Linear-quadratic blind source separating structure for removing show-through in scanned documents

Farnood Merrikh-Bayat, Massoud Babaie-Zadeh, Christian Jutten

IJDAR (2011) 14:319–333
DOI 10.1007/s10032-010-0131-7

ORIGINAL PAPER

Linear-quadratic blind source separating structure for removing show-through in scanned documents

Farnood Merrikh-Bayat · Massoud Babaie-Zadeh · Christian Jutten

Received: 27 October 2009 / Revised: 29 September 2010 / Accepted: 4 October 2010 / Published online: 28 October 2010

Abstract Digital documents are usually degraded during the scanning process due to the contents of the backside of the scanned manuscript. This is often caused by the show-through effect, i.e. the backside image that interferes with the main front side picture due to the intrinsic transparency of the paper. This phenomenon is one of the degradations that one would like to remove especially in the field of Optical Character Recognition (OCR) or document digitalization which require denoised texts as inputs. In this paper, we first propose a novel and general nonlinear model for canceling the show-through phenomenon. A nonlinear blind source separation algorithm is used for this purpose based on a new recursive and extendible structure. However, the results are restricted due to a blurring effect that appears during the scanning process due to the light transfer function of the paper. Consequently, for improving the results, we introduce a refined separating architecture for simultaneously removing the show-through and blurring effects.

This work has been partially funded by the Iran Telecom Research Center (ITRC), by the Iran National Science Foundation (INSF), and also by the Center for International Research and Collaboration (ISMO) and the French Embassy in Tehran, Iran, in the framework of a Gundi-Shapour collaboration program.

F. Merrikh-Bayat (✉) · M. Babaie-Zadeh
Electrical engineering department, Sharif university of technology, Azadi Avenue, Tehran, Iran
e-mail: f_merrikhbayat@ee.sharif.edu

M. Babaie-Zadeh
e-mail: mbzadeh@sharif.edu

C. Jutten
GIPSA-lab, Department of Images and Signal, Grenoble, France
e-mail: christian.jutten@gipsa-lab.grenoble-inp.fr

C. Jutten
Institut Universitaire de France, Paris, France

Keywords Show-through reduction · Blind source separation · Linear-quadratic · Blurring effect

1 Introduction

Archives usually need automatic methods for improving the readability of the digital versions of the ancient or printed documents. This processing is also recommended for the applications concerned with the production of the machine-readable digital document from ancient handwriting for example by using character recognition algorithms; these algorithms require clean versions of the original documents. In this paper, we consider one of the most common degradations, usually appearing in ancient documents that are written or printed on both sides of the page: this phenomenon is generally called print-through. Print-through is an undesired appearance of a printed image or text of the reverse side of the paper and can be divided into three additive components, each of them corresponding to a physical phenomenon [1]:

- The show-through component related to the paper's intrinsic transparency or low thickness;
- The pigment penetration component;
- The vehicle oil component, which is related to the loss of opacity due to the filling of pores with oil.

When the ink of the printer does not penetrate the paper considerably, effects of pigment penetration and of vehicle oil are negligible and print-through can be approximated only by show-through. Show-through can significantly impair the readability of the document and also cause visual fatigue for the reader. When the show-through degradation is significant (the darkness of the show-through is comparable to, or even greater than, that of some parts of the desired writing), then it is practically impossible to remove show-through by only using a simple thresholding operation.

Several approaches for show-through reduction have been already investigated. In [2,3], the authors used various features in the document for distinguishing show-through from the foreground image and presented show-through removal techniques involving only one side of the document. Although these methods certainly perform better than simple thresholding, there is no way to unambiguously differentiate foreground from show-through without comparing both sides of the document especially in grayscale images. Other works process both sides of the document simultaneously, in order to identify regions that are mainly show-through, and replace them by an estimate of the background [4,5]. Most of these works deal with only text or handwriting, and the original images are degraded after the show-through removal procedure. Recent investigations are applying Blind Source Separation (BSS) algorithms for solving this problem. In [6,7], the authors assume that the show-through effect can be modeled by linear superimposition of the back and front sides. Then, the scanned front-side image (corrupted by the backside) and the scanned back-side image (corrupted by the frontside) are linear mixtures of the pure front and back images, assumed to be independent, which can be estimated by BSS techniques. Tonazzini et al. in [8] represent an effective method for removing show-through in color images by using only one side of the paper. However, the method is not applicable for grayscale images. Although these methods give good results, the results are not perfect especially in regions where the images of the front and back sides of the paper overlap and the front side's image is nearly black. In such regions, the recovered front image is whiter than other sections where there is no overlap. The main reason of these poor results is that show-through is a nonlinear effect as we will show in the next section, and as it has also been previously considered and modeled, e.g. in [9]. Sharma in [10] considered a nonlinear model for this phenomenon and proposed to compensate for this effect by using adaptive filters. The main disadvantage of using adaptive filters, as we will see later, is that the Signal to Noise Ratio (SNR) of the outputs cannot be more than the noise to signal ratio of the input images (due to the *power inversion property* in adaptive filters [11, page 78]). Castro et al. [12] used fuzzy classification for detecting degraded sections of the input images. The proposed algorithm is fully dependent on the specific application (only applicable to ancient music notes) and suffers from disadvantages of fuzzy classification. Moreover, they did not consider the blurring effect that we will point out in Sect. 4.1. By taking into account the nonlinearity of show-through, Almeida used in [13] the MISEP method based on Multi-Layer Perceptron (MLP) networks for separating the real-world nonlinear image mixtures. The main drawback of using these kinds of "universal nonlinear networks" for BSS is the separability issue: the Independent Component Analysis (ICA) does not necessarily lead to BSS using such networks. Note that based on the way that the show-through phenomenon is modeled, this or other nonlinear BSS techniques [14–17] may be applicable for solving this problem. In fact, finding a suitable BSS method for separating the show-through from the original sources based on the acquired model is one of the main goals of this paper.

Here, it is worth to mention that, although there are some Markov Random Field (MRF)-based show-through removal methods [18,19], in this paper, we only focus on BSS-based algorithms. This is because of the fact that in the framework of BSS, considering MRF nature of images requires one to take into account the joint distribution of neighbor pixels, which leads to very tricky estimation problems and very cost-consuming algorithms. In fact, the Markovian nature has already been considered in BSS, for instance by Hosseini et al. [20]. In addition in the framework of hyperspectral processing, it has been shown that MRF and BSS, although based on different properties, lead to very similar results [21].

In this paper, we first show that show-through is a nonlinear phenomenon and introduce a nonlinear model based on experiments. Then, we show that, in addition to nonlinear mixing models, we have to take into account a blurring effect. Our method is derived from the nonlinear BSS recursive structure presented by Hosseini and Deville [22,23], in which we add refinements for jointly compensating for the degradation caused by show-through and blurring effects, as it will be demonstrated by some simulation results. Although the principles and preliminary results have briefly been presented in [24], the current paper, both in theoretical and experimental parts, contains many new extensions.

The paper is organized as follows. In Sect. 2, we experimentally show that show-through is a nonlinear phenomenon and introduce our nonlinear model. Section 3 describes a basic blind separation structure for dealing with this nonlinear model. The blurring effect and the usage of adaptive filters for its compensation, and the combination of these filters with the non-linear BSS structure are developed in Sect. 4. Finally, a few experimental results with real and artificial printed or manuscript documents are presented in Sect. 5, before the conclusions and perspectives of Sect. 6.

2 Show-through nonlinear modeling

Show-through appears when a fraction of the verso is mixed with the recto pixel by pixel in the scanning process. However, as we will demonstrate in this section, this fraction is proportional to the grayscale of the front image, i.e. as the front image becomes darker, the show-through will be lower. Therefore, the mixing model cannot be linear. Figure 1 illustrates this claim based on an experiment explained below.

In this experiment, the two images shown in Fig. 1a are first printed on the two sides of a sheet of paper and then scanned as illustrated in Fig. 1b. Then, the grayscales of the

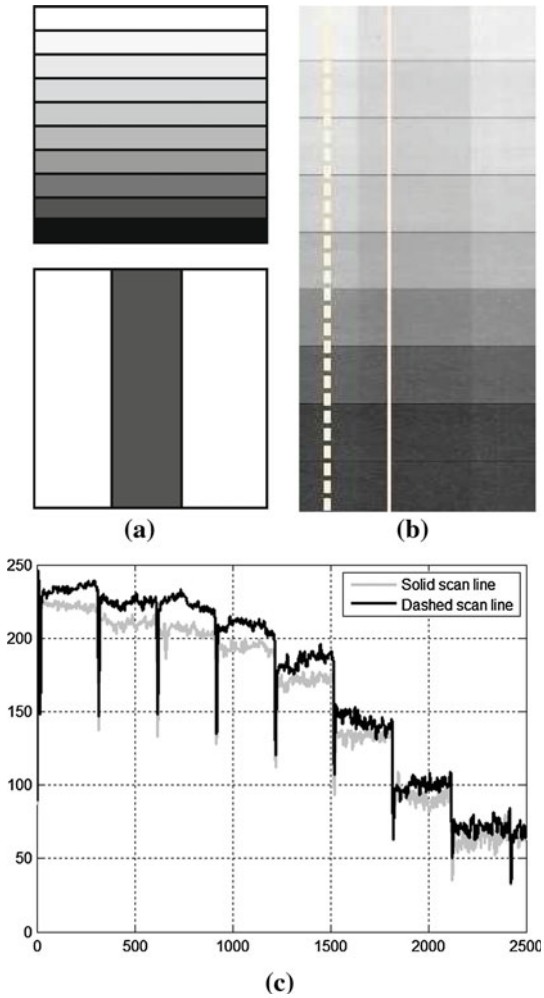

Fig. 1 Experiment used for demonstrating the nonlinearity of the show-through phenomenon. Grayscales of the pixels on two scanning lines shown in Fig. 1b are plotted in Fig. 1c. It shows that, as the front image becomes whiter, show-through will become more apparent. Therefore, show-through depends on the properties of the front side image as well as the back side one. **a** Images which are printed on the sides of a paper. **b** Scanned image with two scanning lines. **c** Plotting the grayscales of the pixels relied on two scanning lines in (b)

pixels on two scanning lines shown in Fig. 1b (solid line in the region with show-through and dashed line in the region without show-through) are plotted in Fig. 1c. It is obvious in this figure that as the front image becomes darker, the difference between the pixels with show-through and the pixels without show-through decreases with the pixel value. This demonstrates that the appearance of the show-through in the scanned image depends on the grayscale of the front image of the paper. This means that the gain by which the backside image is added to the front side image is variable, and it is proportional to the brightness of the image printed on the front side of the paper. We model this variable gain at any pixel of the scanned image of the front side of the paper by $g_1\left(f_r^i(m,n)\right)$ and at any pixel of the back side by $g_2\left(f_v^i(m,n)\right)$, where $g_1\left(f_r^i(m,n)\right)$ and $g_2\left(f_v^i(m,n)\right)$ represent the dependency of the mentioned gain to the pixel values of the front side and back side images, respectively.

Therefore, we model the show-through effect and its dependency to the grayscale of the front image as

$$\begin{aligned} f_r^s(m,n) &= a_1 f_r^i(m,n) + b_1 f_v^i(m,n) \times g_1(f_r^i(m,n)), \\ f_v^s(m,n) &= a_2 f_v^i(m,n) + b_2 f_r^i(m,n) \times g_2(f_v^i(m,n)), \end{aligned} \tag{1}$$

where

- m and n are the 2-D spatial coordinates on the paper being scanned,
- the subscript r is related to the front side (recto),
- the subscript v is related to the back side (verso),
- the superscript i denotes the ideal (without show-through) version of the image,
- the superscript s denotes the scanned image (having show-through);
- the function g_i for $i = 1, 2$ is a nonlinear function representing the effect of the paper,
- and $f(.,.)$ denotes a 2-D signal (image).

Note that if the function g_i is cancelled, (i.e. if $g_i(\cdot) = 1$), Eq. (1) becomes a simple linear mixing model. Since the scanning process of both sides of a paper is symmetric and usually done under the same conditions, in [6] it has been assumed that $a_1 = a_2, b_1 = b_2$ and $g_1(\cdot) = g_2(\cdot) = 1$ (because in [6] the authors assumed the model was linear). However, in this paper and for preserving the generality of our nonlinear model as much as possible, we only assume that the nonlinear functions g_1 and g_2 have the same shape. By rewriting (1) we obtain

$$\begin{aligned} \frac{f_r^s(m,n) - a_1 f_r^i(m,n)}{b_1 f_v^i(m,n) + \epsilon} &= g_1(f_r^i(m,n)), \\ \frac{f_v^s(m,n) - a_2 f_v^i(m,n)}{b_2 f_r^i(m,n) + \epsilon} &= g_2(f_v^i(m,n)), \end{aligned} \tag{2}$$

where ϵ is a small positive number inserted to avoid division by zero. Note that adding ϵ is useful only if $b_1 f_v^i(m,n)$ never becomes negative. $b_1 f_v^i(m,n)$ is the attenuated backside image which appears in the scanned image of the front side of the paper. Since in the scanned image it is possible to see the same image of the backside of the paper (and not its negative), it means that $b_1 f_v^i(m,n)$ has been added and not subtracted from the image of the front side of the paper and therefore, $b_1 f_v^i(m,n)$ is nonnegative. For another reason, if $b_1 f_v^i(m,n)$ is negative, in those areas where the front side and backside printed images are white (have pixel values near 255), the scanned image should be nearly black which is not the case. By plotting the left-hand side of the first

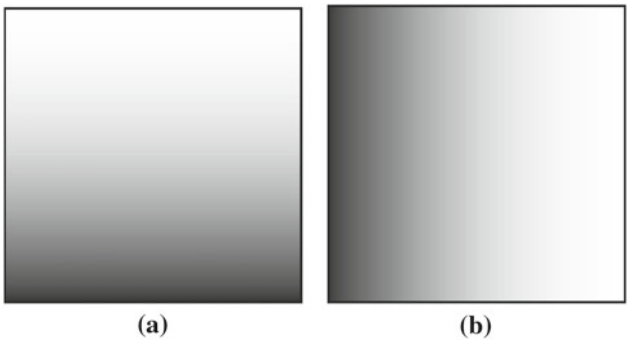

Fig. 2 Images used for experimentally modeling show-through. These two images are printed on back and front sides, respectively, of a sheet of paper and then scanned

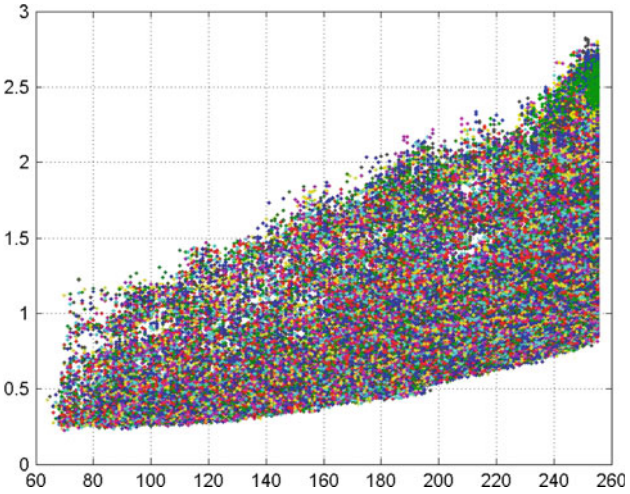

Fig. 3 Scatter plot of $g\left(f_r^i(m,n)\right)$ versus $f_r^i(m,n)$. The shape of $g_i(\cdot)$ suggests an exponential function

equation in (2) versus $f_r^i(m,n)$ for particular images having show-through, it is possible to draw the general shape of the function $g_i(\cdot)$. As an experiment, we used images shown in Fig. 2 as $f_r^i(m,n)$ and $f_v^i(m,n)$. These images were printed on sides of a sheet of paper and then were scanned. By choosing these images, we are sure that in the final scanned pictures, we have all possible combinations of grayscales.

Figure 3 shows the plot of $g_1(f_r^i(m,n))$ versus $f_r^i(m,n)$. Note that for plotting this image, parameters a_1 and b_1 are needed. Since in this experiment, we have the original pure sources, i.e. $f_v^i(m,n)$ and $f_r^i(m,n)$, these parameters are not unknown and by using Eq. (1), they can be estimated from the images used for performing this simulation. For example, a_1 can be approximated by comparing $f_r^i(m,n)$ and $f_r^s(m,n)$ where there is no image printed on the backside. Similarly, b_1 can be estimated by comparing $f_v^i(m,n)$ and $f_r^s(m,n)$ where this time, there is no image printed on the front side.

Based on Fig. 3, we propose to model the function $g_i(u)$ as an exponential function $g_i(u) = \gamma_i e^{\beta_i u}$ for $i = 1, 2$ where β_i and γ_i are constant. This model implies that as the front image becomes whiter, the backside image will be added to the front image with a higher gain. Note that if show-through was a linear phenomenon then g would be a constant, i.e. a horizontal line in Fig. 3.

By replacing $g_1(\cdot)$ by $\gamma_1 e^{\beta_1(\cdot)}$ in (1) we get

$$f_r^s(m,n) = a_1 f_r^i(m,n) + b_1' f_v^i(m,n) \times e^{\left(\beta_1 f_r^i(m,n)\right)}, \quad (3)$$

where $b_1' = b_1 \times \gamma_1$. Equation (3) can be simplified by replacing the exponential function by its first order Taylor expansion. As it will be seen in the experimental results, this approximation works well. Moreover, the experiments also show that β_1 is usually small for documents having show-through, which is consistent with the first order expansion. This can also be justified by Fig. 3, which shows that g is close to being linear, i.e. β_1 is small. Using this approximation, we obtain

$$\begin{aligned} f_r^s(m,n) &\approx a_1 f_r^i(m,n) + b_1' f_v^i(m,n) \times \left[1 + \beta_1 f_r^i(m,n)\right] \\ &= a_1 f_r^i(m,n) + b_1' f_v^i(m,n) \\ &\quad + d_1 f_r^i(m,n) f_v^i(m,n), \end{aligned} \quad (4)$$

where $d_1 = b_1' \times \beta_1$. Similar to (4), the scanned image of the back side can be written as

$$\begin{aligned} f_v^s(m,n) &= a_2 f_v^i(m,n) + b_2' f_r^i(m,n) \times e^{\left(\beta_2 f_v^i(m,n)\right)} \\ &\approx a_2 f_v^i(m,n) + b_2' f_r^i(m,n) \times \left[1 + \beta_2 f_v^i(m,n)\right] \\ &= a_2 f_v^i(m,n) + b_2' f_r^i(m,n) + d_2 f_v^i(m,n) f_r^i(m,n). \end{aligned} \quad (5)$$

Finally, since the gains a_1 and a_2 are undetermined, Eqs. (4) and (5) can be simplified as follows [22,23]:

$$\begin{aligned} f_r^s(m,n) &= f_r^i(m,n) - l_1 f_v^i(m,n) - q_1 f_r^i(m,n) f_v^i(m,n), \\ f_v^s(m,n) &= f_v^i(m,n) - l_2 f_r^i(m,n) - q_2 f_v^i(m,n) f_r^i(m,n). \end{aligned} \quad (6)$$

For separating the sources in the above model, we first have to find the mixing coefficients and then to solve the nonlinear system of Eq. (6) numerically. This will be done using nonlinear BSS techniques, as explained in the next sections.

3 Basic linear-quadratic blind source separation structure

The nonlinear model (6) actually is a linear-quadratic (or bilinear) mixing model, whose blind separation has already been addressed by Hosseini and Deville [22,23] based on the recurrent separating structure of Fig. 4, which is inspired

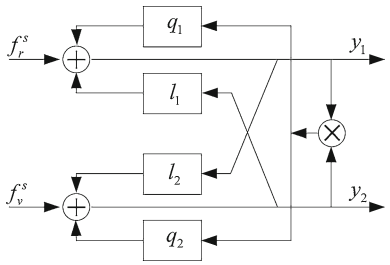

Fig. 4 Basic linear-quadratic blind source separating structure proposed in [22,23]

from the early work of Hérault and Jutten [25]. Note that (i) by setting $q_1 = 0$ and $q_2 = 0$ in (7) and in Fig. 4, this structure is reduced to the basic linear network proposed by Hérault and Jutten and (ii) this structure has the capability to be generalized to arbitrary polynomial models [22,23].

The computation of the outputs of the structure requires the realization of the following recurrent iterative expression

$$\begin{aligned} y_1^{(t+1)}(\cdot,\cdot) &= f_r^s(\cdot,\cdot) + l_1 y_2^{(t)}(\cdot,\cdot) + q_1 y_1^{(t)}(\cdot,\cdot) y_2^{(t)}(\cdot,\cdot), \\ y_2^{(t+1)}(\cdot,\cdot) &= f_v^s(\cdot,\cdot) + l_2 y_1^{(t)}(\cdot,\cdot) + q_2 y_2^{(t)}(\cdot,\cdot) y_1^{(t)}(\cdot,\cdot), \end{aligned} \tag{7}$$

where t is the iteration index. It has been shown in [26] that the necessary and sufficient condition for local stability of this model at the separating point $(y_1, y_2) = \left(f_r^i, f_v^i\right)$ is that the absolute values of the two eigenvalues of $\mathbf{J}_1$ be smaller than one where $\mathbf{J}_1$ is the Jacobian matrix of (7) at $\left[f_v^i, f_r^i\right]^T$. By defining $s_1 = f_r^s(m,n)$ and $s_2 = f_v^s(m,n)$, the Jacobian matrix can be written as

$$\mathbf{J}_1 = \begin{bmatrix} q_1 s_2 & l_1 + q_1 s_1 \\ l_2 + q_2 s_2 & q_2 s_1 \end{bmatrix}, \tag{8}$$

and its eigenvalues are

$$\begin{aligned} \lambda_{1,2} &= \frac{1}{2}(q_1 s_2 + q_2 s_1) \\ &\pm \frac{1}{2}\left(\left[(q_1 s_2 + q_2 s_1)^2 + 4(l_1 l_2 + l_1 q_2 s_2 + l_2 q_1 s_1)\right]^{\frac{1}{2}}\right). \end{aligned} \tag{9}$$

The Maximum Likelihood (ML) estimate of the parameter vector $\mathbf{p} = [l_1, l_2, q_1, q_2]^T$ is obtained using the updating rule

$$\mathbf{p}^{(j+1)} = \mathbf{p}^{(j)} + \mu \frac{\partial L}{\partial \mathbf{p}}, \tag{10}$$

where L denotes the likelihood of $\mathbf{p}$ based on the given samples of the mixtures f_r^s and f_v^s. By defining J_2 as the Jacobian of the mixing model (6), $\frac{\partial L}{\partial \mathbf{p}}$ can be written as [22,23]:

$$\frac{\partial L}{\partial \mathbf{p}} = -E\left[\frac{A}{J_2}, \frac{B}{J_2}, \frac{C}{J_2}, \frac{D}{J_2}\right]^T, \tag{11}$$

where

$$\begin{aligned} A =& \psi_1(s_1)(1 - q_2 s_1)s_2 + \psi_2(s_2)(l_2 + q_2 s_2)s_2 \\ & - (l_2 + q_2 s_2) - (q_2 + l_2 q_1)(1 - q_2 s_1)s_2/J \\ & - (q_1 + l_1 q_2)(l_2 + q_2 s_2)s_2/J, \\ B =& \psi_1(s_1)(l_1 + q_1 s_1)s_1 + \psi_2(s_2)(1 - q_1 s_2)s_1 \\ & - (l_1 + q_1 s_1) - (q_1 + l_1 q_2)(1 - q_1 s_2)s_1/J \\ & - (q_2 + l_2 q_1)(l_1 + q_1 s_1)s_1/J, \\ C =& \psi_1(s_1)(1 - q_2 s_1)s_2 s_1 + \psi_2(s_2)(l_2 + q_2 s_2)s_2 s_1 \\ & - (l_2 s_1 + s_2) - (q_2 + l_2 q_1)(1 - q_2 s_1)s_1 s_2/J \\ & - (q_1 + l_1 q_2)(l_2 + q_2 s_2)s_1 s_2/J, \\ D =& \psi_1(s_1)(l_1 + q_1 s_1)s_1 s_2 + \psi_2(s_2)(1 - q_1 s_2)s_1 s_2 \\ & - (s_1 + l_1 s_2) - (q_1 + l_1 q_2)(1 - q_1 s_2)s_1 s_2/J \\ & - (q_2 + l_2 q_1)(l_1 + q_1 s_1)s_1 s_2/J, \\ J_2 =& 1 - l_1 l_2 - (q_2 + l_2 q_1)s_1 - (q_1 + l_1 q_2)s_2. \end{aligned} \tag{12}$$

In (11), E denotes the spatial averaging operation on all of the samples and $\psi_i(\cdot), i = 1, 2$ are the score functions [22,23] of the source images (original images without show-through, i.e. f_r^i and f_v^i).

Here, it is worth mentioning that as explained and proved in [22,23], the linear-quadratic mixing model of (6) is not bijective in general. Consequently, the recurrent structure shown in Fig. 4 has two equilibrium points corresponding to a non-permuted solution and a permuted one, while only one of them is stable. Moreover, if the Jacobian of the mixing model, J_2, is always positive or always negative for all of the values of the sources, the recurrent structure leads to the entirely permuted or entirely non-permuted sources for all of the image pixels and therefore, it can be used for separating the sources. On the other hand, if J_2 is positive for some values of the sources and negative for other values, the recurrent structure leads to the non-permuted sources for some image pixels and to the permuted sources for other pixels. In this case, although the sources are separated "sample by sample", each retrieved image contains samples of the two image sources. From the explanations provided in this section, it should become clear that the recurrent network generally may diverge in many cases. But fortunately, using this structure for show-through removal adds some constraints (which come from the physics of the problem) on the sources and parameter values which can somehow relax this drawback. For example, since in this particular application, the sources are images, they can only have bounded non-negative values. In addition, by referring to (6) and considering the explanations provided in the previous section about the signs of coefficients, e.g. b_1 and b_2, it is clear that l_1 and l_2 are always negative and their values are close to each other, i.e. $l_1 \approx l_2$ (because of the symmetric nature of the scanning process in scanning two sides of the paper). Moreover, we know that $q_1 \approx q_2$ and both of them have a small

positive value as will be seen in the experimental results. Applying these constraints to the aforementioned conditions may reduce the divergence probability of the recurrent structure when it is used for this application.

Eventually, the show-through cancelation algorithm based on the structure of Fig. 4 can be summarized as follows:

1. align the front side and the back side scans by a registration method.
2. Initialize the structure of Fig. 4 with $\left(y_1^{(t)}, y_2^{(t)}\right)\Big|_{t=0} = (\mathbf{0}, \mathbf{0})$ and $\mathbf{p} = [l_1, l_2, q_1, q_2]^T = [0,\ 0,\ 0,\ 0]^T$.
3. Apply the registered front and back side images to the inputs of the structure of Fig. 4.
4. Compute the outputs of the structure, i.e. y_1 and y_2, by repeating (7) until convergence. The final outputs are assumed to be the independent original sources that are needed for updating the parameter vector in the next step.
5. Update the parameter vector $\mathbf{p}$ by iterating (10). Note that for calculating $\frac{\partial L}{\partial \mathbf{p}}$ in (10), the original sources, i.e. s_1 and s_2, are needed. These sources are approximated by the outputs of the structure which are computed in the previous step.
6. Repeat steps 4 and 5 until convergence is achieved (for instance, if the difference of two successive parameter vectors becomes less than a predetermined value).

For experimentally validating this approach, we applied this structure and the algorithm of Hosseini and Deville [22,23] to the scanned image of Fig. 5a and b. The result of the algorithm is shown in Fig. 5c and d. The result proves a certain capability of the structure of Fig. 4 in removing show-through, but it also emphasizes residual errors, i.e. show-through is not perfectly canceled by the above nonlinear model.

Fig. 5 Results obtained with the structure shown in Fig. 4 for removing show-through. **a** and **b**: Original scanned images obtained from http://www.site.uottawa.ca/~edubois/documents. **c** and **d**: Outputs of the separating structure shown in Fig. 4. **a** Input recto image with show-through. **b** Input verso image with show-through. **c** Recovered recto image. **d** Recovered verso image

4 Adaptive linear-quadratic blind source separation structure

In this section, we first explain that the poor results obtained with the previous nonlinear model can be due to a blurring effect. We then propose a new structure based on a modification on the structure of Fig. 4 for jointly removing the show-through and blurring effects.

4.1 Blurring effect

Figure 5c and d show that the show-through has not been removed perfectly in regions where there is no text or image in one of the two sides.

In fact, during the scanning process, if the page is not completely opaque and the scanner uses a white backing behind the page, the sensor of the scanner receives a light that is reflected from the backing and transmitted back through the paper. Thus, the light may scatter in different directions, especially due to an unsmooth surface of the paper. This scattering phenomenon, which acts as a low-pass filter, is known as the blurring effect and some methods for its compensation have already been studied in the literatures [7,10,27]. However, the blurring effect changes the nonlinear model that we considered previously. In fact, when scanning the recto, the scanned image is not a mixture of the front and back side images, but *it is a mixture of the front side image and of the low-pass filtered image of the back side image*. Clearly, this filtering cannot be canceled with the separating structure of Fig. 4.

Therefore, to take into account the blurring effect in our model, we modify the show-through model of (6) as follows

$$\begin{aligned} f_r^s &\approx f_r^i - l_1 f_v^i * \mathbf{H}_1 - q_1 f_r^i \times (f_v^i * \mathbf{H}_1), \\ f_v^s &\approx f_v^i - l_2 f_r^i * \mathbf{H}_2 - q_2 f_v^i \times (f_r^i * \mathbf{H}_2), \end{aligned} \tag{13}$$

where $*$ denotes the convolution operator, while $\mathbf{H}_1$ and $\mathbf{H}_2$ are the point spread functions of the paper which are unknown. Therefore, the computation of the structure outputs

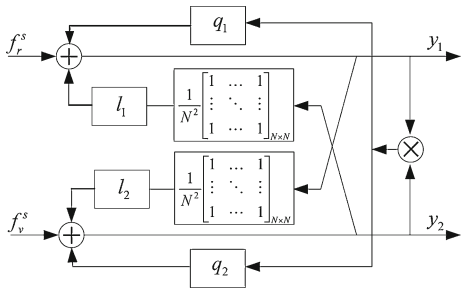

Fig. 6 Simplified bilinear blind source separating structure proposed for removing show-through and enhancing blurring effect. In this separating structure, filtering blocks are simply fixed averaging filters

requires the realization of the following recurrent iterative expression

$$\begin{aligned} y_1^{(t+1)} &= f_r^s + l_1 y_2^{(t)} * \mathbf{H}_1 + q_1 y_1^{(t)} \times (y_2^{(t)} * \mathbf{H}_1),\\ y_2^{(t+1)} &= f_v^s + l_2 y_1^{(t)} * \mathbf{H}_2 + q_2 y_2^{(t)} \times (y_1^{(t)} * \mathbf{H}_2). \end{aligned} \tag{14}$$

Usually, as is confirmed by our simulations, q_1 and q_2 are very small. Moreover, noting also that $y_i^{(t)}$ and $y_i^{(t)} * \mathbf{H}_i$, although different, are in fact very similar, (only here and for illustrating the effectiveness of this new model by an example) we ignore the filtering operation in the quadratic part of (14), and we obtain [24]:

$$\begin{aligned} y_1^{(t+1)} &= f_r^s + l_1 y_2^{(t)} * \mathbf{H}_1 + q_1 y_1^{(t)} \times y_2^{(t)},\\ y_2^{(t+1)} &= f_v^s + l_2 y_1^{(t)} * \mathbf{H}_2 + q_2 y_2^{(t)} \times y_1^{(t)}. \end{aligned} \tag{15}$$

Figure 6 shows the structure associated with this recurrent iterative expression with *fixed* and *known* averaging filters. Here, we simplified the model and used fixed averaging filters as an approximation of the point spread function of the paper only for demonstrating the excellence of this new configuration. The validity of this model will be illustrated by an experiment. Note that, as we used *fixed* and *known* filters and because of the similarities between $y_i^{(t)}$ and $y_i^{(t)} * \mathbf{H}_i$, the parameter vector $\mathbf{p}$ can be updated using equations derived in Sect. 3. Figure 7a and b show the results obtained by using this separating structure on images of Fig. 5a and b. The results demonstrate the ability of this structure to reduce the blurring effect even by using these approximations.

Figure 7c shows one of the outputs of the structure of Fig. 6 when it is configured as a linear separation structure (enforcing $q_1 = 0$ and $q_2 = 0$). An interesting part of the figure is magnified in Fig. 7d to better understand the problem associated with linear modeling of show-through. Arrows emphasize the *whitening* effect produced by using linear mixing models discussed in the introduction. This figure confirms that the averaging filters (in the structure of Fig. 6) are not sufficient and that a nonlinear modeling of show-through is required.

The disadvantage of using fixed filters (and the approximations we did in the above experiment) is that it requires

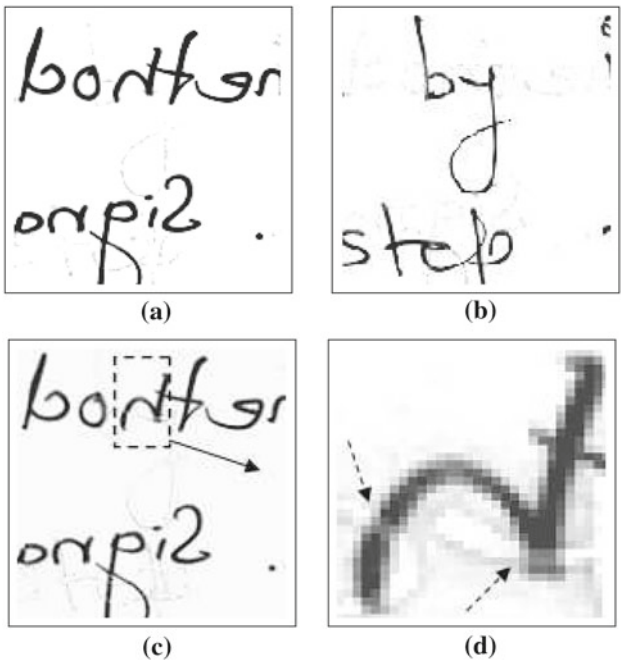

Fig. 7 Cancelation of show-through and blurring effects using the structure of Fig. 6, based on nonlinear structure and fixed filters. The results show that a linear separation model cannot provide perfect results, even if the errors are weak. **a** Output of the structure shown in Fig. 6. **b** Output of the structure shown in Fig. 6. **c** Output of a linear separating structure with averaging filters. **d** Distortion introduced by the use of a linear separation structure

prior knowledge about the shape of these filters. In addition, especially in ancient documents and due to the heterogeneity transparency of the paper, the filter can vary according to the location on the page. Therefore, it will be appropriate to determine the filter coefficients based on the actual shape of the point spread function of the paper which is proportional to the type of the paper. In this case, since the nonlinear mixing model in (13) is different from (6), we must find a new way to estimate the parameter vector, $\mathbf{p}$, and to find the coefficients of the filters, $\mathbf{H}_1$ and $\mathbf{H}_2$. The nonlinear mixing model (13) is not an instantaneous mixture; therefore, its blind separation is much more difficult. However, by adding some simplifying assumptions, it will be possible to estimate the shape of these filters as explained in the next subsection.

4.2 Estimating the shapes of the filters

If we assume that show-through is a linear phenomenon (i.e. $q_1 = q_2 = 0$) and also assume that we are near the convergence which implies $y_1^{(t)}|_{t\to\infty}$ to be almost equal to f_r^i (or f_v^i) and $y_2^{(t)}|_{t\to\infty}$ to be almost equal to f_v^i (or f_r^i) in (14), then the filter coefficients (entries of $\mathbf{H}_1$ and $\mathbf{H}_2$) can be determined through the minimization of the following cost functions with respect to the filter coefficients (note

that in these cost functions, f_r^s, f_v^s, f_v^i and f_r^i are the image matrices)

$$\begin{aligned} F_1 &= \left\| f_r^s + l_1 f_v^i * \mathbf{H}_1 \right\|^2 = \left\| f_r^s + l_1 \mathbf{H}_1 * y_2^{(t)} \Big|_{t\to\infty} \right\|^2 \\ &= \|e_1\|^2, \\ F_2 &= \left\| f_v^s + l_2 f_r^i * \mathbf{H}_2 \right\|^2 = \left\| f_v^s + l_2 \mathbf{H}_2 * y_1^{(t)} \Big|_{t\to\infty} \right\|^2 \\ &= \|e_2\|^2. \end{aligned} \tag{16}$$

These equations are inspired from (13) by setting q_1 and q_2 to zero. In addition, it is worth mentioning that by considering (13) and the fact that f_r^i and f_v^i are independent from each other, the minimum values of these two cost functions are equal to f_r^i and f_v^i respectively. However, since the original images f_v^i and f_r^i are unknown (actually it is our aim to determine them), it will be impossible to estimate the filter coefficients accurately by this way. However, from the show-through model of (13), it is obvious that the amplitude of the filters can be changed by the parameter vector, $\mathbf{p}$, so it will be sufficient to only estimate the shape of these filters up to a gain indeterminacy. By considering the fact that, as the iteration number (t) increases, outputs of the structure are expected to be closer to the original independent sources, we propose to approximately estimate the shape of the filters in each iteration by minimizing the following cost functions

$$\begin{aligned} F_1 &= \left\| f_r^s + l_1 \mathbf{H}_1^{(t)} * y_2^{(t)} \right\|^2 = \|e_1^{(t)}\|^2, \\ F_2 &= \left\| f_v^s + l_2 \mathbf{H}_2^{(t)} * y_1^{(t)} \right\|^2 = \|e_2^{(t)}\|^2. \end{aligned} \tag{17}$$

Note that we ignored the quadratic parts of the show-through model for estimating the shape of the filters because they are only used for modeling the nonlinearity of this phenomenon and therefore, have a weak contribution in the final scanned images (which means q_1 and q_2 have small values as we can check in the experimental results). Thus, using the linear model for show-through provides accurate enough estimation of the shape of the deblurring filters (point spread functions of the paper).

For minimizing the cost functions of the first equation in (17) with respect to the coefficients of the filter $\mathbf{H}_1$, we used the adaptive structure shown in Fig. 8. In this structure, the output is the filtered version of $y_2^{(t)}$, since we need it for calculating the output in (14). A similar structure can be used for estimating $\mathbf{H}_2$. Figure 9 shows the final Least Mean Square (LMS) algorithm for minimizing these cost functions and estimating the filter coefficients. To avoid divergence of the adaptive filters, the input images are normalized (such that they have zero mean with a unit variance) before being applied to the structure.

Since the SNR of $e_1^{(t)}$ $\left(\text{or } e_2^{(t)}\right)$ cannot be more than the noise to signal ratio of the input images (due to the *power inversion property* in adaptive filters) [11, page 78], we cannot use this adaptive method by itself for separating the sources (as done by Sharma [10]), since it will not lead to good separation results.

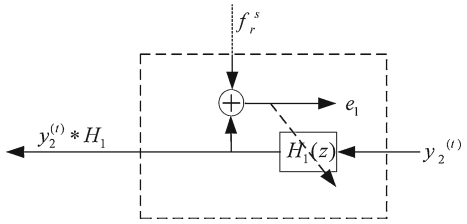

Fig. 8 Adaptive structure for estimating the shape of the point spread function of the paper

4.3 The modified structure

In the previous subsection, we proposed to use an adaptive filter to estimate the point spread functions of the paper in each iteration. The new separating structure associated with (14) is shown in Fig. 10. In this figure, each filtering block is the same as the adaptive filter detailed in Fig. 8. In this structure and at each iteration, we apply first the algorithm shown in Fig. 9 to the output images of the structure in the previous iteration. During the computation of the outputs of the filters, the coefficients of the filters are also updated to be used in the next iteration. Then, using the current value of the parameter vector, the outputs of the structure are calculated for all of the input pixels through the recurrent iterative expression of (14). Finally, the parameter vector of the structure is updated by using these new outputs. As we mentioned earlier, a new method should normally be introduced for updating the parameters vector in each iteration since we have changed the mixing model (compared to the show-through model represented in (6). However, because the effect of regions belonging to the blurring effect is small compared to the image itself, we propose to use the followings approximations: $\mathbf{H}_2 * y_1^{(t)} \approx y_1^{(t)}$ and $\mathbf{H}_1 * y_2^{(t)} \approx y_2^{(t)}$. Thus, the method designed for updating the parameter vector of the structure of Fig. 4 can be used for updating the parameter vector of the structure of Fig. 10. The approximation leads to a simplified algorithm.

Figure 11 shows the final show-through (and blurring) removal algorithm which we used for obtaining the results presented in Sect. 5. Note that since we do normalization at the beginning of this algorithm, there is no need to normalize the inputs of the algorithm described in Fig. 9 again. Applying this algorithm to Fig. 5a and b leads to results shown in Fig. 12a and b and demonstrates the effectiveness of our method in removing the blurring effect as well as the show-through.

1. Initialization:
   - Denote the input images by f_r^s and f_v^s.
   - Choose the size of the 2-D adaptive filters $\mathbf{H}_1$ and $\mathbf{H}_2$ and let the initial value of them to be equal to $\mathbf{0}$.
   - Choose a suitable value for the step-size of the adaptive LMS algorithm, μ_{adaptive}.
2. Compute the errors of the adaptive filters at pixel (m, n) with the current values of $\mathbf{H}_1$ and $\mathbf{H}_2$:

$$e_1^{(t)}(m,n) = f_r^s(m,n) + \sum_{k=-K}^{K}\sum_{l=-L}^{L} l_1 \mathbf{H}_1^{(t)}(k,l) y_2^{(t)}(m-k, n-l),$$

$$e_2^{(t)}(m,n) = f_v^s(m,n) + \sum_{k=-K}^{K}\sum_{l=-L}^{L} l_2 \mathbf{H}_2^{(t)}(k,l) y_1^{(t)}(m-k, n-l),$$

   where $(2K+1)\times(2L+1)$ is supposed to be the size of the 2-D adaptive filters and $y_1^{(t)}$ and $y_2^{(t)}$ are the output images of the structure in the previous iteration.
3. Update the coefficients of the adaptive filters in that pixel:
   for $k = -K \ldots K,\ l = -L \ldots L$:

$$\mathbf{H}_1^{(t+1)}(k,l) = \mathbf{H}_1^{(t)}(k,l) + \mu_{\text{adaptive}} f_r^s(m-k, n-l) e_1^{(t)}(m,n),$$

$$\mathbf{H}_2(k,l) = \mathbf{H}_2(k,l) + \mu_{\text{adaptive}} f_v^s(m-k, n-l) e_2^{(t)}(m,n).$$

4. If there is any other pixel, go to step 2 otherwise finish the loop.
5. $y_1^{(t)} * \mathbf{H}_2$ and $y_2^{(t)} * \mathbf{H}_1$ images are the outputs of the filtering blocks.

Fig. 9 Algorithm for estimating the adaptive filters of Fig. 8

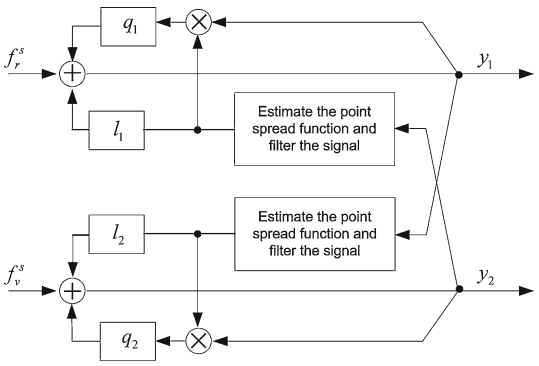

Fig. 10 Adaptive linear quadratic separating structure for removing show-through and blurring effects simultaneously

1. Initialization:
   - Normalize the input images and denote them by $f_r^{(s,N)}(m,n)$ and $f_v^{(s,N)}(m,n)$ where superscript N represents Normalization.
   - Choose the size of the 2-D adaptive filters $\mathbf{H}_1$ and $\mathbf{H}_2$ and let the initial value of them to be equal to $\mathbf{0}$.
   - Let the initial value of parameter vector $\mathbf{p}$ be equal to $\mathbf{0}$.
   - Choose suitable values for the step-size of the LMS adaptive algorithm μ_{adaptive} and for μ_{ML} which is needed for maximizing likelihood cost function in (10).
   - Let $y_i^{(t)}|_{t=0} = \mathbf{0}$ for $i = 1,\ 2$.
2. For $t = 1, 2, \ldots$ until convergence:
   - Apply the adaptive filter algorithm described in Fig. 9 to the images $y_i^{(t)}$ for $i = 1,\ 2$ by using the final values of $\mathbf{H}_1$ and $\mathbf{H}_2$ in the previous iteration as initial values.
   - Compute the output images of the structure shown in Fig. 10 by using Equation (14) and the current value of the parameter vector.
   - Update the parameter vector $\mathbf{p}$ by maximizing the likelihood cost function as follows

$$\mathbf{p}^{(j+1)} = \mathbf{p}^{(j)} + \mu_{ML} \frac{\partial L}{\partial \mathbf{p}}.$$

3. $y_1^{(t)}$ and $y_2^{(t)}$ will be the final results.

Fig. 11 Proposed algorithm for removing show-through and blurring effect

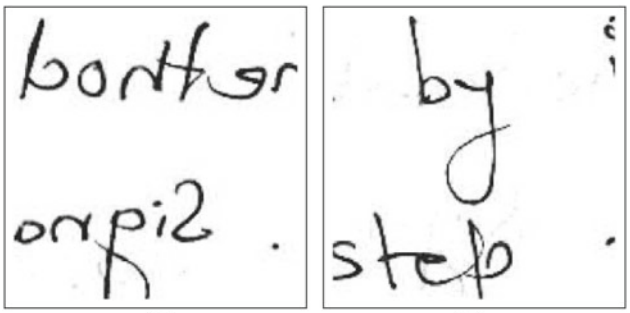

Fig. 12 Output results obtained by applying the algorithm of Fig. 11 to Fig. 5a and b that show the capability of our proposed method in reducing those degradations which are caused by show-through and blurring effect during the scanning process

5 Experimental results

In this section, we will demonstrate the advantages of the proposed algorithm for jointly removing the show-through and the blurring effects by performing experiments on real and artificial images. In the field of image processing, it is well known that usually subjective measures are better than quantitative measures (like SNR) for comparing the quality of two images. In addition, it is very hard to find a unique quantitative criterion for evaluating the performance of show-through removal algorithms and this is because of the fact that in almost all of the experiments performed on real degraded

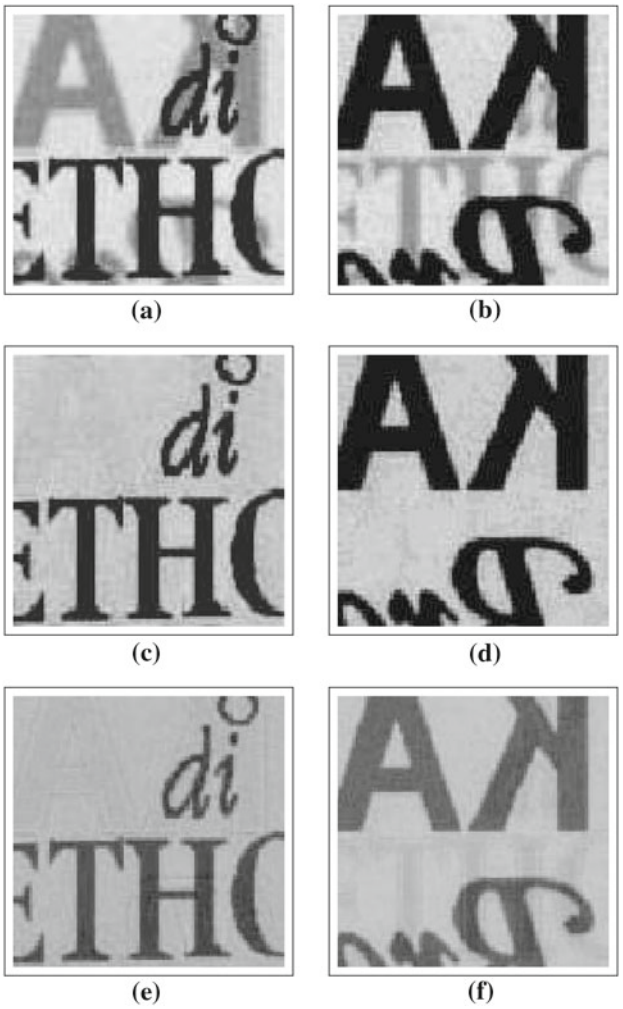

Fig. 13 Show-through and blurring cancelation on simulated linear mixtures. **a** and **b**: Original scanned images. **c** and **d**: outputs of the separating structure of Fig. 10. **e** and **f**: Output images of the structure of Fig. 4. We note that the quality is poor due to blurring effect which is not canceled. **a** Recto of the paper with show-through. **b** Verso of the paper with show-through. **c** Recovered image with adaptive filter. **d** Recovered image with adaptive filter. **e** Recovered image without filtering. **f** Recovered image without filtering

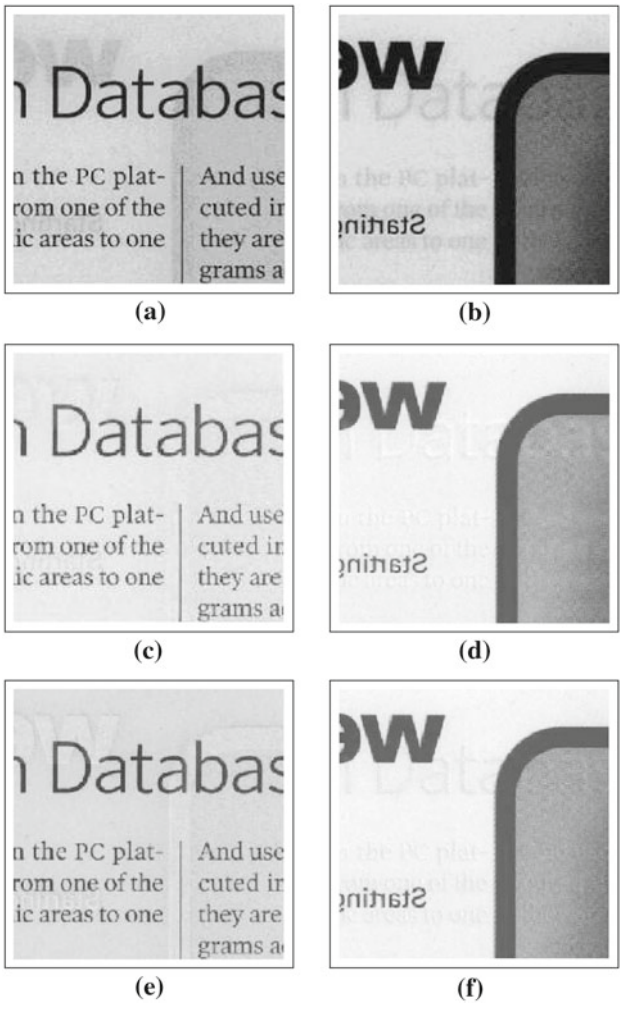

Fig. 14 Show-through and blurring cancelation of real world scanned images, mainly containing text. **a** and **b**: Original scanned images. **c** and **d**: Outputs of the separating structure of Fig. 10. **e** and **f**: Output images of the structure of Fig. 4. **a** Recto of the paper with show-through. **b** Verso of the paper with show-through. **c** Recovered image with adaptive filter. **d** Recovered image with adaptive filter. **e** Recovered image without filtering. **f** Recovered image without filtering

manuscripts or documents, original pure sources are not available. Therefore, quantitative criteria such as SNR are not always applicable. As a consequence, for comparing our method to previously developed algorithms qualitatively, we use the same images utilized in those works. In all of the following experiments, the initial value for parameters vector $\mathbf{p}$ and the coefficients of the adaptive filters are zero. In addition, μ_{adaptive} and μ_{ML} are set to 0.00001 and 0.005, respectively. Registration of the input images is done simply by shifting and rotating one of the input images with respect to the other one and minimizing the mean square error between them. Note also that in nonlinear BSS based on ICA, each source is retrieved at best up to an unknown nonlinear function [14,28]. Since for the recurrent structure utilized in this work the separability analysis has not yet be performed, indeterminacies involved in this problem are unknown. Nevertheless, we know that the two stable points of this recurrent structure differ from each other in scaling, permutation and additive constant [26]. Therefore, at the end of each simulation, we should adjust the DC and scaling of the output images. Since the retrieved images are similar to the scanned images except in those areas having show-through, the DC and scale of the images should not change significantly during the show-through removal process. As a result, in all of the following experiments, we change the DC and the scale of the output images so they will be equal to the DC and the scale of the input scanned images. As it is seen in (11) and (12), maximizing the likelihood needs

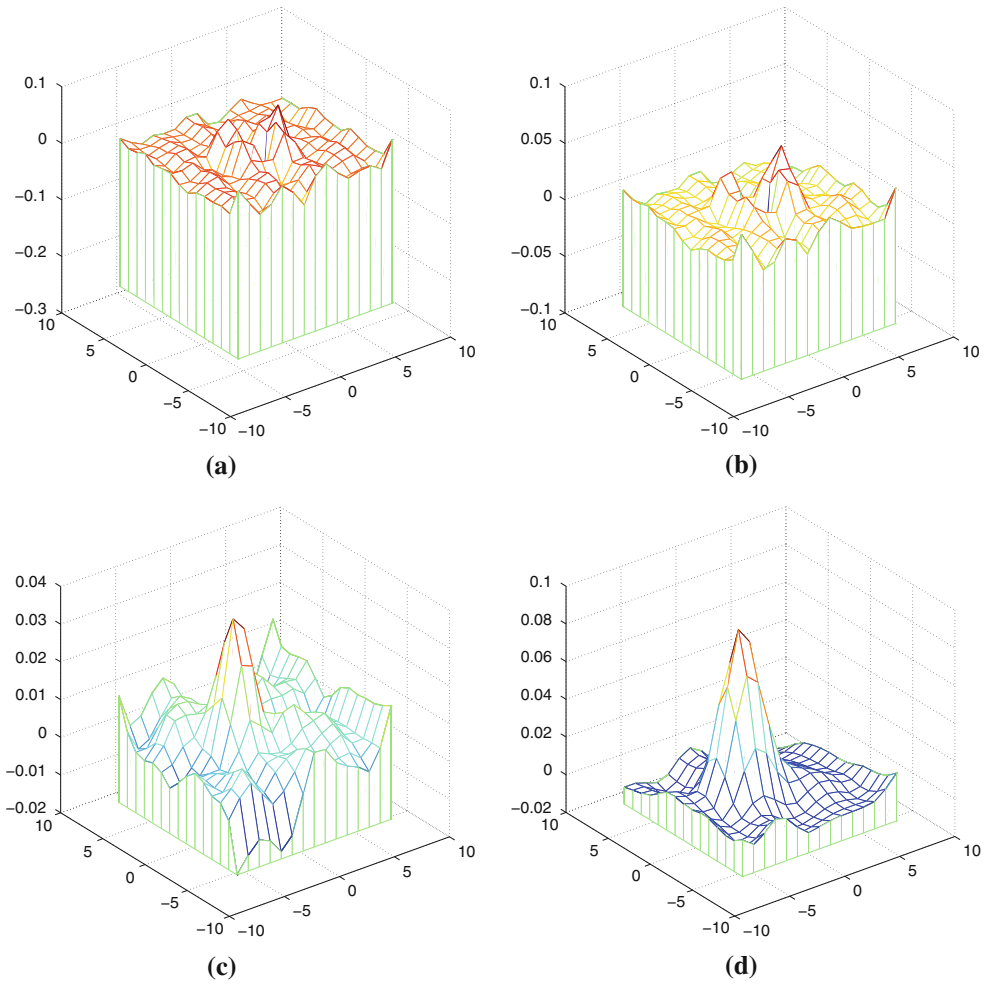

Fig. 15 Evolution of the coefficients of the adaptive filters during the convergence of the structure of Fig. 10 for our second experiment. **a** After 1 iteration. **b** After 3 iterations. **c** After 8 iterations. **d** After 25 iterations

estimation of the score functions. For this purpose, we used polynomial estimation of score functions [29] instead of the kernel estimation method used in [22], to obtain a fast and still accurate algorithm.

We can use two criteria for stopping the iterations: we can stop the algorithm (i) when the changes in the coefficients of the adaptive filters are less than a specific amount or (ii) when the changes in the parameter vector, $\mathbf{p}$, is small enough. Experimentally, the second option works better than the first one (because the second one is related to the show-through removal algorithm). However, since the two optimization algorithms (the one for obtaining the coefficients of the filters and the one for determining the parameter value) are working independently, it is relevant to check both criteria as we have done in our simulations.

In the following experiments, we ran the algorithm described in Fig. 11 in MATLAB 7.1 on a Windows XP PC with a 1.7 GHz CPU and 256 MB of RAM.

The results of the first simulation that we performed by using the structure shown in Fig. 10 are shown in Fig. 13c and d. The original images (Fig. 13a and b) with show-through are obtained from the paper [27]. These images have been produced artificially by adding a blurred version of the original images and then noise. Therefore, it is obvious that here the images have been mixed linearly, and not bilinearly.

The size of the input images is 99×98 pixels. The size of the adaptive filters was 9×9 and after the convergence of the structure, the final parameter vector was $\mathbf{p} = [-0.454, -0.42, 0.007, 0.012]^T$ which confirms the approximate symmetry property assumed in Sect. 2. The very small values of q_1 and q_2 confirms that this mixture is almost perfectly linear. For these particular input images, the algorithm converged after 26 iterations and each iteration took about 14 s. Output results are as perfect as the ones presented in [27] in removing show-through. Outputs of the

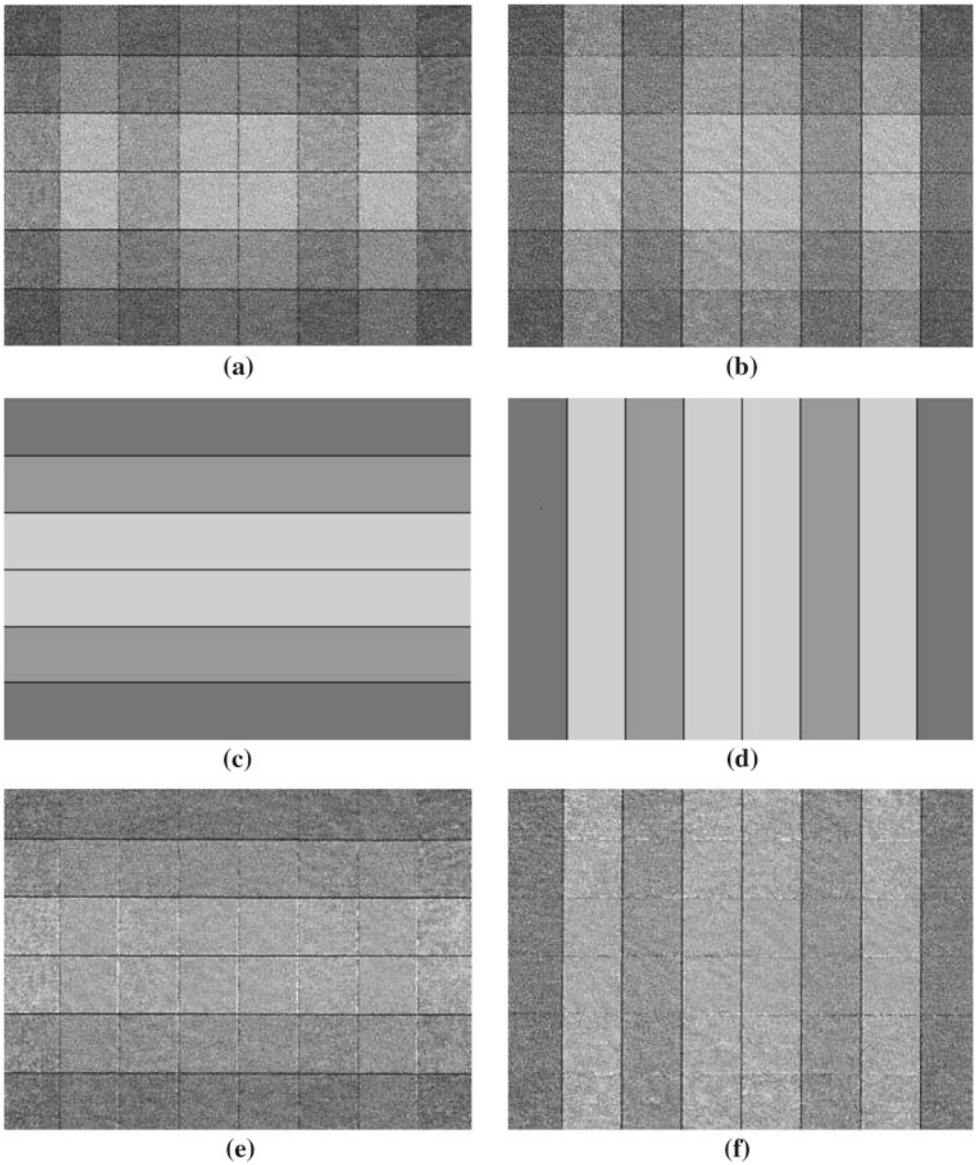

Fig. 16 Show-through and blurring cancelation on actual scanned images. **a** and **b** show the scanned input images applied to the structure of Fig. 10. **c** and **d** are the original images. Images shown in **e** and **f** are the outputs estimated by our algorithm, where show-through has been reduced significantly, even if the nonlinear mixture has not been perfectly canceled. **a** Recto of the paper with show-through after scanning the paper. **b** Verso of the paper with show-through after scanning the paper. **c** Original image printed on one side of the paper. **d** Original image printed on the other side of the paper. **e** First recovered image by using the structure of Fig. 10. **f** Second recovered image by using the structure of Fig. 10

structure of Fig. 4 for these input images are shown in Fig. 13e and f. In these images, show-through has been removed almost completely but the blurring effect does still exist. Note that in this and all of the following experiments, absolute values of l_1 and l_2 are less than unity. By considering (6), this is because of the fact that the strength of the show-through is less than the strength of the image which is printed on the backside of the paper. Moreover, these parameters have negative signs which shows that the back-side image is added to and not subtracted from the front-side image. The most time consuming part of the algorithm is the calculation of the score functions of the sources which are needed in (12). In addition, as the size of the input images or adaptive filters increases, computational time of the proposed algorithm increases as well.

In the next experiment, we considered the images previously used in [7]. These are real data obtained through the scanning process, with an image size equal to 401 × 401 pixels. In this case, the structure of Fig. 10 converged after 29 iterations and the final parameter vector equals to $\mathbf{p} = [-0.311, -0.287, 0.024, 0.031]^T$. In this experiment, the size of the adaptive filters in that structure was 11 × 11. Each iteration took about 57 s. Figure 14c and d show the final results. It is obvious especially in the right black part of the recovered image that the estimated front image has not become whiter in regions where the two (front and back) images have overlap, like what had been seen in Fig. 7c and d. Finally, Fig. 14e and f show the results obtained using the structure of Fig. 4, which demonstrate the relevance of adaptive filters in improving the quality of the output images.

Figure 15 illustrates the convergence of the coefficients of one of the adaptive filters through the run of the algorithm.

In the above experiments, the two input images were almost text and have little overlap with each others.

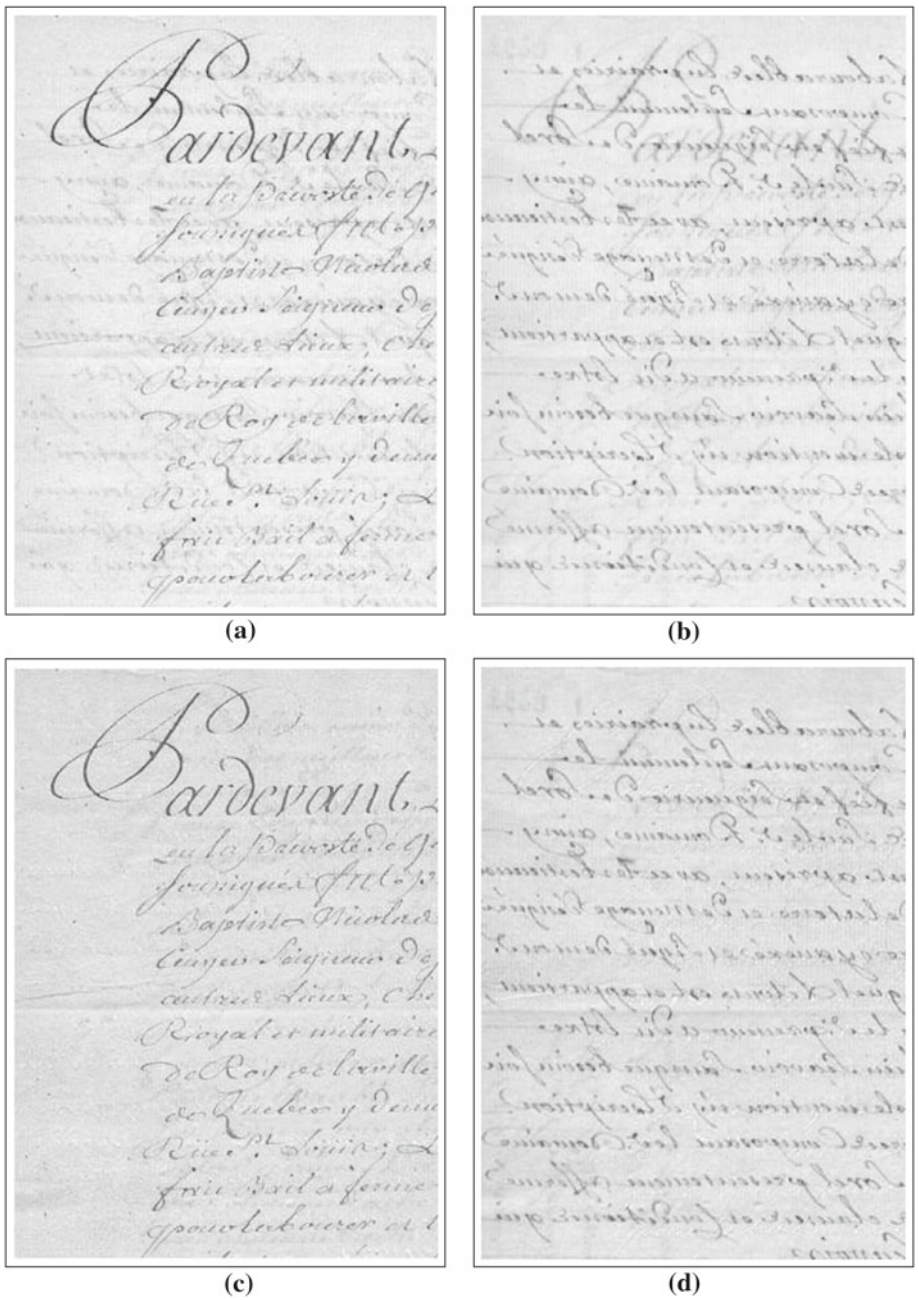

Fig. 17 Removing show-through from ancient real manuscripts. **a** and **b** show the scanned input images applied to Fig. 10. Output images of Fig. 10 in **c** and **d** show that the proposed method can increase the readability of the documents considerably. **a** Recto of the paper with show-through. **b** Verso of the paper with show-through. **c** First recovered image. **d** Second recovered image

Consequently, we performed another experiment on two real pictures strongly distorted by show-through as shown in Fig. 16a and b. These scanned images are obtained by scanning both sides of a paper on which the images shown in Fig. 16c and d had been printed. The image size is 326×242 and the algorithm converged after 71 iterations, and each iteration took about 35 s. The size of the adaptive filters was 5×5. After the convergence of the structure, the estimated parameter vector was $\mathbf{p} = [-0.476, -0.46, 0.041, 0.045]^T$. The outputs of the algorithm are shown in Fig. 16e and f. It is seen that the proposed structure removed the show-through considerably.

Another experiment is presented in Fig. 17. The recto and verso images with the size of 570×698 are obtained from http://www.site.uottawa.ca/~edubois/documents. However, to improve visual quality, only parts of them are shown in Fig. 17a and b. These are real ancient manuscripts whose readability is degraded due to the presence of show-through. The result of applying our show-through removal method to these figures is shown in Fig. 17c and d. The amount of enhancement achieved in the readability of the manuscripts is evident in these figures. In this simulation, adaptive filters of size 3×3 have been used. Execution time of this experiment was about 2,163 s

and the values of parameters vector converged to $\mathbf{p} = [-0.301, -0.318, 0.027, 0.032]^T$.

In all of these simulations, the approximate size of the filters is determined heuristically based on the severity of the blurring effect that we visually observed in scanned images: as the blurring effect becomes worse, the filter size should be increased. Then, the exact size of the filters are determined by trial and error.

6 Conclusion

In this paper, we experimentally showed that show-through is a nonlinear phenomenon and can be approximately represented by a linear-quadratic model. For such a mixing model, based on [22,23], we first used a linear-quadratic separating structure which is able to extract the original sources. Simulations show that this structure can remove show-through but cannot cancel the blurring effect implied by the scanning process due to the paper's intrinsic transparency. Therefore, we proposed a newer adaptive linear-quadratic blind source separating structure, in which one adds adaptive filters to estimate the shape of the point spread function of the paper. Therefore, using this separating structure, one can simultaneously remove show-through and blurring effects. Finally, we justified the effectiveness of this new method with a few experiments on simulated mixtures and real-world scanned images. The results were very satisfactory in comparison with other current show-through removal methods.

The main advantages of this show-through removal method can be summarized as:

- It may be extended to more complicated polynomial models [22,23];
- Filter coefficients are updated adaptively. Therefore, the structure is able to estimate the point spread function of the paper and remove the blurring effect as well, without priors on the paper properties;
- This structure is also applicable to linear mixtures: in this case, the parameters q_1 and q_2 will become zero.

For the near future, we are planning to complete our study about the theoretical invertibility of bilinear mixing models. In addition, investigating of the effects of those simplifying assumptions that we used in our method can be useful for improving the performance of the proposed algorithm. Another problem to be addressed is to find a procedure to evaluate the quality of the output results. By these objective performance indexes, we can provide more consistent comparisons between show-through removal algorithms.

References

1. Larsson, L.O., Trollsås P.O.: Print-through as an ink/paper interaction effect in newsprint, the fundamental properties of paper related to its uses. In : Bolam, F. (ed.) Trans. of the Cambridge Symposium, vol. 1, pp. 600–612. London (1976)
2. Nishida, H., Suzuki, T.: A multi-scale approach to restoring scanned colour documents with show-through effects. In: Proceedings Seventh International Conference on Document Analysis and Recognition, vol. 1, pp. 584–588 (2003)
3. Wang, Q., Xia, T., Li, L., Tan, C.L.: Document image enhancement using directional wavelet. In: Proceedings IEEE Conference Computer Vision Pattern Recognition, vol. 2, pp. 534–539 (2003)
4. Knox, K.: Show-through correction for two-sided documents. United States Patent 5,832,137 (1998, Nov)
5. Wang, Q., Tan, C.L.: Matching of double-sided document images to remove interference. IEEE CVPR2001 (2001, Dec)
6. Tonazzini, A., Salerno, E., Bedini, L.: Fast correction of bleed-through distortion in grayscale documents by a blind source separation technique. Int. J. Doc. Anal. IJDAR(10) (1), 17–25 (2007)
7. Ophir, B., Malah, D.: Show-through cancellation in scanned images using blind source separation techniques. IEEE International Conference on Image Processing, **3**, 233–236 (2007)
8. Tonazzini, A., Salerno, E., Mochi, M., Bedini, L.: Bleed-through removal from degraded documents using a color decorrelation method. In: Proceedings Document Analysis Systems VI: 6th International Workshop, Springer-Verlag GmbH, vol. 3163 of Lecture Notes in Computer Science, pp. 229–240 (2004)
9. Almeida, M.S.C., Almeida, L.B.: Separating nonlinear image mixtures using a physical model trained with ICA. In: Proceedings 2006 IEEE International Workshop Machine Learning for Signal Processing, Maynooth, Ireland (2006, Sep)
10. Sharma, G.: Show-through cancellation in scans of duplex printed documents. In: IEEE Transactions on Image Processing, vol. 10, no. 5, pp. 736–754 (2001, May)
11. Farhang-Boroujeny, B.: Adaptive Filters, Theory and applications. Wiley , London (1998)
12. Castro, P., Almeida, R.J., Caldas Pinto, J.R.: Restoration of double-sided ancient music documents with bleed-through. In: CIARP 2007, pp. 940–949 (2007)
13. Almeida, L.B.: Separating a real-life nonlinear image mixture. J. Mach. Learn. Res. **6**, 1199–1232 (2005)
14. Jutten, C., Babaie-Zadeh, M., Hosseini, S.: Three easy ways for separating nonlinear mixtures? . Signal Process. **84**(2), 217–229 (2004)
15. Hosseini, S., Jutten, C.: On the separability of nonlinear mixtures of temporally correlated sources. IEEE Signal Process. Lett. **10**(2), 43–46 (2003)
16. Almeida, L.B.: Linear and nonlinear ICA based on mutual information. In: Proceedings IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (AS-SPCC), pp. 117–122. Lake Louise, Canada (2000, Oct)
17. Almeida, L.B.: MISEP—linear and nonlinear ICA based on mutual information. J. Mach. Learn. Res. **4**, 1297–1318 (2003)
18. Tonazzini, A., Bedini, L., Salerno, E.: A markov model for blind image separation by a mean-field EM algorithm. IEEE Trans. Image Process. **26**(2), 473–482 (2006)
19. Wolf, C.: Document ink bleed-through removal with two hidden markov random fields and a single observation field. IEEE Trans. Pattern Anal. Mach. Intell. **32**(3), 431–447 (2010)
20. Hosseini, S., Jutten, J., Pham, D.T.: Markovian source separation. IEEE Trans. Signal Process. **51**(12), 3009–3019 (2003)
21. Moussaoui, S., Hauksdottir, H., Schmidt, F., Jutten, C., chanussot, J., Brie, D., Doute, S., Benediktsson, J.: On the

decomposition of mars hyperspectral data by ICA and bayesian positive source separation. Neurocomputing **71**(10–12), 2194–2208 (2008)
22. Hosseini, S., Deville, Y.: Blind maximum likelihood separation of a linear-quadratic Mixture. In: Proceedings 5th International Conference on Independent Component Analysis and Blind Source Separation (ICA'04), pp. 694–701. Granada, Spain (2004, Sep)
23. Hosseini, S., Deville, Y.: Correction to: blind maximum likelihood separation of a linear-quadratic mixture. available at http://arxiv.org/abs/1001.0863
24. Merrikh-Bayat, F., Babaie-Zadeh, M., Jutten, C.: A nonlinear blind source separation solution for removing the show-through effect in the scanned documents. In: 16th European Signal Processing Conference (EUSIPCO-2008). Lausanne, Switzerland (2008, Aug)
25. Jutten, C., Hérault, J.: Blind separation of sources, part I: an adaptive algorithm based on neuromimetic architecture. Signal Process. **24**, 1–10 (1991)
26. Hosseini, S., Deville, Y.: Blind separation of linear-quadratic mixtures of real sources using a recurrent structure. In: Proceedings IWANN, vol. 2, pp. 241–248. Mao, Menorca, Spain (2003, June)
27. Tonazzini, A., Gerace, I.: Bayesian MRF-based blind source separation of convolutive mixtures of images. In: EUSIPCO. Antalya, Turkey (2005, Sep)
28. Babaie-Zadeh, M.: On Blind Source Separation in Convolutive and Nonlinear Mixtures. Ph.D. thesis, INP Grenoble (2002)
29. Babaie-Zadeh, M., Jutten, C.: A general approach for mutual information minimization and its application to blind Source separation. Signal Process. (Elsevier) **85**(5), 975–995 (2005)