intensities from both sides of the document. It would be interesting to use criteria based on a notion of image complexity, but these may not be easy to define, and may be even harder to use as criteria for optimizing a source separation system.

Fig. 9. Fourth pair of tracing paper mixtures separated without using contrast compensation.

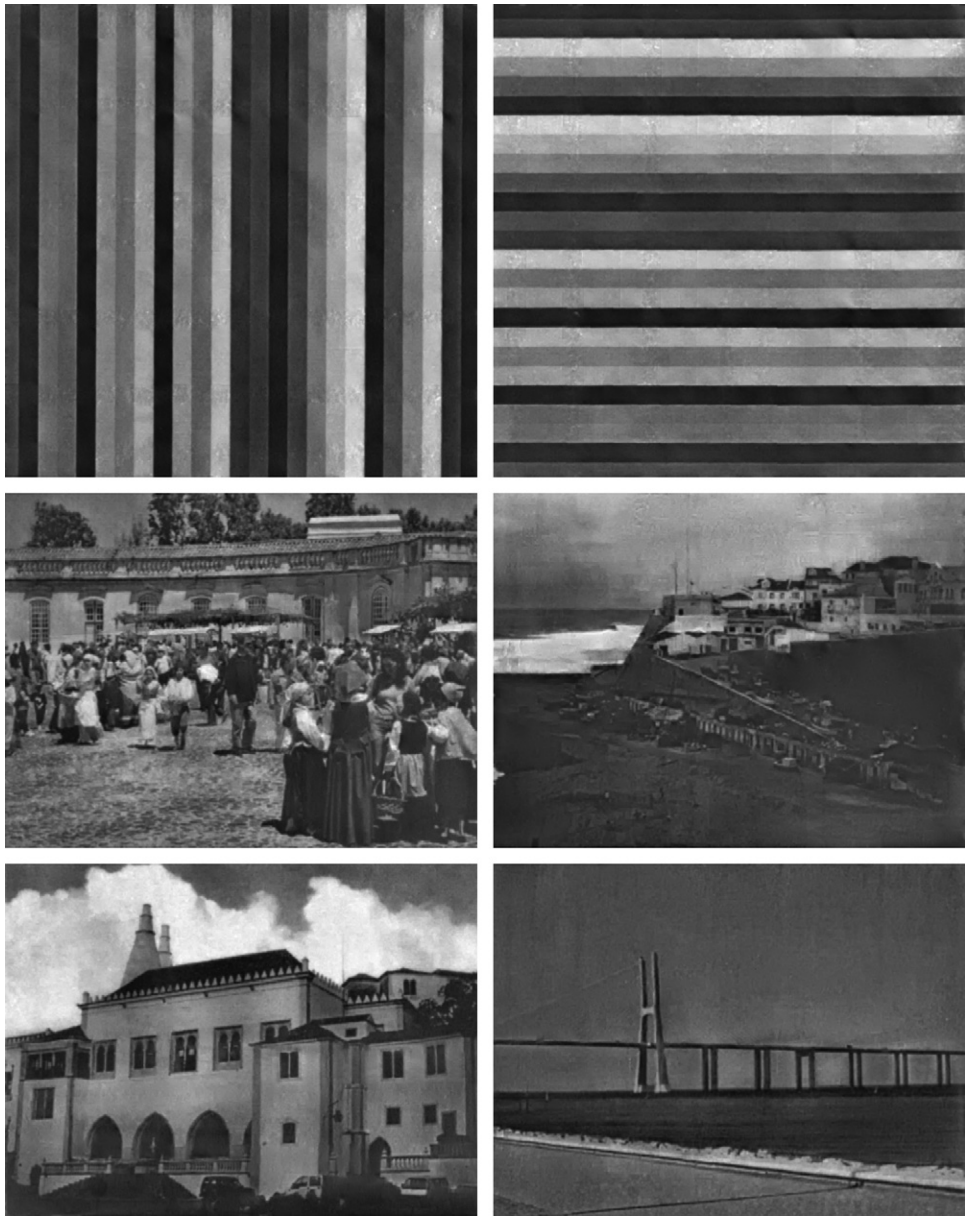


Fig. 10. Best separation results for the three first pairs of tracing paper mixtures.

oscillatory artifacts in the separated images. This difference in performance is probably due, to a large extent, to the fact that the complex wavelet is oscillatory, and to the difference in support size of the complex and Haar wavelets. The latter, having a much smaller support, handles fine details better. It has the disadvantage of being computationally heavier, since the stationary transform uses a redundant representation.

Contrast compensation allows a better restoration of the edge intensities, leading, for the tracing paper mixtures, to substantially better results. On the other hand, the use of this compensation also reinforces imperfections resulting from mis-handling of edges (which was mainly due to small edge misalignments between the two mixture components). The advantage of using contrast compensation in the air mail and partiture mixtures was found to be much smaller than for the tracing paper mixtures, and therefore we did not use it in the results presented ahead.

The results for the air-mail letter, presented in Fig. 13, were obtained using the stationary Haar wavelet transform with a 7-level decomposition, no decorrelation (as justified above) and no contrast compensation. Taking into account the complexity of this mixture, we consider the separation quality to be quite good. It is possible to read the separated letter more easily than the original one. As expected, our method performed better in areas of the letter where the lines of text of both sides were not aligned with one another, since, in those areas, the edges from the two sides coincide much less frequently. The zones where bleed-through appears to have occurred seem to have been well separated.

Regarding the partitures, we selected form each one a block of size 1024×1024 containing the area where the bleed-through effect was stronger (the selected blocks are identified in Fig. 6). These blocks were processed by the proposed separation method, with no decorrelation and no contrast compensation. Fig. 14 shows the separation results obtained using a 7-level stationary transform with the Haar wavelet. Although the separation is not perfect, most of the bleed-through effect was removed. It became possible to read the transcription without the strong interference from the opposite page. For image areas where there was no strong bleed-through, our method behaved almost perfectly. This

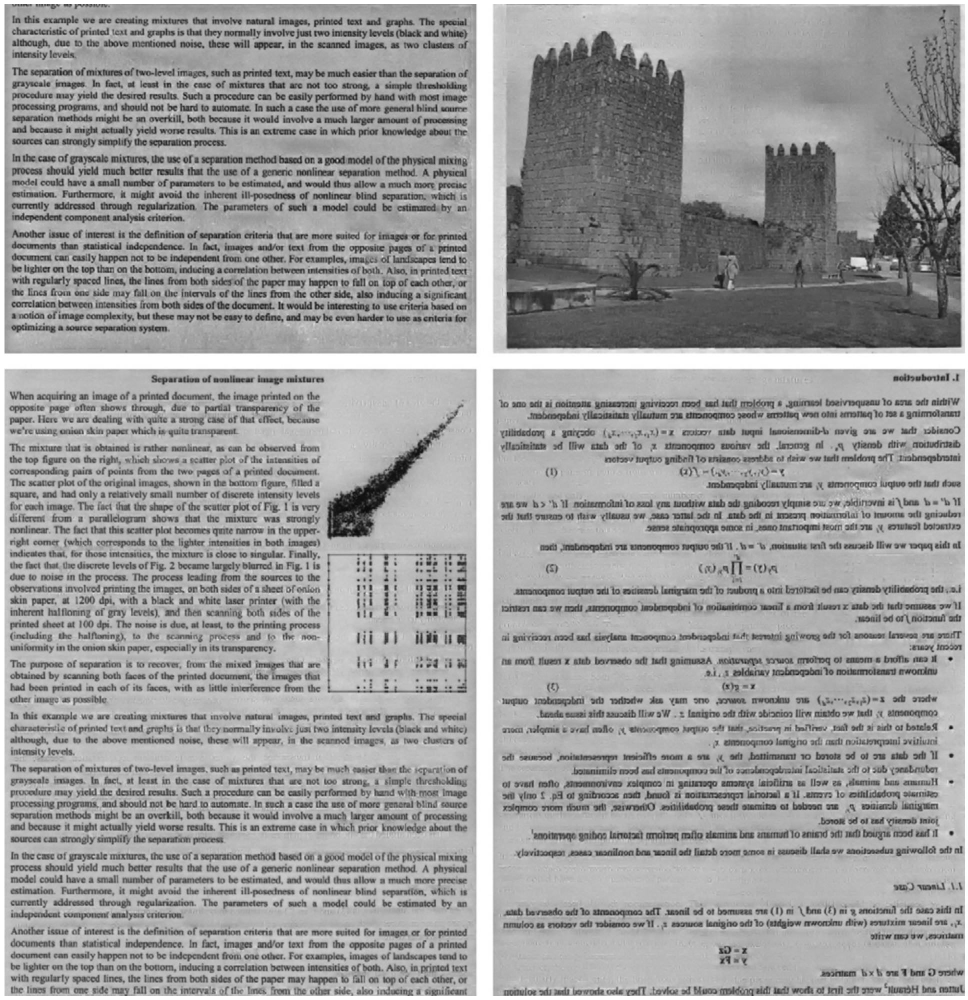


Fig. 11. Best separation results for the fourth and fifth pairs of tracing paper mixtures.

is, as far as we know, the first time that a bleed-through mixture of this kind was separated using the images from both sides of the document.

In a 1.6 MHz Pentium-M processor, images with 512×512 pixels (from the first three tracing paper mixtures) took 3 s to be separated with the complex wavelet transform and 21 s with the stationary transform with Haar wavelet, both of them without contrast compensation. Images with 1024×1024 pixels (from the last two pairs of tracing paper mixtures) took 16 and 70 s, respectively, with the same transforms. The reconstruction of the stationary transform incorporated an optimization which was not compatible with the proposed contrast compensation mechanism. As consequence, the use of contrast compensation led, for results obtained using the stationary transform, to an increase of the separation time (a total separation time of 40 s for images with 512×512 pixels and of 120 s for images with 1024×1024 pixels). In the case of the complex transform, contrast compensation did not cause any significant increase in computation time. We do not have separation times for images larger than 1024×1024. Those images (the ones of the manuscript letter) were processed in blocks of size 1024×1024, due to memory limitations.

4.1. *Subjective evaluation*

Previous works [2,5,3] have used point-wise quality measures to objectively assess the performances of the various methods. Those measures are not suitable for evaluating the separation quality achieved with our method, which is strongly non-point-wise (note that the size of the Haar wavelet, at the 7th level, is 128×128 pixels). There are image quality estimators that are perceptually oriented [26,25,20,15]. These estimators normally use a linear combination of several perceptual measures, and are adjusted to fit mean opinion score results. Most of these perceptual estimators were developed for compression or denoising problems, focusing on distortions that are quite different from the ones found in this work. For these reasons we considered that they were not appropriate for our case, and that a formal subjective evaluation would provide the most reasonable indication of separation quality.

Before the contrast compensation mechanism had been implemented, we submitted the results of our method to an opinion evaluation. Since no contrast compensation was used, these opinions do not consider the best results shown in Figs. 10 and 11 but, instead, consider the ones that correspond to $g_{\max} = 1$ (published in [4]). Nine people (those, among the authors' close colleagues, who volunteered for the test, and none of them working on source separation or on image processing) were asked to order, according to separation quality, the images obtained with our method and with four other methods: best and worst images obtained in [2] with MISEP and a MLP separator, images separated with nonlinear DSS [5], and images separated with an inverse physical model trained with MISEP [3]. The evaluators numbered the images from 1 (best) to 5 (worst). They had access to source

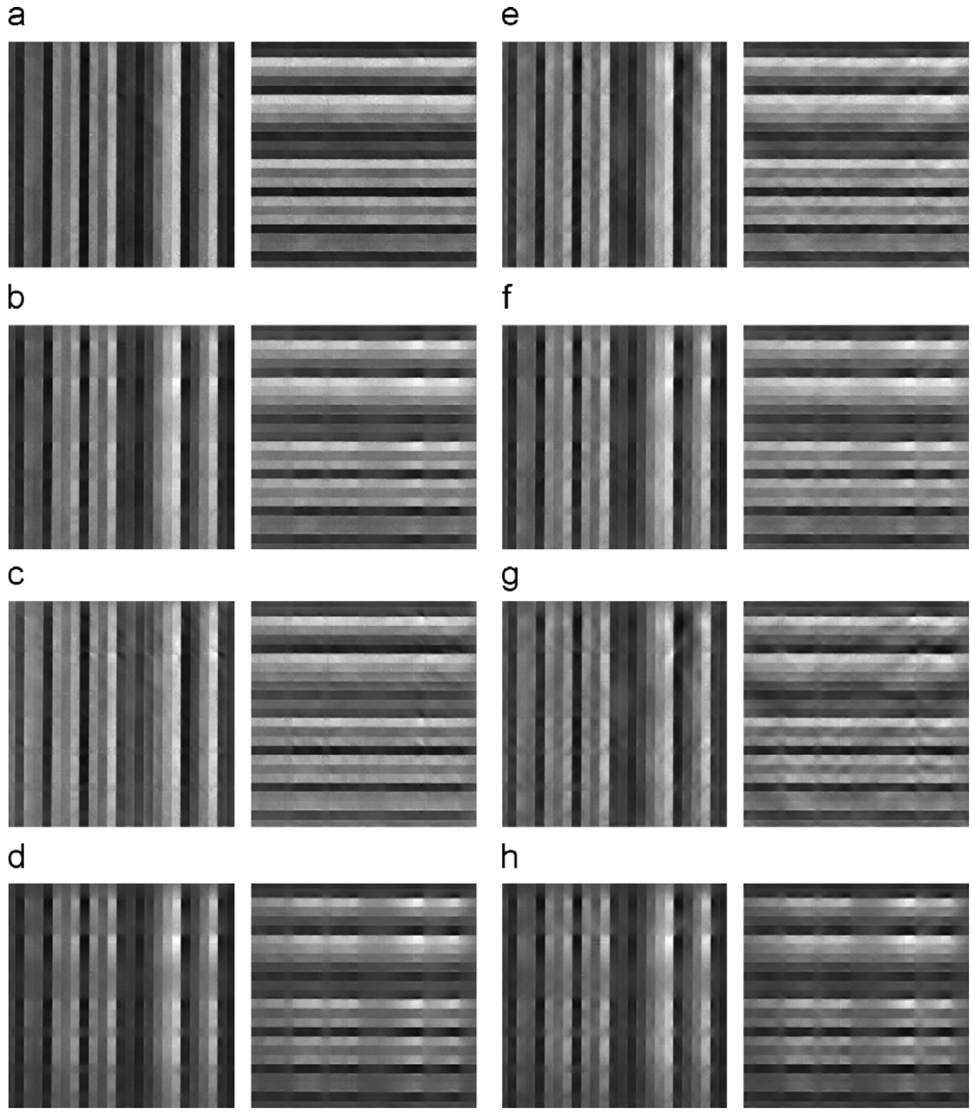


Fig. 12. Effect of different variants of the method on the separation of the bars mixture: (a) Stationary wavelet transform with decorrelation, $g_{max} = 3$ (same as Fig. 10). (b) Stationary wavelet transform with decorrelation, $g_{max} = 1$. (c) Stationary wavelet transform without decorrelation, $g_{max} = 10$. (d) Stationary wavelet transform without decorrelation, $g_{max} = 1$. (e) Complex wavelet transform with decorrelation, $g_{max} = 3$. (f) Complex wavelet transform with decorrelation, $g_{max} = 1$. (g) Complex wavelet transform without decorrelation, $g_{max} = 10$. (h) Complex wavelet transform without decorrelation, $g_{max} = 1$.

and mixture images, and their evaluation considered both the similarity with the corresponding source and the amount of interference from the opposite source. The mean rankings are shown in Table 1. As an indication of the confidence of these rankings, the table also presents, in brackets, the standard deviation computed over the nine rankings corresponding to each method and each pair of images.

The method proposed in this paper was chosen as best for the last three mixture pairs, and was among the three best for the other two pairs. Based on the improvement achieved with the introduction of contrast compensation (compare Fig. 12 with Fig. 10) we expect that, if contrast compensation had been used, the classification of the method would have been better still.

5. Conclusions

A non-iterative method for separating real-life nonlinear mixtures of images was presented. The method is fast and yields images with a perceptual separation quality that is competitive with the one obtained with previous methods.

The proposed method does not assume independence of the sources, but uses other properties of the problem. Therefore the quality of the results is not affected by the possible non-independence of the source images. Since the method processes wavelet coefficients down to a deep level, it performs a strongly non-point-wise separation. In contrast with previous solutions, this method does not assume the mixture to be invariant, and is therefore suitable for mixtures with varying local characteristics, such as those that result from bleed-through or from wrinkled documents.

A contrast compensation mechanism was proposed, to better recover the local contrast of the separated images. While this mechanism resulted in a clear improvement in the separation quality of the tracing paper mixtures, no significant improvements were obtained when it was applied either to the air-mail letter or to the partitures. So far, contrast compensation requires the user to set the value of a parameter. In future we plan to develop a criterion to automatically set this value.

None of the existent objective quality measures for images seems appropriate for evaluating the results of the proposed method. We used a subjective opinion evaluation to assess the

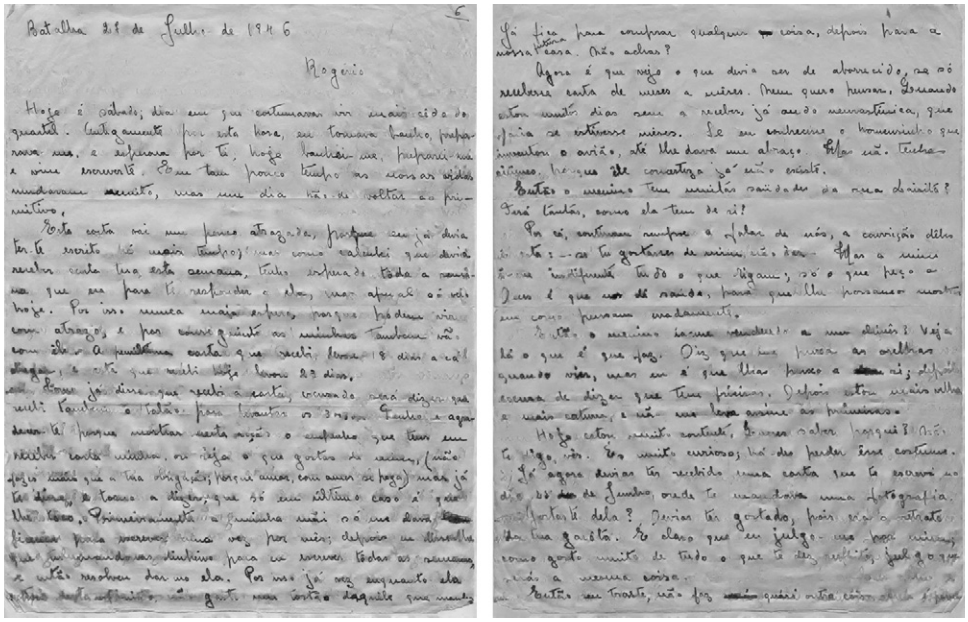


Fig. 13. Manuscript letter after separation.

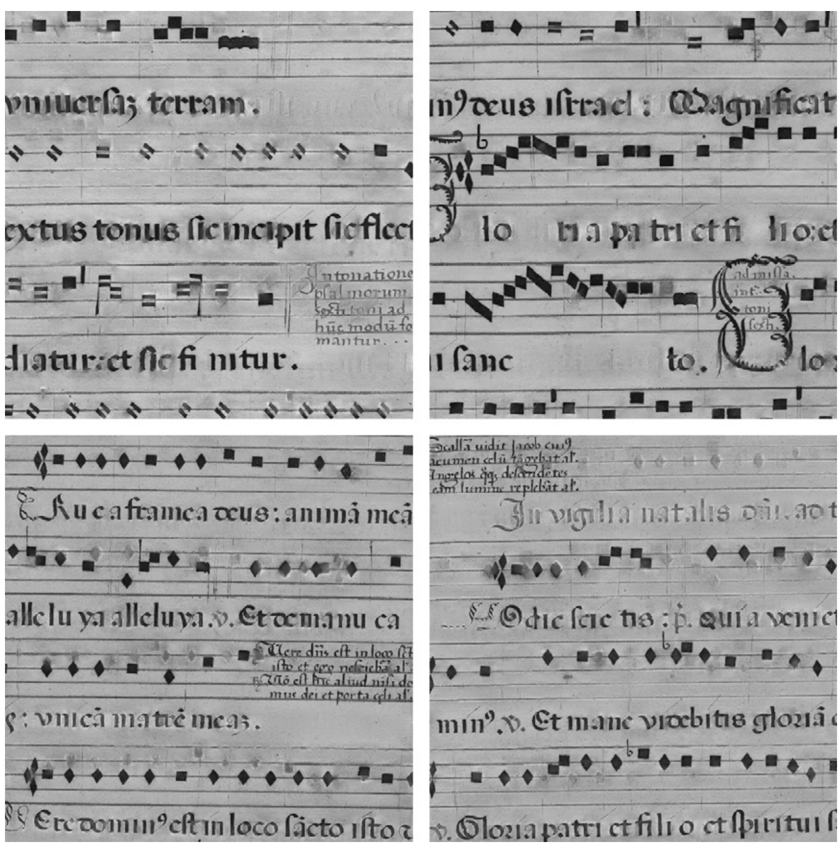


Fig. 14. Partitures after separation.

relative quality of our results. The evaluation of the results without contrast compensation was quite positive. Due to the practical complexity of the evaluation, it was not applied yet to the results obtained with contrast compensation. That evaluation will be addressed in future work.

As was said above, the old letter and partiture mixtures that we used resulted from converting the original color images to grayscale. Quite probably, better results could be obtained by separately processing the three color channels of the original images. This also is a subject of future work.

Table 1
Mean rank, for different methods, of the separated images from the tracing paper mixtures

Image pair	MISEP (best)	MISEP (worst)	Nonlinear DSS	Physical model	Proposed method
1	**1.2** (0.42)	4.7 (0.59)	2.6 (0.70)	3.7 (0.97)	2.9 (1.29)
2	2.8 (1.29)	3.2 (1.11)	4.3 (1.27)	**2.2** (1.15)	2.5 (1.38)
3	3.4 (0.70)	4.6 (0.62)	3.6 (1.34)	2.4 (0.70)	**1.1** (0.23)
4	3.1 (0.68)	2.8 (1.06)	–	2.8 (1.09)	**1.3** (0.69)
5	3.6 (0.50)	2.8 (0.81)	–	2.4 (1.04)	**1.2** (0.43)

Lower ranks are better. For each pair, the best result is shown in bold. Standard deviations are given in brackets.

Acknowledgments

We wish to acknowledge our colleague Rogério C. Pinto for kindly making available the old letter and the partiture images. We also acknowledge the use of the free package available at http://taco.poly.edu/WaveletSoftware/ for computing complex wavelet transforms. This work was partially supported by the Portuguese FCT and by the "Programa Operacional Sociedade do Conhecimento (POS-Conhecimento)", comparticipated by the FEDER European Community fund, under the Project POSC/EEA-CPS/61271/2004 and the PhD fellowship SFRH/BD/23919/2005.

References

[1] L.B. Almeida, MISEP—linear and nonlinear ICA based on mutual information, J. Mach. Learn. Res. 4 (2003) 1297–1318 URL: ⟨http://www.jmlr.org/papers/volume4/almeida03a/almeida03a.pdf⟩.
[2] L.B. Almeida, Separating a real-life nonlinear image mixture, J. Mach. Learn. Res. 6 (2005) 1199–1229 URL: ⟨http://www.jmlr.org/papers/volume4/almeida03a/almeida03a.pdf⟩.
[3] M.S.C. Almeida, L.B. Almeida, Separating nonlinear image mixtures using a physical model trained with ica, in: IEEE International Workshop on Machine Learning for Signal Processing, Maynooth, Ireland, 2006.
[4] M.S.C. Almeida, L.B. Almeida, Wavelet based nonlinear separation of images, in: IEEE International Workshop on Machine Learning for Signal Processing, Maynooth, Ireland, 2006.
[5] M.S.C. Almeida, H. Valpola, J. Särelä, Separation of nonlinear image mixtures by denoising source separation, in: J. Rosca, D. Erdogmus, J. Principe, S. Haykin (Eds.), Independent Component Analysis and Blind Signal Separation, Lecture Notes in Computer Science, vol. 3889, Charleston, SC, USA, Springer, Berlin, 2006, pp. 8–15. URL: ⟨http://www.cis.hut.fi/jaakkos/papers/nonlindss.pdf⟩.
[6] A. Bell, T. Sejnowski, An information-maximization approach to blind separation and blind deconvolution, Neural Comput. 7 (1995) 1129–1159 URL: ⟨ftp://ftp.cnl.salk.edu/pub/tony/bell.blind.ps⟩.
[7] A. Belouchrani, K.A. Meraim, J.-F. Cardoso, E. Moulines, A blind source separation technique based on second order statistics, IEE Trans. Signal Process. 45 (2) (1997) 434–444 URL: ⟨http://www.tsi.enst.fr/~cardoso/Papers.PDF/ieeesobi.pdf⟩.
[8] T. Blaschke, L. Wiskott, Independent slow feature analysis and nonlinear blind source separation, in: Proceedings of the International Workshop on Independent Component Analysis and Blind Source Separation, Lecture Notes in Artificial Intelligence, vol. 3195, Springer, Berlin, 2004, pp. 742–749. URL: ⟨http://itb.biologie.hu-berlin.de/~blaschke/publications/isfa.pdf⟩.
[9] E. Dubois, A. Pathak, Reduction of bleed-through in scanned manuscript documents, in: Proceedings of the IS&T Image Processing, Image Quality, Image Capture Systems Conference, Montreal, Canada, 2001, pp. 177–180.
[10] S. Harmeling, A. Ziehe, M. Kawanabe, B. Blankertz, K.-R. Müller, Nonlinear blind source separation using kernel feature spaces, in: T.-W. Lee (Ed.), Proceedings of the International Workshop on Independent Component Analysis and Blind Source Separation, San Diego, CA, USA, 2001, pp. 102–107. URL: ⟨http://ica2001.ucsd.edu/index_files/pdfs/080-harmeling.pdf⟩.
[11] A. Hyvärinen, P. Pajunen, Nonlinear independent component analysis: existence and uniqueness results, Neural Networks 12 (3) (1999) 429–439 URL: ⟨http://www.cis.hut.fi/~aapo/ps/NN99.ps⟩.
[12] F.A.A. Kingdom, A. Hayes, D.J. Field, Sensitivity to contrast histogram differences in synthetic wavelet-textures, Vision Res. 41 (5) (2001) 585–598.
[13] H. Lappalainen, A. Honkela, Bayesian nonlinear independent component analysis by multi-layer perceptrons, in: M. Girolami (Ed.), Advances in Independent Component Analysis, Springer, Berlin, 2000, pp. 93–121 URL: ⟨http://www.cis.hut.fi/harri/ch7.ps.gz⟩.
[14] G.C. Marques, L.B. Almeida, Separation of nonlinear mixtures using pattern repulsion, in: J.F. Cardoso, C. Jutten, P. Loubaton (Eds.), Proceedings of the First International Workshop on Independent Component Analysis and Signal Separation, Aussois, France, 1999, pp. 277–282. URL: ⟨http://www.lx.it.pt/~lbalmeida/papers/MarquesAlmeidaICA99.ps.zip⟩.
[15] M. Miyahara, K. Kotani, V.R. Algazi, Objective picture quality scale (pqs) for image coding, IEEE Trans. Commun. 46 (9) (1998) 1215–1226.
[16] P. Moulin, J. Liu, Analysis of multiresolution image denoising schemes using generalized-Gaussian and complexity priors, IEEE Trans. Inf. Theory 14 (3) (1999) 909–919.
[17] G.P. Nason, B.W. Silverman, The stationary wavelet transform and some statistical applications, in: Wavelets and Statistics, Lecture Notes in Statistics, Springer, Berlin, 1995, pp. 281–299. URL: ⟨citeseer.ifi.unizh.ch/nason95stationary.html⟩.
[18] I.W. Selesnick, R.G. Baraniuk, N.G. Kingsbury, The dual-tree complex wavelet transform, IEEE Signal Process. Mag. November (2005) 123–151.
[19] G. Sharma, Show-through cancellation in scans of duplex printed documents, IEEE Trans. Image Process. 10 (5) (2001) 736–754.
[20] H. Sheikh, A. Bovik, G. Veciana, An information fidelity criterion for image quality assessment using natural scene statistics, IEEE Trans. Image Process. 14 (12) (2005) 2117–2128.
[21] E.P. Simoncelli, B. Olshausen, Natural image statistics and neural representation, Annu. Rev. Neurosci. 24 (2001) 1193–1216.
[22] C.L. Tan, R. Cao, S. Peiyi, Restoration of archival documents using a wavelet technique, IEEE Trans. Pattern Anal. Mach. Intell. 24 (2002) 1399–1404.
[23] A. Tonazzini, L. Bedini, E. Salerno, Independent component analysis for document restoration, in: IJDAR, vol. 7, 2004, pp. 17–27.
[24] A. Tonazzini, E. Salerno, L. Bedini, Fast correction of bleed-through distortion in grayscale documents by a blind source separation technique, Int. J. Doc. Anal. 10 (1) (2007) 17–25.
[25] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[26] Z. Wang, E.P. Simoncelli, Reduced-reference image quality assessment using a wavelet-domain natural image statistic model, in: 17th Annual Symposium on Electronic Imaging, January 2005, pp. 17–20.
[27] A. Ziehe, K.-R. Müller, TDSEP—an efficient algorithm for blind separation using time structure, in: Proceedings of the International Conference on Artificial Neural Networks, Skövde, Sweden, 1998, pp. 675–680. URL: ⟨http://wwwold.first.gmd.de/persons/Mueller.Klaus-Robert/ICANN_tdsep.ps.gz⟩.

Mariana S.C. Almeida was born in Lisbon, Portugal, in May 30, 1982. She received the engineering degree in electrical and computer engineering from the Instituto Superior Técnico (I.S.T., the Engineering School of the Technical University of Lisbon), Lisbon, Portugal, in 2005. Since 2006 she is working on her PhD in the Institute of Telecommunications, Lisbon, under the supervision of Prof. Luís B. Almeida.

Luís B. Almeida was born in Lisbon, Portugal, in April 15, 1950. He graduated in Electrical Engineering by the Instituto Superior Técnico, Lisbon, in 1972, and obtained a "Doutor" degree by the Technical University of Lisbon, in 1983, with a thesis on nonstationary modelling of speech. Since 1972 he has been with the Instituto Superior Técnico, where he is, since 1995, a full professor in the areas of signal processing and machine learning. From 1984 to 2004 he was head of the Neural Networks and Signal Processing Group of INESC-ID. In the years 2000–2003 he was president of INESC-ID. In 2005 he joined the Instituto de Telecomunicações (Telecommunications Institute). He is the author of many papers on speech modelling and coding, time-frequency representations of signals and the fractional Fourier transform, learning algorithms for neural networks, and independent component analysis/source separation. Currently his research focuses mainly on nonlinear source separation and, more generally, on unsupervised or semi-supervised learning of structure from data. He was the recipient of an IEEE Signal Processing Area ASSP Senior Award and of several national awards.