Blind image separation for document restoration using plug-and-play approach

Xhenis Çoba, Fangchen Feng, Azeddine Beghdadi
Laboratoire de Traitement et Transport de l'Information (L2TI)
Université Sorbonne Paris Nord
Villetaneuse, France
xhenis.coba@etu.univ-grenoble-alpes.fr, fangchen.feng@univ-paris13.fr, azeddine.beghdadi@univ-paris13.fr

***Abstract*—In this paper we propose a new method for document image restoration based on Blind Source Separation. The existing separation methods rely on the general properties of source images such as independence, sparsity, and non-negativity. In this work we show that by exploiting some characteristics of image denoising methods in a play-and-plug scheme, efficient BSS results could be achieved. In particular, we show that the use of BM3D and Non-local Means denoising methods as ingredients in the proposed scheme, which exploits the non-local properties of the image, leads to better image separation in terms of convergence rate and perceptual image quality. We also propose to use the dictionary-learning approach to take the concept of visual chirality into consideration. Finally, we apply the proposed scheme to document image restoration problem and show its advantage through experiments and objective performance evaluation.**

***Index Terms*—Blind image separation, Document restoration, Plug-and-play, Image denoising, Dictionary learning**

I. Introduction

Ancient documents often suffer from various physical degradations over time such as show-through and bleed-through effect (see Fig. 1 as an example). Show-through is a front-to-back interference, mainly due to the scanning process and the transparency of the paper, which causes the text in the verso side of the document to appear in the recto side (and vice versa) [1]. Bleed-through is an intrinsic physical deterioration due to ink seeping and produces an effect similar to that of show-through [2]. Despite the complex physical model of the process which depends on some chemical and physical properties of the paper such as, its porosity, ink diffusion, the reflectance of the verso, and the spreading of light in the paper [3]–[8], the signal acquisition process could be represented through the following model to account for these effects locally [2] :

$$\mathbf{x}[i] = A\mathbf{s}[i] + \mathbf{n}[i], \tag{1}$$

where i is the index of the image pixels. $\mathbf{s}[i] = [s_r[i], s_v[i]]^T$ is the concatenation of the source images in gray-scale where $s_r[i]$ and $s_v[i]$ are the recto and verso sides (with a horizontal flip) of the source images respectively. $\mathbf{x}[i] = [x_r[i], x_v[i]]^T$ with $x_r[i]$ and $x_v[i]$ being the observed recto and verso side images. $A \in \mathbb{R}^{2\times 2}$ is the mixing matrix. $\mathbf{n}[i]$ is the additive white Gaussian noise.

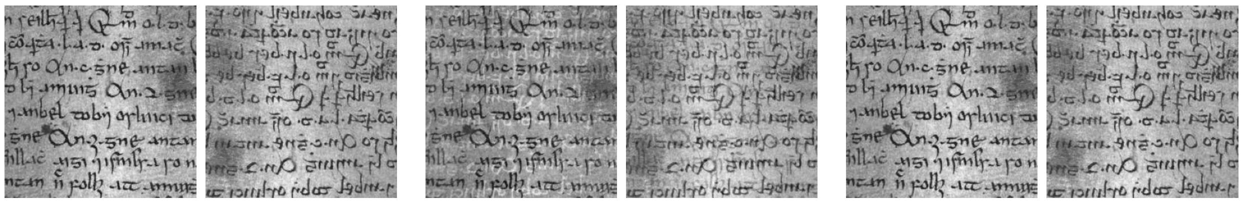
The goal of blind document restoration is to recover the source images $\mathbf{s}[i]$ from the observed image containing recto-verso mixtures $\mathbf{x}[i]$ without knowing the mixing matrix A. This problem falls in the general framework of Blind Source Separation (BSS) [9].

Among the classical methods, the Independent Component Analysis (ICA) [10] could be considered as a BSS solution. The key idea of ICA is to look for a demixing matrix such that the estimated source components are as independent as possible. Despite its relatively good performance, the ICA does not work if more than one of the sources follow the Gaussian distribution since ICA-based methods rely on higher-order statistics [10]. Sparse Component Analysis (SCA) is another well investigated approach for BSS [11]–[13]. This approach assumes that the sources are sparse in the spatial or in the transform domain. The separation problem is then formulated through an optimization framework and the estimation of the mixing matrix and the source images are performed simultaneously. Compared to ICA, SCA-based methods can deal with dependent sources [14] and produce better results in the case of mixtures degraded by additive noise [13]. It is also important to notice that several works have been dedicated to exploiting the links between ICA and SCA methods [15]–[17].

Besides the independence and sparsity properties of the source images, other image representations have also been investigated for source separation such as Non-negative Matrix Factorization (NMF) [18] and Canonical Correlation Analysis (CCA) for document image restoration [2].

For blind image separation (BIS), in the determined scenario [1], the demixing (or the mixing) matrix and the source images are estimated based on different source image characteristics. In the proposed method, instead of searching for other image characteristics for blind image separation, we propose to use the output of image denoising as a part of the estimation process, to improve the efficiency of the BIS algorithm. Indeed, extensive research has been carried out on image denoising over the past three decades, which has led to highly efficient algorithms [19]–[22]. This is what motivated

[1]The number of observations equals the number of sources.



(a) Original sources (b) Mixtures (c) Estimated sources

Fig. 1: From left to the right by pairs: the original sources, the mixed images and the estimated sources images using the proposed Double Dictionary Learning. See Section IV for implementation details.

us to integrate these algorithms as ingredients in our BIS approach.

In addition, visual chirality-the notion of objects distinct to their mirror image, has been recently exploited in a learning framework [23] to detect whether an image has been horizontally flipped or not. This concept is particularly useful in the case of old document images, as the verso page appears to be mirrored, viewed from the recto side because of the bleed-through and show-through effects.

To the best of our knowledge, this paper introduces for the first time the non-local image characteristics and chirality property [23] of document images as key ingredients in blind image separation scheme and particularly in a very challenging real-word problem.

The main contributions of the paper are the following:

- A blind source separation framework, that encompasses several existing denoising methods, is developed and described in the context of document image restoration.
- The idea of using non-local characteristics of the image signal in this BIS scheme is proposed for the first time.
- A dictionary learning strategy which improves the separation by considering the visual chirality of document images is proposed.

The reminder of this paper is organized as follows. In Section II, we describe the proposed separation method with plug-and-play approach. We describe the proposed dictionary learning strategy in Section III. The performance evaluation of the proposed method is presented in Section IV. Section V is dedicated to conclusion and some open problems.

II. Plug-and-play approach for image separation

In the following, we explain the motivation of using plug-and-play for image separation and describe the proposed approach.

A. The proposed optimization framework

The ICA methods can be formulated in an optimization framework [15] as follows:

$$\underset{W,\mathbf{s}}{\operatorname{argmin}} \frac{1}{2}\|W\mathbf{x}-\mathbf{s}\|_F^2 + \sigma\mathcal{P}(\mathbf{s}) + \mathcal{G}(W), \tag{2}$$

where W is the demixing matrix. $\|\cdot\|_F$ denotes the Frobenius norm. σ is the hyperparameter used to control the source penalty term. $\mathcal{P}(\mathbf{s})$ is the independence measure for the sources and $\mathcal{G}(W)$ is the unitary constraint which leads to source decorrelation through a whitening pre-processing.

For SCA, the separation is usually performed based on the mixing matrix instead of the demixing matrix as follows:

$$\underset{A,\mathbf{s}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{x}-A\mathbf{s}\|_F^2 + \sigma\mathcal{P}(\mathbf{s}) + \mathcal{G}(A), \tag{3}$$

$\mathcal{P}(\mathbf{s})$ is the sparse penalty in the spatial or the transform domain and $\mathcal{G}(A)$ is the unit-norm constraint of the mixing matrix to regularize the scaling ambiguity between the mixing matrix and the source images.

The above two formulations are closely related [15], but have some major differences. The formulation (3) is more adapted for mixtures degraded with additive noise and can be easily generalised to under-determined scenario. As we don't consider additive noise in this paper, we focus on the formulation (2). Another reason of such choice is that the formulation (2) can be easily adapted in the plug-and-play scheme.

Indeed, this optimization problem in the general form can be solved with the iterative alternating minimization algorithm [24]. At each iteration, the following sub-problems are solved in an alternating way until convergence:

$$\underset{\mathbf{s}}{\operatorname{argmin}} \frac{1}{2}\|W\mathbf{x}-\mathbf{s}\|_F^2 + \sigma\mathcal{P}(\mathbf{s}), \tag{4}$$

$$\underset{W}{\operatorname{argmin}} \frac{1}{2}\|W\mathbf{x}-\mathbf{s}\|_F^2 + \mathcal{G}(W). \tag{5}$$

The sub-problem (4) is directly related to the classical problem of denoising additive white Gaussian noise. This is the main idea behind the use of the plug-and-play approach [25]–[28]. It is worth noticing that the considered denoising algorithms do not necessarily correspond to any penalty term in a closed form as in (4). In this paper, we illustrate the separation results using Total-Variation (TV) minimization [29], BM3D [21] and Non-local Mean (NLM) [19] denoising approaches and propose a dictionary learning denoiser [30] based on the chirality property of document images as the first step of our algorithm.

In the denoising scheme, the hyper-parameter σ is directly related to the noise variance. In the plug-and-play scheme, σ controls the regularization term of the source images which enforces the desired property on the solution [28] (see Fig. 2 for an illustration). Note that, since we do not consider additive noise in the mixing model (1), the hyper-parameter in the proposed framework only serves as a weighting parameter for the source penalty term $\mathcal{P}(\mathbf{s})$. Theoretically, a small σ should be chosen for noiseless mixtures. However, as evidenced by the experimental results, an appropriate choice of σ leads to a rapid convergence towards the expected separation result.

The sub-problem (5) can be solved with the gradient descent approach with the unitary projection. The proposed algorithm is summarized as follows where Denoise_σ can be any aforementioned denoising method with a fixed hyper-parameter σ and $\mathcal{U}$ is the unitary projection.

Algorithm 1: Plug-and-play for blind image separation

```
Whiten the mixture x;
Initialization : W ∈ R^{2×2}, s ∈ R^{2×I};
repeat
    Update the images with a denoising method:
    s ← Denoise_σ(s);
    Update the mixing matrix by Least-Square
      followed by the unitary projection:
    W ← U(sx^T);
until convergence;
```

B. Related BIS methods

The proposed framework encompasses several existing source separation approaches. When the wavelet-based denoiser is used, the proposed approach is closely related to wavelet-based SCA proposed in [13], [31], [32]. When the TV denoiser is used, the proposed approach is related to the gradient-based sparsity method [11] and when the dictionary learning denoiser [30] is used, the proposed approach is closely related to the method proposed in [33]. It is also important to mention that the proposed framework is not limited to the aforementioned denoisers.

III. Dictionary learning for visual chirality

The use of BM3D and NLM as ingredients leads to better separation compared to TV denoiser, as they exploit the non-local characteristics of images (see IV-D for experimental results). However, they suffer from computational burden due to their complexity. In order to solve this problem and further improve the performance, we propose to use the dictionary learning denoiser which not only takes the non-local image characteristics into account, but also considers the chirality of document images.

For document image restoration from scanned recto-verso document pages, the two original document images are mixed with one of them (verso side) naturally horizontally flipped. It has been shown recently in [23] that text is a strong signal for high-level visual chirality which is a fundamental property of images. By predicting whether an image is mirrored or not using self-supervised learning, they claim that visual chirality exists in many vision datasets and should be taken into account when developing real-world vision systems. Classical denoisers (TV minimization, BM3D and NLM) are not sensitive to such image property and we show by experiments that the chirality property can be well captured by dictionary-learning. More precisely, two different dictionaries (one for document images in the readable direction and another for document images in the unreadable direction) are learned and the learned dictionaries are used in the denoising step for recto and verso images respectively. This strategy takes into account the intrinsic discrepancy of recto and verso images in the BIS process of document images, thus improves the performance. It's important to mention that such learning strategy can be further improved by discriminative learning process [34], [35] and will be part of the future work.

For synthesized mixtures, as we suppose that the scanned recto and verso pages are perfectly aligned, the above process can be simplified by using only one dictionary. The dictionary is first learned from a dataset of document images where all the images are in the readable direction. The learned dictionary is then used as the denoiser for the estimated recto image. For the verso image, the image is horizontally flipped into the readable direction before the denoising process using the same dictionary. We refer to this approach as the Double Dictionary Learning. Compared to classic denoisers, this method benefits from the learning paradigm and the chirality property of document images. For comparison purposes, to negate the advantages of the learning approach and to emphasize the benefits of chirality property, we also develop an ablation method where such flipping is not applied. We refer to this approach as Single Dictionary Learning which is used as a baseline.

IV. Performance evaluation of the proposed method

In this section, the performance of the proposed source separation framework is evaluated. We first consider various mixing scenarios by using the algorithm with different parameters. We then focus on the old document restoration application.

A. Experimental setup

In this preliminary study, the experiments are performed with synthesized image mixtures. In total, we use 30 gray-scale old document images of size 256×256 (15 sets of image pairs) from the data set available in [36]. The image sources are mixed using a 2×2 rotation mixing matrix according to the model (1). We don't add any noise to the observations. For the separation, the Principle Component Analysis (PCA) is performed to whiten the mixtures and the proposed separation algorithm is initialized with the whitened mixtures. We set the stopping criteria as $\|\mathbf{s}^{(k+1)} - \mathbf{s}^{(k)}\|_F^2 < 10^{-8}$ where $\mathbf{s}^{(k)}$ is the

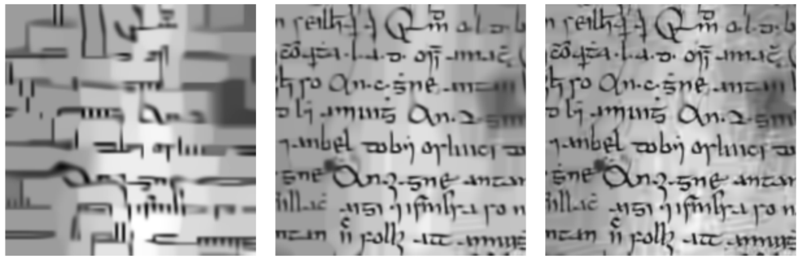


Fig. 2: The results of BM3D denoising on the original recto image in Fig. 1 with different σ values. From left to right: $\sigma = 2 \times 10^{-2}$, 8×10^{-3}, 5×10^{-3}. As a denoiser, BM3D is capable of regularising document images with adapted σ value (5×10^{-3} in this case). Note that high σ value results in an over-smoothing of the image leading thus to loss of fine details.

estimated source image at the k-th iteration. We also limit the maximum number of iterations to 100.

To better analyze the performance of the source separation algorithm, we evaluate both the estimated source images and the estimated demixing matrix. More precisely, we use the Structural Similarity Index Measure (SSIM) [37] to evaluate the perceptual closeness of the estimated images to the original ones. This measure takes values between 0 and 1 and higher SSIM indicates a better performance. Note that for all the experiment, we show the mean value of the SSIM of the two estimated recto and verso source images. To evaluate the demixing matrix, we consider the following Diagonal Ratio (DR):

$$\mathrm{DR}(\hat{W}) = \frac{\sum |[\hat{W}A]_{\text{off}}|}{\sum |[\hat{W}A]_{\text{diag}}|}, \tag{6}$$

where $\hat{W}$ is the estimated demixing matrix and $[\hat{W}A]_{\text{off}}$ and $[\hat{W}A]_{\text{diag}}$ are the off-diagonal and diagonal element of the matrix $\hat{W}A$, respectively. This measure takes values between 0 and 1 and a lower value indicates a better separation result.

B. Source Separation Performance - the role of the hyper-parameter σ

Here we briefly study the effect of both the denoiser and the hyper-parameter on the performance of the proposed BIS scheme. The hyper-parameter σ used in the Algorithm 1 is a key weighting parameter that controls the impact of the source penalty term on the result. We show the results of using BM3D and TV as the denoiser and use a rotation mixing matrix with an angle $\theta = 30°$. Fig. 3 depicts the separation evaluations as a function of the number of iterations.

One can remark that the choice of σ affects the performance of the algorithm significantly. For TV denoiser, with $\sigma = 0.3 \times 10^{-2}$ a convergence is attained in 25 iterations whereas for $\sigma = 0.3 \times 10^{-3}$, 35 additional iterations are needed. For both denoisers, a too-large σ leads to unacceptable results. These observations are coherent with the claim in Section II-A that a small value of σ is expected for noiseless scenarios. It also validates the claim in [28] that a too-small σ leads to slow convergence. Similar behavior is observed with NLM which is not shown here due to the limited space.

In the following, we use the σ value corresponding to the optimal performance of the considered denoisers.

C. Separation performance evaluation - the role of mixing matrix

In the determined scenario, the linear mixing process can be seen as a rotation of the two sources. Indeed, with the whitening pre-processing, the mixing matrix can be reduced into a rotation matrix as follows:

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{7}$$

where θ is the rotation angle. Note that a large rotation angle indicates a hard-to-separate mixture. Here, we test the proposed approach with TV and Double Dictionary denoiser for different mixing conditions, i.e. with different θ. More precisely, we consider $0° \leq \theta \leq 45°$. Note that testing the proposed approach with a θ outside this range does not provide any additional information due to trigonometric properties. Fig. 4 illustrates the separation results as a function of the number of iterations.

As expected, for large θ, more iterations are needed to reach the convergence for the considered denoisers. It is also interesting to notice the difference between the mean SSIM for different mixing conditions which is consistent with the expectation that higher θ yields hard-to-separate mixtures. Such difference is not visible in terms of Diagonal Ratio (DR) and this shows that the mean SSIM is a more sensitive measure and illustrates the importance of using multiple evaluation methods.

D. Separation performance evaluation - the role of the denoisers

The performance of the proposed plug-and-play BIS using different denoising techniques as ingredients is evaluated on ancient document images with the optimized hyper-parameter σ. The mixture images are synthesized by using the rotation matrix with $\theta = 45°$ corresponding to the hardest-to-separate mixture. We deliberately choose this condition to test the limit of the proposed approach. For dictionary learning denoiser (Double Dictionary and Single Dictionary), the dictionary is learned from a document image data set filtered by BM3D

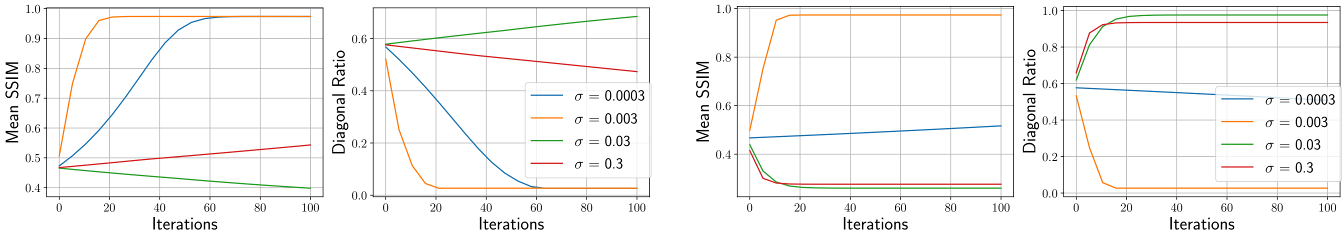


Fig. 3: Mean SSIM and Diagonal Ratio (DR) for TV [29] (two leftmost) and BM3D [21] (two rightmost) denoising approach using different σ.

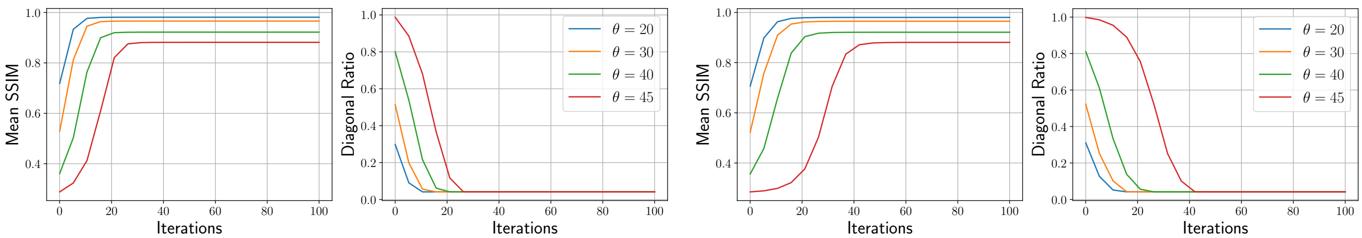


Fig. 4: Mean SSIM and Diagonal Ratio (DR) for Double Dictionary Learning (two leftmost) and TV denoising [29] (two rightmost) methods using different θ.

denoiser with $\sigma = 5 \times 10^{-3}$, and we then perform the separation based on the learned dictionary as described in Section III. Fig. 5 depicts the performance of the proposed method. In this scenario, ICA method does not perform well [2].

It can be noticed that Double Dictionary Learning and the NLM approach are clearly the two methods that yield the best results in terms of convergence rate. However, NLM denoising takes a considerable amount of time compared to Double Dictionary Learning, making the latter preferable in terms of practical use. Moreover, Double Dictionary Learning performs better than Single Dictionary Learning, emphasizing the advantage of exploiting visual chirality property of document images for the design of BIS scheme.

Fig. 1 illustrates an example of the process through the results obtained with the optimal hyper-parameter using Double Dictionary Learning for document images.

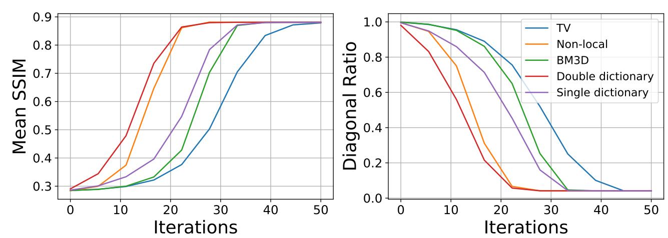


Fig. 5: Separation performance of the proposed plug-and-play framework with TV [29], NLM [19], BM3D [21] and the proposed Single and Double dictionary learning.

V. Conclusion and Future Work

In this paper, we demonstrated that the blind image separation problem could be addressed in an efficient way by incorporating in the source estimation scheme some of the existing image denoising methods. Indeed, through this study, it has been shown that the use of TV minimization, BM3D, and NLM denoisers as ingredients in the BIS scheme leads to promising results in the context of old document images. Moreover, we demonstrated that thanks to the visual chirality property, satisfied in the case of document images and exploited through the dictionary learning approach, a substantial computational time saving is achieved while improving the perceptual quality of the estimated image sources. This first work could be considered as a first step to put forward the proposed strategy to deal with various challenging problems related to the multidimensional signal separation context.

As for future work, we will extend this framework to color images and multi-spectral images. Discriminative learning process will also be considered to improve the separation. Furthermore, besides the experiments on simulated mixtures from real images, the next step would be to use real old document images containing specific mixtures corresponding to show-through and bleed-through degradation. Another challenging future work is to develop performance evaluation metrics for perceptual image quality assessment dedicated to BIS.

[2]Mean SSIM $\leq 0.89 \times 10^{-3}$. More results with ICA can be found on our GitHub page https://github.com/ffc28/ImageSeparationL2TI

References

[1] Gaurav Sharma, "Show-through cancellation in scans of duplex printed documents," *IEEE Transactions on Image Processing*, vol. 10, no. 5,

pp. 736–754, 2001.

[2] Antonio Boccuto, Ivan Gerace, and Valentina Giorgetti, "A blind source separation technique for document restoration," *SIAM Journal on Imaging Sciences*, vol. 12, no. 2, pp. 1135–1162, 2019.

[3] Anna Tonazzini, Pasquale Savino, and Emanuele Salerno, "A non-stationary density model to separate overlapped texts in degraded documents," *Signal, Image and Video Processing*, vol. 9, no. 1, pp. 155–164, 2015.

[4] Francesca Martinelli, Emanuele Salerno, Ivan Gerace, and Anna Tonazzini, "Nonlinear model and constrained ML for removing back-to-front interferences from recto–verso documents," *Pattern Recognition*, vol. 45, no. 1, pp. 596–605, 2012.

[5] Reza Farrahi Moghaddam and Mohamed Cheriet, "Low quality document image modeling and enhancement," *International Journal of Document Analysis and Recognition (IJDAR)*, vol. 11, no. 4, pp. 183–201, 2009.

[6] Mariana SC Almeida and Luís B Almeida, "Wavelet-based separation of nonlinear show-through and bleed-through image mixtures," *Neurocomputing*, vol. 72, no. 1-3, pp. 57–70, 2008.

[7] Anna Tonazzini, Ivan Gerace, and Francesca Martinelli, "Multichannel blind separation and deconvolution of images for document analysis," *IEEE Transactions on Image Processing*, vol. 19, no. 4, pp. 912–925, 2009.

[8] Farnood Merrikh-Bayat, Massoud Babaie-Zadeh, and Christian Jutten, "A nonlinear blind source separation solution for removing the show-through effect in the scanned documents," in *2008 16th European Signal Processing Conference*. IEEE, 2008, pp. 1–5.

[9] Pierre Comon and Christian Jutten, *Handbook of Blind Source Separation: Independent component analysis and applications*, Academic press, 2010.

[10] Aapo Hyvärinen, Juha Karhunen, and Erkki Oja, *Independent component analysis*, vol. 46, John Wiley & Sons, 2004.

[11] Wided Souidène, Abdeldjalil Aïssa-El-Bey, Karim Abed-Meraim, and Azeddine Beghdadi, "Blind image separation using sparse representation," in *2007 IEEE International Conference on Image Processing*. IEEE, 2007, vol. 3, pp. III–125.

[12] Michael Zibulevsky and Barak A Pearlmutter, "Blind source separation by sparse decomposition in a signal dictionary," *Neural computation*, vol. 13, no. 4, pp. 863–882, 2001.

[13] Jérôme Bobin, J-L Starck, Jalal Fadili, and Yassir Moudden, "Sparsity and morphological diversity in blind source separation," *Image Processing, IEEE Transactions on*, vol. 16, no. 11, pp. 2662–2674, 2007.

[14] Frederic Abrard and Yannick Deville, "A time–frequency blind signal separation method applicable to underdetermined mixtures of dependent sources," *Signal processing*, vol. 85, no. 7, pp. 1389–1403, 2005.

[15] Fangchen Feng and Matthieu Kowalski, "Revisiting sparse ICA from a synthesis point of view: Blind source separation for over and underdetermined mixtures," *Signal Processing*, vol. 152, pp. 165–177, 2018.

[16] Alexander M Bronstein, Michael M Bronstein, Michael Zibulevsky, and Yehoshua Y Zeevi, "Sparse ICA for blind separation of transmitted and reflected images," *International Journal of Imaging Systems and Technology*, vol. 15, no. 1, pp. 84–91, 2005.

[17] Zois Boukouvalas, Yuri Levin-Schwartz, Vince D Calhoun, and Tülay Adalı, "Sparsity and independence: Balancing two objectives in optimization for source separation with application to fMRI analysis," *Journal of the Franklin Institute*, vol. 355, no. 4, pp. 1873–1887, 2018.

[18] Farnood Merrikh-Bayat, Massoud Babaie-Zadeh, and Christian Jutten, "Using non-negative matrix factorization for removing show-through," in *International Conference on Latent Variable Analysis and Signal Separation*. Springer, 2010, pp. 482–489.

[19] Antoni Buades, Bartomeu Coll, and J-M Morel, "A non-local algorithm for image denoising," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. IEEE, 2005, vol. 2, pp. 60–65.

[20] Wided Souidene, Azeddine Beghdadi, and Karim Abed-Meraim, "Image denoising in the transformed domain using non local neighborhoods," in *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*. IEEE, 2006, vol. 2, pp. II–II.

[21] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Transactions on image processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[22] Quoc Bao Do, Azeddine Beghdadi, and Marie Luong, "Combination of closest space and closest structure to ameliorate non-local means method," in *2011 IEEE Symposium On Computational Intelligence For Multimedia, Signal And Vision Processing*. IEEE, 2011, pp. 134–141.

[23] Zhiqiu Lin, Jin Sun, Abe Davis, and Noah Snavely, "Visual chirality," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 12295–12303.

[24] Paul Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *Journal of optimization theory and applications*, vol. 109, no. 3, pp. 475–494, 2001.

[25] Singanallur V Venkatakrishnan, Charles A Bouman, and Brendt Wohlberg, "Plug-and-play priors for model based reconstruction," in *2013 IEEE Global Conference on Signal and Information Processing*. IEEE, 2013, pp. 945–948.

[26] Stanley H Chan, Xiran Wang, and Omar A Elgendy, "Plug-and-play ADMM for image restoration: Fixed-point convergence and applications," *IEEE Transactions on Computational Imaging*, vol. 3, no. 1, pp. 84–98, 2016.

[27] Shunsuke Ono, "Primal-dual plug-and-play image restoration," *IEEE Signal Processing Letters*, vol. 24, no. 8, pp. 1108–1112, 2017.

[28] Kai Zhang, Wangmeng Zuo, and Lei Zhang, "Deep plug-and-play super-resolution for arbitrary blur kernels," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 1671–1681.

[29] Antonin Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical imaging and vision*, vol. 20, no. 1-2, pp. 89–97, 2004.

[30] Michael Elad and Michal Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image processing*, vol. 15, no. 12, pp. 3736–3745, 2006.

[31] Mahieddine M Ichir and Ali Mohammad-Djafari, "Wavelet domain blind image separation," in *Wavelets: Applications in Signal and Image Processing X*. International Society for Optics and Photonics, 2003, vol. 5207, pp. 361–370.

[32] Gitta Kutyniok and Wang-Q Lim, "Image separation using wavelets and shearlets," in *International Conference on Curves and Surfaces*. Springer, 2010, pp. 416–430.

[33] Xiaochen Zhao, Guangyu Zhou, Wei Dai, and Wenwu Wang, "Blind source separation based on dictionary learning: a singularity-aware approach," in *Blind Source Separation*, pp. 39–59. Springer, 2014.

[34] Julien Mairal, Francis Bach, Jean Ponce, Guillermo Sapiro, and Andrew Zisserman, "Discriminative learned dictionaries for local image analysis," in *2008 IEEE conference on computer vision and pattern recognition*. IEEE, 2008, pp. 1–8.

[35] Meng Yang, Lei Zhang, Xiangchu Feng, and David Zhang, "Fisher discrimination dictionary learning for sparse representation," in *2011 International Conference on Computer Vision*. IEEE, 2011, pp. 543–550.

[36] Róisín Rowley-Brooke, François Pitié, and Anil Kokaram, "A ground truth bleed-through document image database," in *International Conference on Theory and Practice of Digital Libraries*. Springer, 2012, pp. 185–196.

[37] Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.