



INSTITUTO  
SUPERIOR  
TÉCNICO

UNIVERSIDADE TÉCNICA DE LISBOA

INSTITUTO SUPERIOR TÉCNICO

# Blind Separation and Blind Deblurring of Natural Images

Mariana Sá Correia Leite de Almeida

(Licenciada)

**Supervisor:** Doctor Luís Henrique Martins Borges de Almeida

Thesis approved in public session to obtain the PhD Degree in  
Electrical and Computer Engineering

**Jury final classification:** Pass With Merit (unanimous)

## Jury

<b>Chairperson:</b>	Chairman of the IST Scientific Board
<b>Members of the Committee:</b>	Doctor Luís Henrique Martins Borges de Almeida
	Doctor Armando José Formoso de Pinho
	Doctor Mário Alexandre Teles de Figueiredo
	Doctor José Manuel Biucas Dias
	Doctor Pedro Manuel Quintas Aguiar
	Doctor Javier Portilla





INSTITUTO  
SUPERIOR  
TÉCNICO

UNIVERSIDADE TÉCNICA DE LISBOA

INSTITUTO SUPERIOR TÉCNICO

# Blind Separation and Blind Deblurring of Natural Images

**Mariana Sá Correia Leite de Almeida**

(Licenciada)

**Supervisor:** Doctor Luís Henrique Martins Borges de Almeida

**Thesis approved in public session to obtain the PhD Degree in  
Electrical and Computer Engineering**

**Jury final classification:** Pass With Merit (unanimous)

## **Jury**

<b>Chairperson:</b>	Chairman of the IST Scientific Board
<b>Members of the Committee:</b>	Doctor Luís Henrique Martins Borges de Almeida
	Doctor Armando José Formoso de Pinho
	Doctor Mário Alexandre Teles de Figueiredo
	Doctor José Manuel Biucas Dias
	Doctor Pedro Manuel Quintas Aguiar
	Doctor Javier Portilla

<b>Funding Institutions:</b>	Fundação para a Ciência e a Tecnologia
	Instituto de Telecomunicações - polo de Lisboa



## Abstract

This thesis addresses two important nonlinear inverse problems in image processing: the separation of *show-through* and the *bleed-through* mixtures and the blind deblurring of images. New solutions to cope with their high levels of indetermination are proposed.

Two separation methods are developed for the first problem. In a first approach, the indeterminacy of nonlinear Independent Component Analysis (ICA) is reduced through the use of a physical model with only four parameters. Based on other properties, a wavelet-based method is also developed. This non-iterative approach performs space-variant non pixel-wise separation. Both techniques reach separation results competitive with those of other methods.

Regarding blind deblurring, the technique that is developed does not impose strong restrictions on the blurring filter, overcoming the ill-posedness of Blind Image Deconvolution (BID) by initially considering the main image edges and, progressively, handling fainter and smaller ones. The BID technique is extended for deblurring shift-variant degradations in which the blurred image consists of two layers that were subjected to different degradations. The approach is successfully tested on several images, with a variety of synthetic and real-life blurs, both in shift-invariant and two-layer problems. The deblurring results are visually and quantitatively better than those obtained with other state-of-the-art methods.

**Keywords:** Nonlinear image separation, *Show-through*, *Bleed-through*, Independent Component Analysis (ICA), Blind image deconvolution, Image enhancement, Image restoration, Sparse distributions, Space-variant blur, Object segmentation.



## Resumo

Esta tese aborda dois problemas não lineares em processamento de imagem: a separação de misturas resultantes dos fenómenos de *show-through* e *bleed-through* e a focagem cega de imagens. Apresentam-se novas soluções para ultrapassar os seus elevados níveis de indeterminação.

Propõem-se dois métodos para o primeiro problema. Numa primeira abordagem, a indeterminação existente em análise não-linear em componentes independentes é praticamente eliminada através do uso de um modelo físico com apenas quatro parâmetros. Com base noutras características do problema, desenvolve-se um método não-iterativo que realiza separação não pontual e variante no espaço. Ambos os métodos de separação alcançam resultados competitivos com os obtidos por outros métodos.

Desenvolve-se um método de focagem cega de imagens. Sem impor restrições fortes no filtro de desfocagem, o método ultrapassa a indeterminação da focagem cega começando por estimar apenas os contornos principais da imagem, tomando gradualmente em consideração os seus detalhes. A técnica proposta é estendida de forma a também recuperar imagens constituídas por duas camadas que tenham sofrido desfocagens diferentes. O método é testado com sucesso em várias imagens, várias degradações sintéticas e reais, ambas em desfocagens invariantes no espaço e multi-camada. Os resultados obtidos são visual e quantitativamente superiores aos alcançados por outros métodos.

**Palavras-chave:** Separação não-linear de imagens, *Show-through*, *Bleed-through*, Análise em componentes independentes, Desconvolução cega de imagens, Melhoramento de imagem, Restauração de imagem, Distribuições esparsas, Desfocagens variantes no espaço, Segmentação de objectos.



## Acknowledgments

The accomplishment of this work was only possible thanks to the help and support of many people and institutions.

First of all I would like to thank my supervisor, Professor Luís Borges de Almeida, for his guidance, scientific rigor, constructive criticism and encouragement throughout my PhD training. I am also grateful for his support, especially in the most difficult parts of this journey.

This work was performed at the “Instituto de Telecomunicações” (IT) of Lisbon, Portugal, as a PhD student of the Department of Electrical and Computing Engineering (DEEC) of the “Instituto Superior Técnico” (IST), Technical University of Lisbon (UTL). I acknowledge IT for hosting me and providing me all the required human and physical resources. I also acknowledge the additional 6 months scholarship (BID/no5/2010-MarianaAlmeida) and funds to support the attendance of international scientific meetings (under the project POSC/EEA-CPS/61271/2004).

The Ph.D. grant SFRH/BD/23919/2005 from “Fundação para a Ciência e a Tecnologia” (FCT) is also gratefully acknowledged.

The initial part of my Ph.D. studies has benefited from the work performed during my final graduation project. I am grateful to that experience, which was carried out under the Socrates-Erasmus program.

I greatly appreciated the opportunity to take the courses of my Ph.D. program, whose content was definitely related to the work that I have developed. I am also grateful for the opportunity to work during one semester in the Computational NeuroEngineering Laboratory (CNEL) at the University of Florida (UF), where I was very well received by Prof. José C. Principe and his group. Although the work developed there was not included in the final dissertation, I believe that this was an important experience for my scientific training.

I am also grateful to Prof. José M. Bioucas-Dias and to Prof. Pedro M. Q. Aguiar, members of my “Comissão de Acompanhamento de Tese” (CAT), for several useful comments to improve this dissertation.

Finally, I would like to thank my friends and family for their support, especially to my dear parents.



*To my Parents*



# Contents

<b>List of Acronyms</b>	<b>xix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation and problem formulation	1
1.1.1 Separation of show-through and bleed-through image mixtures	2
1.1.2 Automatic image deblurring	3
1.2 Organization of the thesis	5
1.3 Nonlinear image separation	5
1.3.1 Nonlinear physical model trained with ICA	7
1.3.2 Wavelet-based nonlinear separation	10
1.4 Shift-invariant Image Deblurring	13
1.4.1 Proposed blind deblurring method	15
1.4.2 Main results	18
1.4.3 Overview of the publications	19
1.5 Shift-variant Image Deblurring	19
1.5.1 Proposed blind deblurring method	20
1.5.2 Main results	21
1.5.3 Overview of the publications	22
1.6 List of original contributions	22
1.6.1 Nonlinear images separation	22
1.6.2 Image deblurring	23
1.7 List of publications	23
1.7.1 Conference papers	23
1.7.2 Journal papers	24
1.7.3 Patents	24
<b>2 Nonlinear separation of show-through and bleed-through image mixtures</b>	<b>27</b>
2.1 Separating Nonlinear Image Mixtures Using a Physical Model Trained With ICA	29
2.1.1 Introduction	29
2.1.2 Overview of the MISEP method	31
2.1.3 Mixing Model	32
2.1.4 Experiments	35
2.1.5 Conclusions	37
2.2 Wavelet based nonlinear separation of images	41
2.2.1 Introduction	41
2.2.2 Separation Method	43
2.2.3 Experimental results	45
2.2.4 Discussion and future work	49

2.2.5	Acknowledgment . . . . .	50
2.3	Wavelet-based separation of nonlinear show-through and bleed-through image mixtures . . . . .	51
2.3.1	Introduction . . . . .	51
2.3.2	Experimental setup . . . . .	53
2.3.3	Separation method . . . . .	56
2.3.4	Experimental results . . . . .	64
2.3.5	Conclusions . . . . .	71
2.3.6	Acknowledgment . . . . .	71
2.4	Separation of nonlinear show-through image mixtures using a physical model trained with ICA . . . . .	73
2.4.1	Introduction . . . . .	73
2.4.2	The MISEP method . . . . .	75
2.4.3	Physical model of the mixture process . . . . .	78
2.4.4	Experimental setup . . . . .	82
2.4.5	Separation measures . . . . .	84
2.4.6	Experimental results . . . . .	86
2.4.7	Conclusions . . . . .	89
2.4.8	Acknowledgments . . . . .	89
2.4.9	Appendix . . . . .	90
<b>3</b>	<b>Shift-invariant blind image deblurring</b>	<b>99</b>
3.1	Blind deblurring of natural images . . . . .	101
3.1.1	Introduction . . . . .	101
3.1.2	Deblurring method . . . . .	102
3.1.3	Experiments . . . . .	105
3.1.4	Conclusions and future work . . . . .	107
3.1.5	Acknowledgments . . . . .	109
3.2	Blind and semi-blind deblurring of natural images . . . . .	111
3.2.1	Introduction . . . . .	111
3.2.2	Deblurring method . . . . .	114
3.2.3	Quality measure . . . . .	122
3.2.4	Experimental results . . . . .	124
3.2.5	Conclusions . . . . .	143
3.2.6	Acknowledgment . . . . .	144
3.2.7	Appendix . . . . .	144
<b>4</b>	<b>Two-layer blind image deblurring</b>	<b>149</b>
4.1	Blind deblurring of foreground-background images . . . . .	151
4.1.1	Introduction . . . . .	151
4.1.2	Single-layer method . . . . .	153
4.1.3	Two-layer method . . . . .	154
4.1.4	Experiments . . . . .	156
4.1.5	Conclusions and future work . . . . .	157
4.2	Blind deblurring of two-layer images . . . . .	161
4.2.1	Introduction . . . . .	161
4.2.2	Two-layer method . . . . .	162
4.2.3	Experiments . . . . .	163

4.2.4	Conclusions . . . . .	163
<b>5</b>	<b>Conclusions and future work</b>	<b>165</b>
5.1	Conclusions . . . . .	165
5.2	Future perspectives . . . . .	166



# List of Figures

1.1	Inverse Problems. Scheme and deblurring example. . . . .	2
1.2	<i>Show-through</i> and <i>bleed-through</i> effects in an old air-mail letter. Acquired front-page. . . . .	2
1.3	Examples of Show-through and bleed-through mixtures addressed in the thesis. a) Strong show-through mixture obtained using a tracing paper. b) Bleed-through mixture existing in an old partiture. . . . .	3
1.4	Examples of actual blurred photos that are addressed in the thesis. a) Out-of-focus blur. b) Blur due to the horizontal motion of the camera. . . . .	4
1.5	Examples of actual two-layered blurred photos. Both photos are restored, ahead, in the thesis. . . . .	5
1.6	Third pair of real-life tracing paper mixtures: sources, mixtures and some separation results obtained without contrast compensation. . . . .	9
1.7	Box-plots of the 10 values of Signal to Noise Ratio (SNR) (after being averaged across the 5 pairs of images) obtained for the best three separation strategies: MISEP with a symmetric Multi-Layer Perceptron (MLP) with (“MND-Sym”) and without (“Sym”) the Minimal Nonlinear Distortion (MND) principle and our model-based approach (“Model”). . . . .	10
1.8	Wavelet-based separation of <i>bleed-through</i> mixtures: mixtures and separation results. . . . .	12
1.9	BID Results. Left-side: “Barbara” synthetically blurred with a $9 \times 9$ randomly generated filter. Right-side: real photo with motion blur. . . . .	25
1.10	Variant BID results. a) Blurred photo. b) Sharp image estimate. c) Foreground filter estimate. d) Background filter estimate. . . . .	26
1.11	Shift-variant deblurring results. The foreground filter was imposed to be the identity. a) Blurred photo. b) Sharp image estimate. c) Background filter estimate. . . . .	26
2.1	Network structure used in INFOMAX and in MISEP. In INFOMAX, $\mathbf{F}$ is an adaptive linear block and the $\psi_i$ are fixed a priori. In MISEP, $\mathbf{F}$ can be nonlinear and both $\mathbf{F}$ and $\psi_i$ are adaptive. . . . .	31
2.2	Scatter plots corresponding to the “bars” pair. . . . .	36
2.3	Sources and mixtures of the first pair of images. . . . .	38
2.4	Images separated using the proposed model. From top to down: Pair #1, Pair #2, Pair #3, Pair #4 and Pair #5. . . . .	39
2.5	Schematic representation of the wavelet-based separation method. . . . .	44
2.6	Images used in the tests. Left: Source images. Right: Mixture images. From top to down: Pair #1, Pair #2, Pair #3, Pair #4 and Pair #5. . . . .	46
2.7	Results obtained with the stationary wavelet transform with preprocessing. . . . .	47
2.8	Results obtained with the complex wavelet transform with preprocessing. . . . .	48

2.9	Results obtained with the stationary wavelet transform without preprocessing.	48
2.10	Results obtained with the complex wavelet transform without preprocessing.	49
2.11	The first three first pairs of tracing paper source images. In this and all subsequent figures containing images, one of the images of each pair has been horizontally flipped.	56
2.12	The fourth and fifth pairs of tracing paper source images.	57
2.13	First three pairs of tracing paper mixtures.	58
2.14	Fourth and fifth pairs of tracing paper mixtures.	59
2.15	Acquired images of the old air-mail letter.	60
2.16	Acquired images of the old partitures. The squares indicate the areas that were selected for separation.	61
2.17	Behavior of the mask $m_1$ for different values of parameter $a$ . Focusing on the case where $x_1^2 + x_2^2 = 100$ , which is of the order of magnitude found in our images, mask $m_1$ is plotted against $\theta$ , with $x_1 = 10\sin\theta$ and $x_2 = 10\cos\theta$ .	62
2.18	Schematic representation of the wavelet-based separation method.	62
2.19	Fourth pair of tracing paper mixtures separated without using contrast compensation.	63
2.20	Best separation results for the three first pairs of tracing paper mixtures.	65
2.21	Best separation results for the fourth and fifth pairs of tracing paper mixtures.	66
2.22	Effect of different variants of the method on the separation of the bars mixture.	67
2.23	Manuscript letter after separation.	68
2.24	Partitures after separation.	69
2.25	Network structure used in INFOMAX and in MISEP. In INFOMAX, $\mathbf{F}$ is an adaptive linear block and the $\psi_i$ are fixed a priori. In MISEP, $\mathbf{F}$ can be nonlinear and both $\mathbf{F}$ and $\psi_i$ are adaptive.	76
2.26	Geometric scheme of the bilinear transformation of rectangles (contiguous sources).	81
2.27	The first four pairs of source images. In this and all subsequent figures containing image pairs, one of the images of each pair has been horizontally flipped.	91
2.28	Fifth pair of tracing paper source images.	92
2.29	First four pairs of tracing paper mixtures.	93
2.30	Fifth pair of tracing paper mixtures.	94
2.31	First four pairs of separated images.	95
2.32	Fifth pair of separated images.	96
2.33	Variation of MND's regularizing parameter along the training.	96
2.34	Box plots of the values of $Q_1$ obtained for the MLP-based separator with various regularization strategies, and for the proposed model-based separator.	97
2.35	Box plots of the values of $Q_2$ obtained for the MLP-based separator with various regularization strategies, and for the proposed model-based separator.	97
2.36	Box plots of the values of $Q_3$ obtained for the MLP-based separator with various regularization strategies, and for the proposed model-based separator.	98
3.1	Schematic representation of the deblurring method. Block $\mathbf{F}$ extracts edge intensities.	103
3.2	The edge detection filters in the four orientations that were used.	104
3.3	Results of synthetic experiments. a) Original image. b) Estimate for the first $\lambda$ value, for the square blur. Next rows: left, square blur; right, motion blur. Second row: Blurred images. Third row: Deblurred images. Fourth row: Filter estimates.	107

3.4	Results of synthetic experiments. a) Original image. b) and c) Filter estimates. Next rows: Left, without noise; Right, with noise. Second row: Blurred images. Third row: Deblurred images. . . . .	108
3.5	Results with actual blurred photos. a) Sharp photo. b) and c) Filter estimates. Next rows: Left, defocus blur; Right, motion blur. Second row: Blurred images. Third row: Deblurred images. . . . .	109
3.6	Left: “Lena” image blurred with a $9 \times 9$ square blur. Right: Image estimate obtained for the first value of $\lambda$ . . . . .	117
3.7	The set of edge detection filters, in the four orientations that were used. The leftmost filter is the base filter $d_0$ . . . . .	117
3.8	Edges of “Lena” computed using the proposed edge detector. . . . .	118
3.9	Plot of $R( f_i )$ , computed for the values of the parameters used in the experimental tests. . . . .	119
3.10	Filter estimate at an early stage of the method. a) With the safety zone. b) After discarding the safety zone. . . . .	120
3.11	Set of images used for synthetic experiments. a) “Lena”. b) “Cameraman”. c) “Satellite”. d) “Testpat1”. e) “Barbara”. . . . .	128
3.12	Set of blurring filters used in synthetic experiments. #1 - Out-of-focus blur. #2 - Linear motion blur. #3 - Uniform square. #4 - Random square. #5 - Nonlinear motion blur. #6 - Circular motion blur. #7 - Gaussian. . . . .	129
3.13	“Lena” image with blur #5. a) Sharp image. b) Blurred image. Next rows: Deblurred images. Left: Without noise. Right: With noise. Second row: Without filter constraints or regularization. Third row: Using TV regularization on the blurring filter. . . . .	130
3.14	“Barbara” image with blur #2. a) Sharp image. b) Blurred image. Next rows: Deblurred images. Left: Without noise. Right: With noise. Second row: Without filter constraints or regularization. Third row: With constraints. . . . .	131
3.15	“Cameraman” image with blur #1. a) Sharp image. b) Blurred image. Next rows: Deblurred images. Left: Without noise. Right: With noise. Second row: Without filter constraints or regularization. Third row: With constraints. . . . .	132
3.16	Filter estimates, for non-noisy experiments. First row: without restrictions. Second row: with restrictions. Third row: with TV regularization. . . . .	133
3.17	Filter estimates, for noisy experiments. First row: without restrictions. Second row: with restrictions. Third row: with TV regularization. . . . .	133
3.18	Deblurring of a color image. a) Sharp image. b) Image degraded with blur #2 and 30dB of noise. c) Image estimate without noise. d) Image estimate with noise. . . . .	134
3.19	Deblurring of a color image. a) Sharp image. b) Image degraded with blur #5 and 30dB of noise. c) Image estimate without noise. d) Image estimate with noise. . . . .	135
3.20	Multi-frame and single-frame estimates obtained with 30dB BSNR. First row: Degraded images. Second row: multi-frame image estimate. Third row: single-frame image estimates. . . . .	136
3.21	Results with real-life blurred photos. a) Sharp photo of the scene. b) and c) Filter estimates. Next rows: Left - out-of-focus blur. Right - motion blur. Second row: Blurred photos. Third row: Deblurred images. . . . .	137
3.22	Results with an actual blurred photo. a) Sharp photo of the scene. b) Filter estimate. c) Blurred photo. d) Deblurred image. . . . .	138

3.23	Deblurring of an actual photo with several methods. a) Blurred photo. b) Image deblurred with APEX. c) Image deblurred with the YK method. d) Image deblurred with our method. . . . .	141
3.24	Deblurring of an actual photo with several methods. a) Blurred photo. b) Image deblurred with APEX. c) Image deblurred with the YK method. d) Image deblurred with our method. . . . .	142
4.1	The set of edge detection filters, in the four orientations that were used. . . . .	154
4.2	First synthetic degradation. a) Background image. b) Foreground object. c) Background blurring filter. d) Foreground blurring filter. e) Region segmentation. f) Blurred image. . . . .	155
4.3	Results of the first synthetic experiment. a) Estimated background image, b) estimated foreground image, c) segmentation, d) estimated sharp scene, e) estimated background blur, f) estimated foreground blur. . . . .	158
4.4	Second synthetic experiment. a) Blurred image. b) Estimated sharp scene. c) Estimated background blur. d) Estimated foreground blur. . . . .	158
4.5	Third synthetic experiment. a) Blurred image. b) Estimated sharp scene. c) Estimated foreground blur. d) Estimated background blur. . . . .	159
4.6	Manual supervision. a) Synthetic experiments. b) Photo. White: foreground. Black: Background. Gray: Unspecified. . . . .	159
4.7	Deblurring a real photo. a) Blurred photo. b) Estimated sharp scene. c) Estimated background blur. d) Estimated foreground blur. . . . .	160
4.8	Synthetic degradation. a) Background image. b) Foreground object. c) Background blurring filter. d) Foreground blurring filter. e) Region segmentation. f) Blurred image. . . . .	163
4.9	Synthetic experiment. a) Estimated sharp scene. b) Estimated background blur. c) Estimated foreground blur. . . . .	164
4.10	a) Blurred photo. b) Estimated sharp scene. c) Estimated background blur. d) Estimated foreground blur. . . . .	164

# List of Tables

2.1	Parameters obtained after training the model using, as training set, each of the five pairs of mixtures. . . . .	35
2.2	Values of the quality measures for the results obtained with the “own” model and with the “bars” model. The best results are shown in bold. . . . .	37
2.3	Values of the quality measures, averaged across all image pairs, for three separation methods. The best results are shown in bold. For comparison, column MSE shows the quality of what could be considered an “ideal” separation. . . .	38
2.4	Mean rank, for different methods, of the separated images from the tracing paper mixtures. Lower ranks are better. For each pair, the best result is shown in bold. Standard deviations are given in brackets. . . . .	70
2.5	Separation measures averaged, for each image pair, across the two separated images and the 10 randomly initialized experiments. Standard deviations are shown in parentheses. . . . .	87
2.6	Model parameters obtained for the five mixtures, averaged over the 10 random experiments. Standard deviations are given in parentheses. . . . .	87
2.7	Quality measures averaged across the 10 random runs and the five pairs of images. The standard deviation across the 10 random initializations is shown in brackets. For each quality measure, the best method is shown in bold. . . . .	89
2.8	Average values of the separation measures computed for each image pair. Separation measures were averaged across the 10 random runs. The standard deviations are shown in brackets. For each line, the best method is shown in bold. .	90
3.1	Deblurring method . . . . .	103
3.2	Deblurring method . . . . .	116
3.3	Summary of the ISNR values obtained with our method, with no constraints and no regularization on the estimated filter. Each entry gives the average of the ISNRs obtained for the five test images, under the indicated conditions. . .	127
3.4	Comparison of the results obtained without and with constraints on the estimated filter. Each entry gives the average of the ISNRs obtained for the five tested images, under the indicated conditions. The best result for each case is shown in bold. . . . .	134
3.5	Multi-frame performance versus single-frame performance (“Lena” image with motion blurs of 11 pixels). The last column shows the SNR (in dB) of the deblurred images, relative to the original sharp one. The best results for each case are shown in bold. . . . .	135
3.6	ISNR values (in dB) obtained for several methods. For each degradation, the best results are shown in bold. . . . .	146
3.7	ISNR (in dB) computed for experiments performed with the “Lena” image. For each degradation of the grayscale image, the best value is shown in bold. . . . .	147

3.8	ISNR (in dB) computed for experiments performed with the “Cameraman” image. For each degradation, the best value is shown in bold. . . . .	147
3.9	ISNR (in dB) computed for experiments performed with the “Satellite” image. For each degradation, the best value is shown in bold. . . . .	147
3.10	ISNR (in dB) computed for experiments performed with the “Barbara” image. For each degradation of the grayscale image, the best value is shown in bold. . . . .	148
3.11	ISNR (in dB) computed for experiments performed with the “Testpat1” image. For each degradation, the best value is shown in bold. . . . .	148
4.1	Method for deblurring a single image composed by two differently blurred areas.	156

# List of Acronyms

<b>SS</b>	Source Separation
<b>ICA</b>	Independent Component Analysis
<b>MLP</b>	Multi-Layer Perceptron
<b>DSS</b>	Denoising Source Separation
<b>BID</b>	Blind Image Deconvolution
<b>MND</b>	Minimal Nonlinear Distortion
<b>MOS</b>	Mean Opinion Score
<b>SNR</b>	Signal to Noise Ratio
<b>PSF</b>	Point Spread Function
<b>TV</b>	Total Variation
<b>ML</b>	Maximum Likelihood
<b>MAP</b>	Maximum-A-Posteriori
<b>ISNR</b>	Improvement in Signal to Noise Ratio
<b>MOS</b>	Mean Opinion Score
<b>MISEP</b>	Mutual Information SEParation



# Chapter 1

## Introduction

### 1.1 Motivation and problem formulation

A considerable amount of the world's information is organized in the form of images. Despite they are originally a function of continuous variables, images are usually stored in computers in the digital format. Once the possibility of storing information in computers appeared, (digital) image processing has rapidly emerged as an area of computer science. Nowadays, image processing has applications in several important engineering areas, such as document analysis [1–6], photography and video analysis [7–9], astronomical imaging [10], tomography [11–14] and other medical imaging [15–17], remote sensing [18], etc.

In imaging, such as in other areas, the underlying data are not usually directly accessible. On the contrary, the sensors' outputs typically give a corrupted version of the exact desired information. Up to the final acquisition, the data can pass either through simple degradations, such as the additive interference of noise that comes from the sensors, or through more complex signal interactions, such as a convolution. Several image processing application will thus fall into the class of inverse problems [8], i. e., problems in which one aims at recovering the original data by inverting a direct degradation process (see Fig. 1.1). In inverse problems, the goal can either be recovering the original data or characterizing the corrupting process. According to the nature of the degradation, different inverse problems have appeared in imaging: denoising [19], deblurring [20, 21], separation [22], super-resolution [23], in-painting [24], compressed sensing [25], etc.

This dissertation addresses two inverse problems in imaging. In the field of document recovery, we start by studying a nonlinear image separation problem (see Section 1.1.1). With the aim of recovering blurred images, the blind deblurring problem is later addressed (Section

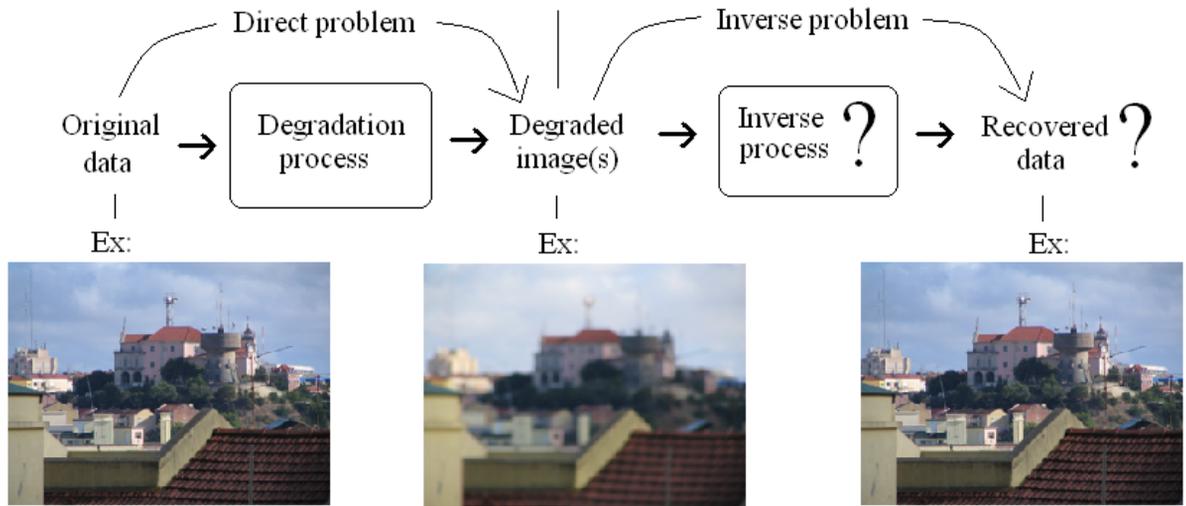


Figure 1.1: Inverse Problems. Scheme and deblurring example.

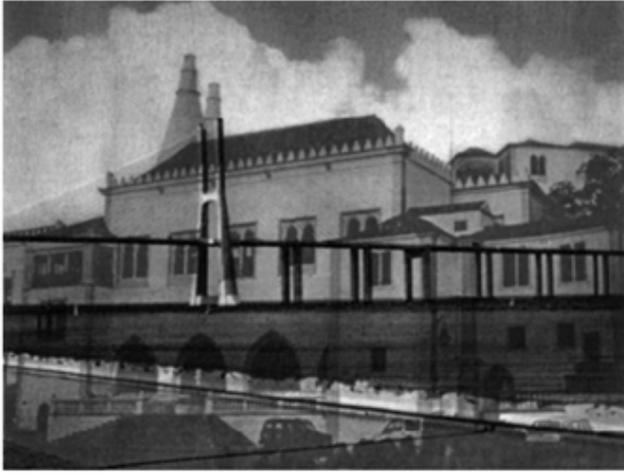


Figure 1.2: *Show-through* and *bleed-through* effects in an old air-mail letter. Acquired front-page.

1.1.2).

### 1.1.1 Separation of show-through and bleed-through image mixtures

When a two-sided document is imaged through a copier or a scanner, it is common that the image of the back page shows through (see an example in Fig. 1.2). This interference is often due to the partial transparency of the paper (the *show-through* effect) and gets stronger as the paper transparency increases. Another possible cause for the mixture is the bleeding of ink through the paper. This phenomenon is called *bleed-through* and it is more common in old documents, in which the ink has had more time to bleed. Both bleed-through and show-through effects may simultaneously be present in the same document, as happens in Fig. 1.2.



a) “Tracing paper mixture”.



b) “Partiture mixture”.

Figure 1.3: Examples of Show-through and bleed-through mixtures addressed in the thesis. a) Strong show-through mixture obtained using a tracing paper. b) Bleed-through mixture existing in an old partiture.

The goal in the show/bleed-through inverse problem is to undo the mixing process, i. e., to recover the original images from the acquired mixed ones. This is a restoration problem of great interest for both academy [1–6, 22] and industry [26, 27]. For example, Xerox has already registered several patents on the subject [26, 27]. Besides application in the area of photocopy, show/bleed-through restorations play a special role in the analysis of archives [5, 28, 29] and old documents, such as old letters [30–32], palimpsests [1], partitures [32] and other historical documents [31], etc.

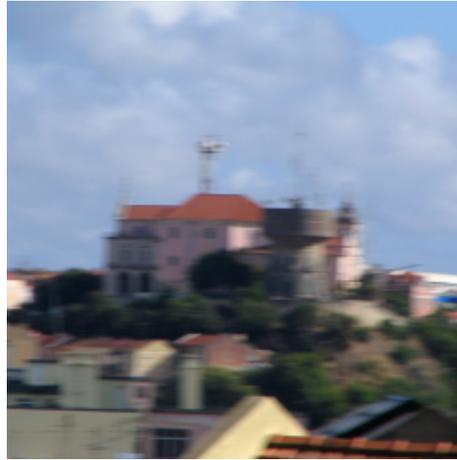
In this thesis, different types of show/bleed-through degraded documents are restored (see examples in Figs. 1.2 and 1.3). Although they have different origins, all documents that were processed suffered nonlinear real-life mixtures, for which a simple linear separation showed to be insufficient.

### 1.1.2 Automatic image deblurring

Acquired images are commonly affected by blur. This kind of degradation happens in various physical processes and is usually reasonably modeled by a mathematical convolution. An example of this type of degradation is the well known case of a blurred photo (see Figs. 1.1 and 1.4). Blurring degradations lead us to the inverse problem of Blind Image Deconvolution (BID) or blind deblurring. In this problem, the goal is to recover the original sharp image form the blurred one. A wide range of physical processes lead to blurring degradations of practical interest. Automatic image deblurring is then an inverse problem of great practical interest in several applications areas, such as photography and video [7–9], astronomy [10], remote sensing



a) Out-of-focus blur.



b) Motion blur.

Figure 1.4: Examples of actual blurred photos that are addressed in the thesis. a) Out-of-focus blur. b) Blur due to the horizontal motion of the camera.

[18], tomography [11–14] and other biomedical imaging [15–17], etc.

Similar to the document restoration problem of Section 1.1.1, BID is a problem of both academic and industrial interest. Several deblurring techniques have been patented by different companies and research centers [33–35].

This dissertation focus on the everyday problem of focusing blurred photographs. Despite that, the methods that are presented in this thesis are also applicable to other deblurring problems. Regarding the BID problem, we have initially studied blurring degradations which are invariant along the image, i. e., shift-invariant degradations. Figure 1.4 shows the two most common types of shift-invariant blurring effects that appear in photography: out-of-focus blur and motion blur. Both photos shown in Fig. 1.4 are restored, ahead, in this thesis.

### Shift-variant image deblurring

Shift-variant image deblurring is an extension of the shift-invariant deblurring problem, in which the characteristics of the blurring degradation change across the observed image. Shift-variant deblurring has applications in several engineering problems [36–39]. Besides that, in areas such as photography, shift-variant models can be more realistic and appropriate than the shift-invariant one [40–42].

Remaining in the photography area, the thesis addresses a practical situation in which the blur is not constant across the image. We study the case in which the photographed scene consists of two layers (a closer object and a farther background) that have been corrupted by different blurring filters. Two typical degradations of this kind occur when: (1) the closer object has motion blur and the background scene is static and in focus (or vice-versa); (2) the

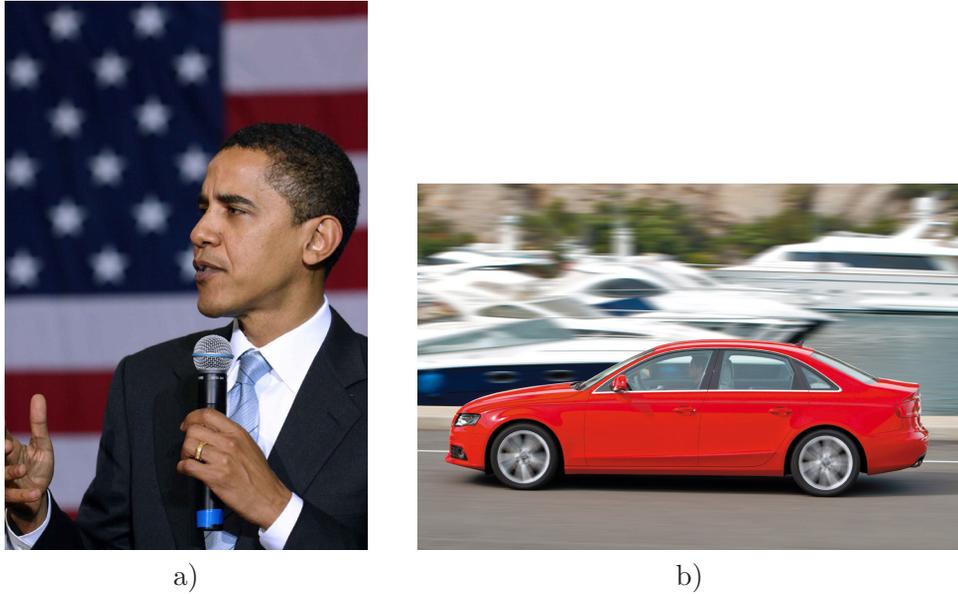


Figure 1.5: Examples of actual two-layered blurred photos. Both photos are restored, ahead, in the thesis.

closer object and the farther background are both stationary, but have different focus blurs. Examples of these types of blurred photos are shown in Fig. 1.5. Both of these photos are restored, ahead, in this dissertation.

## 1.2 Organization of the thesis

The thesis is organized in five chapters. Chapter 1 is an introductory chapter in which the problem under study is described. This chapter also gives a general overview of the techniques that were proposed, the results that were attained and the papers that were published. Chapters 2, 3 and 4 contain the publications regarding the three inverse problems that were addressed: show/bleed-through image separation (Chapter 2), shift-invariant blind image deblurring (Chapter 3) and blind deblurring of images consisting of two layers (Chapter 4). Finally, general conclusions and future perspectives are the focus of Chapter 5.

## 1.3 Nonlinear image separation

Considering that we aim at recovering two source images from two mixed ones, the bleed/show-through inverse problem falls into the class of the Source Separation (SS) problems. Problems of this class are often addressed by assuming that the sources are independent from each other and performing Independent Component Analysis (ICA).

The fact that linear ICA has a unique solution apart from a permutation and a scale

factor [43] has allowed the development of several efficient methods for the separation of linear mixtures [44–47]. However, this development has not occurred for nonlinear mixtures, in which ICA has an infinite number of solutions, almost all consisting in mixtures of the original sources [48–50]. Despite being ill-posed, nonlinear ICA has been addressed by several approaches [51–59]. While the first developed methods [51, 52, 55] are mainly based on the independence criterion, the latter ones [53, 56, 58, 59] try to cope with the indeterminacies inherent to nonlinear ICA. In [53, 54], the temporal structure of the signal is exploited, [56] exploits the smoothness of the signal and [58] the smoothness of the nonlinear map. Despite their punctual advantages, none of these methods guarantees separation. Uniqueness of the solution was only assured in a very recent method [59], which forces the sources to be independent not just in the signal domain, but in the state space of the signals and their derivatives. This method was, however, only applied to one-dimensional signals. Besides that, the technique is computationally demanding.

In spite of its indeterminacies, nonlinear ICA has been applied to a few real-life separation problems. Besides the show-through and bleed-through problems, nonlinear ICA was successfully applied to perform nonlinear denoising in [60]. Nonlinear-ICA has been applied to other real-life separation problems in [58, 61, 62]. However, these publications did not confirm whether the extracted sources were actually close to real ones or not.

During the past years, bleed/show-through mixtures have been addressed by various approaches [1–6, 28–31, 63–66]. A considerable part of these works [5, 28–31] was developed for text documents. The fact that text documents can be seen as binary images makes them easier to separate than general grayscale images. Text images are typically separated resorting to threshold and/or segmentation techniques. Despite they were developed for text, most of the mentioned works address actual old manuscripts that present both bleed-through and show-through effects.

Regarding grayscale (non-text) mixtures, only the show-through effect has been addressed. The methods presented in [1, 63, 65] achieve separation through linear point-wise models, whose parameters are trained with linear ICA techniques. Though it works reasonably well in some cases, linear point-wise separation showed to be inappropriate to separate stronger mixtures, such as the tracing paper mixtures that are addressed in this thesis.

Tracing paper has a high level of transparency, which leads to nonlinear mixtures [64, 67] stronger than the ones that are typically treated in the literature. Previously to this thesis, tracing paper mixtures have already been addressed by other nonlinear approaches [64, 66, 68]. In [68], a multiplicative separation model is trained though the nonlinear ICA Mutual Informa-

tion SEPARation (MISEP) method [55]. MISEP was also applied using a symmetric MLP in the separation map [64]. In a different approach, the Denoising Source Separation (DSS) method of linear separation [69] was extended to the nonlinear scenario of image mixtures. This approach [66], however, does not guarantee the convergence and stability of the separation result.

The first physical model that was proposed for the show-through effect is actually nonlinear and non pixel-wise [2]. In this and some succeeding works [6, 22, 70], this physical model is linearized and adaptive linear filtering schemes are used for removing the show-through effect. Also based on physical principles, the present thesis uses a different pixel-wise nonlinear model [71]. In a latter publication [72], a similar model was also used. However, the latter work reached the show-through model based on empirical observations. Based on non point-wise degradation models, other separation strategies [73] were also proposed for the show-through inverse problem. In contrast with show-through, bleed-through is a shift-variant process, for which we are not aware of any accurate model.

Show-through and/or bleed-through mixtures have also been addressed by other methods. Self-organizing maps (SOM) were used, in [4], in an attempt to compensate for the nonlinearity of the show-through separation map. Another separation technique [74] only requires the front page image to perform the separation. This method was, however, only developed for color or multi-spectral images, and was only tested on relatively weak mixtures.

### 1.3.1 Nonlinear physical model trained with ICA

In order to studying the show-through effect, a set of five pairs of tracing paper mixtures was used (the set is shown in Fig. 2.6, on page 46 of the thesis). The nonlinear character of the mixture under consideration is quite strong and a linear separation model does not lead to good separation results [64, 67]. It is necessary to resort to nonlinear maps and nonlinear ICA, which has the handicap of being an ill-posed problem for which there are an infinite number of solutions, almost all consisting in mixtures of the original sources. The separation problem that we intend to solve is challenging because it involves a noisy nonlinear mixture. Furthermore, some pairs of sources are not fully independent [64, 67].

#### Proposed method

In a first approach, the difficulties of the present problem were overcome through the use of a physical model that fits the mixture process well and that, similarly to linear ICA, has a small number of parameters. The small number of parameters of this model restricts the flexibility of the separating system, eliminating the undesired indeterminacies which characterize unre-

stricted nonlinear ICA. After being trained with the MISEP method of nonlinear ICA [55], the inverse of the mixing model is used to perform the separation. The approach manages to estimate both the original sources and the parameters of the degradation model.

Based on the halftoning process, the present work represents the show-through effect by a *bi-affine* model (the mixtures are affine functions of each of the sources, if the other source is kept constant):

$$\begin{aligned} y_1^i &= \alpha x_1^i + \beta x_2^i + \gamma x_1^i x_2^i + \delta \\ y_2^i &= \alpha x_2^i + \beta x_1^i + \gamma x_1^i x_2^i + \delta \end{aligned} \quad (1.1)$$

in which  $y_1$  and  $y_2$  are the observed images that result from the mixture of the source images  $x_1$  and  $x_2$ ;  $i$  indexes the pixel; and  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  are the four parameters of the *bi-affine* mixture model, which are trained through a nonlinear ICA method [55].

## Main results

The structure of the bilinear physical model allowed us to reach good visual separation results. An example of the quality of the method's results is shown in Fig. 1.6 c). For more separation results, see Figs. 2.31–2.32 on pages 95–96. Based on three quantitative separation measures (see Section 2.4.5 on page 84), the proposed method proved to be better than the techniques [64, 66] that had, until there, been applied to tracing paper mixtures (see Table 2.3 , on page 38).

The model-based approach was more accurately compared with the use of an MLP separator. In this comparison the MISEP method was applied using various regularization strategies: using a symmetric MLP in the separation block [64], either or not addressing the recent MND principle [58]. For each pair of images and for each separation method, 10 experiments were run with different sets of training pixels. Regarding the show-through problem, the model-based approach proved to be, irrespective of the training set, the best separation strategy. The box-plot of Fig. 1.7 shows a representative example of the comparison results. A summary of the comparison can be found in Table 2.7, on page 89.

## Overview of the publications

The model-based approach was first introduced in a conference paper [71]. The same technique was further addressed in an article that was recently submitted to a scientific journal [75].

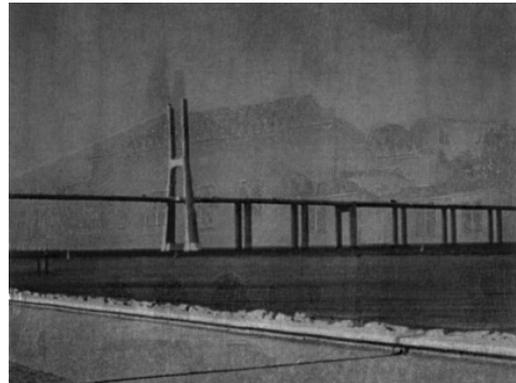
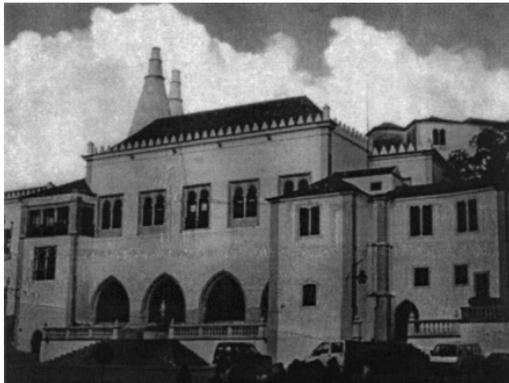
In [71], the physical model that was developed by Miguel Faria and Luís Almeida [76] is derived, the *bi-affine* function is inverted and the model-based method is proposed. As far as we know, this was the first time that a nonlinear physical model was trained with an



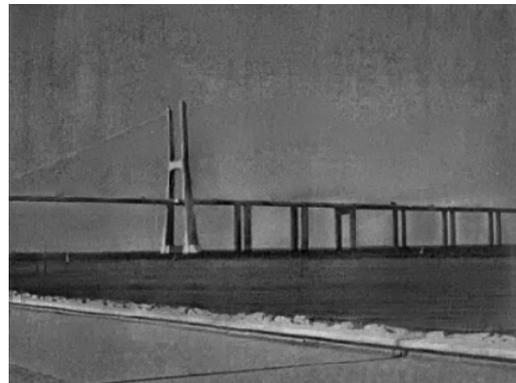
a) Sources.



b) Mixtures.



c) Model-based separation.



d) Wavelet-based separation without contrast compensation.

Figure 1.6: Third pair of real-life tracing paper mixtures: sources, mixtures and some separation results obtained without contrast compensation.

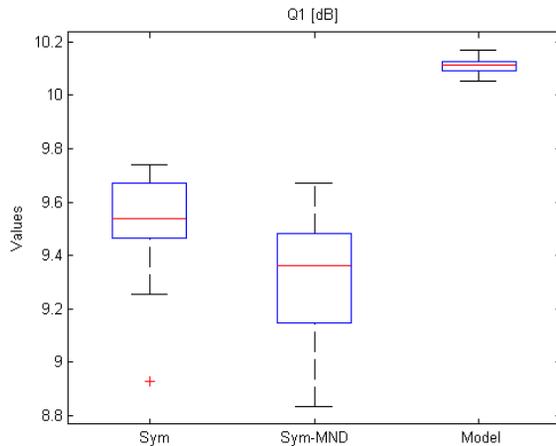


Figure 1.7: Box-plots of the 10 values of SNR (after being averaged across the 5 pairs of images) obtained for the best three separation strategies: MISEP with a symmetric MLP with (“MND-Sym”) and without (“Sym”) the MND principle and our model-based approach (“Model”).

ICA technique. Considering the quality of the separation results, this work confirms that the *bi-affine* model fits, in the low resolution scenario, the show-through process well.

The submitted journal paper [75] addresses the same separation technique. However, a richer set of experiments and comparisons is carried out. Furthermore, this publication also shows that, for non-noisy bounded independent sources, the *bi-affine* mixture (1.1) is separable through the independence criterion (see Section 2.4.3, on page 80).

### 1.3.2 Wavelet-based nonlinear separation

In most of the real-life Source Separation (SS) problems, the sources are not fully independent. In spite of that, linear ICA methods have been successfully used to separate linear mixtures. This is justified with the rigidity and simplicity of the linear mixing model, which has a small number of parameters. However, for more flexible mixtures (such as nonlinear ones), ICA methods become less robust with respect to the independence assumption. The fact that the sources of opposite pages are not usually fully independent can then be a disadvantage for nonlinear separation approaches whose cost function is only determined by the independence criterion. Despite having a reduced number of parameters, the previous model-based approach has already shown to reduce its effectiveness when the images pairs are only approximately independent. Furthermore, that technique does not take into consideration variations of the mixture process along the image. Being specially designed for the show-through degradation, the model-based method adds the drawback of not accounting for the bleed-through phenomenon.

## Proposed method

In contrast to the previous approach, the second separation method that was proposed is not directly based on the independence criterion. Instead, the solution consists of a non-iterative procedure that is based on two simple observations: (1) the high frequency content of images is sparse, and (2) the image printed on each side of the paper appears more strongly in the mixture acquired from that side than in the mixture acquired from the opposite side.

The method starts by decomposing the mixtures into wavelet coefficients. Afterward, the wavelet coefficients are assigned to the corresponding source by a competition process that consists of a *soft-winner-take-all* operation. Finally, after assigning the wavelet coefficients, the estimated sources are reconstructed using the inverse wavelet transform. The method can also be applied with a contrast compensation mechanism, whose aim is to compensate for the reduction of intensity of the edges of one side in the areas where the opposite image is darker. The method does not require the sources to be independent nor the mixture to be space-invariant. That makes the method also suitable for separating mixtures such as those produced by bleed-through, for which we do not have an adequate physical model. This wavelet-based separation method returns separated estimates without directly inferring the structure of the degradation process.

## Main results

The separation results attained with the proposed wavelet-based approach were visually good and competitive with those of other existing methods. An example of the method's separation quality is shown in Fig. 1.6 d). More tracing paper results can be seen in Fig. 2.20 and Fig. 2.21 on pages 65 and 66. The wavelet-based method was also tested on other kinds of old documents which exhibit a considerable amount of bleed-through: an old air-mail letter and old manual transcriptions of music (partitures). An example of a bleed-through separation result is given in Fig. 1.8. More bleed-through results can be found in Figs. 2.23 – 2.24, on pages 68 – 69.

The contrast compensation mechanism that was developed typically improves the separation quality of the method. The advantage of the compensation mechanism is clear if we compare the tracing paper results obtained with the basis method (Fig. 2.7 on page 47) with those obtained using the compensation mechanism (Fig. 2.20 and Fig. 2.21 on pages 65 and 66). In addition to the visual inspection, a comparison based on observers' opinions is summarized in Table 2.4, on page 70. Note that the values of that table were obtained before the contrast compensation mechanism had been implemented.



a) Acquired partitions' pages.



b) Wavelet-based separation.

Figure 1.8: Wavelet-based separation of *bleed-through* mixtures: mixtures and separation results.

## Overview of the publications

The wavelet-based approach was first introduced in a conference paper [77]. The same technique was further extended and published in a journal paper [32].

In the conference paper [77], the wavelet-based separation method is proposed. The method is similar to the denoising step used in nonlinear DSS [66], but incorporates two important improvements. One corresponds to the use of a more suitable wavelet transform, which is shift-invariant or almost shift-invariant. The other improvement is based on the use of a better form of competition. Together, these two improvements led to a one-step procedure that is, by itself, sufficient to separate the images. The proposed method avoids the use of both the MLP and the iterations that were required in nonlinear DSS [66]. Being non-iterative, the method is much more efficient than both nonlinear DSS and the nonlinear ICA-based methods that have previously been proposed [64, 71, 75].

In the journal paper [32], the contrast compensation scheme is added to the basis method. This compensation strategy visually improves the separation quality obtained for the tracing paper mixtures. The journal paper also provides a formal qualitative evaluation, showing that the separation results attained with the wavelet-based approach are visually competitive with those of other existing methods [32, 77].

Besides addressing the tracing paper mixtures, the latter publication treats two new kinds of image mixtures: the mixture existing in an old air-mail letter, and the mixture present in old partitures. These two kinds of documents exhibit a considerable amount of bleed-through which was reduced with the proposed method. As far as we know, this is the only method applicable to grayscale (non-text) images that is simultaneously suitable for both show-through and bleed-through degradations.

## 1.4 Shift-invariant Image Deblurring

In the scope of image deblurring, this Section focuses on shift-invariant degradations, i. e., degradations in which the blurring filter is the same across the whole image. Shift-variant deblurring is a more complex problem that will be addressed ahead, in Section 1.5.

Image deblurring methods can be divided into two classes: *non-blind*, in which we assume the blurring operator to be known, and *blind*, in which the blurring operator is unknown. The method proposed in this thesis (Section 1.4.1) belongs to the latter class. The application range of non-blind methods is much narrower than the one of blind methods: in most situations of practical interest the blurring filter's impulse response, also called Point Spread Function (PSF), is not known with good accuracy. Since non-blind deblurring methods are very sensitive to mismatches between the PSF used by the method and the true blurring PSF, a poor knowledge of the blurring PSF normally leads to poor deblurring results.

Despite its narrower applicability, non-blind deblurring is a challenging inverse problem. The main difficulty faced by non-blind deblurring methods has to do with the presence of noise in the blurred image. Since the blurring operator typically is very ill-conditioned, this noise, even if very weak, can strongly contaminate the deblurred image. The problem is serious in situations in which the blurring PSF is exactly known, and gets worse if there is even a slight mismatch between the PSF used for deblurring and the one that has caused the blur. A considerable amount of the non-blind deblurring methods [78–83] overcome this difficulty through the use of prior information about the image to be recovered, often doing this within a Bayesian or Maximum-A-Posteriori (MAP) framework.

In BID, not only the degradation operator is ill-conditioned, but the problem also is, inherently, severely ill-posed: there is an infinite number of solutions (original image + blurring filter) that are compatible with the degraded image. An overview of BID methods and applications can be found in [84, 85].

Most of the previously published blind deblurring methods are very limited, since they do not allow the use of a generic PSF. Many of them are based, instead, on PSF models with a small number of parameters [86–91]. For example, to model an out-of-focus blur, they normally use a circle with uniform intensity, having as single parameter the circle’s radius [86]. Similarly, to model a motion blur, they normally use straight-line segments with uniform intensity, the only parameters being the segment length and slope [86–88]. A Gaussian blur, which is completely defined by its variance parameter, is normally used for modeling atmospheric turbulence [86, 90, 91]. These approaches are very limited, because such models rarely fit actual blurring PSFs well. For example, the out-of-focus blurring PSF generally is more complex than a simple uniform circle, and the camera motion that causes a motion blur generally is much more complex than an uniform, straight-line motion. And, as was said above, even a slight mismatch between the deblurring PSF and the blurring PSF strongly degrades the quality of the deblurred image.

In an attempt to encompass less restrictive blurs, a fuzzy technique that uses several pre-specified PSF models has been considered in [92]. Another blind deconvolution method, which is fast and has a proof of convergence, is described in [93]. However, this method assumes that the PSF is zero-phase and, furthermore, depends on the existence of a good initial estimate of the PSF. References [94] and [95] present a method called APEX. Although this method covers some blurs which can be found in real-life, it is limited to blurring PSFs modeled by a symmetrical Lévy distribution with just two parameters.

Some methods have been proposed, which do not impose strict restrictions on the blurring filter [7, 20, 96–107]. Instead of using rigid model structures, these methods impose priors over the blurring filter, and do not seem to be able to handle a wide variety of blurs and scenes. Most of these methods use smoothing regularizers on the blurring filter [20, 96–104, 107, 108]. Total Variation (TV)-based regularizes, which are appropriate for piece-wise-constant blurs such as out-of-focus and motion blurs, were used in [20, 97, 102, 104]. Besides considering space-invariant blurs, the method described in [20] was also applied with success in a synthetic image with a space-variant blur. Other works [96, 98, 101, 103, 107] address priors that are appropriate for smoother blurs, such as Gaussian filters. The methods presented in [7, 40, 105, 106] are much less restrictive than parameterized ones and yield good results. However, these approaches

were specially designed for motion blurs. The technique of [109] was developed for motion degradations, but it does not make use of any regularization over the blurring filter.

Prior information about the image and/or the blur is typically used by means of regularizing terms that have regularizing parameters whose values must be set. Parameter estimation is a recurrent and challenging issue in various inverse problems. Some strategies have been developed for automatically estimating the value of the regularizing parameters [83, 110, 111]. However, they were not developed for the BID problem, not being appropriate and robust for addressing generic types of images, blurring degradations and noise levels. Regarding the BID problem, most of the existing methods require the regularizing parameters to be somehow tuned or empirically selected. There are, however, some recent BID approaches [101, 102, 107] that, under the Bayesian framework, manage to incorporate the estimation of the regularizing parameters on their methodologies. Despite this advantage, these techniques were specially designed for and tested on smooth filters, only.

In some BID cases, one has access to more than one degraded image of the same original scene, a fact which can be used to reduce the ill-posedness of the BID problem [112–116]. There are also BID solutions like the ones presented in [117–119], which cannot be considered completely blind, since they require the use of additional data for preliminary training.

#### 1.4.1 Proposed blind deblurring method

Blind Image Deconvolution (BID) or blind deblurring of a single-frame image is an ill-posed problem with an infinite number of possible solutions. In order to reach reasonable deblurring results, previous BID methods typically restrict the searching range of the blurring filter. These restrictions have been applied either in a hard way, through the use of parametric models [86–91], or in a soft way, through the use of priors or regularizing terms [7, 20, 96–107].

Despite of the blurring regularization which have been used in previous BID method, one can visually guess, from a blurred image, how should the shape image look and how should the shape of the blurring filter be. For example, it is visually perceptible that Fig. 1.4 b) suffered an horizontal motion blur, while Fig. 1.4 a) was most probably blurred by an isotropic filter. This happens due to the sharp and sparse edges of natural images that retain, after getting blurred, a trace of the blurring filter. These observations indicate that the shape of the blurring filter may be automatically learned, and that a less restrictive deblurring method should be able to be developed.

## Basic method

Based on the previous observations, we have developed a BID method [9, 120] that only makes very few assumptions on the blurring filter and on the original image: the blurring filter is assumed to have limited support, and the original image is assumed to be a sharp natural image, which typically has a strong piece-wise-smooth component. The method can be interpreted under the regularization and/or the Bayesian framework. Through a MAP approach, we have obtained a cost function of the form of

$$C(x, h) = \|y - h * x\|_2^2 + \lambda R_f(x), \quad (1.2)$$

in which  $y$  and  $x$  are images which represent, respectively, the degraded image and the recovered image;  $h$  is the PSF of the blurring operator and  $*$  denotes the mathematical operation of convolution.  $R_f(x)$  is a regularizing term which favors solutions in which the edges detected by an edge detector  $f(\cdot)$  are sparse, and  $\lambda$  is a regularization parameter. The edge detector that was used was specially designed for the proposed method. The cost function (1.2) has a data connection term (left-hand term of the right-hand side of (1.2)) and a regularizing term (right-hand term of the right-hand side of (1.2)). Cost functions with a structure similar to (1.2) have been widely used in image processing. However, a new regularizing term was developed. Besides that, we have used a new learning strategy which allows to avoid many of the poor local minima in which a simple optimization would fall. The learning technique starts by minimizing (1.2) using a large value of  $\lambda$ , proceeding with  $\lambda$  values which are progressively decreased. Initial filter estimates are far from the desired ones and a stronger regularization is required. Over-regularized images are typically piece-wise constant images with sharp edges that (when compared with the edges of the blurred image) have enough information to start to learn the filter shape and to improve its estimate. As the filter estimate improves, the regularization intensity can be reduced, and fainter image details can be gradually learned. Guided by the piece-wise-smooth component of the images, the proposed optimization strategy manages to reach a reasonable local minimum, being able to overcome, for a large range of problems, the indeterminacies of BID. Although designed to blindly undo several blurring degradations, the method can also be shaped to specific degradations, allowing the inclusion of parametrized blurs or blur-regularizing terms.

The BID method estimates both the original sharp image and the blurring filter. In this basis version, the method is only suitable for single-channel and single-frame images. Nevertheless, the technique is easily extensible to multi-channel images and to the multi-frame case,

as described in the following paragraphs.

### Color images

The main motivation for developing a multi-channel version of our BID method was the interest in addressing color images, whose information is distributed through three different channels. Besides that, a multi-channel version of the method would allow us to generally treat any multi-spectral images. In order to avoid spectral misalignments (chromatic aberrations in color images) a simple and effective approach was implemented: the regularizer,  $R(\cdot)$ , was applied to the sum of the edge images obtained from the three channel images, instead of being applied separately to each channel:

$$R_f(x) = R\left(\sum_c f(x_c)\right), \quad (1.3)$$

in which  $f_c$  is the edge image computed by applying  $f(\cdot)$  to the  $c^{\text{th}}$  color channel of  $x$ . In (1.3),  $\sum_c f(x_c)$  represents the edge image of a colored image  $x$ , in which the edge extractor  $f(\cdot)$  has been separately applied to each of the channels, and the results have been added up together. Assuming that all channels have suffered the same blurring degradation, only a single blurring filter  $h$  needs to be estimated. In this case, the method's cost function is given by

$$C(x, h) = \sum_c \|y_c - h * x_c\|_2^2 + \lambda R_f(x), \quad (1.4)$$

in which  $x_c$  is the  $c^{\text{th}}$  channel of the estimated image  $x$ ,  $y_c$  is the  $c^{\text{th}}$  channel of the degraded image  $y$  and  $R_f(x)$  is computed as in (1.3). Apart from a slightly different cost function (1.4), all the method's details remain unaltered. A suitable pair of image ( $x$ , composed by its channels  $x_c$ ) and filter estimates ( $h$ ) are reached by starting with a strong regularization, which is gradually decreased.

### Multi-frame scenarios

In multi-frame deblurring one has access to several blurred images (frames), each one degraded by its own blur, all of them obtained from the same sharp scene. Methods that operate in this mode can typically take advantage of the extra information which results from the existence of more than one blurred image of the same scene.

The simplest way to extend the method's cost function in order to address multi-frame scenarios is to use a data term which adds up the data terms of all frames. This corresponds

to assuming that all frames have additive Gaussian noise  $n$ , with the same variance. Assuming that all image channels have suffered the same degradation, the method’s multi-channel, multi-frame cost function is given by

$$C(x, h) = \sum_{d,c} \lambda_d \|y_{d,c} - h_d * x_c\|_2^2 + \lambda R_f(x), \quad (1.5)$$

in which  $d$  indexes frames and  $c$  indexes channels (color),  $y_{d,c}$  represents each acquired frame degraded by a different blurring operator,  $h_d$ ;  $\lambda_d$  are parameters which weights each frame model. As in the basic method, suitable image ( $x$ , composed by its channels  $x_c$ ) and filter estimates ( $h_d$ ) are reached by optimizing the cost function with a strong regularization, which is later decreased.

We would like to add that, in the initial iterations of the method, the majority of the model’s data term error does not result from the additive noise  $n$ , but from the mismatch between the blur estimate and the real blur. This should be taken in account in the choice of the parameters  $\lambda_d$ , e.g., a less blurred frame should, at least in the earlier iterations, be more strongly weighted than a more blurred one.

### 1.4.2 Main results

The BID method was successfully tested on a wide range of images (Fig . 3.11 on page 128) degraded with strong blurs (Fig. 3.12 on page 129), with and without additive noise. As an example of the performance and flexibility of our method, Fig. 1.9 shows the results obtained for a real-life photo with motion blur and for the “Barbara” image degraded with a randomly generated blur of size of  $9 \times 9$  pixels. More restoration results can be seen in Figs. 3.13-3.22 on pages 130-138. The method was also tested on multi-channel and multi-frame situations, and with both hard (parametric) and soft blurring regularizations. The results can be seen in Figs. 3.18-3.22 on pages 134-138.

The Improvement in Signal to Noise Ratio (ISNR) measure was adapted to the BID scenario and its value was computed for each experiment (Tables 3.7-3.11 on pages 147-148). Besides giving an indication of the method’s performance, ISNR values were also used for a comparison of methods. The proposed approach was compared with the APEX method [94, 95], whose PSF estimate only has two parameters, and with a method [20] that uses a TV-prior on the blurring filter. Our method was clearly the best one (see Table 3.6 on page 146 and Fig. 3.23-3.24 on pages 141-142), as well as the most robust with respect to the blurring degradations.

### 1.4.3 Overview of the publications

The BID method was first introduced in a conference paper [9]. The same technique was patented [121] and extended in a journal article [120].

In the first publication [9], the basic BID method is proposed. Along with the method, the new optimization scheme and the new image prior are introduced. The deblurring results that are reported show that, in contrast to other existing methods, the proposed technique reaches reasonable results for a wide range of images (“Lena”, “Cameraman” and other natural scenes), blurring filters (simulated motion blurs, simulated out-of-focus blurs, quadrangular filters, and real-life blurring degradations) and noise levels.

The extension of the BID method in order to address multi-channel and multi-frame images is done in the journal article [120]. In this article, the method is tested in a wider range of images (Fig. 3.11 on page 128) and blurs (Fig. 3.12 on page 129), with and without additive noise. Experiments including hard (parametric) and soft blurring regularizations are also reported. To our knowledge, this BID approach is the only one that is able to produce reasonable results for such a wide set of scenarios.

The latter publication also extends the ISNR to the BID scenario (see Section 3.2.3 on page 122). In agreement with our visual inspection, the quality measure shows that the deblurring results are good. These values are also used for comparing the proposed approach with two other BID methods: APEX [94, 95] and the method from [20]. Our approach is shown to be the best, and the most robust with respect to the blurring degradations.

## 1.5 Shift-variant Image Deblurring

In shift-variant BID one aims at inverting a blurring degradation that, contrary to the shift-invariant case (Section 1.4), varies across the image. Considering that the blurring filter is not shift-invariant, the number of parameters under estimation in the shift-variant case is considerably larger than in the shift-invariant one, and the deblurring problem gets more challenging.

In the non-blind setting [36, 122, 123], the challenge of shift-variant deblurring arises not only from the presence of additive noise, but also from the degradation model, which is more complex and difficult to invert than the shift-invariant one. The method proposed in [122] addresses the same two-layer problem of this thesis (Fig. 1.5), in which the blur is approximately constant within each layer. Blurring degradations that slowly vary across the image are typically addressed by partitioning the image into small regions in which the blur is assumed

to be constant [36]. A more recent approach [123] restores shift-varying degradations through the use of a transform designed for that purpose.

In spite of being highly ill-posed, shift-variant BID has been addressed by a few methods [20, 21, 37, 38, 124–131]. In most of the existing approaches [21, 37, 38, 124–127, 131], the problem’s difficulty is reduced by using multiple degraded images of the same scene. Regarding the single-frame scenario, some methods concentrate on blurring degradations that slowly vary across the image [20, 132, 133]. Other approaches [128, 129, 134] address problems that are similar to the two layer degradation that we address here. The methods presented in [128, 129] are either restricted to motion or to out-of-focus blurs, and do not accurately treat the layers’ boundaries. A recent method [134] manages, through the application of previously published parametric methods [132, 135], to be suitable for both motion and out-of-focus degradations. This technique uses, similarly to our shift-variant method, the precise blurring degradation model. As a consequence of that, the boundary regions are accurately handled in [134].

There are also shift-variant methods [130, 132, 135, 136] that aim at estimating the filter’s parameters along the image, without performing deblurring. Some other techniques [137–140] perform shift-variant BID based on camera apparatuses that were specially designed in order to manipulate the image acquisition process.

### 1.5.1 Proposed blind deblurring method

Similarly to shift-invariant deblurring methods, existing shift-variant BID solutions are not flexible enough to adapt to a wide range of blurring degradations. On the contrary, these methods are typically designed for parametric blurring filters [128, 129, 129, 135]. Given the flexibility of the shift-invariant BID method that we have developed, it is interesting to extend the basic method for the two-layer shift-variant problem.

#### The method

In order to extend the shift-invariant BID method of Section 1.4 to the variant case, the two-layer degradation process (Fig. 1.5) was first modeled. Let us consider a mask image  $o$ , with the opacity of the foreground layer: if the foreground only contains opaque objects,  $o$  is a binary mask having ones in the pixels corresponding to the foreground objects and zeros in those corresponding to the background scene. The sharp image  $x_s$  is given by

$$x_s = x_f \cdot \hat{o} + x_b \cdot (1 - \hat{o}) + n, \quad (1.6)$$

in which the operator  $\cdot$  denotes the Hadamard product,  $x_b$  is the background image,  $x_f$  is the foreground image, and  $n$  represents the additive noise. After being blurred with the foreground filter  $h_f$ , the blurred mask  $\hat{o} = h_f * o$ , will measure the contribution of each of the blurred layers to the degraded image  $y$

$$y = (h_f * x_f) \cdot \hat{o} + (h_b * x_b) \cdot (1 - \hat{o}) + n, \quad (1.7)$$

in which  $h_b$  is the PSF of the blurring operator that has degraded the background image  $x_b$ .

The two-layer deblurring method is similar to the single-layer one described in Section 1.4. However, the method has more unknowns ( $x_f$ ,  $x_b$ ,  $h_f$ ,  $h_b$  and  $o$ ), which are estimated through the minimization of an extended cost function

$$C = \|y - \hat{y}\|_2^2 + \lambda_x [R_f(x_r) + R_f(x_f) + R_f(x_b)] + \lambda_o R_f(o), \quad (1.8)$$

in which  $\hat{y}$  is the estimate of the blurred image, constructed from the image and the filter estimates according to the model (1.7), and  $x_r$  is the estimated sharp image computed as in (1.6).  $R_f$  is the regularizing function developed for the single-layer method.  $\lambda_x$  and  $\lambda_o$  are regularizing parameters that control the regularization applied to the estimated images (complete image, background and foreground) and to the opacity mask image, respectively.

Like the single-layer method [120], this two-layer method only restricts the blurring filters to have a limited support. A suitable deblurring solution is reached by starting with a large value for the regularizing parameters ( $\lambda_x$  and  $\lambda_o$ ) and progressively reducing them.

The method needs to be initialized with a small amount of prior information, which is easily provided by a human. Two compact image regions must be indicated: one that surely corresponds to a background area, and another one corresponding to a foreground area. Once these regions have been assigned, a segmentation algorithm [141] is applied to find an initial estimate for the  $o$  mask.

### 1.5.2 Main results

The two-layer shift-variant BID method was tested on several synthetic degradations of natural images as well as on actual blurred photographs, with satisfactory results. Figures 1.10 and 1.11 show the method's performance on two blurred photos which were downloaded from the Internet. More deblurring results can be seen in Figs. 4.4-4.7 and in Figs. 4.9-4.10, on pages 158-160 and 164-164.

### 1.5.3 Overview of the publications

The shift-variant BID method was introduced in a national [142] and in an international conference [143]. The method is planned to be published in a scientific journal, after some improvements.

Being an extension of our shift-invariant BID method, the proposed shift-variant technique can deal with rather unrestricted blurs in both the foreground and the background layers. The method is satisfactorily tested on several synthetic degradations of natural images and on actual blurred photographs.

The approach makes use of an accurate degradation model (1.7), which leads to a more complex cost function around the layers' boundaries. The accuracy of the degradation model is specially important in the single-frame scenario, in which the available information is extremely limited. Apart from a very recent work [134], we do not know any other single-frame method that models the layers's boundaries with the same accuracy.

## 1.6 List of original contributions

### 1.6.1 Nonlinear images separation

#### **Model-based approach**

Implementation of a new separation technique in which a nonlinear physical model is trained with an ICA technique. Proof that the bilinear model is separable with the ICA criterion.

Experimental validation of the method's performance. Comparison showing that the model-based method clearly outperforms other show-through separation methods.

#### **Wavelet-based approach**

Development of a new separation method which is non-iterative and not directly based on the ICA criterion. The method performs a space-variant separation and is suitable for non-independent source images. Development of a contrast compensation mechanism.

Experimental validation of the method's performance in grayscale images with both show-through and bleed-through degradations. Formal subjective evaluation that assessed the relative quality of the results: our results were visually competitive with those obtained with two other state-of-the-art methods. Separation of new kinds of real-life mixtures that exhibit the bleed-through effect: two old partitures and an old air-mail letter.

## 1.6.2 Image deblurring

### Shift-invariant deblurring

Development of a new blind deblurring method which does not impose strong restrictions on the blurring filter, being suitable for inverting a wide range of blurring degradations. Development of a new image prior and an optimization strategy for the BID method. Extension of the basic BID method in order to address multichannel images, such as color images, and multi-frame scenarios.

Experimental validation of the method's efficiency with a wide range of synthetic blurring degradation and with some actual blurred photos. Comparison of the BID method with two other deblurring approaches, clearly showing that our solution is the most accurate one and the most flexible with respect to the blurring degradation.

### Shift-variant deblurring

Extension of the single-layer blind deblurring method in order to address images consisting of two differently blurred layers. The method uses an accurate degradation model on the layers' boundaries.

Experimental validation of the method's performance in several synthetic degradations and in some actual blurred photos.

## 1.7 List of publications

### 1.7.1 Conference papers

Mariana S. C. Almeida and Luís B. Almeida, "Separating nonlinear image mixtures using a physical model trained with ICA," in *IEEE Int. Worksh. on Machine Learning for Signal Processing - MLSP*, Maynooth, Ireland, 2006.

Mariana S.C. Almeida and Luís B. Almeida, "Wavelet based nonlinear separation of images," in *IEEE Int. Worksh. on Machine Learning for Signal Processing - MLSP*, Maynooth, Ireland, 2006.

Mariana S. C. Almeida and Luís B. Almeida, "Blind deblurring of natural images," in *IEEE Int. Conf. Acoustics, Speech, and Signal Processing - ICASSP*, Las Vegas, 2008, pp. 1261-1264.

Mariana S. C. Almeida and Luís B. Almeida, "Blind deblurring of foreground-background

images,” in *IEEE International Conference on Image Processing - ICIP*, Cairo, Egypt, 2009, pp. 1301-1304.

### **National conferences**

Mariana S. C. Almeida and Luís B. Almeida, “Blind deblurring of two-layer images,” in *Portuguese Conf. on Pattern Recognition - RecPad*, Aveiro, Portugal, October 2009.

### **1.7.2 Journal papers**

Mariana S.C. Almeida and Luís B. Almeida, “Wavelet-based separation of nonlinear show-through and bleed-through image mixtures,” *Neurocomputing*, vol. 72, pp. 57-70, December 2008.

Mariana S. C. Almeida and Luís B. Almeida, “Blind and semi-blind deblurring of natural images,” *IEEE Trans. on Image Processing*, vol. 19, no. 1, pp. 36-52, January 2010. (in the top 10 accessed articles of March 2010)

Mariana S.C. Almeida and Luís B. Almeida, “Separation of nonlinear show-through image mixtures using a physical model trained with ICA”, *Journal of Machine Learning Research - JMLR*, 2010. (submitted)

### **1.7.3 Patents**

Mariana S. C. Almeida and Luís B. Almeida. “Processo de focagem cega de imagens,” *Portuguese patent*, 2009.



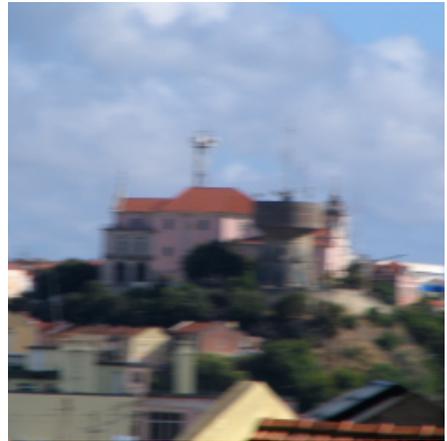
a) "Barbara" image.



b) Sharp photo.



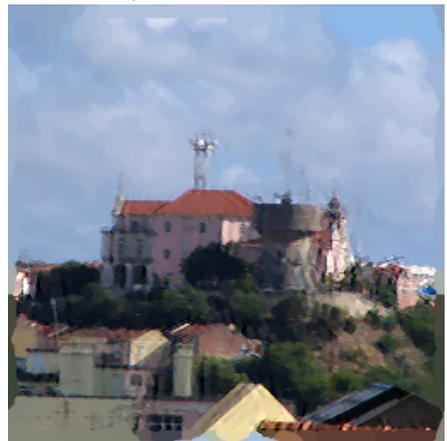
c) Blurred "Barbara".



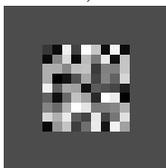
d) Blurred photo.



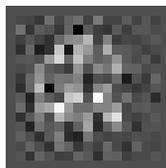
e) "Barbara" estimate.



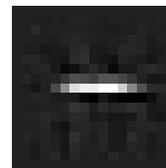
f) Scene estimate.



g) Blurring filter.



h) Filter estimate.



i) Filter estimate.

Figure 1.9: BID Results. Left-side: "Barbara" synthetically blurred with a  $9 \times 9$  randomly generated filter. Right-side: real photo with motion blur.

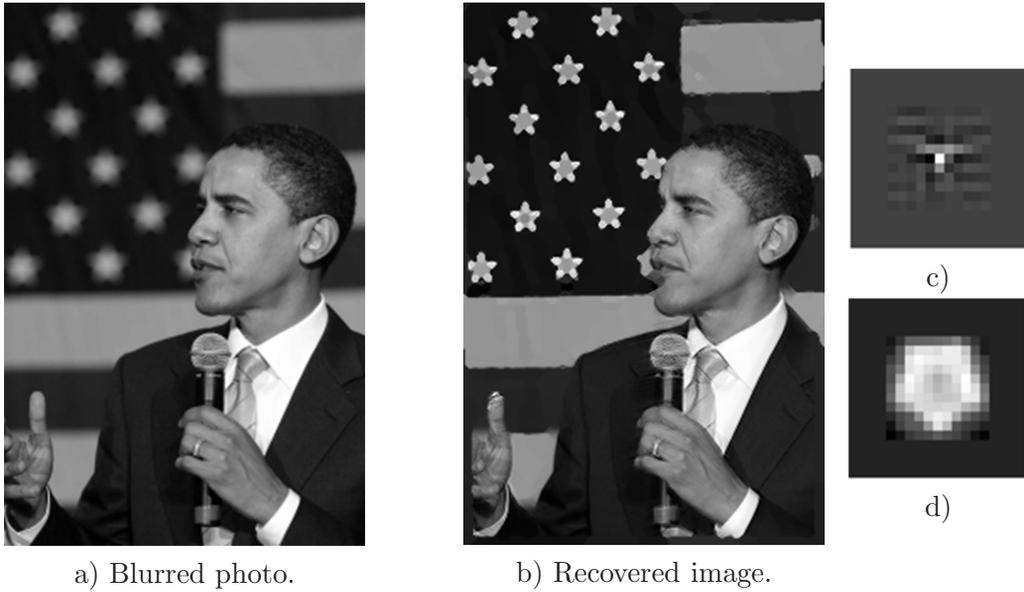


Figure 1.10: Variant BID results. a) Blurred photo. b) Sharp image estimate. c) Foreground filter estimate. d) Background filter estimate.

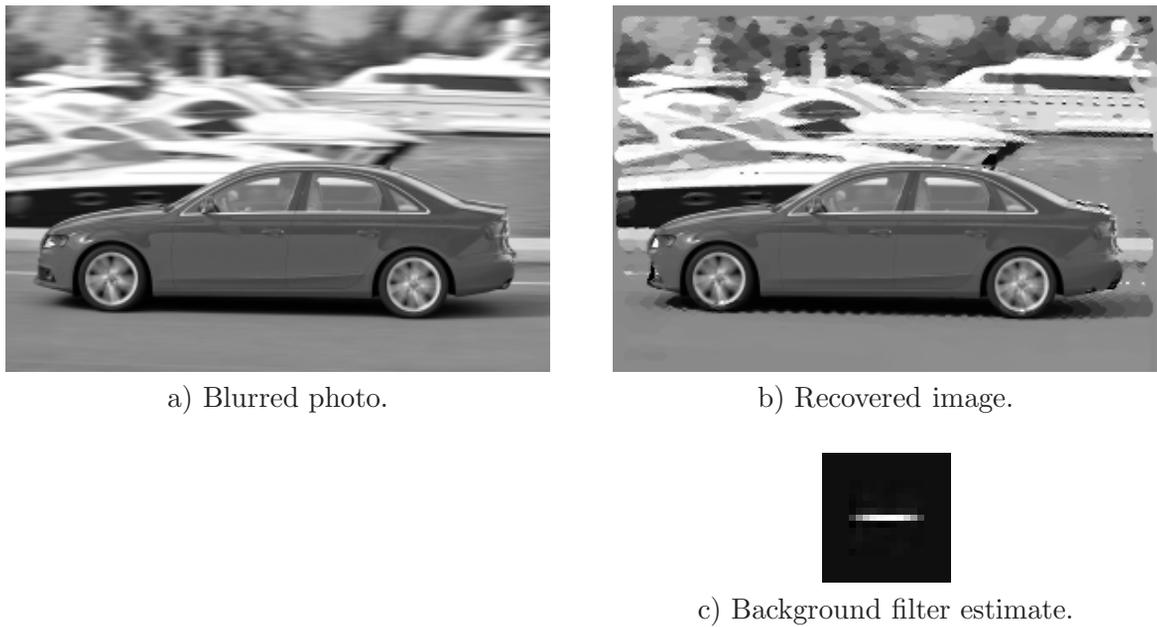


Figure 1.11: Shift-variant deblurring results. The foreground filter was imposed to be the identity. a) Blurred photo. b) Sharp image estimate. c) Background filter estimate.

## Chapter 2

# Nonlinear separation of show-through and bleed-through image mixtures

### Conference papers

Mariana S. C. Almeida and Luís B. Almeida, “Separating nonlinear image mixtures using a physical model trained with ICA,” in *IEEE Int. Worksh. on Machine Learning for Signal Processing - MLSP*, Maynooth, Ireland, 2006.

Mariana S.C. Almeida and Luís B. Almeida, “Wavelet based nonlinear separation of images,” in *IEEE Int. Worksh. on Machine Learning for Signal Processing - MLSP*, Maynooth, Ireland, 2006.

### Journal papers

Mariana S.C. 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, December 2008.

Mariana S.C. Almeida and Luís B. Almeida, “Separation of nonlinear show-through image mixtures using a physical model trained with ICA”, *Journal of Machine Learning Research - JMLR*, 2010. (submitted)



## 2.1 Separating Nonlinear Image Mixtures Using a Physical Model Trained With ICA

### Abstract

This work addresses the separation of real-life nonlinear mixtures of images, which occur when a paper document is scanned and the image from the back page shows through. A physical model of the mixing process, based on the consideration of the halftoning process used to print grayscale images, is presented. The corresponding inverse model is then used to perform image separation. The parameters of the inverse model are optimized through the MISEP technique of nonlinear ICA, which uses an independence criterion based on minimal mutual information.

The quality of the separated images is competitive with the one achieved by other techniques, namely by MISEP with a generic MLP-based separation network and by Denoising Source Separation. The separation results show that MISEP is an appropriate technique for training the parameters and that the model fits the mixing process well, although not perfectly. Prospects for improvement of the model are presented.

### 2.1.1 Introduction

When scanning or photographing a paper document, interference of the back page image on the front page one is a common problem, especially if the paper is thin or rather transparent. In this paper we focus on a difficult version of this problem, in which the paper is of the onion skin type, which creates a strong, significantly nonlinear mixture. The mixtures that we use were obtained by printing images and/or text on both sides of a sheet of onion skin, which was then scanned, on both sides, with a desktop scanner. The scanned images of each pair were then aligned with each other. The source images were also aligned with the mixture ones, for quality assessment. A more complete description of the image preparation procedure is given in [64].

These images have already been used to test other separation methods, mentioned ahead, and are available at <http://www.lx.it.pt/~lbalmeida/ica/seethrough/index.html>. Due to lack of space, we only show the first pair of source and mixture images (Fig. 2.3). The other four pairs can be found in [64] and in the mentioned web location.

Reconstructing two sources from two mixtures can be seen as a blind source separation (BSS) problem. BSS is often achieved by assuming that the sources are statistically independent

from each other and performing independent component analysis (ICA). Linear ICA is a well studied problem with essential uniqueness of the solution [43]. Nonlinear ICA is still much less studied. With no additional constraints it is an ill-posed problem, having an infinite number of solutions that are not related to one another in any simple way [48] [49].

The problem under study is especially challenging because it involves a real-life nonlinear, noisy mixture and, furthermore, some pairs of source images do not satisfy the independence assumption. Due to the small number of parameters under estimation and to the simplicity of the mapping, linear ICA often recovers the sources satisfactorily from linear mixtures, even if they are not completely independent. However, in nonlinear mixtures the quality of the separation can easily be impaired when the independence assumption is not met [64].

In this paper we first present a physical model of the mixture process. Then, the inverse of that model is used to perform separation. The parameters of the inverse model are estimated through an ICA criterion, using the MISEP method [55]. The results show that a separation with a good quality is achieved. The small number of degrees of freedom of the model eliminates the ill-posedness that is normally associated with less constrained nonlinear ICA.

To our knowledge this is the first time that a nonlinear, physically based model is trained, with an ICA criterion, to perform source separation. The same image separation problem, with the same dataset, has previously been addressed by two other methods. One was based on the use of the MISEP method of nonlinear ICA with a multilayer perceptron as separating system and with regularization constraints to deal with the ill-posedness of nonlinear ICA [64]. The other approach was the use of the nonlinear denoising source separation (DSS) method [66], which is not based on an independence criterion, but instead uses some basic prior knowledge about images to perform separation. The separation results presented in this paper are competitive with those obtained with those methods, as evidenced by objective quality measures that we include ahead.

The same separation problem is also addressed in another paper in this conference [77] That paper presents a non-iterative separation method, based on the sparsity of the coefficients of the wavelet decomposition of images.

This manuscript is organized as follows: Section II briefly describes the MISEP method. Section III describes the mixing model and its inverse. Section IV presents experimental results, including objective quality measures. Section V concludes and presents future research directions.

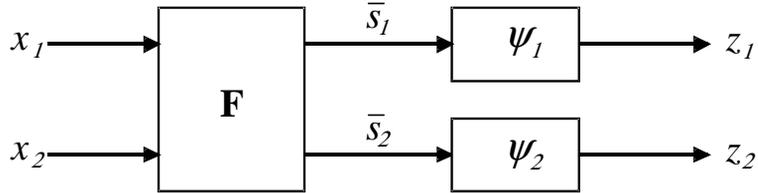


Figure 2.1: Network structure used in INFOMAX and in MISEP. In INFOMAX,  $\mathbf{F}$  is an adaptive linear block and the  $\psi_i$  are fixed a priori. In MISEP,  $\mathbf{F}$  can be nonlinear and both  $\mathbf{F}$  and  $\psi_i$  are adaptive.

### 2.1.2 Overview of the MISEP method

MISEP is a generalization of the well known INFOMAX technique of linear ICA [44]. INFOMAX maximizes the entropy of the output of the network depicted in Fig. 2.1. Block  $\mathbf{F}$  performs the linear separation. The separated components are  $s_i$ . Blocks  $\psi_i$  are auxiliary, being used only during the training phase. Each of these blocks implements an invertible, increasing transformation  $z_i = \psi_i(s_i)$ , whose co-domain is the interval  $[0, 1]$ . Ideally each of these blocks should implement the cumulative distribution function (cdf) of the corresponding input  $s_i$  (**estimated sources**). In that case, maximizing the output entropy corresponds to minimizing the mutual information (MI) between the extracted components  $s_i$ . Thus, INFOMAX performs linear ICA by indirectly minimizing the mutual information of the sources. Due to the small number of parameters under estimation, linear separation can often be achieved by INFOMAX even if the  $\psi_i$  blocks implement only crude approximations of the cdfs of the sources. In nonlinear ICA, however, the correct estimation of the cdfs plays a more crucial rule.

MISEP extends INFOMAX in two directions. First, MISEP handles nonlinear mixtures, by allowing block  $\mathbf{F}$  to be nonlinear. Second, MISEP uses output nonlinearities that adapt to the statistical distributions of the extracted components. In MISEP, the maximization of the entropy of the output of the network of Fig. 2.1 simultaneously optimizes the  $\psi_i$  functions and the separation mapping.

MISEP can use any parameterized, linear or nonlinear block in  $\mathbf{F}$ . In previous tests with the present dataset [64], this block was implemented by means of a multilayer perceptron with suitable regularization. In this paper, block  $\mathbf{F}$  will consist of the inverse of the mixture model, to be presented in the next section. For more details on MISEP see [55, 64].

### 2.1.3 Mixing Model

The physical mixture model that we use was originally developed by Miguel Faria and Luís B. Almeida [68], but its parameters had only been manually adjusted, having never been estimated in a form similar to the one presented in this paper. We present the model in some detail here because it had not been previously described in any widely available publication.<sup>1</sup>

The model takes into account that the printer produces only black dots, using a halftoning process to produce gray tones. Halftoning consists of using a very large number of tiny black dots, whose intensities are averaged out by our eyes, giving the appearance of gray. The level of gray depends on the fraction of area covered by black dots. With the low scanning resolution that was used in the dataset (100 dpi), each scanned pixel encompasses a large number of halftoning dots, and therefore the pixel’s intensity also depends on the fraction of area covered by the dots.

We represent the actual printed intensity at a given point in the page by  $\hat{s}$ . Since the printer only produces black and white,  $\hat{s} \in \{0, 1\}$ , with 0 representing black and 1 representing white. The halftoning process is modeled by considering  $\hat{s}$  to be a random variable which takes independent values in different locations of the image, with a distribution defined by the probability  $P(\hat{s} = 1)$ . This probability is equal, at each point, to the intensity of the image being printed (the source image). The mean intensity at a given point is given by the expected value

$$s = E(\hat{s}) = P(\hat{s} = 1) \tag{2.1}$$

at that point. We denote it by  $s$  since it corresponds to the intensity of the source image. Labeling the two sides of the paper with subscripts 1 and 2 respectively, we have the following relationships for the two sources:

$$\begin{aligned} s_1 &= P(\hat{s}_1 = 1) \\ s_2 &= P(\hat{s}_2 = 1). \end{aligned} \tag{2.2}$$

With a semi-transparent paper like onion skin, the observed intensity on each side of the paper depends on what is printed on both sides. Assume that we are observing the document

---

<sup>1</sup>An equivalent model was developed by Stefan Harmeling without any physical considerations, based only on the observation of the source and mixture data from the “bars” images (S. Harmeling, private communication).

from side number 1. The observed intensity at each point can take only four levels:

$$\hat{x}_i = \begin{cases} l_1 & \text{if } \hat{s}_1 = 0 \text{ and } \hat{s}_2 = 0 \\ l_2 & \text{if } \hat{s}_1 = 0 \text{ and } \hat{s}_2 = 1 \\ l_3 & \text{if } \hat{s}_1 = 1 \text{ and } \hat{s}_2 = 0 \\ l_4 & \text{if } \hat{s}_1 = 1 \text{ and } \hat{s}_2 = 1 \end{cases} \quad (2.3)$$

The values of  $l_1, \dots, l_4$  depend on the physical properties of the paper and of the scanner, and also on the printing process. Due to physical constraints, we know that  $l_1 \leq l_2 \leq l_3 \leq l_4$ , with strict inequality holding if the paper is not completely opaque and not completely transparent.

The mean intensity at each point, observing from side 1 of the paper, is given by the expected value

$$\begin{aligned} x_1 &= E(\hat{x}_1) \\ &= l_1 P(\hat{s}_1 = 0, \hat{s}_2 = 0) + l_2 P(\hat{s}_1 = 0, \hat{s}_2 = 1) + \\ &\quad l_3 P(\hat{s}_1 = 1, \hat{s}_2 = 0) + l_4 P(\hat{s}_1 = 1, \hat{s}_2 = 1). \end{aligned}$$

This is what corresponds, in our model, to the intensity acquired by the scanner.

We shall assume that  $\hat{s}_1$  and  $\hat{s}_2$  are independent from each other. Taking (2.2) into account,

$$x_1 = l_1(1 - s_1)(1 - s_2) + l_2(1 - s_1)s_2 + l_3s_1(1 - s_2) + l_4s_1s_2. \quad (2.4)$$

Assuming that the printing and acquisition systems are symmetrical, i.e., that they treat both sides of the paper in the same way, we have

$$x_2 = l_1(1 - s_1)(1 - s_2) + l_2s_1(1 - s_2) + l_3(1 - s_1)s_2 + l_4s_1s_2. \quad (2.5)$$

In order to simplify these equations we can define the following parameters:

$$\begin{aligned} \alpha &= l_3 - l_1 \\ \beta &= l_2 - l_1 \\ \gamma &= l_4 + l_1 - l_2 - l_3 \\ \delta &= l_1 \end{aligned} \quad (2.6)$$

Substituting into (2.4) and (2.5) we get the following equations

$$x_1 = \alpha s_1 + \beta s_2 + \gamma s_1 s_2 + \delta \quad (2.7)$$

$$x_2 = \alpha s_2 + \beta s_1 + \gamma s_1 s_2 + \delta, \quad (2.8)$$

This shows that the mixture is *bi-affine* (it is affine as a function of each of the sources, if the other source is kept constant). The parameters of the mixture  $(\alpha, \beta, \gamma, \delta)$  have a direct correspondence with the intensity levels  $l_i$  of the four possible combinations that result from the halftoning process. If the paper is not perfectly transparent,  $l_3 \neq l_2$  and consequently  $\alpha \neq \beta$ , which is required for the model to be invertible.

To recover the sources from the mixtures, we must now invert the model (2.7, 2.8). Subtracting (2.7) from (2.8) we see that  $s_1$  and  $s_2$  are related by

$$s_2 = s_1 + (x_2 - x_1)/(\alpha - \beta). \quad (2.9)$$

Substituting now (2.9) into (2.7) we get a quadratic equation,

$$\gamma s_1^2 + \left[ \alpha + \beta + \frac{\gamma(x_2 - x_1)}{\alpha - \beta} \right] s_1 - x_2 + \delta + \frac{\alpha(x_2 - x_1)}{\alpha - \beta} = 0 \quad (2.10)$$

which can be explicitly solved. We first define

$$\begin{aligned} a &= \gamma \\ b &= \alpha + \beta + \frac{\gamma(x_2 - x_1)}{\alpha - \beta} \\ c &= -x_2 + \delta + \frac{\alpha(x_2 - x_1)}{\alpha - \beta}. \end{aligned} \quad (2.11)$$

Source  $s_1$  is then given by

$$s_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}. \quad (2.12)$$

One can check that using a minus sign before the square root, in the latter expression, would not yield a solution of (2.7, 2.8). Only the plus sign yields a valid solution.

It would be possible to find an equation similar to (2.12) for  $s_1$ . However, once we have computed  $s_1$ , we can use (2.9) to calculate  $s_2$ . This simplifies the calculation of both  $s_2$  and its derivatives, which are needed in MISEP. The use of the intermediate variables  $a$ ,  $b$  and  $c$  not only helps to invert the system but also greatly simplifies the computation of the derivatives that are required in MISEP, allowing a significant increase in optimization speed.

The model that we have described doesn't take into account any lateral diffusion of light

Image pair	$\alpha$	$\beta$	$\gamma$	$\delta$
1	0.426	0.030	0.020	14.242
2	0.412	0.083	0.011	13.349
3	0.433	0.137	0.010	21.938
4	1.040	0.603	0.023	50.69
5	1.905	1.784	-0.015	0.946

Table 2.1: Parameters obtained after training the model using, as training set, each of the five pairs of mixtures.

in the onion skin paper. At the low scanning resolution that was used this seems to be a reasonable approximation, as evidenced by the results presented ahead.

### 2.1.4 Experiments

The inverse model (2.12, 2.9) was used as the separation block  $\mathbf{F}$  of Fig. 2.1. The MLPs used in the  $\psi_i$  blocks had 10 hidden units each. All of these blocks were optimized using the MISEP method. The parameters of block  $\mathbf{F}$  were initialized close to the identity function ( $\alpha = 1$ ,  $\beta = 0.01$ ,  $\gamma = 0.001$ ,  $\delta = 0.001$ ).

For each pair of mixture images, 1000 pairs of pixels were randomly chosen as training set. One separation model was trained, during 1000 epochs, for each pair of images, leading to the parameter values shown in Table 2.1.<sup>2</sup> It is interesting to note that the estimated parameters differ somewhat among the various pairs of images, although the printing and acquisition was performed in as similar a manner as possible for all images, leading one to expect that the same model would fit all mixtures.

Each mixture was separated using the model trained for that mixture (which we call the mixture’s “own model”). The results are shown in the left half of Fig. 2.4. The model trained for the “bars” images can, in some sense, be considered to be the most basic and most “universal” one, because in that case the source images are independent from each other by construction, and have almost uniform intensity distributions. For that reason we also tried using that model (which we call the “bars” model) to separate the other four mixtures. The results are shown in the right half of Fig. 2.4. The results obtained with the two models are similar in all cases except for the last image pair, which corresponds to a mixture of images containing mostly text. This is also the pair for which the estimated parameters differ most from those of the “bars” images (see Table 2.1).

The scatter plots of the sources, mixture components and separated components, for the

---

<sup>2</sup>In the tests, the range of image intensity values that was used was  $[0, 255]$ , instead of the range  $[0, 1]$  used in the derivations of Section 2.1.3. The parameter values shown in the table correspond to the range  $[0, 255]$ .

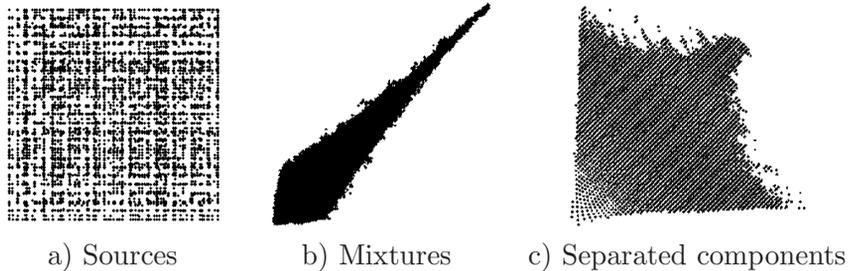


Figure 2.2: Scatter plots corresponding to the “bars” pair.

“bars” pair, are shown in Fig. 2.2. We can see that the model achieved a good, but not perfect separation. The fact that the scatter plot of the separated components shows curved boundaries is probably due to some imperfection of the model. We discuss this further in the Conclusions.

### Quality measures

To analyze the quality of the separated images in a more objective way we computed three quality measures that had already been used for the same mixing problem in [64], [66]. The first quality measure,  $Q_1$ , is simply the signal to noise ratio (SNR) between each extracted component and the corresponding source. The second quality measure,  $Q_2$ , is the signal to noise ratio, compensated for possible nonlinear transformations of the intensity scales of the estimated sources. The third measure,  $Q_3$ , is the mutual information between each extracted component and the corresponding source. The mutual information was estimated, in each case, from a set of 5000 randomly selected pixel pairs, chosen independently from those forming the training set, and was computed using the  $I(1)$  estimator described in [144], with  $k = 3$ . More details about these measures can be found in [64]. We didn’t use measure  $Q_4$ , from that reference, because it had shown, in previous tests, not to be a reliable measure of separation quality [64, 66].

Table 2.2 contains the values of the quality measures of the components obtained, for each pair, with the “own” model and with the “bars” model. Table 2.3 shows the results obtained with the “own” model, together with results obtained with MISEP using an MLP as a separator [64], and with results obtained with nonlinear DSS [66]. The table also shows, for comparison, in column MSE, the quality values of what could be considered an “ideal” separation: the result obtained by training an MLP with the two mixture pixels as inputs and with the two source pixels as desired outputs. The results for the fourth and fifth image pairs are not shown in the table because they were not available for both of the other separation methods. The model proposed in this paper performed better, on average, than both MLP-based MISEP and

Img. pair	Quality measure	“Own” model		“Bars” model	
		src 1	src 2	src 1	src 2
1	$Q_1$ (dB)	13.1	13.2	13.1	13.2
	$Q_2$ (dB)	15.2	14.8	15.2	14.8
	$Q_3$ (bit)	2.55	2.47	2.55	2.47
2	$Q_1$ (dB)	10.6	15.3	<b>10.7</b>	<b>15.4</b>
	$Q_2$ (dB)	11.5	<b>15.9</b>	<b>11.6</b>	<b>15.9</b>
	$Q_3$ (bit)	<b>1.90</b>	2.06	<b>1.90</b>	<b>2.09</b>
3	$Q_1$ (dB)	<b>14.9</b>	7.6	14.7	<b>8.6</b>
	$Q_2$ (dB)	<b>15.8</b>	8.5	15.7	<b>9.4</b>
	$Q_3$ (bit)	<b>2.27</b>	1.46	2.25	<b>1.59</b>
4	$Q_1$ (dB)	<b>4.6</b>	<b>13.7</b>	4.5	13.3
	$Q_2$ (dB)	<b>9.1</b>	<b>14.4</b>	<b>9.1</b>	14.0
	$Q_3$ (bit)	<b>0.90</b>	<b>2.23</b>	0.89	2.19
5	$Q_1$ (dB)	4.4	2.3	<b>4.9</b>	<b>3.2</b>
	$Q_2$ (dB)	6.0	4.6	<b>8.9</b>	<b>7.8</b>
	$Q_3$ (bit)	0.75	0.53	<b>0.88</b>	<b>0.66</b>

Table 2.2: Values of the quality measures for the results obtained with the “own” model and with the “bars” model. The best results are shown in bold.

nonlinear DSS.

### 2.1.5 Conclusions

The inverse of a physical model of the mixture process was used to perform the separation of a nonlinear real-life mixture of images. The model’s parameters were estimated by the MISEP method, which uses an ICA criterion.

The separation results are competitive with those obtained with other methods, namely with MLP-based MISEP and with nonlinear DSS. They show, on the one hand, that the mixture model is appropriate for the problem being addressed and, on the other hand, that MISEP is an adequate technique for estimating the model’s parameters.

The model that was used showed not to be perfect. One possible cause could be the existence of gamma correction in the scanning process, which was not accounted for in the model, and which would produce a nonlinear distortion of the gray scales of the mixture images. This could explain the curvature observed in the boundaries of the scatter plot of the separated components. Incorporating gamma correction in the model will involve the addition of just one more parameter and may lead to a more perfect separation.

Another planned improvement involves the explicit incorporation of noise in the model. Noise is clearly present in the mixture process, as evidenced by the scatter plots that were presented. The strongest source of noise probably is the inhomogeneity of the paper. The incorporation of noise in the model may lead to a better estimation of the source images.

Img. pair	Quality measure	“Own” model	MISEP MLP	Nonl. DSS	MSE
1	$Q_1$ (dB)	13.2	13.5	<b>14.4</b>	14.9
	$Q_2$ (dB)	<b>15.0</b>	14.5	<b>15.0</b>	15.3
	$Q_3$ (bit)	2.51	2.42	<b>2.54</b>	2.55
2	$Q_1$ (dB)	<b>13.0</b>	11.6	10.0	13.4
	$Q_2$ (dB)	<b>13.7</b>	13.0	12.3	13.8
	$Q_3$ (bit)	<b>1.98</b>	1.89	1.77	2.00
3	$Q_1$ (dB)	11.2	10.3	<b>11.4</b>	12.5
	$Q_2$ (dB)	12.1	11.6	<b>12.7</b>	13.1
	$Q_3$ (bit)	1.87	1.74	<b>1.93</b>	1.95
Mean	$Q_1$ (dB)	<b>12.4</b>	11.8	11.9	13.6
	$Q_2$ (dB)	<b>13.6</b>	13.0	13.3	14.1
	$Q_3$ (bit)	<b>2.12</b>	1.98	2.07	2.18

Table 2.3: Values of the quality measures, averaged across all image pairs, for three separation methods. The best results are shown in bold. For comparison, column MSE shows the quality of what could be considered an “ideal” separation.

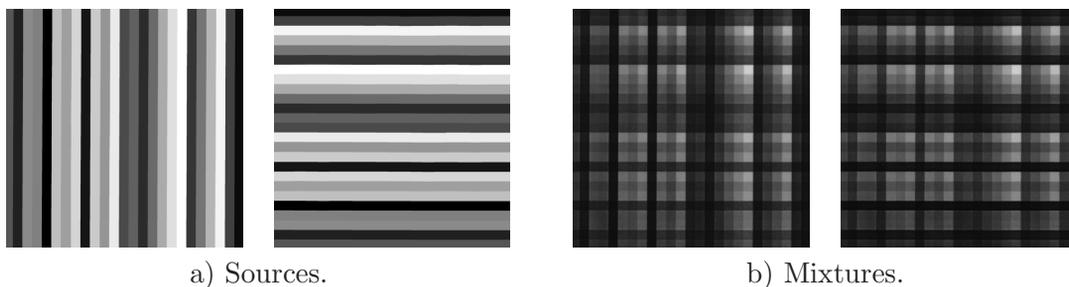


Figure 2.3: Sources and mixtures of the first pair of images.

A more complex improvement will consist of taking into account the non-local character of the mixture. This will become important as the scanning resolution is increased above the one used in this work.





## 2.2 Wavelet based nonlinear separation of images

### Abstract

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

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

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

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

### 2.2.1 Introduction

If we scan or photograph a paper document, the image from the back page often appears superimposed on the image from the front page, especially if the paper is thin or rather transparent. Here we deal with a difficult version of this problem, corresponding to the use of onion skin paper. This creates a strong nonlinear mixture. In our case the mixture images were obtained by printing photos and/or text on both sides of a sheet of onion skin, and then using a common desktop scanner to scan both sides of the onion skin. A more complete description of the way in which the mixture images were obtained is given in [64]. The source and mixture images are available online at <http://www.lx.it.pt/~lbalmeida/ica/seethrough/index.html>, and are shown in Fig. 2.6.

Extracting the source images from the scanned mixtures is a nonlinear blind source separation (BSS) problem. This kind of problem is often approached by assuming that the sources are statistically independent from each other, which then justifies the use of an Independent

Component Analysis (ICA) method. Linear ICA is a well understood problem. One of its main properties is the essential uniqueness of the solution [43]. Nonlinear ICA is still much less studied, however, and suffers, in general, from non-uniqueness of the solution. The problem studied in this paper involves real-life nonlinear, noisy mixtures of images, some of which are not independent from each other. This non-independence, together with the nonlinear character of the mixture, affects the quality of the results of ICA-based methods [64].

The image separation problem that we study here has been previously addressed by two other methods. The one described in [64] was based on the MISEP method of nonlinear ICA [55], using a multilayer perceptron as separating system. The other one [66] was based on the nonlinear denoising source separation (DSS) method, which uses some of the basic ideas that are used in this paper, albeit in a less efficient manner. The same image separation problem is also addressed in another paper in this conference [71]. That paper presents a solution based on a physical model of the mixture process, the model's parameters being trained by means of an ICA criterion through the MISEP method.

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

The separation method proposed here is similar to the denoising step of the nonlinear DSS method described in [66]. However, while that denoising step yielded only a partial separation which had to be improved by iterating within the DSS scheme, the improvements proposed here allow a rather complete separation to be performed in a single step, without the need for any iterative DSS procedure. In contrast to nonlinear DSS, and also to the ICA-based methods mentioned above, which perform a pixel-by-pixel mapping, the method proposed here is non-pointwise, due to its use of wavelets.

This manuscript is organized as follows: Section 2.2.2 describes the separation method. Section 2.2.3 presents experimental results. Section 2.2.4 concludes and presents future research directions.

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

## 2.2.2 Separation Method

### High frequency competition

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

1. The high frequency components of common images are sparse. In a wavelet based representation this is translated into the fact that wavelet coefficients have sparse distributions [145]. As a consequence, the wavelet coefficients from two different source images will seldom both have significant values in the same image location.
2. Each source is more strongly represented in the image acquired from the side of the paper in which that source is printed than in the image acquired from the opposite side.

The separation method is summarized in Fig. 2.5. After performing wavelet decomposition for a certain number of levels, we perform a competition between the corresponding wavelet coefficients from the two mixture images, according to

$$\sigma_i = \sqrt{\langle x_i^2 \rangle} \quad (2.13)$$

$$m_i = \frac{1}{1 + \exp\left(-a \frac{x_i^2 - x_{3-i}^2}{x_i^2 + x_{3-i}^2}\right)} \quad (2.14)$$

$$y_i = \sigma_i \frac{x_i m_i}{\sqrt{\langle x_i^2 m_i^2 \rangle}} \quad (2.15)$$

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

This competition was applied to all horizontal, vertical and diagonal wavelet coefficients at all decomposition levels (represented, in Fig. 2.5, by blocks  $H_j$  – horizontal coefficients at level  $j$ ,  $V_j$  – vertical coefficients at level  $j$ , and  $D_j$  – diagonal coefficients at level  $j$ ). The competition computes mask  $m_i$  through a soft *winner take all* function given by Eq. (2.14), and then applies

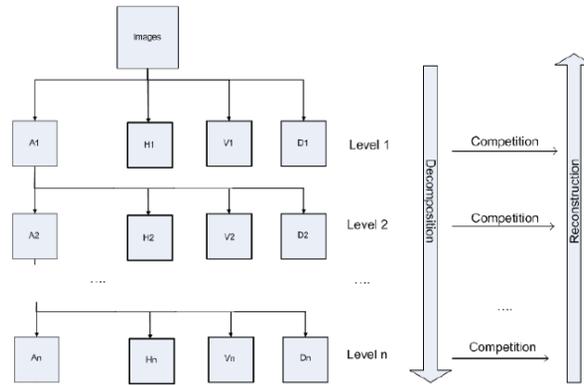


Figure 2.5: Schematic representation of the wavelet-based separation method.

this mask to the wavelet coefficients of the mixture image, so that the coefficients are much intensified in the image in which they originally were strongest and much weakened in the image in which they were weakest.

The separated images were synthesized using, at all levels (except for the low frequency coefficients of the deepest level), the wavelet coefficients computed by this form of competition. For the low frequency coefficients at the deepest level ( $A_n$  in Fig. 2.5) we have used, in some cases, the coefficients from the corresponding mixture image and, in other cases, a value of zero, as detailed ahead.

### Improvements in the separation method

One of the important aspects of the separation method has to do with the choice of the wavelet representation. The commonly used decimated wavelet transform showed not to be very appropriate for the task at hand, leading to a rather incomplete separation. This was probably due to its shift-varying character, which makes it represent edges better or worse depending on their exact location. We tried two different wavelet transforms, to circumvent this limitation:

- The discrete complex wavelet transform [146], which is almost shift-invariant and is rather directionally selective. This transform has the disadvantage of having to use wavelets with a relatively large support, which make it less effective in handling the finer details of images.
- The stationary discrete wavelet transform [147], which is shift-invariant and can use short-support wavelets (e.g. Haar), at the cost of using a largely redundant representation, which translates into a somewhat higher computational cost.

An important improvement of the proposed separation method relative to the denoising step used in nonlinear DSS [66] consists of the use of a shift-invariant (or almost shift-invariant) wavelet transform instead of a decimated one. Another important improvement consists of the use of the competition procedure described in Eqs. (2.13-2.15). Together, these two improvements make it possible to perform the separation in a single step, just by means of the wavelet-based competition. The method avoids both the use of the multilayer perceptron and of the iteration that were needed in nonlinear DSS. Being non-iterative, the method is much more efficient than both nonlinear DSS and the ICA-based methods mentioned above.

### Preprocessing

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

$$Q = \begin{bmatrix} \frac{a_{11}+a_{22}}{2} & a_{12} \\ a_{21} & \frac{a_{11}+a_{22}}{2} \end{bmatrix}, \quad (2.16)$$

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

### 2.2.3 Experimental results

The separation method described in Section 2.2.2 was applied to the five pairs of mixtures shown in Fig. 2.6, both with and without the preprocessing described in Section 2.2.2. In the reconstruction of each separated image we used, for the low frequency coefficients ( $A_n$  in Fig. 2.5) the coefficients from the corresponding preprocessed image, in the cases in which preprocessing was used. In the cases in which no preprocessing was used, we have set  $A_n = 0$  for the reconstruction. This yielded somewhat better results than keeping the coefficients from the corresponding mixture image. The value of parameter  $a$  used in Eq (2.14) was 510, which corresponds to a strong competition, the mask implementing almost a hard winner take all function.



Figure 2.6: Images used in the tests. Left: Source images. Right: Mixture images. From top to down: Pair #1, Pair #2, Pair #3, Pair #4 and Pair #5.

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

The best separation results were obtained with the stationary wavelet transform with pre-



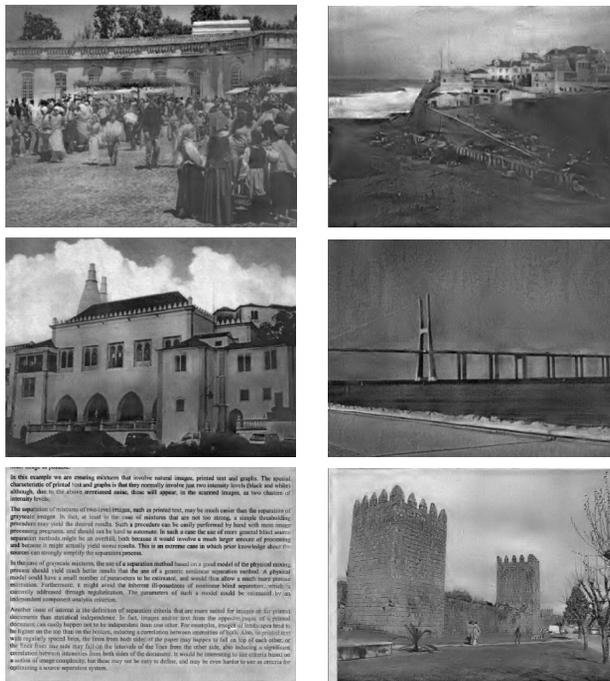


Figure 2.8: Results obtained with the complex wavelet transform with preprocessing.

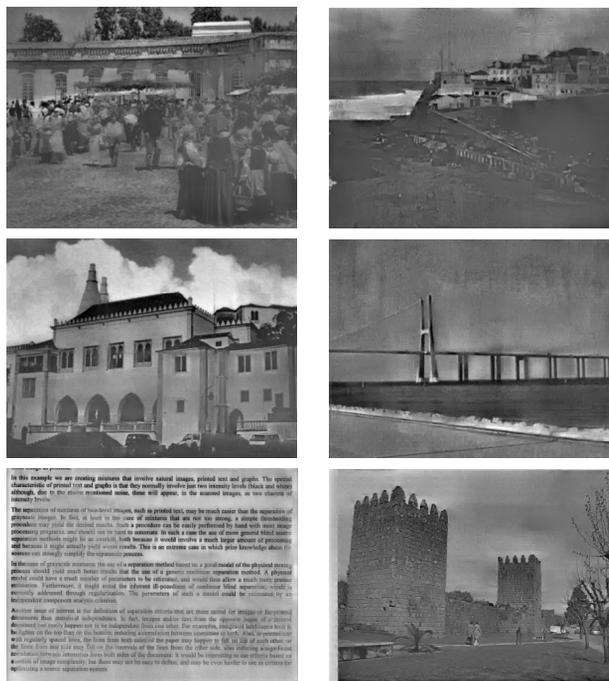


Figure 2.9: Results obtained with the stationary wavelet transform without preprocessing.

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

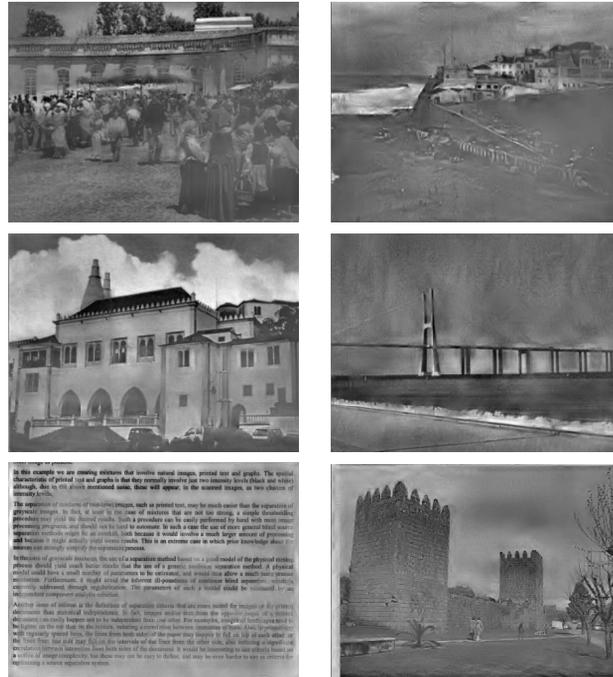


Figure 2.10: Results obtained with the complex wavelet transform without preprocessing.

The results obtained with complex wavelets and preprocessing (Fig. 2.8) are still quite good, although they show somewhat stronger imperfections. The results obtained without preprocessing (Figs. 2.9 and 2.10) show significantly stronger degradations.

## 2.2.4 Discussion and future work

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

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

Contrary to other separation methods that have been presented in the literature, which are pointwise (in the sense that the intensity of each pixel of a separated image depends only on the intensities of the corresponding pixels of the mixture images), the method presented here is not pointwise. If we take into account that the largest Haar wavelets have  $128 \times 128$  pixels in a 7 level decomposition and  $256 \times 256$  pixels in an 8 level one, we see that the separation that is performed is strongly non-pointwise. Therefore, pointwise quality measures such as those

that have been used with other separation methods [64, 66] are not appropriate for evaluating the perceptual quality of this method's results. We plan to assess the separation quality with other more perceptually oriented measures, such as those proposed in [148, 149].

### **2.2.5 Acknowledgment**

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

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

### Abstract

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

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

**keywords:** Nonlinear separation, image mixture, show-through, bleed-through, image enhancement.

### 2.3.1 Introduction

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

In this work we use, as test examples, three different kinds of mixtures. The first kind essentially only contains the show-through effect: five pairs of images were printed on the two sides of five sheets of tracing paper<sup>3</sup> which, due to its high transparency, creates very strong

---

<sup>3</sup>In previous publications, this tracing paper has been improperly called "onion skin paper". The latter is, for example, the very thin paper commonly used, some decades ago, for air mail letters, and actually appears in the "air mail letter" mixture that we used in this work. Tracing paper is the semi-transparent paper often used

mixtures. The second type corresponds to an old manuscript letter written in very thin “air mail” paper (also called *onion skin paper*, which is rather transparent, causing show-through to occur). This document also has some areas in which bleed-through appears to have occurred. The third kind of mixture corresponds to images of old manual transcripts of music (partitures), which mostly contain the bleed-through effect. For each document, scanning or photographing both sides allowed us to obtain two different mixtures of the contents of the two pages. In this paper we address the source separation problem whose aim is to recover, from the two acquired images of each document, the original page images.

Show-through is known to lead to nonlinear mixtures [64, 66, 71]. A physical model of the show-through mixture of gray-level images printed with halftoning has been presented in [71]. Bleed-through probably is a much more complex phenomenon, which is much harder to model.

Source separation is often performed by assuming that the sources are statistically independent from each other, an assumption which leads to the use of independent component analysis (ICA) techniques. While linear ICA is a well studied problem for which several efficient solutions exist [44, 45, 47], nonlinear ICA is a much less studied problem [52, 53, 55, 56]. Nonlinear ICA has the additional difficulty of being ill-posed, having an infinite number of solutions without any simple relationship with one another [48, 49]. The mixtures addressed in this work are nonlinear and noisy, and the letter and partiture mixtures are spatially variant. Besides these challenging properties, most of the sources studied in this work don’t completely obey the independence assumption, a fact which affects the quality of the results obtained through ICA-based methods [64, 66].

Instead of assuming independence of the source images, we propose a solution that uses other properties of images and of the mixture process. We use the well known fact that high-frequency components of images are sparse (and that high-frequency wavelet coefficients are also sparse), and we formulate a competition based on the observation that each source is more strongly represented in one of the mixture components than in the other one. Making assumptions that are suited to the present problem, our method achieves a good perceptual separation quality even when the sources are non-independent and the mixture is spatially variant. The separation method that we propose is similar to the denoising step used by nonlinear DSS [66]. However, we use an improved form of competition, and also a wavelet transform that is more suited to the problem at hand. These improvements lead to a method that performs the separation in a single step, without the iterative procedure required by nonlinear DSS.

---

in professional drawing.

Both the old manuscript letter and the old partitures with the bleed-through effect are addressed for the first time in this paper. On the other hand, the tracing paper mixtures have already been studied in other works [64, 66, 71]. Contrasting with the method proposed here, which, due to its use of wavelets, performs a non-point-wise transformation, all the other mentioned methods performed point-wise separation. One of them [64] used the MISEP method of nonlinear ICA [55] to train a regularized MLP which performed the separation. In another one [71], MISEP was used to train a nonlinear physical model of the mixture process. Nonlinear DSS has also been applied to some of the tracing paper mixtures [66]. Nonlinear DSS doesn't assume independence of the sources, but assumes spatial invariance of the mixture. It uses the same basic ideas that are used in this paper, albeit in a less efficient manner. Show-through and/or bleed-through mixtures have also been addressed in [1, 2, 29, 31, 63], but in different settings from the one considered here. In [2] and [1] separation is archived through linear models, which were shown to be too restrictive to separate tracing paper mixtures [64]. [29] and [31] focus only on the restoration of text documents, for which linear separation yields relatively good results (see [64]). In [63], the contents of both pages are assumed to consist of text, and separation is linear and is based on a single color image from one side of the document.

This manuscript is structured as follows: Section 2.3.2 describes the three kinds of mixtures that were studied, as well as the processes of image acquisition and alignment. Section 2.3.3 describes the proposed separation method. Section 2.3.4 presents experimental results, and Section 2.3.5 concludes.

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

## 2.3.2 Experimental setup

### Mixtures and acquisition

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

- **Tracing paper images** - Five different pairs of images (including synthetic bars, photos and text) were used as sources of five pairs of mixtures. These pairs of sources, shown in

Figs. 2.11 and 2.12, were printed on opposite pages of tracing paper, which was chosen for its high transparency. Image printing and scanning were performed as symmetrically as possible, regarding the two images of each pair. Printing was performed with a 1200 dpi laser printer, using the printer's default halftoning system. Both pages were then scanned with a desktop scanner at a low resolution of 100 dpi, which was chosen so that the printer's halftoning grid would not be apparent in the scanned images. Figures 2.13 and 2.14 show the acquired images after the alignment procedure that is described below. The mixtures involved here are highly nonlinear and, except for the pair of text images, linear approaches do not archive reasonable separation results (see [64] for a detailed analysis).

- **Air mail letter** - The document used in this case was an old, handwritten air mail letter. Old air mail paper was very thin, having a high transparency that, as can be seen in the acquired pages shown in Fig. 2.15, leads to strong mixtures. Since the letter was very thin and old, it was hard to place in the scanner in a good position, without any wrinkles. The letter had three folds and, although the acquisition was performed as carefully as possible, these folds and some wrinkles could not be eliminated (the wrinkles are especially visible in the bottom corners of the acquired images). This led to a mixture that is not space-invariant, but instead has characteristics that vary from one place to another. This mixture is not compatible with a global mixture/separation model. Also, differently from the tracing paper images, whose middle-tone levels were generated by a halftoning process, the middle-tones in these sources are due to variable transparency of the ink. The acquired images show that the ink wasn't homogeneously distributed in different words: its density depended on the pressure applied by the writer. Furthermore, in a few areas, the ink seems to have bled through the paper.

Many of the lines of text of both pages are aligned with one another, which makes the original images non-independent. This is a disadvantage for ICA-based methods, and also a challenge for our method due to the increase in the number of superimposed edges from both pages. The images were acquired using the same desktop scanner that was used for the tracing paper mixtures. However, since the writing ink was blue, the images were acquired in color. They were then converted to grayscale, specifically for this work. Contrary to the tracing paper mixtures, which were specially produced for studies of this kind, we couldn't access the actual sources corresponding to the air mail letter.

- **Old partitures** - These mixtures were obtained from very old handwritten partitures,

in which bleed-through is the main effect. We used two partiture sheets (see Fig. 2.16) which were chosen, among those that we had available, for presenting the strongest bleed-through effects. Bleed-through is present, to a significant level, in a few areas of these documents. The images were acquired by a photographic process, and only the photographs were available to us, not the original documents. As with the air mail letter, they were originally acquired in color and then converted to grayscale. They also show some evidence of wrinkles. And, as in the case of the air mail letter, we didn't have access to the original source images.

## **Alignment**

Image separation methods usually require that the components of the mixture be precisely registered with one another. This alignment is essential both in ICA-based methods and in our method, which explicitly assumes edges to be in the same spatial position in the two mixture components. To achieve this correspondence, one of the images of each pair had first to be horizontally flipped. After that, an alignment procedure was applied, to correct misalignments due to the different positions of the paper during the two scanning acquisitions. In order for it to be precise, the alignment had to be performed locally. This local alignment was needed even for documents that were not wrinkled, probably due to some geometrical imperfections of the scanner. In the case of the air mail letter, which was significantly wrinkled, the local alignment was even more important.

All mixture pairs were subject to an initial crude alignment, which was performed by hand, with the aid of an image editing program. After this manual alignment, the tracing paper and air mail letter mixtures were subject to a local alignment procedure, which was specially developed for this problem. Mixture images were first expanded in resolution, by a factor of four in each direction, using bicubic interpolation. Then, one of the images of the pair was divided into 100x100 pixel squares (corresponding to 25x25 pixels in the original image) and, for each square, the best displacement was found, based on the maximum of the cross-correlation with the other image. Aligned images were then reduced to the original resolution. Therefore this alignment method performed a local alignment with a resolution of 1/4 pixel. More information about the tracing paper images and the alignment process is available in [64]. Since the partitures only presented significant bleed-through in a few areas of the images, showing a weak mixture on most other areas, the alignment routine based on small-block correlations didn't work satisfactorily for this case. Therefore, these mixtures were only subject to a careful manual alignment.

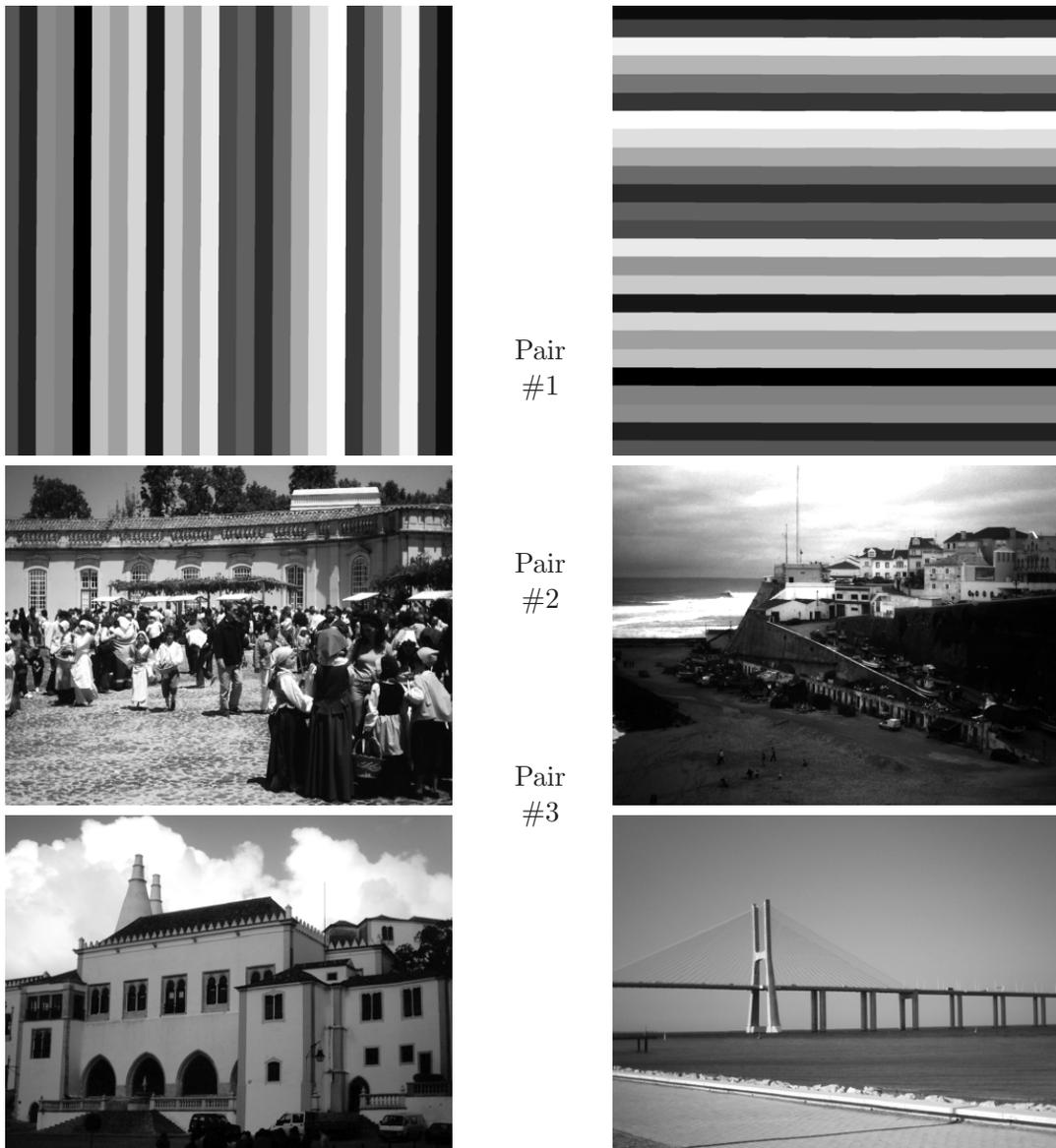


Figure 2.11: The first three first pairs of tracing paper source images. In this and all subsequent figures containing images, one of the images of each pair has been horizontally flipped.

### 2.3.3 Separation method

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

1. High-frequency components of common images are sparse. This translates into the fact that high-frequency wavelet coefficients have sparse distributions [145]. Consequently, the high-frequency wavelet coefficients from two different source images will seldom both have significant values in the same image location.

source image or page.

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

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

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

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

**Separation of nonlinear image mixtures**

When acquiring an image of a printed document, the image printed on the opposite page often shows through, due to partial transparency of the paper. Here we are dealing with a strong case of that effect, because we're using onion-skin paper which is quite transparent.

The mixture that is obtained is rather nonlinear, as can be observed from the top figure on the right, which shows a scatter plot of the intensities of corresponding pairs of points from the two pages of a printed document. The scatter plot of the original images, shown in the bottom figure, filled a square, and had only a relatively small number of discrete intensity levels for each image. The fact that the shape of the scatter plot of Fig. 1 is very different from a parallelogram shows that the mixture was strongly nonlinear. The fact that this scatter plot becomes quite narrow in the upper right corner (which corresponds to the lighter intensities in both images) indicates that, for those intensities, the mixture is close to singular. Finally, the fact that the discrete levels of Fig. 2 become largely blurred in Fig. 1 is due to noise in the process. The process leading from the sources to the observations involved printing the images, on both sides of a stack of onion skin paper, at 1200 dpi, with a black and white laser printer (with the inherent halftoning of gray levels), and then scanning both sides of the printed sheet at 100 dpi. The noise is due, at least, to the printing process (including the halftoning), in the scanning process and to the non-uniformity in the onion skin paper, especially in its transparency.

The purpose of separation is to recover, from the mixed images that are obtained by scanning both faces of the printed document, the images that had been printed in each of its faces, with as little interference from the other image as possible.

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

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

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

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

Pair #4



nonlinear .1 1

While the case of mixtures that involve natural images, printed text and graphs is that they normally involve just two intensity levels (black and white) although, due to the above mentioned noise, these will appear, in the scanned images, as two clusters of intensity levels.

Pair #5

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

Figure 2.12: The fourth and fifth pairs of tracing paper source images.

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

**High-frequency competition**

The separation method, which is summarized in Fig. 2.18, manipulated the images through a wavelet-based representation. First, both images were subject to a wavelet decomposition down to a certain depth. Then, the corresponding high-frequency wavelet coefficients of both mixture images were subject to the following competition process:

$$m_i = \frac{1}{1 + \exp\left(-a \frac{x_i^2 - x_{3-i}^2}{x_i^2 + x_{3-i}^2}\right)} \tag{2.17}$$

$$y_i = x_i m_i \tag{2.18}$$

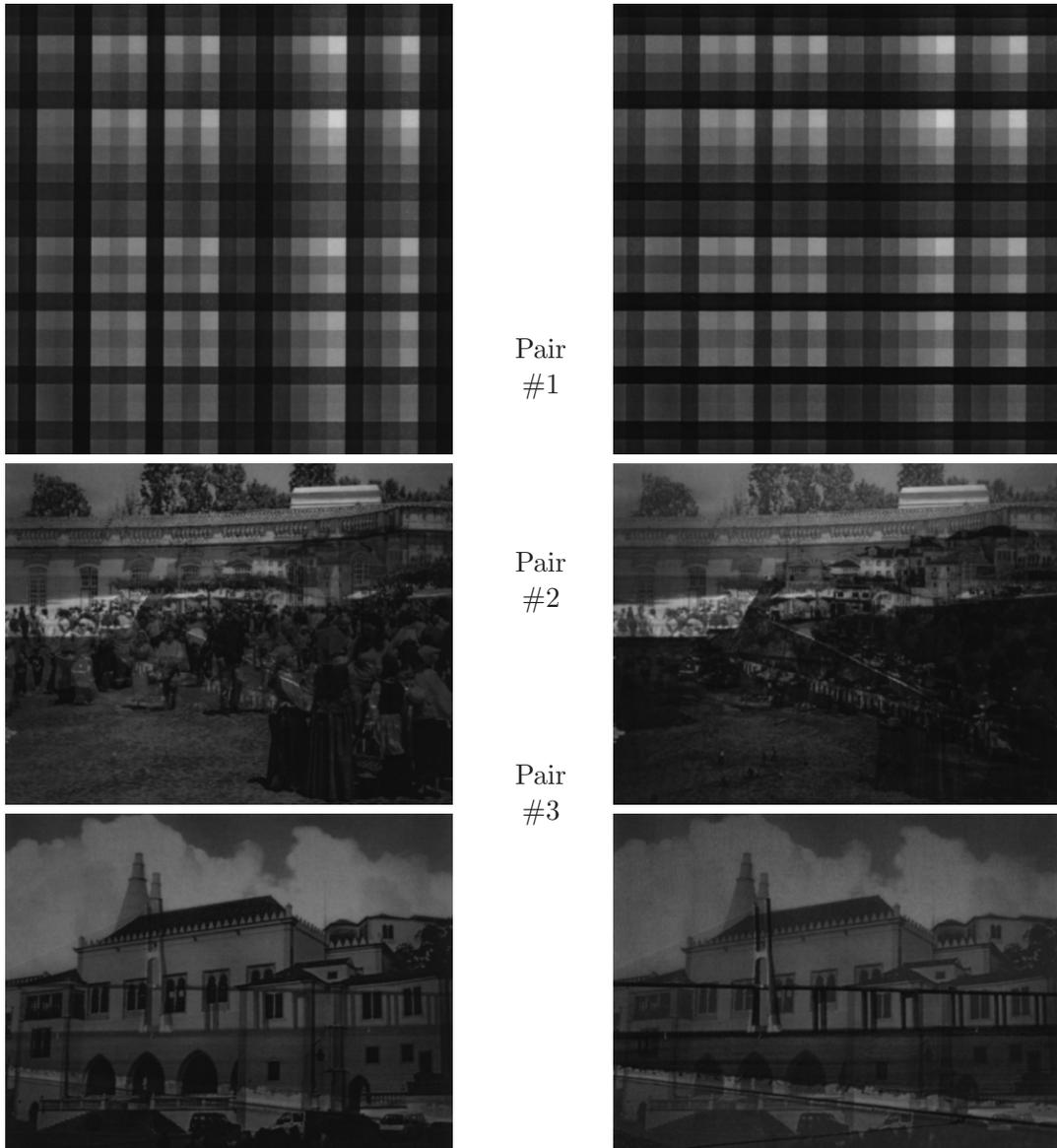


Figure 2.13: First three pairs of tracing paper mixtures.

where  $i \in \{1, 2\}$  indexes the two sides of the paper,  $x_i$  are the wavelet coefficients of a given type (for example, vertical coefficients at the first decomposition level) from the  $i^{\text{th}}$  mixture image,  $x_{3-i}$  are the corresponding coefficients from the other image of the same mixture, and  $y_i$  are the coefficients that are used for synthesizing the  $i^{\text{th}}$  separated image;  $a$  is a parameter that controls the strength of the competition. The first equation computes a soft winner-take-all mask  $m_i$  which is then used, in the second equation, to control the strength of the corresponding high-frequency component in the separated image. This procedure preserves the coefficients that are stronger in the mixture component under consideration than in the opposite component, and weakens the coefficients that are weaker than in the opposite component. Fig. 2.17 illustrates how the soft winner-take-all mask (2.17) rules the competition process

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

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

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

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

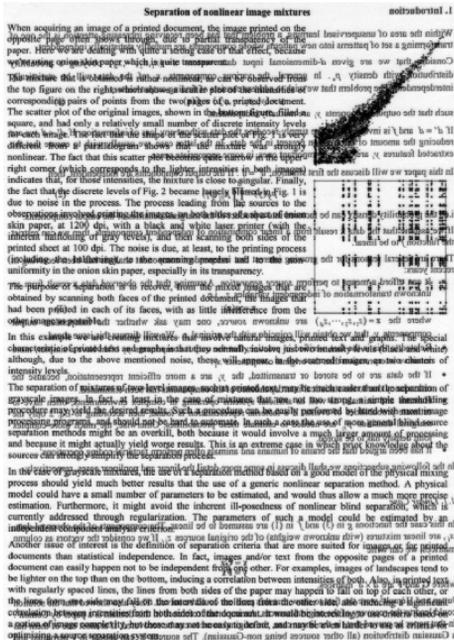
In this example we are creating mixtures that involve natural images, printed text and graphs. The special characteristic of printed text and graphs is that they normally involve just two intensity levels (black and white) although, due to the above mentioned noise, these will appear, in the scanned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

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

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

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

Pair #4



Pair #5

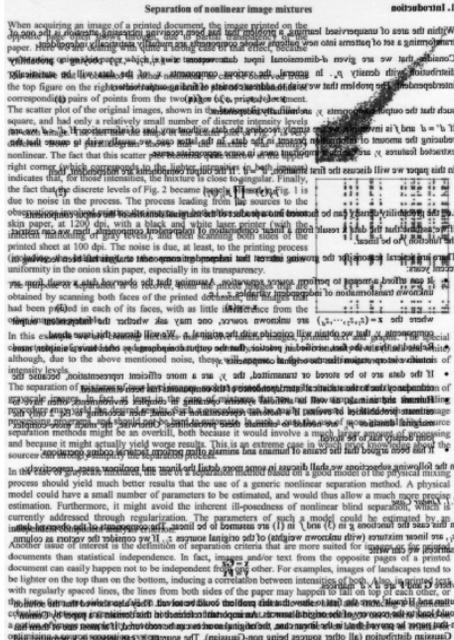


Figure 2.14: Fourth and fifth pairs of tracing paper mixtures.

described in (2.17-2.18). We should note that the exact definition of the competition mask is not important, as long as it has the general behavior illustrated in Fig. 2.17.

The competition process described in (2.17-2.18) was applied to all horizontal, vertical and diagonal wavelet coefficients at all decomposition levels (represented, in Fig. 2.18, by blocks  $H_j$  – horizontal coefficients at level  $j$ ,  $V_j$  – vertical coefficients at level  $j$ , and  $D_j$  – diagonal coefficients at level  $j$ ).

The separated images were obtained by wavelet reconstruction using the high-frequency coefficients ( $H_j, V_j, D_j$ ) after competition. For the low frequency coefficients ( $A_n$  in Fig. 2.18) we used the coefficients obtained from the decomposition of the corresponding mixture image, with no change.

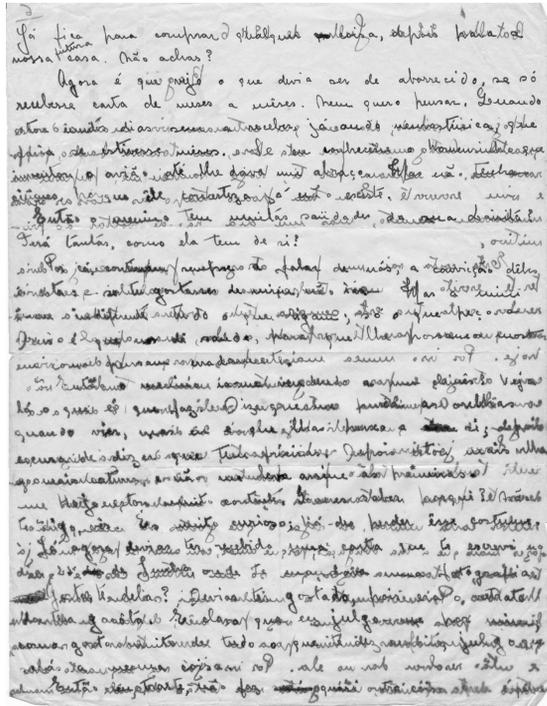
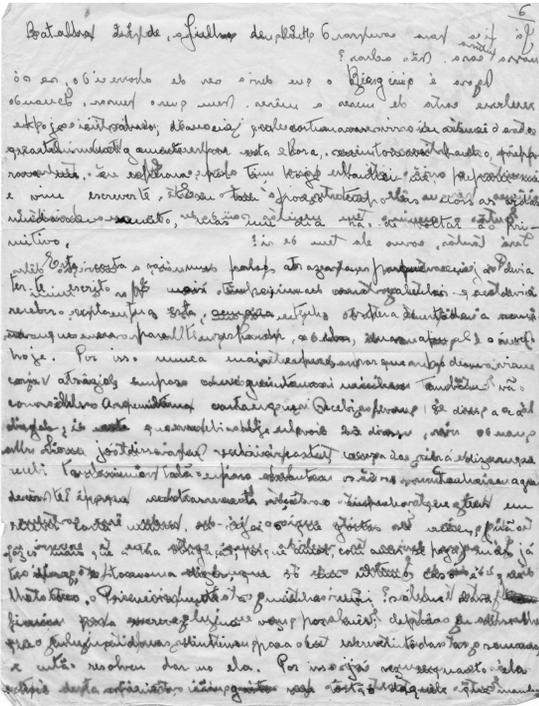


Figure 2.15: Acquired images of the old air-mail letter.

## Wavelet representation

An important aspect of the proposed method has to do with the choice of the wavelet representation. The commonly used decimated wavelet transform showed not to be very appropriate for the task at hand, leading to a rather incomplete separation. This was probably due to its shift-varying character, which makes it represent edges better or worse depending on their exact locations. Since this representation didn't yield satisfactory results, we tried two different wavelet transforms which circumvent this limitation:

- The stationary discrete wavelet transform [147], which is shift-invariant and can use short-support wavelets (e.g. Haar). This transform uses a very redundant representation, which translates into a somewhat higher computational cost.
- The discrete complex wavelet transform [146], which is almost shift-invariant and is rather directionally selective. This transform has the disadvantage of having to use wavelets with a relatively large support, making it less effective in handling the finer details of the images.



Partiture #1



Partiture #2

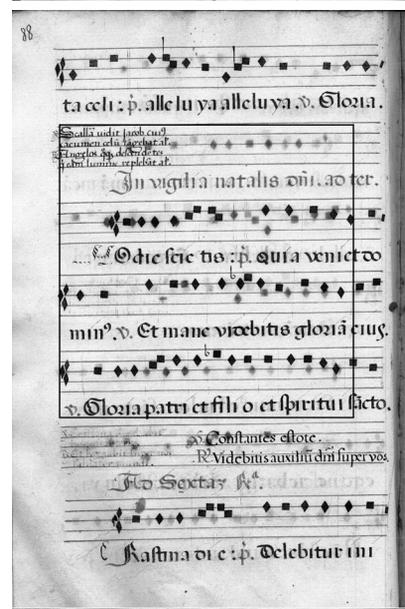


Figure 2.16: Acquired images of the old partitures. The squares indicate the areas that were selected for separation.

### Comparison with nonlinear DSS

The method that we propose is similar to the denoising step used in nonlinear DSS [66], but incorporates two important improvements. One corresponds to the use of a more suitable wavelet transform, which is shift-invariant or almost shift-invariant. The other improvement has to do with the use of a better form of competition. Together, these two improvements led to a one-step procedure that is, by itself, sufficient to separate the images. The proposed method avoids the use of both the multilayer perceptron and the iteration that were required in nonlinear DSS. Being non-iterative, the method is much more efficient than both nonlinear

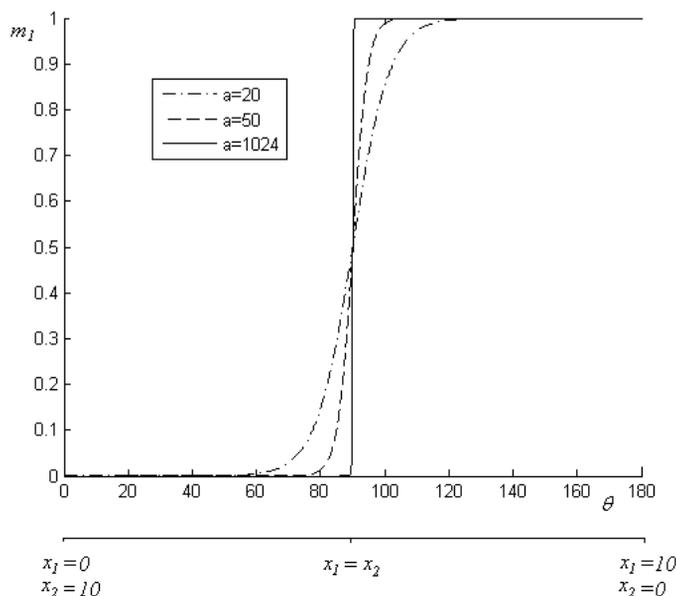


Figure 2.17: Behavior of the mask  $m_1$  for different values of parameter  $a$ . Focusing on the case where  $x_1^2 + x_2^2 = 100$ , which is of the order of magnitude found in our images, mask  $m_1$  is plotted against  $\theta$ , with  $x_1 = 10\sin\theta$  and  $x_2 = 10\cos\theta$ .

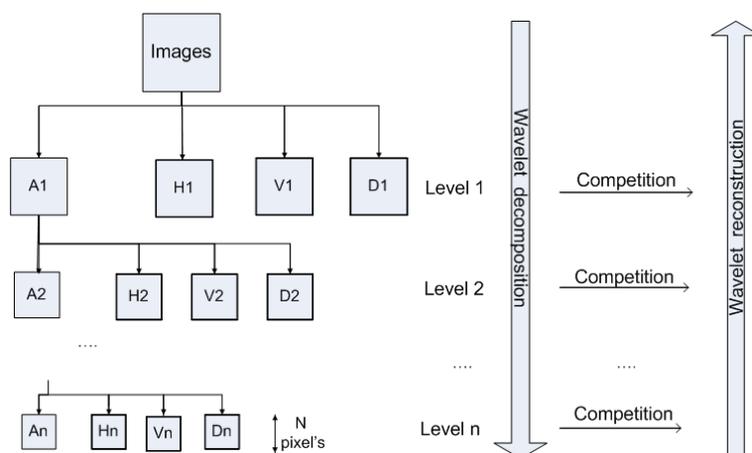


Figure 2.18: Schematic representation of the wavelet-based separation method.

DSS and the ICA-based methods that have previously been proposed.

### Linear preprocessing

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

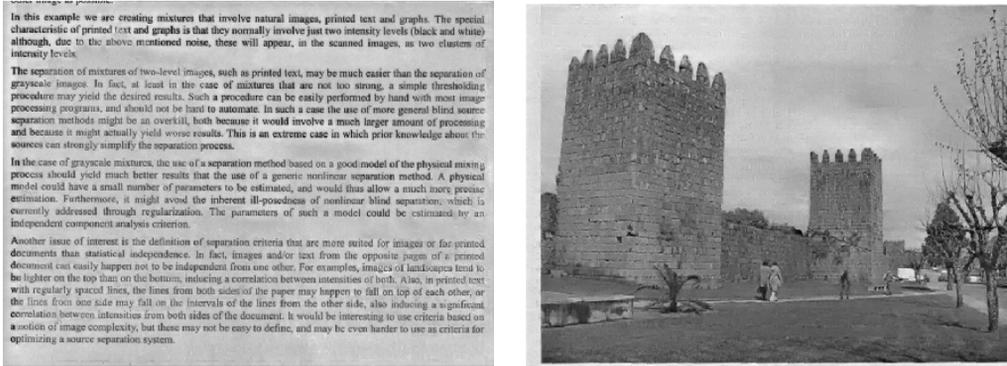


Figure 2.19: Fourth pair of tracing paper mixtures separated without using contrast compensation.

general, to perform an exact decorrelation that is symmetrical in this sense. Designating by  $A$  the square root of the autocovariance matrix of the mixture vector,  $A$  satisfies  $a_{12} = a_{21}$  but, in general,  $a_{11} \neq a_{22}$ . The linear preprocessing that we performed used a matrix  $Q$  defined by

$$Q = \begin{bmatrix} \frac{a_{11}+a_{22}}{2} & a_{12} \\ a_{21} & \frac{a_{11}+a_{22}}{2} \end{bmatrix} \quad (2.19)$$

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

### Contrast compensation

The previously described method tries to identify, through the competition detailed in Section 2.3.3, which edges correspond to each source. Since edges contain the main information that a human being extracts from an image [150, 151], edge identification leads to a good perceptual separation of images. However, the intensity of the reconstructed edges may not be correct. When the separation method described above was applied to the tracing paper mixtures, the contrast of each recovered source image was found to be reduced in the areas where the other source image was darker. This effect is quite visible in Fig. 2.19, which shows the separation obtained for the fourth pair of tracing paper mixtures without contrast compensation. One can observe a reduction of the contrast of the separated text in the areas where the opposite component is darker (castle and ground areas). This imperfection was to be expected since, in each mixture component, the contrast of each source is lower where the other source is darker, and the competition that we described above doesn't compensate for this effect: it simply assigns the high-frequency components to the sources, but does not correct their intensities.

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

The contrast compensation uses a parameter ( $g_{max} \geq 1$ ) that controls how strong the maximal compensation will be. The gain applied to each estimated source is an affine function of the other source's estimated intensity, varying between 1 (when the other source is white) and  $g_{max}$  (when the other source is black). For this contrast compensation we need to use an estimate of the other source's intensity. At each wavelet reconstruction level  $j$ , we have access to the low-frequency image of that level ( $A_j$ ), obtained from the wavelet reconstruction of the lower levels, which were already subject to the competition process. In the proposed method, this low-frequency image is used for estimating the other source's intensity, to control the gain to be applied to each of the high-frequency wavelet coefficients in the reconstruction at level  $j$  (referring to Fig. 2.18, the reconstructed image  $A_j$  from one component is used to control the gain applied to  $H_j$ ,  $V_j$  and  $D_j$  of the other component). This contrast compensation mechanism is thus applied while the sources are being reconstructed, using the information of lower levels that have already been reconstructed. It leads to a moderate increase of the computational complexity of the separation method, as will be detailed ahead.

### 2.3.4 Experimental results

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

The separation method presented in Section 2.3.3 was applied to the three mixture sets that were described in Section 2.3.2. For all experiments, the value of the competition parameter  $a$  (see (2.17)) was set to 1024, which yields a mask that is almost a hard winner-take-all function.

The tracing paper mixture pairs were separated using the proposed method with and without decorrelation preprocessing, with and without contrast compensation, and using both the complex wavelet transform and the stationary discrete wavelet transform with the Haar wavelet. All separations were performed with a 7-level wavelet analysis. The results that we considered to be best are shown in Figs. 2.20 and 2.21, and correspond to using the stationary discrete wavelet transform with the Haar wavelet, with decorrelation preprocessing and with contrast compensation using  $g_{max} = 3$ . The separation results of other variants of the method, for the

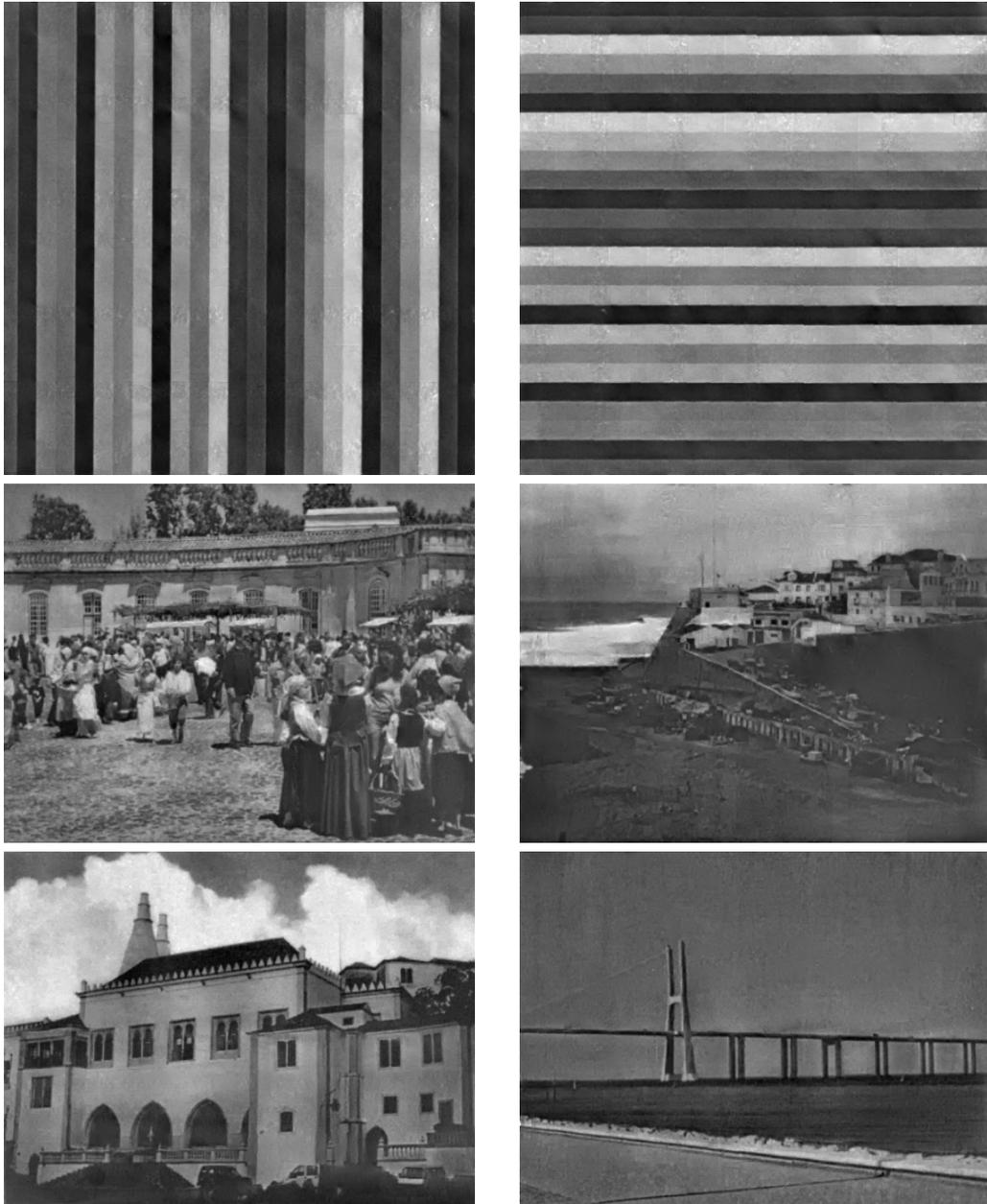


Figure 2.20: Best separation results for the three first pairs of tracing paper mixtures.

bars mixture, are shown in Fig. 2.22. The impact of the various options (preprocessing, contrast compensation and choice of wavelet transform) is clearly visible. The results of Figs. 2.20 and 2.21 have a perceptual separation quality that we consider better than the one of the results obtained with the other existing methods (see [64, 66, 71]. In Section 2.3.4 we present a formal evaluation of some of these results.

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

When images are printed...

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

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

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

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



**Separation of nonlinear image mixtures**

When acquiring an image of a printed document, the image printed on the opposite page often shows through, due to partial transparency of the paper. Here we are dealing with quite a strong case of that effect, because we're using onion skin paper which is quite transparent.

The mixture that is obtained is rather nonlinear, as can be observed from the top figure on the right, which shows a scatter plot of the intensities of corresponding pairs of points from the two pages of a printed document. The scatter plot of the original images, shown in the bottom figure, filled a square, and had only a relatively small number of discrete intensity levels for each image. The fact that the shape of the scatter plot of Fig. 1 is very different from a parallelogram shows that the mixture was strongly nonlinear. The fact that this scatter plot becomes quite narrow in the upper-right corner (which corresponds to the lighter intensities in both images) indicates that, for those intensities, the mixture is close to singular. Finally, the fact that the discrete levels of Fig. 2 became largely blurred in Fig. 1 is due to noise in the process. The process leading from the sources to the observations involved printing the images, on both sides of a sheet of onion skin paper, at 1200 dpi, with a black and white laser printer (with the inherent halftoning of gray levels), and then scanning both sides of the printed sheet at 100 dpi. The noise is due, at least, to the printing process (including the halftoning), to the scanning process and to the non-uniformity in the onion skin paper, especially in its transparency.

The purpose of separation is to recover, from the mixed images that are obtained by scanning both faces of the printed document, the images that had been printed in each of its faces, with as little interference from the other image as possible.

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

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

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

Another issue of interest is the definition of separation criteria that are more suited for images or for printed documents than statistical independence. In fact, images and/or text from the opposite pages of a printed document can easily happen not to be independent from one other. For examples, images of landscapes tend to be lighter on the top than on the bottom, inducing a correlation between intensities of both. Also, in printed text with regularly spaced lines, the lines from both sides of the paper may happen to fall on top of each other, or the lines from one side may fall on the intervals of the lines from the other side, also inducing a significant

**Nonlinear image mixtures**

When acquiring an image of a printed document, the image printed on the opposite page often shows through, due to partial transparency of the paper. Here we are dealing with quite a strong case of that effect, because we're using onion skin paper which is quite transparent.

The mixture that is obtained is rather nonlinear, as can be observed from the top figure on the right, which shows a scatter plot of the intensities of corresponding pairs of points from the two pages of a printed document. The scatter plot of the original images, shown in the bottom figure, filled a square, and had only a relatively small number of discrete intensity levels for each image. The fact that the shape of the scatter plot of Fig. 1 is very different from a parallelogram shows that the mixture was strongly nonlinear. The fact that this scatter plot becomes quite narrow in the upper-right corner (which corresponds to the lighter intensities in both images) indicates that, for those intensities, the mixture is close to singular. Finally, the fact that the discrete levels of Fig. 2 became largely blurred in Fig. 1 is due to noise in the process. The process leading from the sources to the observations involved printing the images, on both sides of a sheet of onion skin paper, at 1200 dpi, with a black and white laser printer (with the inherent halftoning of gray levels), and then scanning both sides of the printed sheet at 100 dpi. The noise is due, at least, to the printing process (including the halftoning), to the scanning process and to the non-uniformity in the onion skin paper, especially in its transparency.

The purpose of separation is to recover, from the mixed images that are obtained by scanning both faces of the printed document, the images that had been printed in each of its faces, with as little interference from the other image as possible.

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

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

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

Another issue of interest is the definition of separation criteria that are more suited for images or for printed documents than statistical independence. In fact, images and/or text from the opposite pages of a printed document can easily happen not to be independent from one other. For examples, images of landscapes tend to be lighter on the top than on the bottom, inducing a correlation between intensities of both. Also, in printed text with regularly spaced lines, the lines from both sides of the paper may happen to fall on top of each other, or the lines from one side may fall on the intervals of the lines from the other side, also inducing a significant

Figure 2.21: Best separation results for the fourth and fifth pairs of tracing paper mixtures.

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

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

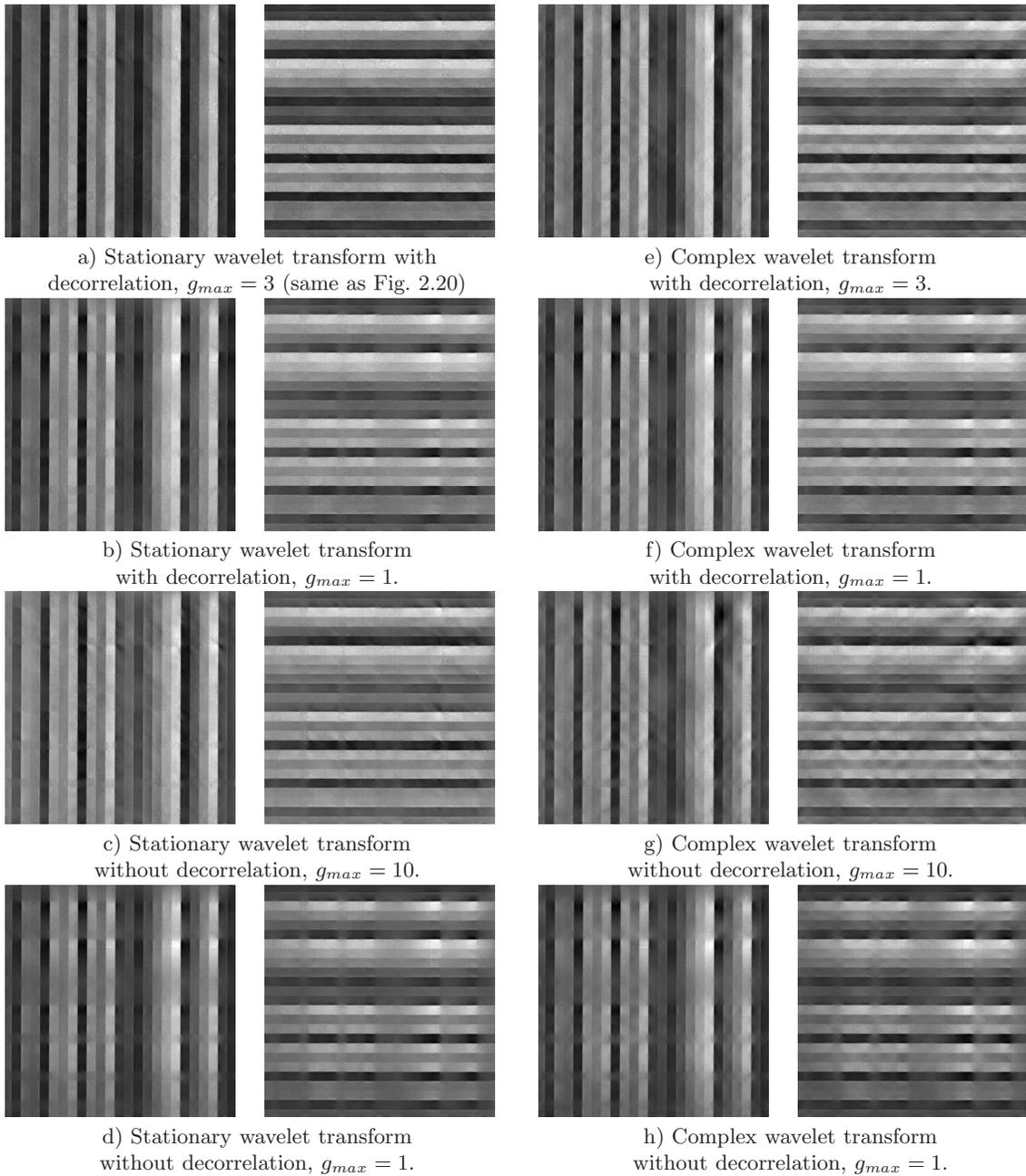


Figure 2.22: Effect of different variants of the method on the separation of the bars mixture.

The results for the air mail letter, presented in Fig. 2.23, were obtained using the stationary Haar wavelet transform with a 7-level decomposition, no decorrelation (as justified above) and no contrast compensation. Taking into account the complexity of this mixture, we consider the separation quality to be quite good. It is possible to read the separated letter more easily than the original one. As expected, our method performed better in areas of the letter where the lines of text of both sides were not aligned with one another, since, in those areas, the edges from the two sides coincide much less frequently. The zones where bleed-through appears to have occurred seem to have been well separated.

Batalla 21 de Julho de 1946 6  
 Rogério

Hoje é sábado, dia em que costumava vir aqui, e da do  
 quarto. Integramente por esta hora, eu tomava banho, preparava  
 roupa, e deixava por lá, hoje bati-me, preparei-me  
 e vim novamente. Como tem pouco tempo as coisas estão  
 mudando muito, mas um dia há de voltar ao pri-  
 mitivo.

Esta carta vai um pouco atrasada, porque eu já devia  
 ter te escrito há mais tempo; mas como calculei que devia  
 receber carta tua esta semana, tudo enfundei toda a semana  
 que me era para te responder a ela, mas afinal o meu  
 hoje. Por isso nunca mais escrevo, porque há de vir  
 com o atraso, e por isso quero ao mesmo tempo  
 com a tua a primeira carta que recebi, há 18 dias a cá  
 chegar, e esta que recebi hoje há 23 dias.

Como já disse que recebi a carta, quando recebi digas que  
 recebi também a tua para lembrar os dias. Então a apa-  
 recer lá porque receber estas cartas o culpado que tem em  
 receber tanta coisa, eu não o que gosto de receber, (isto  
 digo isto que a tua obrigado; porque assim eu não se hoje) mas já  
 te disse, e tocou a dizer, que só em alguns casos a que  
 me dá. Primeiramente a primeira carta só me dá, e  
 depois para receber mais ou por mim; depois eu recebo  
 que eu recebo mais dinheiro para eu receber todas as semanas  
 e até receber das tuas. Por isso já me esqueço de  
 te escrever cartas, mas não te esqueço de que me  
 escrevas cartas.

Já fico há comprando qualquer coisa, depois há a  
 minha casa. Não achas?

Agora é que vejo o que devia ser de aborrido, se só  
 receber carta de tuas a receber. Meus queridos amigos, quando  
 estou muito dias sem a receber, já me dá um nervosismo, que  
 não se entende mais. Se eu receber o homenagem que  
 invento o artigo, até lhe dou um abraço. Ela está. Teu  
 amigo, porque de contigo já não está.

Então o mesmo tem muitas coisas da tua direita?  
 Será tanto, como da tua de ti?

Por aí, continuam sempre a falar de nós, a corrigir de  
 nós; — se tu gostares de mim, dá — e a minha  
 forma indiferente tudo o que ligas, só o que faço —  
 Deus é que nos dá saúde, para que não pensemos muito  
 em como pensamos madamente.

Então o mesmo parece receber a tua direita? Já  
 há o que é que faz. Diz que me puses as orelhas  
 quando vias, mas eu é que não puses a tua; depois  
 estava de dizer que tu recebas. Depois entre outros  
 a mais calma, e não me dá assim as palavras.

Hoje estou muito contente, porque saber há que é  
 te digo, não. É muito bonito; há de perder esse costume.  
 Já agora devia ter recebido uma carta que te escrevi no  
 dia 20 de Junho, onde te mandava uma fotografia.  
 Então a dela? Devias ter gostado, pois eu a escrevi  
 da tua carta. É claro que eu julgo — me por mim  
 como gosto muito de tudo o que te dá, e que  
 não a recebo mais.

Então me trazes, não faz mais qual carta, e a tua  
 e a tua.

Figure 2.23: Manuscript letter after separation.

Regarding the partitures, we selected from each one a block of size  $1024 \times 1024$  containing the area where the bleed-through effect was stronger (the selected blocks are identified in Fig. 2.16). These blocks were processed by the proposed separation method, with no decorrelation and no contrast compensation. Fig. 2.24 shows the separation results obtained using a 7-level stationary transform with the Haar wavelet. Although the separation is not perfect, most of the bleed-through effect was removed. It became possible to read the transcription without the strong interference from the opposite page. For image areas where there was no strong bleed-through, our method behaved almost perfectly. This is, as far as we know, the first time that a bleed-through mixture of this kind was separated using the images from both sides of the document.

In a 1.6 MHz Pentium-M processor, images with  $512 \times 512$  pixels (from the first three tracing paper mixtures) took 3 seconds to be separated with the complex wavelet transform and 21 seconds with the stationary transform with Haar wavelet, both of them without contrast compensation. Images with  $1024 \times 1024$  pixels (from the last two pairs of tracing paper mixtures) took 16 seconds and 70 seconds, respectively, with the same transforms. The reconstruction of the stationary transform incorporated an optimization which was not compatible with the proposed contrast compensation mechanism. As consequence, the use of contrast compensation led, for results obtained using the stationary transform, to an increase of the separation time

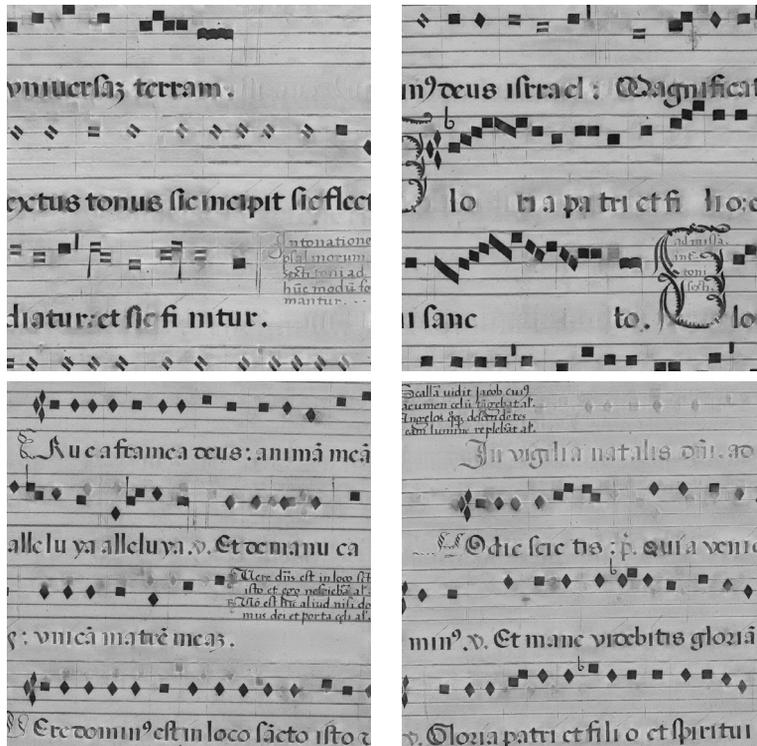


Figure 2.24: Partitures after separation.

(a total separation time of 40 seconds for images with  $512 \times 512$  pixels and of 120 seconds for images with  $1024 \times 1024$  pixels). In the case of the complex transform, contrast compensation did not cause any significant increase in computation time. We don't have separation times for images larger than  $1024 \times 1024$ . Those images (the ones of the manuscript letter) were processed in blocks of size  $1024 \times 1024$ , due to memory limitations.

### Subjective evaluation

Previous works [64, 66, 71] have used point-wise quality measures to objectively assess the performances of the various methods. Those measures are not suitable for evaluating the separation quality achieved with our method, which is strongly non-point-wise (note that the size of the Haar wavelet, at the 7<sup>th</sup> level, is  $128 \times 128$  pixels). There are image quality estimators that are perceptually oriented [148, 149, 152, 153]. These estimators normally use a linear combination of several perceptual measures, and are adjusted to fit mean opinion score results. Most of these perceptual estimators were developed for compression or denoising problems, focusing on distortions that are quite different from the ones found in this work. For these reasons we considered that they were not appropriate for our case, and that a formal subjective evaluation would provide the most reasonable indication of separation quality.

Before the contrast compensation mechanism had been implemented, we submitted the

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

Table 2.4: Mean rank, for different methods, of the separated images from the tracing paper mixtures. Lower ranks are better. For each pair, the best result is shown in bold. Standard deviations are given in brackets.

results of our method to an opinion evaluation. Since no contrast compensation was used, these opinions don't consider the best results shown in Figs. 2.20 and 2.21 but, instead, consider the ones that correspond to  $g_{max} = 1$  (published in [77]). Nine people (those, among the authors' close colleagues, who volunteered for the test, and none of them working on source separation or on image processing) were asked to order, according to separation quality, the images obtained with our method and with four other methods: best and worst images obtained in [64] with MISEP and a MLP separator, images separated with nonlinear DSS [66], and images separated with an inverse physical model trained with MISEP [71]. The evaluators numbered the images from 1 (best) to 5 (worst). They had access to source and mixture images, and their evaluation considered both the similarity with the corresponding source and the amount of interference from the opposite source. The mean rankings are shown in Table 2.4. As an indication of the confidence of these rankings, the table also presents, in brackets, the standard deviation computed over the nine rankings corresponding to each method and each pair of images.

The method proposed in this paper was chosen as best for the last three mixture pairs, and was among the three best for the other two pairs. Based on the improvement achieved with the introduction of contrast compensation (compare Fig. 2.22 with Fig. 2.20) we expect that, if contrast compensation had been used, the classification of the method would have been better still.

### 2.3.5 Conclusions

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

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

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

None of the existent objective quality measures for images seems appropriate for evaluating the results of the proposed method. We used a subjective opinion evaluation to assess the relative quality of our results. The evaluation of the results without contrast compensation was quite positive. Due to the practical complexity of the evaluation, it was not applied yet to the results obtained with contrast compensation. That evaluation will be addressed in future work.

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

### 2.3.6 Acknowledgment

We wish to acknowledge our colleague Rogério C. Pinto for kindly making available the old letter and the partiture images. We also acknowledge the use of the free package available at <http://taco.poly.edu/WaveletSoftware/> for computing complex wavelet transforms. This work was partially supported by the Portuguese FCT and by the "Programa Operacional

Sociedade do Conhecimento (POS-Conhecimento)”, comparticipated by the FEDER European Community fund, under the project POSC/EEA-CPS/61271/2004 and the PhD fellowship SFRH/BD/23919/2005.

## 2.4 Separation of nonlinear show-through image mixtures using a physical model trained with ICA

### Abstract

Often, when we scan a document, the image from the back page shows through, due to partial transparency of the paper, giving rise to a mixture of two images. We address the problem of separating these images through the use of a physical model of the mixture process. The model is nonlinear but invertible, and we use the inverse model to perform the separation. The model is trained through the MISEP technique of nonlinear ICA. Similarly to linear ICA, bounded independent sources can be recovered apart from scale and permutation indeterminacies.

We compare our results with results previously obtained with other approaches, namely with nonlinear denoising source separation and with the use of a generic MLP-based separator trained with MISEP. For the latter case we try both the previously used, symmetry-based regularization and the more recently proposed minimal nonlinear distortion regularization. Quantitative quality measures show that the approach that we propose is superior to the other methods.

**keywords:** Independent Component Analysis (ICA), nonlinear separation, image mixture, MISEP method, Minimal Nonlinear Distortion (MND).

### 2.4.1 Introduction

When we scan a document, it is common that the image of the back page shows through, due to partial transparency of the paper. This results in a nonlinear mixture of the images from both pages. This is the so-called *show-through* effect. It is desirable to be able to obtain the image from the front page without the interference from the back page.

If we scan the document on both sides, we'll obtain two mixture components, from which we would like to recover the two original page images. This is a nonlinear source separation problem. If the images on the two pages can be considered statistically independent from each other, one can address the problem through independent component analysis (ICA) techniques. However, nonlinear source separation through ICA is a severely ill-posed problem, since there is an infinite number of ICA solutions, and almost all of them yield non-separated sources [48–50]. In spite of this, nonlinear ICA has been studied by several authors [51–59] and applied to a few real-life separation problems. Besides show-through image separation problems such as

the one addressed in this paper, nonlinear ICA has been successfully used to perform nonlinear denoising [60]. It has also been applied to other real-life data [58, 61, 62], although, in these cases, it was not confirmed whether the extracted sources were actually close to the real ones or not.

A method that has been used to handle the indeterminacy of nonlinear ICA is the use of regularization: show-through mixtures can be assumed to be smooth, and, depending on the scanner that was used, may be symmetrical. While the corresponding regularization generally yields good results, it is known to have some shortcomings, the main one being that it doesn't always yield a good separation, especially if the independence assumption is not closely met by the images to be separated.

In this paper we describe a different method to perform the separation. We first develop a physical model of the mixture process. This model is invertible, and we use the inverse to perform separation. The model's parameters are estimated through the MISEP technique of nonlinear ICA. The model has just four parameters, and has only the same indeterminacies as linear ICA: scale and permutation. Therefore, by strongly constraining the type of allowed mixture in a way that is appropriate to the problem at hand, we are able to overcome the strong indeterminacies that normally exist in nonlinear ICA.

The same pairs of images presented in this paper have previously been separated using nonlinear ICA with a symmetric multilayer perceptron (MLP) as separator, with smoothness and symmetry-based regularization [64]. In this paper we also test a modification of that method: the use of the MLP-based separator with the recently proposed Minimal Nonlinear Distortion (MND) regularization strategy, which favors solutions corresponding to mixtures that are closer to linear. The same image pairs have also been separated using nonlinear denoising source separation [66]. According to objective quality measures presented ahead, the separation quality achieved with the model-based solution proposed in this paper is better than the one obtained using those other techniques.

Another separation method that has been proposed [32, 77] is not based on ICA, but rather on the observation that the high frequency features of images normally are sparse and can be assigned, in a relatively simple manner, to the appropriate source image. That method yields a separation that is excellent in perceptual quality and that can also handle other situations, such as *bleed-through*, in which the image superposition is due to bleeding of ink through the paper, and not to paper transparency. However, the separation quality of that method cannot be easily assessed through quantitative means, because the separation is not point-wise. Furthermore, if the assumption of sparse high frequency features is not met, the method cannot perform a

good separation.

The method that we describe here was first proposed in a conference paper [71]. The physical model that is used had been developed earlier [76], but separation was then performed only by manual adjustment of the model parameters. At about the same time, the same mixture model was proposed by Stefan Harmeling (private communication), based only on the observation of mixture data, without any physical considerations. In [72] a similar model was proposed, again based only on observation of mixture data, without a physical basis, and was only tested on text images involving much weaker mixtures than the ones considered here. Other separation models have been developed for the see-through problem [2, 6, 22]. In those approaches, the parameter selection and image separation were not performed through nonlinear ICA, and those models were also only tested on mixtures whose strength was considerably lower than what we consider in this paper.

Show-through and/or bleed-through mixtures have also been addressed by other methods. In some of those works [1, 63, 65, 154] the separation is achieved through linear models, which have been shown in [64] not to be adequate for strong mixtures. Regarding nonlinear techniques, the separation is not typically achieved through ICA. In [3, 5, 6, 29, 31], the separation is based based on histogram information and/or segmentation techniques. Most of these methods were specially developed for text documents, for which linear separation yields relatively good results (see [64]). In [4], self-organizing maps were used in an attempt to compensate for the nonlinearity of the separation map. The technique proposed in [74] only requires the front page image to perform the separation, but only deals with weak mixtures of color images.

This paper is organized as follows: Section 2.4.2 briefly describes the MISEP method of nonlinear ICA. Section 2.4.3 introduces the physical model used to represent the mixture, and derives the inverse model to be used as separator. In Section 2.4.4, we describe the tracing paper mixtures to be used in the experimental tests. Section 2.4.5 describes the quality measures that will be used to assess the experimental results, which are shown in Section 2.4.6. Section 2.4.7 concludes.

## 2.4.2 The MISEP method

The MISEP method [55], whose name stands for Mutual Information-based SEParation, is an extension of the well known INFOMAX technique of linear ICA. INFOMAX performs ICA by maximizing the Shannon entropy of the output  $\mathbf{z}$  of the network depicted in Fig. 2.25:

$$H(\mathbf{z}) = - \int p_{\mathbf{z}}(\mathbf{u}) \log p_{\mathbf{z}}(\mathbf{u}) d\mathbf{u}, \quad (2.20)$$

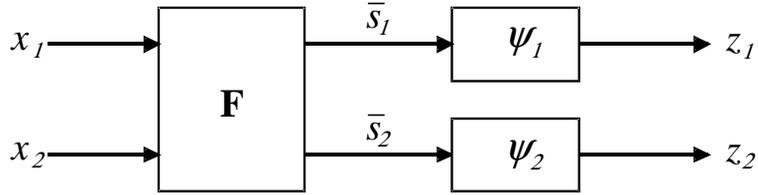


Figure 2.25: Network structure used in INFOMAX and in MISEP. In INFOMAX,  $\mathbf{F}$  is an adaptive linear block and the  $\psi_i$  are fixed a priori. In MISEP,  $\mathbf{F}$  can be nonlinear and both  $\mathbf{F}$  and  $\psi_i$  are adaptive.

where  $p_{\mathbf{z}}(\cdot)$  is the probability density function (PDF) of the vector-valued output random variable  $\mathbf{z}$ . The separation is performed by the  $\mathbf{F}$  block, whose outputs  $\bar{s}_i$  are, at the end of the training, the separated components. The auxiliary blocks  $\psi_i$  implement monotonically increasing transformations whose co-domain is the interval  $[0, 1]$ , and which should equal the cumulative distribution functions (CDFs) of the corresponding  $\bar{s}_i$ . If the  $\psi_i$  blocks do implement the CDFs of these components, all the  $z_i$  will have uniform distributions in  $[0, 1]$ , and thus  $H(z_i) = 0$ . In this case,

$$\begin{aligned}
 H(\mathbf{z}) &= \sum H(z_i) - I(\mathbf{z}) \\
 &= -I(\mathbf{z}) \\
 &= -I(\bar{\mathbf{s}}),
 \end{aligned}$$

where  $I(\bar{\mathbf{s}})$  is the mutual information of the components of  $\bar{\mathbf{s}}$  (see [55] for further details). Therefore, maximizing the entropy of the network output  $\mathbf{z}$  will correspond to minimizing the mutual information of the components  $\bar{s}_i$ . Minimization of this mutual information will yield maximally independent components.

Due to the small number of parameters under estimation in linear source separation, INFOMAX is able to achieve good results even if the  $\psi_i$  blocks implement only a crude approximation of the CDFs of the components. In contrast, nonlinear separation, to be described ahead, generally requires the  $\psi_i$  blocks to yield a fairly good estimate of the CDFs.

MISEP uses the same basic structure of Fig. 2.25, extending it in two directions. The  $\mathbf{F}$  block in the MISEP method can implement a nonlinear function, allowing the separation of nonlinear mixtures. On the other hand, in the MISEP technique, the  $\psi_i$  blocks are adaptive, which allows them to fit the CDFs of the data. As mentioned above, a good estimate of the CDFs of the sources is crucial to achieve nonlinear ICA. By maximizing the entropy of  $\mathbf{z}$ , MISEP

learns the separation map in the  $\mathbf{F}$  block, as explained for INFOMAX, and simultaneously learns the CDFs of the separated components in the  $\psi_i$  blocks. To see why, note that the maximum of  $H(z_i)$  is reached when  $z_i$  has a uniform distribution, which occurs when each  $\psi_i$  implements the CDF of the corresponding separated component. For details see [55].

As noted above, despite performing nonlinear ICA, MISEP alone cannot guarantee that the extracted components will be the original sources, due to the basic ill-posedness of nonlinear ICA. Often, regularization and/or constraints on the separation block  $\mathbf{F}$  are imposed to overcome this indeterminacy.

MISEP can use essentially any parameterized function in the  $\mathbf{F}$  block. In the method proposed in this paper, the separator block ( $\mathbf{F}$ ) implements the inverse of the physical model of the mixture, which is developed in Section 2.4.3. For comparison, we also implement a separation through MISEP using, as the  $\mathbf{F}$  block, a multilayer perceptron (MLP) with a few different forms of regularization. More details about MISEP can be found in [55, 155].

## Entropy maximization

Both INFOMAX and MISEP perform ICA by maximizing the Shannon entropy of the output of the network depicted in Fig. 2.25. If we have the same number of outputs (or sources), as inputs (mixture components) and the map  $\mathbf{z} = \Psi[\mathbf{F}(\mathbf{x})]$  is invertible, the output entropy (2.20) can be written

$$H(\mathbf{z}) = - \int p_{\mathbf{x}}(\mathbf{u}) \log [p_{\mathbf{x}}(\mathbf{u})/|\det\mathbf{J}|] d\mathbf{u} \quad (2.21)$$

$$= H(\mathbf{x}) + \langle \log|\det\mathbf{J}| \rangle, \quad (2.22)$$

in which  $\langle \cdot \rangle$  denotes the expected value and

$$\mathbf{J} = \frac{\partial \mathbf{z}}{\partial \mathbf{x}} \quad (2.23)$$

is the Jacobian of the transformation  $\Psi[\mathbf{F}(\cdot)]$ . Considering (2.22), and taking into account that  $H(\mathbf{x})$  doesn't depend on  $\mathbf{F}$  or  $\Psi$ , we see that we can maximize the entropy of  $H(\mathbf{z})$  by minimizing the cost function

$$C = \frac{1}{K} \sum_{k=1}^K \log|\det\mathbf{J}_k| \approx \langle \log|\det\mathbf{J}| \rangle, \quad (2.24)$$

in which  $\mathbf{J}_k$  is the value of  $\mathbf{J}$  for the  $k^{\text{th}}$  training pattern, and  $K$  is the number of training patterns.

### 2.4.3 Physical model of the mixture process

The physical model takes into account the halftoning process used by printers. Halftoning consists of printing tiny black dots which create different gray levels according to the percentage of area covered by the dots. At the scanning resolution used in our dataset (100 dpi) each pixel covers many black dots, and the pixel's intensity depends on the percentage of area covered by the dots within the area of the pixel.

In this model each point of a page is represented by a binary random variable  $\hat{s} \in \{0, 1\}$  which, at the halftoning resolution, indicates whether that point is black (value 0) or white (value 1). The intensity of a pixel is given by the expected value of the random variable  $\hat{s}$  at the pixel location,

$$s = \langle \hat{s} \rangle = P(\hat{s} = 1). \quad (2.25)$$

We have two binary random variables  $\hat{s}_1$  and  $\hat{s}_2$ , for pages 1 and 2 respectively. The intensity of each pixel of each page is then given by:

$$\begin{aligned} s_1 &= P(\hat{s}_1 = 1) \\ s_2 &= P(\hat{s}_2 = 1). \end{aligned} \quad (2.26)$$

At a high resolution, enough to discriminate the halftoning dots, the mixing process yields four possible combinations of the variables  $\hat{s}_1$  and  $\hat{s}_2$ . When the document is observed, say, from side number 1, these four combinations will result in four different acquired intensities (still at the high resolution),

$$\hat{x}_1 = \begin{cases} l_1 & \text{if } \hat{s}_1 = 0 \text{ and } \hat{s}_2 = 0 \\ l_2 & \text{if } \hat{s}_1 = 0 \text{ and } \hat{s}_2 = 1 \\ l_3 & \text{if } \hat{s}_1 = 1 \text{ and } \hat{s}_2 = 0 \\ l_4 & \text{if } \hat{s}_1 = 1 \text{ and } \hat{s}_2 = 1. \end{cases} \quad (2.27)$$

The mixing process is fully determined by the four parameters  $l_1, l_2, l_3, l_4$ , whose values depend on the characteristics of the material that was used (paper and scanner). This model does not account for lateral diffusion of light within the paper, which is a reasonable approximation at the scanning the resolution that was used (100 dpi).

Still assuming that the page is being observed from side 1, the intensity of each pixel at the scanning resolution,  $x_1$ , is given by

$$\begin{aligned} x_1 &= \langle \hat{x}_1 \rangle \\ &= l_1 P(\hat{s}_1 = 0, \hat{s}_2 = 0) + l_2 P(\hat{s}_1 = 0, \hat{s}_2 = 1) + \\ &\quad l_3 P(\hat{s}_1 = 1, \hat{s}_2 = 0) + l_4 P(\hat{s}_1 = 1, \hat{s}_2 = 1). \end{aligned}$$

Assuming that  $s_1$  and  $s_2$  are mutually independent, we have

$$x_1 = l_1(1 - s_1)(1 - s_2) + l_2(1 - s_1)s_2 + l_3s_1(1 - s_2) + l_4s_1s_2. \quad (2.28)$$

If both sides of the paper are processed in the same way, we obtain a similar equation for the intensity of the pixels of mixture  $x_2$ :

$$x_2 = l_1(1 - s_1)(1 - s_2) + l_2s_1(1 - s_2) + l_3(1 - s_1)s_2 + l_4s_1s_2. \quad (2.29)$$

Defining the parameters

$$\begin{aligned} \alpha &= l_3 - l_1 \\ \beta &= l_2 - l_1 \\ \gamma &= l_4 + l_1 - l_2 - l_3 \\ \delta &= l_1 \end{aligned} \quad (2.30)$$

and substituting them in (2.28) and (2.29) we obtain the bi-affine<sup>4</sup> mixture

$$\begin{aligned} x_1 &= \alpha s_1 + \beta s_2 + \gamma s_1 s_2 + \delta \\ x_2 &= \alpha s_2 + \beta s_1 + \gamma s_1 s_2 + \delta. \end{aligned} \quad (2.31)$$

For the model to be invertible, it is required that  $\alpha \neq \beta$ , or equivalently, that  $l_2 \neq l_3$  (otherwise we would have  $x_1 = x_2$ ). The condition  $l_2 \neq l_3$  is always true if the paper is not fully transparent.

### Separation model

To perform separation, we use the inverse of the mixing model (2.31) in block **F** of Fig. 2.25. The mixing model has four parameters, and therefore its inverse also does. These four parameters are to be estimated by means of the MISEP method.

---

<sup>4</sup>Affine in each of the independent variables,  $s_1$  and  $s_2$ .

To obtain the explicit form of the inverse model, we first subtract the second equation of (2.31) from the first,

$$s_2 = s_1 + (x_2 - x_1)/(\alpha - \beta). \quad (2.32)$$

Substituting (2.32) in the first equation of (2.31), we obtain a quadratic equation for  $s_1$ ,

$$\gamma s_1^2 + \left[ \alpha + \beta + \frac{\gamma(x_2 - x_1)}{\alpha - \beta} \right] s_1 - x_2 + \delta + \frac{\alpha(x_2 - x_1)}{\alpha - \beta} = 0. \quad (2.33)$$

If we define

$$\begin{aligned} a &= \gamma \\ b &= \alpha + \beta + \frac{\gamma(x_2 - x_1)}{\alpha - \beta} \\ c &= -x_2 + \delta + \frac{\alpha(x_2 - x_1)}{\alpha - \beta}, \end{aligned} \quad (2.34)$$

source  $s_1$  will be given by

$$s_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}. \quad (2.35)$$

It is possible to show that for images with positive-intensity pixels, such as the ones used in this work, the positive sign in the square root of Eq. (2.35) is the one that leads to the original source. The other source could be obtained from an equation similar to (2.35). However, once  $s_1$  is known, it is easier to use (2.32). The relationship (2.32) and the auxiliary variables  $a, b, c$  help to simplify the inverse model and also to simplify the computation of the derivatives that are required in the MISEP method.

## Model separability

In this Section we shall show that, for bounded sources, any bi-affine mixture that is non-singular (i.e., that is invertible) is separable through the independence criterion, apart from a possible permutation and from indeterminations of scale and additive offset. The proof will be given for two sources for simplicity, but is easily extendable to any number of sources.

Consider the convex hull of the support of the joint distribution of the two sources. If these are bounded and independent, that hull will be a rectangle with its sides parallel to the coordinate axes. Figure 2.26-a) depicts such a rectangle, and shows how we identify a generic point,  $X$ , of the rectangle. We use two parameters,  $\lambda_1, \lambda_2 \in [0, 1]$ , which indicate the fractions of the total width and height of the rectangle, respectively, at the point where  $X$  is located. Considering points as two-dimensional vectors, we can write

$$X = \lambda_2[\lambda_1 A + (1 - \lambda_1)B] + (1 - \lambda_2)[\lambda_1 C + (1 - \lambda_1)D]. \quad (2.36)$$

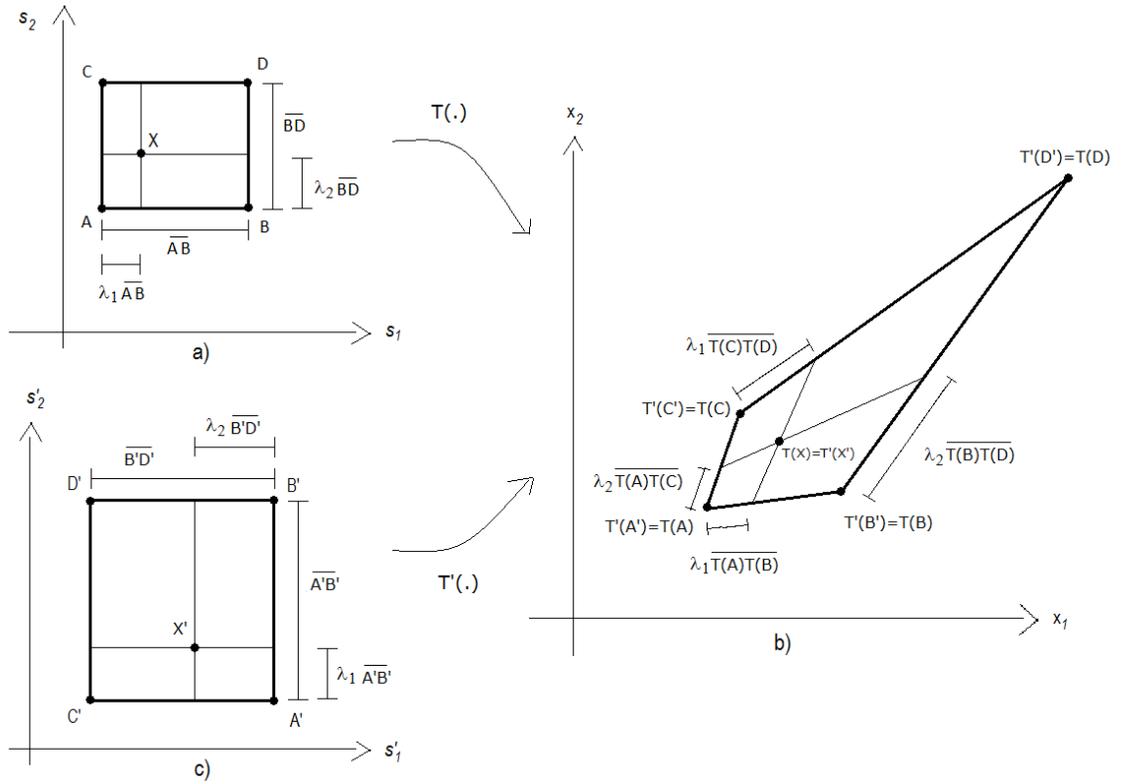


Figure 2.26: Geometric scheme of the bilinear transformation of rectangles (contiguous sources).

Figure 2.26-b) shows, schematically, the result of the bi-affine transformation of the rectangle  $ABCD$ . Denoting the transformation by  $T$  and recalling that it is affine along the directions of both coordinate axes, we can say that, if two points  $U$  and  $V$  are aligned parallel to one of the coordinate axes, then

$$T[\lambda U + (1 - \lambda)V] = \lambda T(U) + (1 - \lambda)T(V). \quad (2.37)$$

Applying this property along each of the coordinate directions, we have

$$\begin{aligned} T(X) &= T\{\lambda_2[\lambda_1 A + (1 - \lambda_1)B] + (1 - \lambda_2)[\lambda_1 C + (1 - \lambda_1)D]\} \\ &= \lambda_2 T[\lambda_1 A + (1 - \lambda_1)B] + (1 - \lambda_2) T[\lambda_1 C + (1 - \lambda_1)D] \\ &= \lambda_2 [\lambda_1 T(A) + (1 - \lambda_1)T(B)] + (1 - \lambda_2) [\lambda_1 T(C) + (1 - \lambda_1)T(D)]. \end{aligned} \quad (2.38)$$

Consider a candidate separating transformation,  $T'^{-1}$ , of the mixture, such that  $T'$  also is bi-affine, and that the result of the transformation  $T'^{-1}$  of the mixture has components that are mutually independent. Figure 2.26-c) shows, schematically, the convex hull of the support of the resulting distribution. This convex hull will again be a rectangle with its sides parallel

to the coordinate axes. The orientation of the rectangle can, however, be different from the orientation of the original rectangle in the source space. The change in orientation can involve translation, rotation by a multiple of 90° and possibly flipping. In the figure we have marked  $A' = T'^{-1}[T(A)]$ , and similarly for the other vertexes. Of course,  $A'$  and  $B'$  must be on the same side of the rectangle, and similarly for  $(A', C')$ ,  $(B', D')$  and  $(C', D')$ .

Let us now consider the point  $X'$  of this rectangle that corresponds, in terms of fractions of length in both directions, to  $X$  in the source rectangle:

$$X' = \lambda_2[\lambda_1 A' + (1 - \lambda_1) B'] + (1 - \lambda_2)[\lambda_1 C' + (1 - \lambda_1) D']. \quad (2.39)$$

Using the same reasoning as above, we have

$$T'(X') = \lambda_2[\lambda_1 T(A) + (1 - \lambda_1) T(B)] + (1 - \lambda_2)[\lambda_1 T(C) + (1 - \lambda_1) T(D)], \quad (2.40)$$

Therefore,  $T'(X') = T(X)$  and  $X' = T'^{-1}[T(X)]$ , i.e.,  $X'$  is the point that results from  $X$  through the mixture followed by separation. Since  $X'$  is the point of the rectangle  $A'B'C'D'$  that corresponds to  $X$  in terms of fractions of length, the mapping from  $X$  to  $X'$  can only involve scaling and translation of each of the coordinates, and possibly also permutation of the two coordinates. Therefore, the output components must be related to the sources through scaling, additive offset and a possible permutation.

## 2.4.4 Experimental setup

### Mixtures

We used, in our tests, the same mixtures that were used in [64], which we briefly describe here. The mixtures were obtained by printing images on both sides of sheets of tracing paper, and then scanning those sheets with a common desktop scanner. Five different pairs of source images (including synthetic bars, photos and text) were used. They are shown in Figs. 2.27 and 2.28, and have different properties:

1. The first pair is formed by synthetic images. The first one contains 25 randomly ordered vertical bars with intensities uniformly distributed between white and black. The second source is simply the first one rotated by 90°. In this pair, by construction, the intensity distribution of each source is close to uniform and the two source images are independent.
2. The second pair consists of two natural scenes, containing tiny details. The two source images are approximately independent.

3. The third pair consists of two natural scenes which have large areas with approximately constant intensities. This often happens in natural images, and, in this case, causes the pair of sources not to follow the independence assumption as closely as the other pairs [64, 66].
4. The fourth pair consists of a natural scene and an image of printed text. While the natural image presents a wide variety of intensity values, the text image is, approximately, a binary image with just black and white. The images are approximately independent.
5. The fifth pair consists of two text images. Both source images are essentially binary. Some source dependence exists due to the alignment of the lines on the two sides of the paper which occurs in several places because lines are spaced equally in both images.

These five pairs of images were printed on opposite pages of sheets of tracing paper, which was chosen for its high transparency. Image printing and scanning were performed as symmetrically as possible, regarding the two images of each pair. Printing was performed with a 1200 dpi laser printer, using the printer's default halftoning system. Both pages were scanned with a desktop scanner at a low resolution of 100 dpi, which was chosen so that the printer's halftoning dots would not be apparent in the scanned images. Figures 2.29 and 2.30 show the acquired images after the alignment procedure that is described ahead (Section 2.4.4).

This set of image mixtures was intentionally created for testing separation methods, and has some important characteristics:

- The mixtures are among the strongest that can be found in show-through situations.
- The set contains degradations (noise, misalignments, etc.) that are typical of real-life scenarios.
- The set includes source images of a wide variety of types.

From a more theoretical perspective, the set also has some interesting properties:

- The mixtures are highly nonlinear, and, except for the pair of text images, linear separation does not lead to good results (see [64])
- Some image pairs (#3 and #5) are not completely independent.
- We have access to the original sources, which allows the computation of objective measures of separation quality.

## Image alignment

Image separation methods usually require the components of the mixture to be precisely registered with one another. To achieve this registration, one of the images of each pair had first to be horizontally flipped. After that, an alignment procedure was applied, to correct misalignments due to the different positions of the paper during the two scanning operations. In order to be precise, the alignment had to be performed locally, probably due to some geometrical imperfections of the scanner.

All mixture pairs were subjected to an initial crude alignment, which was manually performed with the aid of an image editing program. After this manual alignment, the mixture images were subjected to a local alignment procedure which was specially developed for this problem. The images were first expanded in resolution by a factor of four in each direction, using bicubic interpolation. Then, one of the images of the pair was divided into  $100 \times 100$  pixel squares (corresponding to  $25 \times 25$  pixels in the original image) and, for each square, the best displacement was found, based on the maximum of the cross-correlation with the other image. Aligned images were then brought back to the original resolution. Therefore this procedure performed a local alignment with a resolution of  $1/4$  pixel.

More information about the tracing paper mixtures and the alignment procedure is available in [64]. The routines for performing image alignment are available at <http://www.lx.it.pt/~lbalmeida/ica/seethrough/>.

## Minimal Nonlinear Distortion (MND)

The Minimal Nonlinear Distortion (MND) principle [58, 156] is a regularization technique whose underlying idea is to favor maps that are closer to linear, in order to remove the ill-posedness of nonlinear ICA. Among the infinite number of solutions of nonlinear ICA, MND, as described in the above references, favors the one that corresponds to a mixture that is closest to linear. It starts with a linear mixing map and progressively allows the map to learn nonlinearities. In the experimental tests described ahead, we used the MND implementation described in [58].

### 2.4.5 Separation measures

Objective measures of separation quality have the advantage, over visual evaluation, of not depending on the viewer nor on the inspection conditions. Objective measures presented in this section compare the separated components with the original sources. Therefore, they can only be computed for situations in which one has access to the source images, as in the present

case.

The three quality measures presented ahead were initially defined for to the see-through problem in [64], and were later used in several other works [32, 66, 67, 71, 77]. Denoting the original source by  $s_i$  and the corresponding separated source by  $\bar{s}_i$ , the three objective separation measures are the following:

- $Q_1$  – SNR between the original and the estimated source image, compensated for possible differences in brightness and contrast:

$$Q_1(s_i, \bar{s}_i) = 10 \log_{10} \frac{\langle (s_i - \langle s_i \rangle)^2 \rangle}{\langle [n_1(s_i, \bar{s}_i) - \langle n_1(s_i, \bar{s}_i) \rangle]^2 \rangle}, \quad (2.41)$$

in which  $n_1$  is the error between  $s_i$  and  $\bar{s}_i$  computed after an affine transformation of  $\bar{s}_i$ ,

$$n_1(s_i, \bar{s}_i) = s_i - (a\bar{s}_i + b), \quad (2.42)$$

in which  $a$  and  $b$  are the parameters that minimize  $\langle (n_1)^2 \rangle$ .

- $Q_2$  – SNR between the original and the estimated source image, compensated for a possible nonlinear transformation of the intensity scale of the estimated source:

$$Q_2(s_i, \bar{s}_i) = \frac{\langle (s_i - \langle s_i \rangle)^2 \rangle}{\langle [n_2(s_i, \bar{s}_i) - \langle n_2(s_i, \bar{s}_i) \rangle]^2 \rangle}, \quad (2.43)$$

in which  $n_2$  is the error between  $s_i$  and  $\bar{s}_i$  computed after a nonlinear transformation of  $\bar{s}_i$ ,

$$n_2(s_i, \bar{s}_i) = s_i - l(\bar{s}_i), \quad (2.44)$$

in which  $l(\cdot)$  is the nonlinear monotonic function that minimizes  $\langle (n_1)^2 \rangle$ .

- $Q_3$  – Mutual information between the source image and the estimated one:

$$Q_3(s_i, \bar{s}_i) = I(s_i, \bar{s}_i). \quad (2.45)$$

For computing  $Q_3$ , the mutual information  $I(\cdot, \cdot)$  was estimated from a set of 5000 randomly selected pixel pairs, and was computed using the  $I(1)$  estimator described in [144], with  $k = 3$ .

While  $Q_1$  is insensitive only to changes in the brightness and contrast of the estimated image,  $Q_2$  and  $Q_3$  are insensitive to all monotonic transformations of the image's intensity scale. For

all three measures, higher values, are better. In [64] a fourth quality measure was also used. However, it was found to be less significant than the three measures indicated above, and therefore we didn't use it in this work.

## 2.4.6 Experimental results

In this Section we describe the tests that were performed to assess the performance of the proposed separation method. The method was compared to the one presented in [64], which uses an MLP with symmetry constraints as a separator, and also to the use of that same MLP-based separator with other forms of regularization, as described ahead.

### Model-based separation

The model-based separation was implemented using the inverse of the physical model as separator block  $\mathbf{F}$  and an MLP with 20 hidden units in each of the  $\psi_i$  blocks (see Fig. 2.25). For each pair of mixture images, the system was trained 10 times, using, as training set, 5000 randomly selected pairs of pixels, with a different training set for each of the 10 runs. The inverse model was initialized near the identity function ( $\alpha = 1$ ,  $\beta = 0.01$ ,  $\gamma = 0.001$ ,  $\delta = 0.001$ , which is equivalent to  $l_1 = 0.001$ ,  $l_2 = 0.011$ ,  $l_3 = 1.001$  and  $l_4 = 1.012$ ) and the network was trained for 500 epochs.

For each pair of mixtures and each of the training runs, the quality measures described in Section 2.4.5 were computed. Table 2.5 shows the means of the 10 values for each pair of mixtures, averaged over the two estimated source images. The table also shows the standard deviations of the quality measures. These standard deviations are relatively low, showing that the method is rather stable with a training set of 5000 pixels. In a previous work [71], the physical model was trained using only 1000 training points, and the results that were obtained were consistent with the ones presented here.

Table 2.6 shows the model parameters obtained with each of the five pairs of mixtures, averaged over the 10 tests. Although the five mixtures have been generated using essentially the same system, the five sets of estimated parameters are not the same, suggesting that they are somewhat specialized for the images they were estimated from. The parameters generated by the first three pairs of mixtures (which don't include any text image) are relatively similar to one another, while the last two are the most different ones. This indicates a stronger specialization of the last two models, probably due to the small diversity of gray levels of the text images.

Figures 2.31 and 2.32 show examples of separation results obtained with the proposed

Image pair	$Q_1$	$Q_2$	$Q_3$
1	<b>13.36</b> (0.07)	<b>15.09</b> (0.03)	<b>2.530</b> (0.009)
2	<b>12.97</b> (0.06)	<b>13.66</b> (0.05)	<b>1.974</b> (0.005)
3	<b>11.27</b> (0.05)	<b>12.19</b> (0.05)	<b>1.860</b> (0.008)
4	<b>9.16</b> (0.03)	<b>11.73</b> (0.02)	<b>1.561</b> (0.006)
5	<b>3.80</b> (0.08)	<b>5.94</b> (0.13)	<b>0.682</b> (0.009)
Mean	<b>10.11</b> (0.03)	<b>11.72</b> (0.04)	<b>1.721</b> (0.003)

Table 2.5: Separation measures averaged, for each image pair, across the two separated images and the 10 randomly initialized experiments. Standard deviations are shown in parentheses.

Image pair	$l_1$	$l_2$	$l_3$	$l_4$
1	<b>0.080</b> (0.009)	<b>0.258</b> (0.038)	<b>0.478</b> (0.056)	<b>1.751</b> (0.296)
2	<b>0.049</b> (0.006)	<b>0.187</b> (0.016)	<b>0.414</b> (0.023)	<b>1.468</b> (0.122)
3	<b>0.159</b> (0.020)	<b>0.360</b> (0.035)	<b>0.608</b> (0.055)	<b>1.572</b> (0.300)
4	<b>0.202</b> (0.012)	<b>0.600</b> (0.038)	<b>1.842</b> (0.051)	<b>2.742</b> (0.281)
5	<b>0.185</b> (0.070)	<b>0.759</b> (0.068)	<b>0.830</b> (0.074)	<b>1.020</b> (0.059)

Table 2.6: Model parameters obtained for the five mixtures, averaged over the 10 random experiments. Standard deviations are given in parentheses.

approach. These images are representative of the average performance of the model-based separation, since the values of their quality measures are close to the average values listed in Table 2.5. We see that the proposed approach led to good results, in which the interference of the opposite source is strongly reduced.

### Comparison with MLP-based separation

In this Section we compare the results of the proposed model-based separation with those obtained using an MLP as separator. For the latter, we test a few regularization schemes:

- No regularization, beyond the one that is implicitly performed by the MLP (denoted “No Reg”, ahead).
- Symmetry constraints and regularization as described in [57] (denoted “Sym”, ahead).

- Regularization through Maximum Nonlinear Distortion (MND), as described in [58] (denoted “MND”, ahead).
- Combination of the two previous systems: symmetry constraints and MND regularization (denoted “Sym-MND”, ahead)

In all of these situations, the  $\mathbf{F}$  block of Fig. 2.25 was implemented using an MLP with 20 hidden units and with linear output units. Of the 20 hidden units, 10 were connected to one of the output units, and the other 10 to the other output unit. “Shortcut” connections between input and output units were also used. The  $\psi_i$  blocks were implemented with MLPs with 20 hidden units each, as in the model-based separator. For each of the four types of regularization and for each pair of images, the system was trained during 500 epochs using each of the 10 training sets that had been used with the model-based separator. As in [57], during the first 100 epochs, the separator MLP was constrained to be linear, by keeping the weights of the hidden units set to zero. In the two settings involving MND regularization, the regularizing parameter  $\lambda$  was set to 10 during the first 150 epochs, and was then reduced according to the schedule shown in Fig. 2.33. This logarithmic schedule showed to be slightly better, in our problem, than the linear schedule proposed by MND’s authors.

The three quality measures  $Q_1 - Q_3$  were computed for all separated images. Figures 2.34, 2.35 and 2.36 show box plots of the results, encompassing all image pairs and all initializations. Table 2.7 gives numerical values. The model-based approach was the best in all situations, both in terms of the quality values and of their variability: the method yielded the highest values of the separation measures and, simultaneously, was the most robust regarding variations in the training set. In what concerns the MLP-based separators, the advantage of imposing symmetrical separation constraints was also clear, since it led to higher values of separation measures, with a lower dispersion. The use of regularization based on minimum nonlinear distortion was somewhat advantageous when no symmetry constraints were used, but was disadvantageous when those constraints were used.

A more detailed set of values is given in Table 2.8, in the Appendix. In those values, we see that the proposed approach was the best for all images pairs except the fifth. That was the pair composed of two text images, which can be reasonably separated with a linear system [64].

Quality measure	No Reg	MND	Sym	Sym-MND	Model
$Q_1$ [dB]	6.80 (4.069)	8.18 (4.70)	9.49 (1.22)	9.30 (1.31)	<b>10.11</b> <b>(0.24)</b>
$Q_2$ [dB]	9.24 (2.85)	10.36 (4.07)	11.41 (1.33)	11.24 (1.02)	<b>11.82</b> <b>(0.75)</b>
$Q_3$ [bit]	1.41 (0.34)	1.53 (0.45)	1.64 (0.15)	1.63 (0.12)	<b>1.73</b> <b>(0.03)</b>

Table 2.7: Quality measures averaged across the 10 random runs and the five pairs of images. The standard deviation across the 10 random initializations is shown in brackets. For each quality measure, the best method is shown in bold.

### 2.4.7 Conclusions

We have proposed a method for the separation of images affected by show-through degradations. The method is based on the use of a physical model of the degradation process, the inverse of that model being used for separation. The model is nonlinear, and its parameters are estimated through an ICA criterion, using the MISEP technique. The number of parameters is just four, which allows the method to have only the same kinds of indeterminations that exist in linear ICA (scale and permutation), despite the fact that the mixture is nonlinear.

The proposed approach was compared with separation based on an MLP with a few different kinds of regularizations. The results showed that the model-based approach improved over the MLP-based one in almost all cases. We conclude that MISEP is a useful method for estimating the parameters of nonlinear separation models, and that the proposed physical model fits the mixture process reasonably well.

In spite of the good results that were obtained, there is still room for improvement in the method. For example, not all scanners treat both faces of the sheet of paper in the same way. Such scanners would require a non-symmetric model. Some scanners (or the associated software) introduce nonlinearities such as gamma correction, which could be accounted for in the model. And for higher-resolution scans, the model would also need to incorporate lateral diffusion of light within the paper.

### 2.4.8 Acknowledgments

We would like to thank Kun Zhang and Laiwan Chan for providing us the MatLab code of the MND-based regularization proposed in [58].

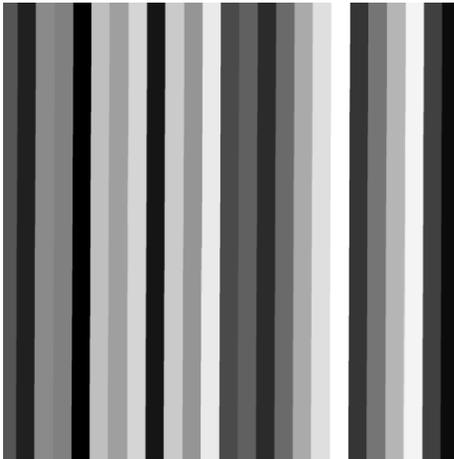
This work was partially supported by the Portuguese FCT and by the “Programa Operacional Sociedade do Conhecimento (POS-Conhecimento)”, comparticipated by the FEDER

Image pair	Quality measure	No Reg	MND	Sym	Sym-MND	Model
1	$Q_1$ [dB]	2.93 (4.31)	8.12 (4.09)	<b>13.64</b> (0.45)	12.94 (1.24)	13.46 <b>(0.07)</b>
	$Q_2$ [dB]	6.05 (3.27)	9.90 (3.36)	14.43 (0.54)	14.05 (0.71)	<b>15.09</b> <b>(0.03)</b>
	$Q_3$ [dB]	1.41 (0.34)	1.86 (0.40)	2.417 (0.079)	2.363 (0.104)	<b>2.531</b> <b>(0.009)</b>
2	$Q_1$ [dB]	8.43 (1.59)	9.90 (1.05)	10.88 (0.57)	11.05 (0.44)	<b>12.97</b> <b>(0.06)</b>
	$Q_2$ [dB]	10.62 (1.46)	11.82 (1.06)	12.73 (0.51)	12.82 (0.18)	<b>13.66</b> <b>(0.05)</b>
	$Q_3$ [bit]	1.628 (0.131)	1.734 (0.097)	1.817 (0.052)	1.832 (0.032)	<b>1.973</b> <b>(0.004)</b>
3	$Q_1$ [dB]	9.96 (0.75)	9.92 (0.21)	10.27 (0.68)	10.22 (0.23)	<b>11.27</b> <b>(0.05)</b>
	$Q_2$ [dB]	11.28 (0.66)	11.53 (0.13)	11.49 (0.62)	11.42 (0.13)	<b>12.18</b> <b>(0.04)</b>
	$Q_3$ [dB]	1.754 (0.062)	1.773 (0.022)	1.783 (0.063)	1.771 (0.019)	<b>1.860</b> <b>(0.009)</b>
4	$Q_1$ [dB]	8.52 (0.70)	8.86 (0.66)	8.34 (0.38)	7.96 (0.38)	<b>9.16</b> <b>(0.03)</b>
	$Q_2$ [dB]	10.80 (0.56)	11.12 (0.67)	10.88 (0.50)	10.39 (0.48)	<b>11.73</b> <b>(0.02)</b>
	$Q_3$ [dB]	1.519 (0.037)	1.538 (0.041)	1.447 (0.044)	1.410 (0.030)	<b>1.561</b> <b>(0.006)</b>
5	$Q_1$ [dB]	4.180 (0.357)	4.097 (0.635)	<b>4.322</b> <b>(0.009)</b>	<b>4.322</b> (0.014)	3.802 (0.076)
	$Q_2$ [dB]	7.46 (0.08)	7.45 (0.10)	<b>7.52</b> (0.05)	7.51 <b>(0.04)</b>	5.94 (0.13)
	$Q_3$ [dB]	0.747 (0.009)	0.747 (0.024)	0.752 <b>(0.006)</b>	<b>0.754</b> <b>(0.006)</b>	0.682 (0.009)

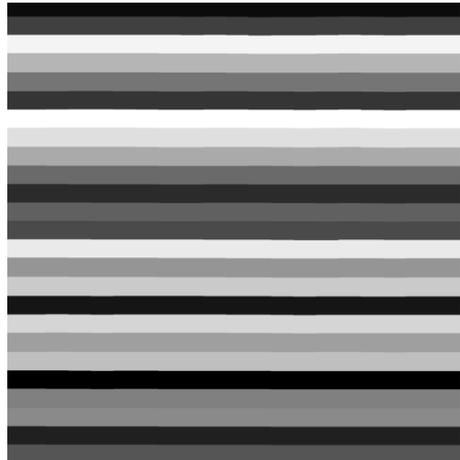
Table 2.8: Average values of the separation measures computed for each image pair. Separation measures were averaged across the 10 random runs. The standard deviations are shown in brackets. For each line, the best method is shown in bold.

European Community fund, under project POSC/EEA-CPS/61271/2004 and under the PhD fellowships SFRH/BD/23919/2005 and BID/n?5/2010-MarianaAlmeida.

## 2.4.9 Appendix



Pair #1



Pair #2



Pair #3



*source image as position.*

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

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

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

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

Pair #4



Figure 2.27: The first four pairs of source images. In this and all subsequent figures containing image pairs, one of the images of each pair has been horizontally flipped.

**Separation of nonlinear image mixtures**

When acquiring an image of a printed document, the image printed on the opposite page often shows through, due to partial transparency of the paper. Here we are dealing with quite a strong case of that effect, because we're using onion skin paper which is quite transparent.

The mixture that is obtained is rather nonlinear, as can be observed from the top figure on the right, which shows a scatter plot of the intensities of corresponding pairs of points from the two pages of a printed document. The scatter plot of the original images, shown in the bottom figure, filled a square, and had only a relatively small number of discrete intensity levels for each image. The fact that the shape of the scatter plot of Fig. 1 is very different from a parallelogram shows that the mixture was strongly nonlinear. The fact that this scatter plot becomes quite narrow in the upper-right corner (which corresponds to the lighter intensities in both images) indicates that, for those intensities, the mixture is close to singular. Finally, the fact that the discrete levels of Fig. 2 became largely blurred in Fig. 1 is due to noise in the process. The process leading from the sources to the observations involved printing the images, on both sides of a sheet of onion skin paper, at 1200 dpi, with a black and white laser printer (with the inherent halftoning of gray levels), and then scanning both sides of the printed sheet at 300 dpi. The noise is due, at least, to the printing process (including the halftoning), in the scanning process and to the non-uniformity in the onion skin paper, especially in its transparency.

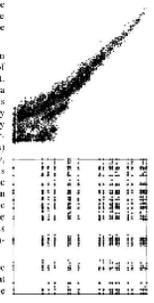
The purpose of separation is to recover, from the mixed images that are obtained by scanning both faces of the printed document, the images that had been printed in each of its faces, with as little interference from the other image as possible.

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

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

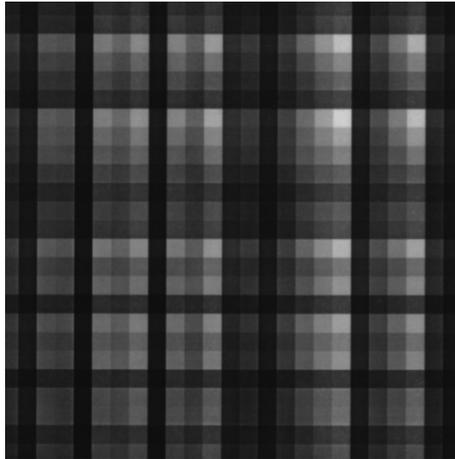
In the case of grayscale mixtures, the use of a separation method based on a good model of the physical mixing process should yield much better results than the use of a generic nonlinear separation method. A physical model could have a small number of parameters to be estimated, and would thus allow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

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

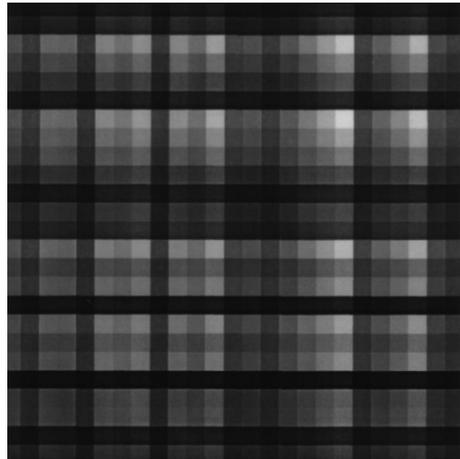


Pair #5

Figure 2.28: Fifth pair of tracing paper source images.



Pair #1



Pair #2



Pair #3



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

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

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

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

Pair #4

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

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

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

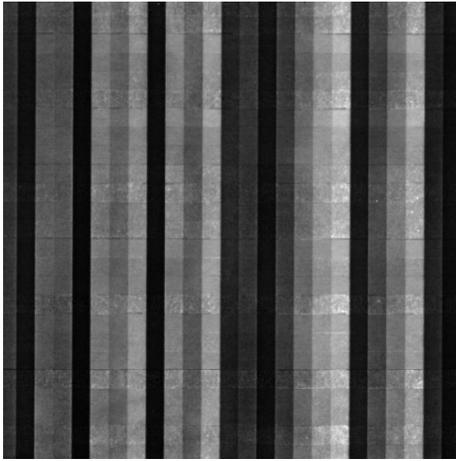
Another issue of interest is the definition of separate criteria that are more suited for images of heterogeneous documents than statistical independence. In fact, images and/or text from the opposite sides of a printed document can easily happen not to be independent from one another. For example, one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

Figure 2.29: First four pairs of tracing paper mixtures.

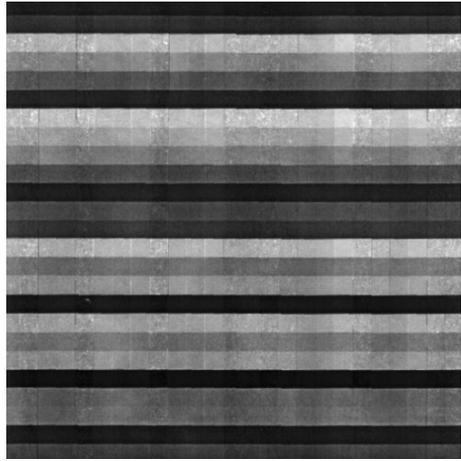


Pair #5

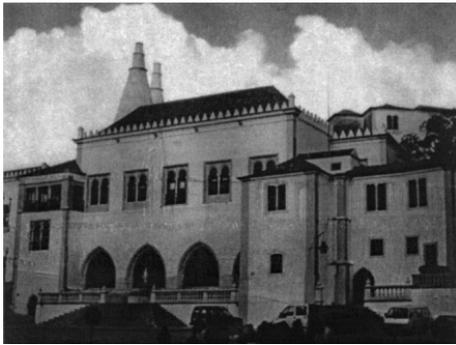
Figure 2.30: Fifth pair of tracing paper mixtures.



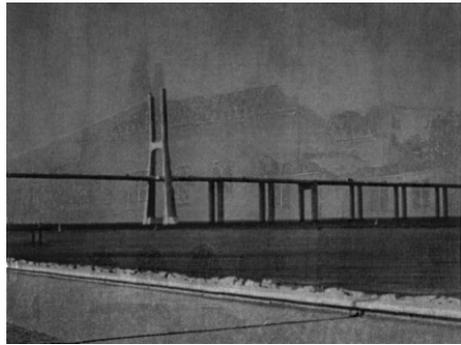
Pair  
#1



Pair  
#2



Pair  
#3



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

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

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

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

Pair  
#4



Figure 2.31: First four pairs of separated images.

**Separation of nonlinear image mixtures**

When acquiring an image of a printed document, the image printed on the opposite page often shows through, due to partial transparency of the paper. Here we are dealing with quite a strong case of that effect, because we're using onion skin paper which is quite transparent.

The mixture that is obtained is rather nonlinear, as can be observed from the top figure on the right, which shows a scatter plot of the intensities of corresponding pairs of points from the two pages of a printed document. The scatter plot of the original images, shown in the bottom figure, filled a square, and had only a relatively small number of discrete intensity levels for each image. The fact that the shape of the scatter plot of Fig. 1 is very different from a parallelogram shows that the mixture was strongly nonlinear. The fact that this scatter plot becomes quite narrow in the upper-right corner (which corresponds to the lighter intensities in both images) indicates that, for those intensities, the mixture is close to singular. Finally, the fact that the discrete levels of Fig. 2 became largely blurred in Fig. 1 is due to noise in the process. The process leading from the sources to the observations involved printing the images, on both sides of a sheet of onion skin paper, at 1200 dpi, with a black and white laser printer (with the inherent halftoning of gray levels), and then scanning both sides of the printed sheet at 100 dpi. The noise is due, at least, to the printing process (including the halftoning), to the scanning process and to the non-uniformity in the onion skin paper, especially in its transparency.

The purpose of separation is to recover, from the mixed images that are obtained by scanning both faces of the printed document, the images that had been printed in each of its faces, with as little interference from the other image as possible.

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

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

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

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

Pair #5

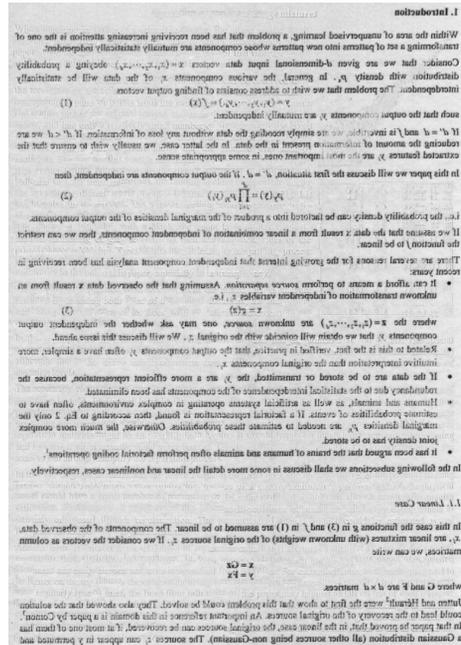


Figure 2.32: Fifth pair of separated images.

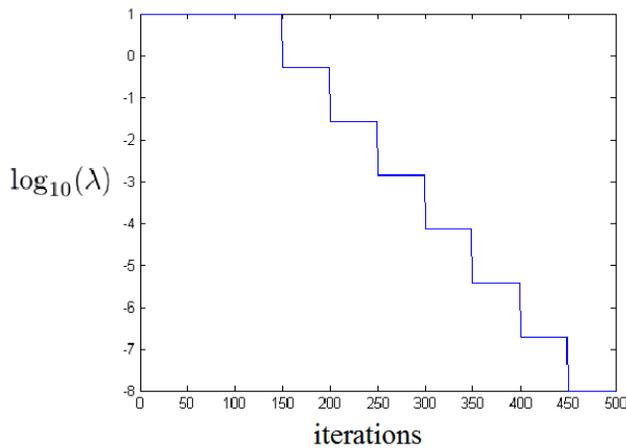


Figure 2.33: Variation of MND's regularizing parameter along the training.

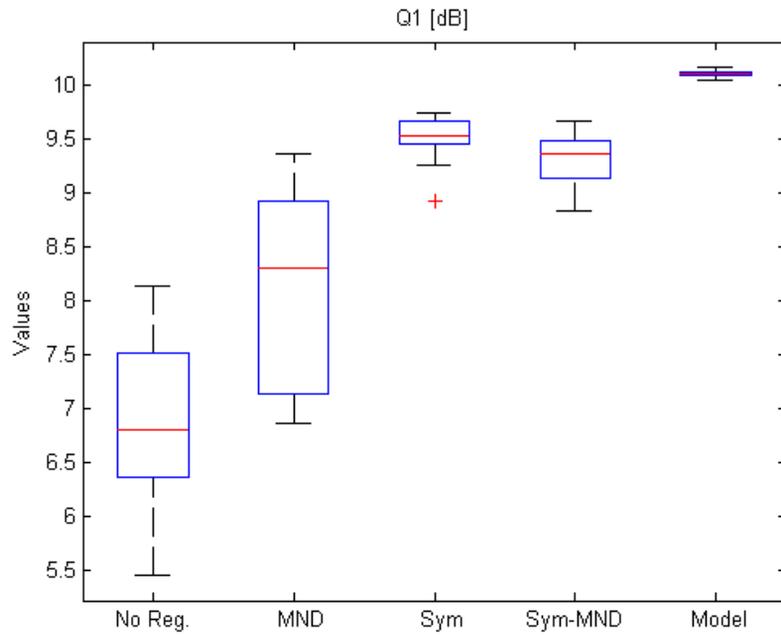


Figure 2.34: Box plots of the values of  $Q_1$  obtained for the MLP-based separator with various regularization strategies, and for the proposed model-based separator.

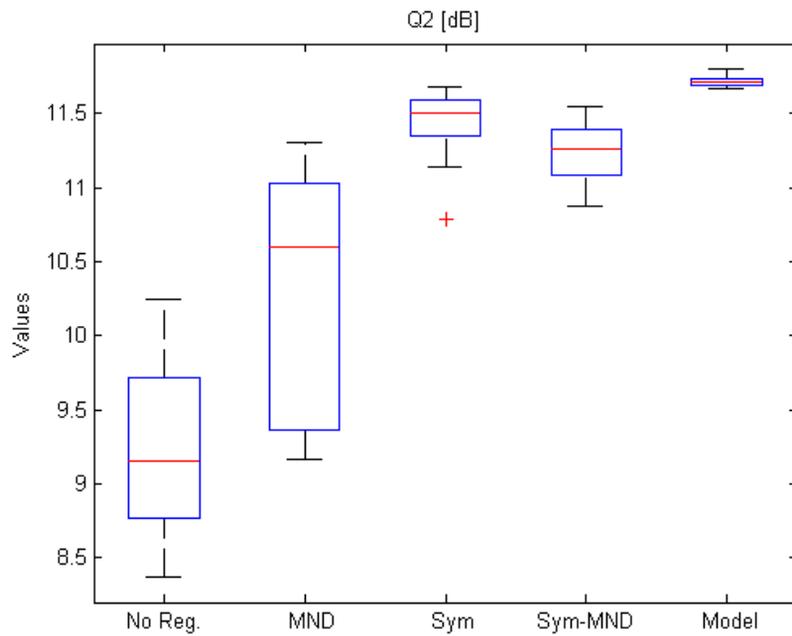


Figure 2.35: Box plots of the values of  $Q_2$  obtained for the MLP-based separator with various regularization strategies, and for the proposed model-based separator.

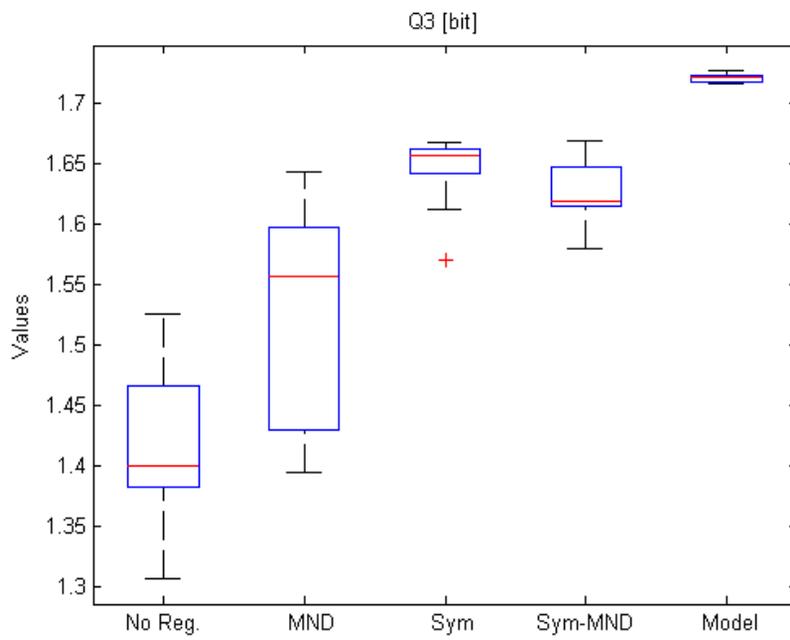


Figure 2.36: Box plots of the values of  $Q_3$  obtained for the MLP-based separator with various regularization strategies, and for the proposed model-based separator.

## Chapter 3

# Shift-invariant blind image deblurring

### Conference papers

Mariana S. C. Almeida and Luís B. Almeida, “Blind deblurring of natural images,” in *IEEE Int. Conf. Acoustics, Speech, and Signal Processing - ICASSP*, Las Vegas, 2008, pp. 1261-1264.

### Journal papers

Mariana S. C. Almeida and Luís B. Almeida, “Blind and semi-blind deblurring of natural images,” *IEEE Trans. on Image Processing*, vol. 19, no. 1, pp. 36-52, January 2010. (in the top 10 accessed articles of March 2010)

### Patents

Mariana S. C. Almeida and Luís B. Almeida. “Processo de focagem cega de imagens,” *Portuguese patent*, 2009.



## 3.1 Blind deblurring of natural images

### Abstract

A new method to perform blind image deblurring is proposed. Very few assumptions are made on the blurring filter and on the original image: the blurring filter is assumed to have limited support and the original image is assumed to be a sharp natural image. A new prior is used, which gives higher probability to images with sharp edges. The estimation of both the deblurred image and the blurring filter is made in a progressive way, first taking into account the main features of the image, and then proceeding to smaller details.

The results obtained with synthetically blurred images are good, even when the blur operator is rather ill-conditioned and the blurred image is noisy. The method also yields improvements in real-life photographs with focus and motion blurs.

**keywords:** Blind image deconvolution, Image enhancement, Image restoration, Image edge analysis, Sparse distributions.

### 3.1.1 Introduction

Image deblurring consists of attempting to recover an image which has been degraded by a linear shift-invariant filtering operation, possibly with noise. This has applications in fields such as astronomy [10], remote sensing [18] and biomedical imaging [16], as well as in everyday life, for the enhancement of blurred photos.

Part of the previous work on image deconvolution has been done in the non-blind setting, in which the blur operator is exactly known. Although that is not a valid assumption in many real-life situations, the problem is still hard, because the blur operator typically is very ill-conditioned. Several approaches to this problem, using prior information on the estimated image, can be found in [78, 79, 157].

In blind image deconvolution (BID), both the image and the blur operator are unknown. The problem is ill-posed, having an infinite number of solutions. Furthermore, as above, the blur operator often is rather ill-conditioned. For an overview of BID see [84].

To the authors' knowledge, there is no previous general solution for blind deblurring without making relatively strong assumptions on the blurring operator and/or on the image. A common approach is to restrict the problem to some blur model, such as motion blur or focus blur [87, 88]. Among recent works on BID we emphasize [92, 93, 117, 118]. Both [117] and [118]

require extra data for a preliminary training. [92] attempts to encompass less restrictive blurs through a fuzzy technique, which is applied under the output of blur models known a priori. A non-iterative, fast method with proof of convergence is presented in [93]. This method assumes the blur to be zero-phase and depends on a good initial estimate of the blur.

The approach that we propose in this paper makes very few assumptions on the blurring filter: only that it has a support of size no larger than a given value and that it generally has a low-pass character. We use maximum a posteriori approach with a new prior which favors images with sharp edges. This leads to a regularizer which generalizes the well known total variation (in its discrete form) [158].

Directly estimating the image using the new prior doesn't easily lead to good solutions. We guide the optimization to a good solution by first concentrating on the main features of the image, and progressively dealing with smaller details. This leads to quite good results, even for blur operators that are rather ill-conditioned.

This paper is structured as follows: Section 3.1.2 describes the proposed method and introduces the new regularizer. Results are presented in Section 3.1.3. Section 3.1.4 concludes.

### 3.1.2 Deblurring method

The image degradation can be formulated as

$$y = Hx + n, \tag{3.1}$$

where  $H$  is a square matrix corresponding to the linear blurring operator, and  $y$ ,  $x$  and  $n$  are column vectors that represent, respectively, the degraded image, the original image and additive noise, all vectorized in lexicographic order.

The method that we propose can be explained with the aid of Fig. 3.1, where  $x$  is the estimated deblurred image,  $H$  the estimated blur filter and  $\hat{y}$  a reconstructed approximation to the blurred image. Block  $F$  represents an edge extractor: for each pixel of the image  $x$ , it computes the intensity of a possible edge passing through that pixel. The estimation of both  $x$  and  $H$  is performed through the minimization of the objective function

$$C(x, H) = \|y - Hx\|_2^2 + \lambda R(F(x)), \tag{3.2}$$

where  $\lambda$  in (3.2) is a regularization parameter and  $R(F(x))$  is a regularizing term which favors solutions in which the deblurred image  $x$  has a sparse response to the edge extractor  $F$ . More

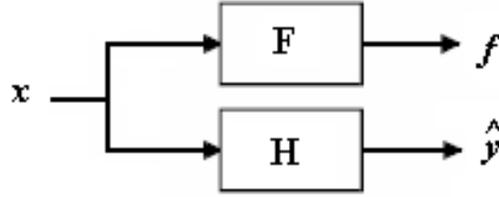


Figure 3.1: Schematic representation of the deblurring method. Block F extracts edge intensities.

*Initialization:*

- 1 – Set  $H$  to the identity operator.
- 2 – Set  $x$  equal to  $y$ .
- 3 – Set  $\lambda$  and the prior’s sparsity to the initial values of the corresponding sequences.

*Optimization loop:*

- 4 – Find new  $x$  estimate:  $x = \operatorname{argmin}_x C(x, H)$  ( $H$  fixed).
- 5 – Find new  $H$  estimate:  $H = \operatorname{argmin}_H C(x, H)$  ( $x$  fixed).
- 6 – Set  $\lambda$  and the prior’s sparsity to the next values in sequence.
- 7 – If  $\lambda \geq \lambda_{min}$  go back to 4; otherwise stop.

Table 3.1: Deblurring method

details on  $F$ , on the prior and on the regularizer  $R$  are given in Sections 3.1.2 and 3.1.2.

The blurring filter  $H$  is restricted to a limited support. The learning is guided to the desired solution by starting with a large regularization parameter  $\lambda$ , which is then progressively reduced, and by starting with a less sparse prior, which is then made sparser. For efficiency reasons, optimizations relative to  $x$  and  $H$  are performed in alternation. An outline of the method is shown in Table 3.1. We assume that we have chosen a decreasing sequence of values for  $\lambda$  and a non-decreasing sequence of values for the prior’s sparsity.

In the beginning of the optimization, with a large  $\lambda$ , only the main features survive in the estimated image. It makes sense to start by considering only these features. In fact, while the estimate of  $H$  is poor, an image estimate with little regularization would contain many wrong high frequencies, and only the stronger and largest features would remain approximately intact. By using a strong regularization we force the process to only consider these features, eliminating the wrong high frequencies. As the optimization proceeds and the estimate of the blurring filter becomes better, smaller and fainter features can progressively be used for the estimation. By progressively lowering  $\lambda$  during the optimization, we guide the method to progressively consider smaller features, leading it to a good solution. An early image estimate, with a large  $\lambda$ , is shown in Fig. 3.3 b).

A very sparse prior doesn’t easily allow edges to move. If, initially, the estimated position of an edge is slightly offset relative to the correct position, it will be hard for the edge to move to the right position during the optimization. For this reason, we often start the optimization

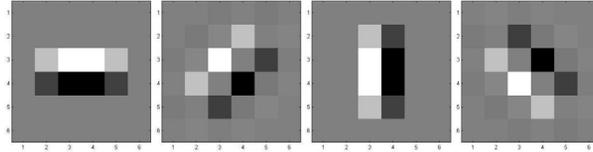


Figure 3.2: The edge detection filters in the four orientations that were used.

process with a prior that is less sparse than the one used in the final phases of the optimization.

### Edge detector

The edge detector uses a set of edge detection filters, which are rotated versions of a basic filter (Fig. 3.2). Filter rotation is performed by rotating the basic filter’s point spread function, using bicubic interpolation. For each pixel, the edge detector computes the outputs of the filters for all orientations under consideration,  $g_\theta$ , where  $\theta$  denotes the filter’s orientation. The detector’s output is given by

$$f = \sqrt{\sum_{\theta \in \Theta} g_\theta^2}, \quad (3.3)$$

where  $\Theta$  is the set of orientations under consideration.

### Image prior

Edges, in natural images, are known to be sparse. They are sparser in sharp images than in blurred ones, because, in the latter, edges are “spread” over a larger width. Therefore, a sparse prior on edges will favor sharp images over blurred ones. The prior that we use assumes that edge intensities at different pixels are independent from one another (which obviously is a large simplification, but still leads to good results). The edge intensity at each pixel is assumed to follow a sparse prior with density

$$p(f) \propto e^{-k(f+\epsilon)^q}, \quad (3.4)$$

where  $k$  adjusts for the scale of edge intensities and  $q$  controls the prior’s sparsity;  $\epsilon$  allows us to obtain finite lateral derivatives at  $f = 0$  (with  $0 < q < 1$ ), making the prior closer to actual observed distributions and also making the optimization easier.

Assuming, for the noise  $n$  in (3.1), a Gaussian prior with zero mean and variance  $\sigma^2$ , the likelihood of the estimated pair (image + filter) is given by

$$p(x, H|y) \propto e^{-\frac{1}{2\sigma^2}\|Hx-y\|_2^2} \prod_i e^{-k(f_i+\epsilon)^q}, \quad (3.5)$$

where  $i$  is an index running through all pixels. The log-likelihood is, apart from an irrelevant

constant,

$$L(x, H|y) = -\frac{1}{2\sigma^2} \|Hx - y\|_2^2 - k \sum_i (f_i + \epsilon)^q. \quad (3.6)$$

Maximizing this is equivalent to minimizing the cost function

$$C(x, H) = \|Hx - y\|_2^2 + \lambda \sum_i (f_i + \epsilon)^q, \quad (3.7)$$

where  $\lambda = 2k\sigma^2$ . This cost function is of the form given in (3.2).

We recognize, in this equation, the data connection term,  $\|Hx - y\|_2^2$ , and the regularizer,  $\sum_i (f_i + \epsilon)^q$ . The latter has the well known total variation regularizer (in its discrete form) as a special case, which is obtained by using just two filters which compute horizontal and vertical differences, and setting  $\epsilon = 0$  and  $q = 1$ .

As was said above, we decrease  $\lambda$  during the optimization. Therefore, except for the last phase of the optimization,  $\lambda$  is not given by  $2k\sigma^2$ . And in fact, even during that last phase,  $\lambda$  still is not given by that expression because the noise  $n$ , besides allowing for possible noise in the blurred image, also allows for a mismatch between the estimated filter  $H$  and the true one. This mismatch leads to a difference between the reconstructed blurred image and the actual one, this difference being treated by the method as noise.

### 3.1.3 Experiments

In this section we present several practical examples illustrating the performance of the proposed method. The first four experiments involve synthetic blurs and the last two involve actual blurred photos.

We have used, in the experiments, the method described above, outlined in Table 3.1. The support of the blurring filter estimate was limited to a square of size  $s \times s$  pixels (the specific values of  $s$  are given ahead for each case). In step 5 of the method, where we estimate the blurring filter, we used a safety margin of width 3 around that square. This means that the filter's support was limited to a size  $(s + 6) \times (s + 6)$  and that, at the end of step 5, the filter estimate was truncated to the central square of size  $s \times s$ . This safety margin was used because we found that there were some undesirable border effects in the filter estimate, in the initial phases of the optimization.

The sequences of values of  $\lambda$  and  $q$  were set, for all experiments, at 2, 0.6, 0.2, 0.06,  $\dots$ ,  $2 \times 10^{-7}$  and 0.8, 0.8, 0.6, 0.6, 0.6, 0.6, 0.4, 0.4,  $\dots$ , 0.4, respectively (a total of 15 values each). In the tests that involved noise (artificial or natural) these sequences were truncated,

so that the final value of  $\lambda$  was larger. Parameter  $\epsilon$  was set to 0.002. Pixel intensities were represented on a scale of 0 (black) to 1 (white). All images had a size of  $256 \times 256$  pixels. The edge detection filters that we used are shown in Fig. 3.2. They correspond to the basic filter with point spread function (PSF)  $F_{\theta=0} = \begin{bmatrix} 1 & 2 & 2 & 1 \\ -1 & -2 & -2 & -1 \end{bmatrix} / 12$ , with rotations of  $45^\circ$ .

The results of the first four experiments are shown in Figs. 3.3 and 3.4. In the printed version of the paper the images are small, but in the electronic version one can zoom in on the images to see finer details. In the first experiment the blur PSF was a uniform square of size  $9 \times 9$ . In the second one it corresponded to a simulated motion blur of length 9 pixels, in a direction at  $45^\circ$  with the horizontal. In the third one it was a simulated defocus blur (a uniform circle of diameter 11 pixels). The size of the support of the estimated filter ( $s$ , above) was 15, 15 and 17, respectively. The fourth experiment was similar to the third, but Gaussian noise with  $\sigma = 0.01$  (SNR of  $25dB$ ) was added to the blurred image. In this case, the sequence of values of  $\lambda$  was truncated at 0.0006, followed by a last iteration with  $\lambda = 0.0004$ . This last value was found, by visual inspection, to provide the best trade-off between sharpness and noise in the deblurred image.

The last two experiments, whose results are shown in Fig. 3.5, involved actual photographs, intentionally taken with a focus and a motion blur, respectively. The sequence of values of  $\lambda$  was truncated at 0.0002. In the case of the motion blur, a further iteration with  $\lambda = 0.0001$  was performed. Again, these values were experimentally found. The filter size  $s$  was set to 15 in both cases.

These results show that, in synthetic blurs with no noise, the estimates of the original image and of the blurring filter were rather good. Even in the noisy case, the recovered image was significantly sharper than the blurred one and didn't have too much noise, although the filter estimate was rather noisy. In the case of the photographs, the recovered images had somewhat a synthetic look and presented some artifacts, but they were significantly sharper and had more details than the blurred ones. The somewhat lower quality of the results with photos probably means that the blurs didn't exactly follow the model (3.1), perhaps due to nonlinearities in the camera's processing and to the JPEG encoding of the photos.

The optimization of the deblurred image  $x$  was performed by gradient descent with adaptive step sizes (150 iterations for each value of  $\lambda$ ). The filter  $H$  was optimized by conjugate gradients (100 iterations for each value of  $\lambda$ ). On an Intel Core 2 Duo system running at 2 GHz, programmed in Matlab and running only on one of the chip's processors, the optimization took about 30 seconds for each value of  $\lambda$  (7.5 minutes for the 15 values).

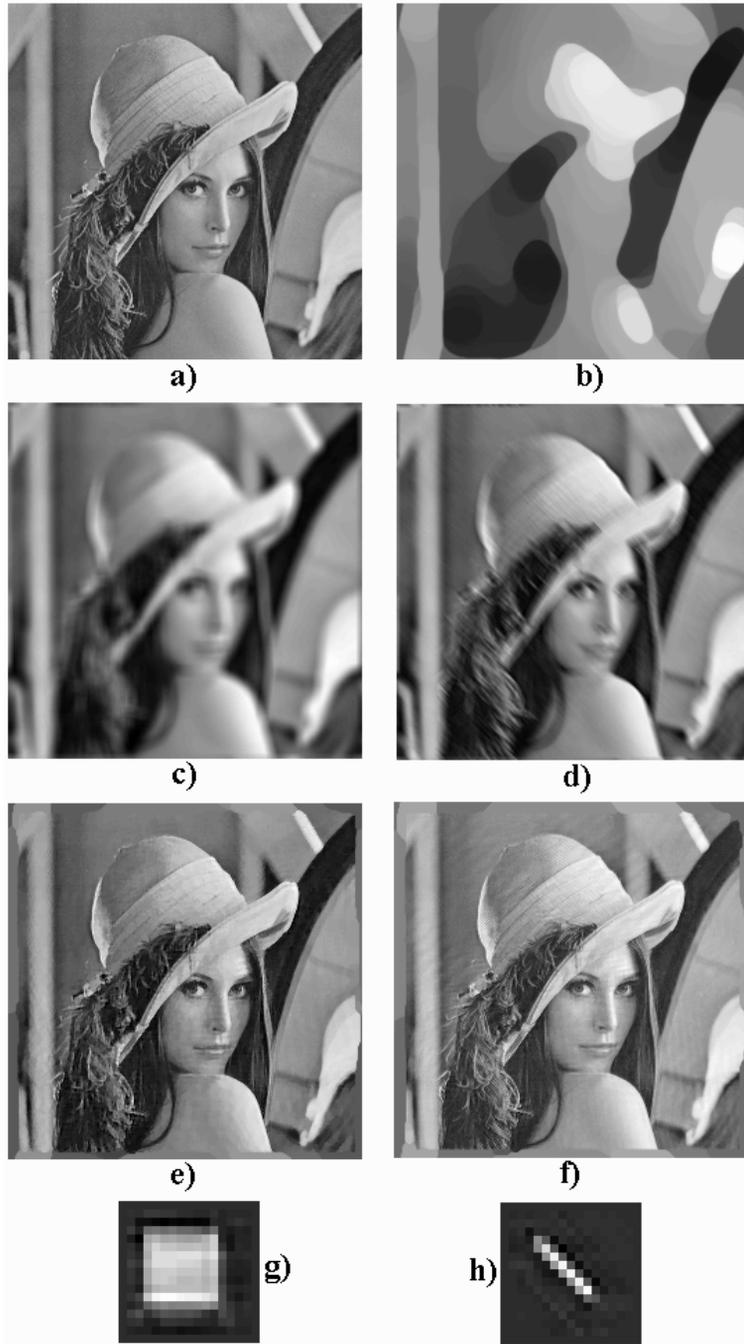


Figure 3.3: Results of synthetic experiments. a) Original image. b) Estimate for the first  $\lambda$  value, for the square blur. Next rows: left, square blur; right, motion blur. Second row: Blurred images. Third row: Deblurred images. Fourth row: Filter estimates.

### 3.1.4 Conclusions and future work

A new method for blind image deblurring was proposed. The only strict assumption of the method is that the blurring filter has limited support, no larger than a given size. A new prior is used, which tends to enforce the sharpness of edges, by making them sparse. Edges are found by means of a new detector, which uses several rotated versions of a basic edge-detection filter.

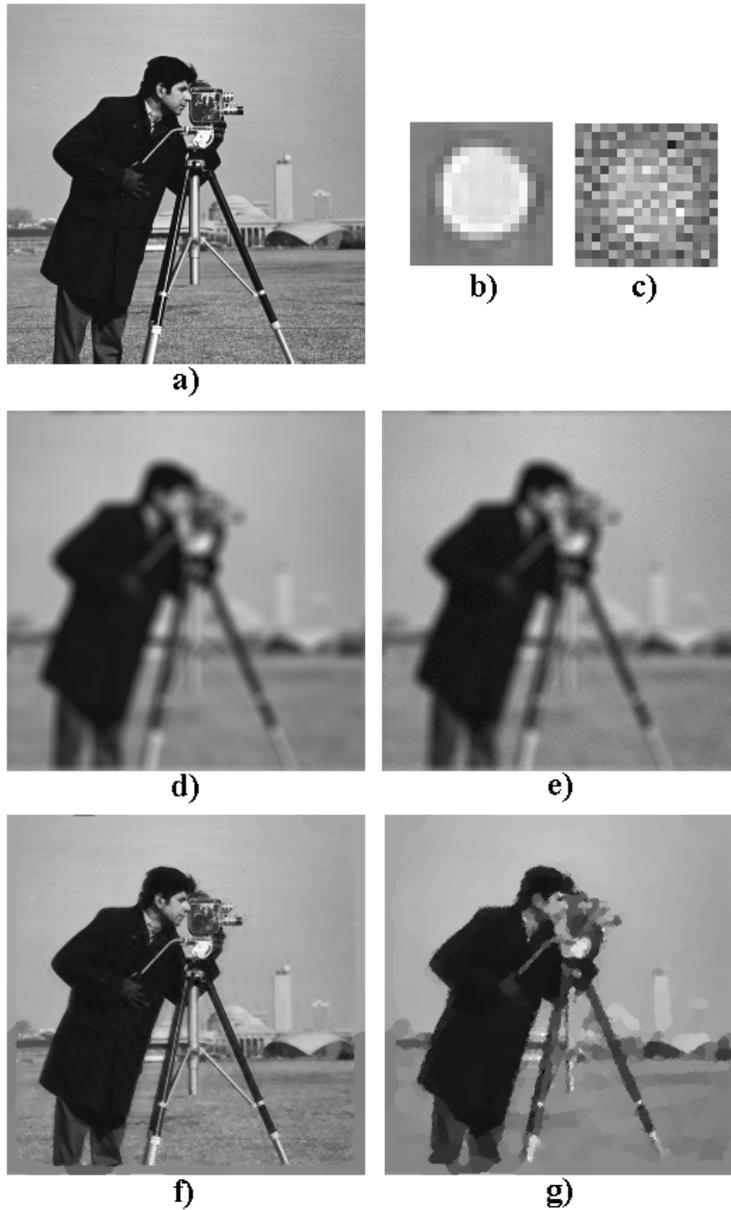


Figure 3.4: Results of synthetic experiments. a) Original image. b) and c) Filter estimates. Next rows: Left, without noise; Right, with noise. Second row: Blurred images. Third row: Deblurred images.

The learning is guided to a reasonable solution by first concentrating on the main features of the image, and progressively taking into account smaller details. Experimental results, both on synthetic blurs and on actual blurred photographs, show the good performance of the method.

The prior that was introduced in this paper will probably be useful in other image restoration problems as well. Quite probably, it will also be possible to improve the prior itself. We also plan to explore efficient techniques, such as majoration-minimization, to improve the optimization speed of the method.

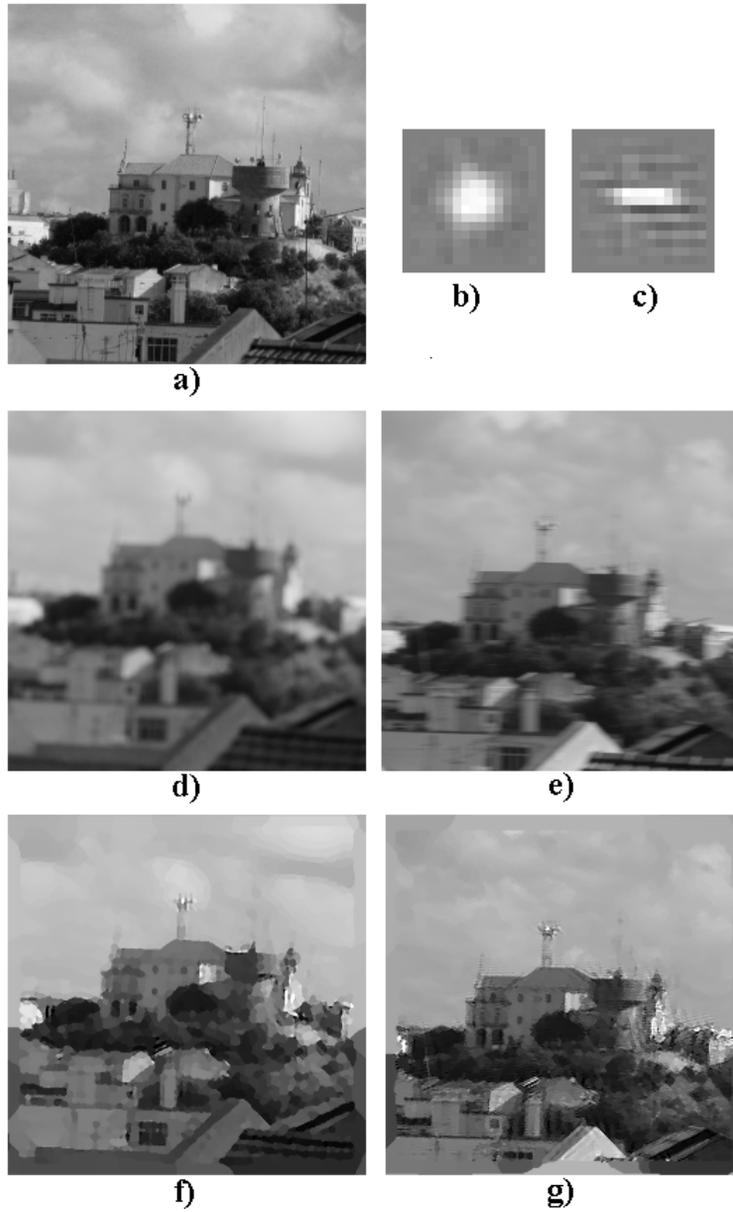


Figure 3.5: Results with actual blurred photos. a) Sharp photo. b) and c) Filter estimates. Next rows: Left, defocus blur; Right, motion blur. Second row: Blurred images. Third row: Deblurred images.

### 3.1.5 Acknowledgments

We wish to acknowledge very useful discussions with our colleagues J. Bioucas-Dias and M. Figueiredo.



## 3.2 Blind and semi-blind deblurring of natural images

### Abstract

A method for blind image deblurring is presented. The method only makes weak assumptions about the blurring filter, and is able to undo a wide variety of blurring degradations. To overcome the ill-posedness of the blind image deblurring problem, the method includes a learning technique which initially focuses on the main edges of the image and gradually takes details into account. A new image prior, which includes a new edge detector, is used.

The method is able to handle unconstrained blurs, but also allows the use of constraints or of prior information on the blurring filter, as well as the use of filters defined in a parametric manner. Furthermore, it works in both single-frame and multi-frame scenarios. The use of constrained blur models appropriate to the problem at hand, and/or of multi-frame scenarios, generally improves the deblurring results.

Tests performed on monochrome and color images, with various synthetic and real-life degradations, without and with noise, in single-frame and multi-frame scenarios, showed good results, both in subjective terms and in terms of the increase of signal to noise ratio (ISNR) measure. In comparisons with other state of the art methods, our method yields better results, and shows to be applicable to a much wider range of blurs.

**keywords:** Blind image deconvolution, Image enhancement, Image restoration, Image deblurring, Sparse distributions.

### 3.2.1 Introduction

Image deblurring is an inverse problem whose aim is to recover an image from a version of that image which has suffered a linear degradation, with or without noise. This blurring degradation can be shift-variant or shift-invariant. Although there have been some proposed methods for recovering shift-variant linear degradations [20, 21, 37, 38, 126, 143, 159, 160], the majority of existing deblurring methods was developed for invariant degradations, and the blind recovery from shift-invariant degradations is still considered a rather challenging problem. This paper focuses on shift-invariant blurs, and, in the context of this paper, “blur” will refer to a linear, shift-invariant degradation, i.e., a convolution, with or without noise, unless stated otherwise.

Automatic image deblurring is an objective of great practical interest for the enhancement of images in photo and video cameras [7–9], in astronomy [10], in remote sensing [18], in

tomography [12, 13], in other biomedical imaging techniques [15–17], etc.

Image deblurring methods can be divided into two classes: *non-blind*, in which we assume the blurring operator to be known, and *blind*, in which we assume that the blurring operator is unknown. The method that we describe here belongs to the latter class. The application range of non-blind methods is much narrower than the one of blind methods: in most situations of practical interest the blurring filter’s impulse response, also called Point Spread Function (PSF), is not known with good accuracy. Since non-blind deblurring methods are very sensitive to mismatches between the PSF used by the method and the true blurring PSF, a poor knowledge of the blurring PSF normally leads to poor deblurring results.

Despite its narrower applicability, non-blind deblurring already is a difficult problem. The main difficulty faced by non-blind deblurring methods has to do with the presence of noise in the blurred image. Since the blurring operator typically is very ill-conditioned, this noise, even if very weak, can strongly contaminate the deblurred image. The problem is serious in situations in which the blurring PSF is exactly known, and gets worse if there is even a slight mismatch between the PSF used for deblurring and the one that caused the blur. Most non-blind deblurring methods [78–82] overcome this difficulty through the use of prior information about the image to be recovered, often doing this within a Bayesian or maximum a posteriori framework.

In blind image deblurring (BID), not only the degradation operator is ill-conditioned, but the problem also is, inherently, severely ill-posed: there is an infinite number of solutions (original image + blurring filter) that are compatible with the degraded image. For an overview of BID methods see [84, 85].

Most previously published blind deblurring methods are very limited, since they do not allow the use of a generic PSF. Most of them are based, instead, on PSF models with a small number of parameters [86–89, 107]. For example, to model an out-of-focus blur, they normally use a circle with uniform intensity, having as single parameter the circle’s radius [86]. Similarly, to model a motion blur, they normally use a straight-line segment with uniform intensity, the only parameters being length and slope [86–88]. These approaches are very limited, because such models rarely fit actual blurring PSFs well. For example, the out-of-focus blurring PSF generally is more complex than a simple uniform circle, and the camera motion that causes a motion blur generally is much more complex than a uniform, straight-line motion. And, as was emphasized above, even a slight mismatch between the deblurring PSF and the blurring PSF strongly degrades the quality of the deblurred image.

A recent work [107] manages to estimate the blur under the variational Bayesian approach.

However, this method models the blur by means of a Gaussian filter, which is completely defined by a single parameter (the Gaussian’s variance), and is a very weak model for real-life blurs.

In an attempt to encompass less restrictive blurs, a fuzzy technique that uses several pre-specified PSF models has been considered in [92]. Another blind deconvolution method, which is fast and has a proof of convergence, is described in [93]. However, this method assumes that the PSF is zero-phase and, furthermore, depends on the existence of a good initial estimate of the PSF.

References [94] and [95] present a method called APEX. Although this method covers some blurs which can be found in real-life, it is limited to blurring PSFs modeled by a symmetrical Lévy distribution with just two parameters. In Section 3.2.4 we present an experimental comparison of our method with APEX.

Some methods have been proposed, which impose no strong restrictions on the blurring filter [20, 96, 97, 99, 101]. These methods typically impose priors over the blurring filter, and don’t seem to be able to handle a wide variety of blurs and scenes. In [97] and [20], total variation (TV) is used to regularize the blurring filters. Besides being used for space-invariant blurs, the method described in [20] was also applied with success in a synthetic image with a space-variant blur. We present an experimental comparison with that method in Section 3.2.4. The method recently presented in [7] is much less restrictive than parameterized ones and yields good results, but is only designed for motion blurs.

An interesting method for blind deblurring of color images was proposed in [100]. This method appears not to pose any strong restrictions on the blurring filter. In the cited paper, several experimental results on synthetic blurs are shown, but little information is provided about them. From the information that is given, it appears that the blurring filters that were used in the experiments were either circularly symmetric (including simulated out-of-focus blurs), or corresponded to straight-line motion blurs. There seems to be no reason for the method not to be able to successfully deal with other kinds of blurs, however. The blurring PSFs that are shown in that paper appear to have a maximum size of about  $5 \times 5$  pixels (or a length of 3 pixels, in the case of the motion blur). The improvements in signal to noise ratio (ISNR, see Section 3.2.3 for a definition) seem to be between 2dB and 4dB for the circularly symmetric blurs, and of 7dB for the motion blur. The experimental results presented in Section 3.2.4 show that, with much stronger blurs (much larger blurring PSFs), our method normally yielded larger improvements than the method of that paper.

In some cases, one has access to more than one degraded image from the same original

scene, a fact which can be used to reduce the ill-posedness of the problem [37, 112–116]. There are also solutions like the ones presented in [117–119], which cannot be considered completely blind, since they require the use of additional data for preliminary training.

Contrary to previously published blind deconvolution methods such as those mentioned above, the method that we propose only makes a weak assumption on the blurring PSF: it must have a limited support. The method also assumes that the leading (most important) edges of the original image, before the blur, are sharp and sparse, as happens in most natural images. To the authors’ knowledge, this is the first method to be proposed, which is able to yield results of good quality in such a wide range of situations.

The method uses a new prior which depends on the image’s edges, and which favors images with sparse edges. This prior leads to a regularizing term which generalizes the well known total variation (TV), in its discrete form [158]. The estimation is guided to a good solution by first concentrating on the main edges of the image, and progressively dealing with smaller and/or fainter details. Though the method allows the use of a very generic PSF, it can also take into account prior information on the blurring PSF, if available. If a parameterized model of the PSF is known, the method allows the estimation of the model’s parameters. Although initially developed for the single-frame scenario, the method can also be used in multi-frame cases, benefiting from the existence of the additional information from the multiple frames.

The performance and the robustness of the method were tested in various experiments, with synthetic and real-life degradations, without and with constraints on the blurring filter, without and with noise, using monochrome and color images, and under the single-frame the multi-frame paradigms. The quality of the results was evaluated both visually and in terms of ISNR. Detailed comparisons with two other methods available in the literature [20, 94] were performed, and show that the proposed method yields significantly better results than these other methods.

This paper is organized as follows: Section 3.2.2 describes the proposed method and presents the new prior for images. Section 3.2.3 describes some special aspects of the computation of the ISNR measure in the blind deblurring case. Experimental results are presented in Section 3.2.4. Section 3.2.5 presents conclusions and future research directions.

### 3.2.2 Deblurring method

The degradation that we aim to recover from, is modeled by

$$y = h * x + n, \tag{3.8}$$

in which  $y$ ,  $x$  and  $n$  are images which represent, respectively, the degraded image, the original image and additive noise;  $h$  is the PSF of the blurring operator, and  $*$  denotes the mathematical operation of convolution.

The deblurring method is based on two simple facts:

- In a natural image, leading edges are sparse.
- Edges of a blurred image are less sparse than those of a sharp image, because they occupy a wider area.

Due to these facts, a prior which tends to make the detected edges sparser will tend to make the image sharper, while preventing it from becoming unnatural (i.e., from presenting noise or artifacts).

Let us designate by  $f(\cdot)$  an edge detection operation, such that  $f(x)$  is an image with the intensities of the edges that exist in the image  $x$ . The deblurring method that we propose finds a local minimum, with respect to both the image  $x$  and the blur  $h$ , of the cost function

$$C(x, h) = \frac{1}{2} \|y - h * x\|_2^2 + \lambda R[f(x)], \quad (3.9)$$

where  $R[f(x)] = R_f(x)$  is a regularizing term which favors solutions in which the edges present in  $f(x)$  are sparse, and  $\lambda$  is a regularization parameter. More details on the edge detector  $f(\cdot)$  are given in Section 3.2.2. Both the regularizer  $R_f(\cdot)$  and the prior from which it is obtained are described in Section 3.2.2. The estimate of the blurring filter's PSF,  $h$ , is restricted to have a limited support, which should not be smaller (but may be larger) than the support of the actual blur.

A local minimum of the cost function, that corresponds to a good deblurring result, is reached by starting with a large value of the regularizing parameter  $\lambda$  and progressively reducing it. That local minimum usually is not the global one. For example, for images with a significant amount of texture, we have found that the (image + filter) pair formed by the blurred image and the identity filter often yields a lower value of the cost function than the value that we obtain at the end of the deblurring process, when the estimated image is much sharper than the blurred one. We don't know whether the global minimum of the cost function would yield a good deblurred image, but we have reasons to believe that it might not do so. Of course, this leads us to think that there should be a better cost function, whose global minimum will yield a good deblurred image. We don't presently know such a function, however, and this clearly is an important topic for further research.

The deblurring method is outlined in Table 3.2, for which a decreasing sequence of values of  $\lambda$  and a non-increasing sequence of values of  $q$  are assumed to have been previously chosen ( $q$  controls the regularizer’s sparsity, as will be discussed ahead). In our experiments, we have used a geometric progression for the sequence of values of  $\lambda$  ( $\lambda_{n+1} = \lambda_n/r$ ). The filter estimation, performed in step 5 of the method, can take into account constraints or prior information on the filter, if these are available.

*Initialization:*

- 1 – Set  $h$  to the identity operator.
- 2 – Set  $x$  equal to  $y$ .
- 3 – Set  $\lambda$  and  $q$  to the initial values of the respective sequences.

*Optimization loop:*

- 4 – Find a new  $x$  estimate:  $x = \operatorname{argmin}_x C(x, h)$
- 5 – Find a new  $h$  estimate:  $h = \operatorname{argmin}_h C(x, h)$
- 6 – Set  $\lambda$  and  $q$  to the next values in sequence.
- 7 – If  $\lambda \geq \lambda_{min}$  go back to 4; otherwise stop.

Table 3.2: Deblurring method

In the early stages of the deblurring process, the estimate of  $h$  is still poor, and a strong regularization is required to make edges sharp in the estimated image, and to eliminate the wrong high frequency components that could otherwise appear. During these early iterations,  $\lambda$  is large and only the main edges of the estimated image survive (see an example in Fig. 3.6). The surviving edges are sharp, however, due to the strong edge-sparsifying regularization, and these sharp edges allow the estimate of the blurring filter to be improved. In the next iteration, the improved filter, together with a somewhat lower value of  $\lambda$ , allows some smaller and/or fainter edges to be extracted. As the iteration proceeds and the filter estimate improves, smaller and/or fainter features are progressively handled, at a rate which is controlled by the rate of decrease of the value of  $\lambda$ . This procedure results in a guidance of the optimization, which leads it to a local minimum that normally corresponds to a good deblurring result.

**Edge detector**

In order to be able to apply a prior over the image edges, an edge detector was developed. This edge detector showed to yield better deblurring results than those obtained using detectors described in the literature, such as [161–164]. The edge detector is based on a set of filters obtained, from a base filter  $d_0$ , through successive rotations (see Fig. 3.7). The base filter (leftmost filter in Fig. 3.7) is a detector of edges in a given direction and, naturally, its rotated versions are detectors of edges in the corresponding rotated directions. Designating by  $d_\theta$  the

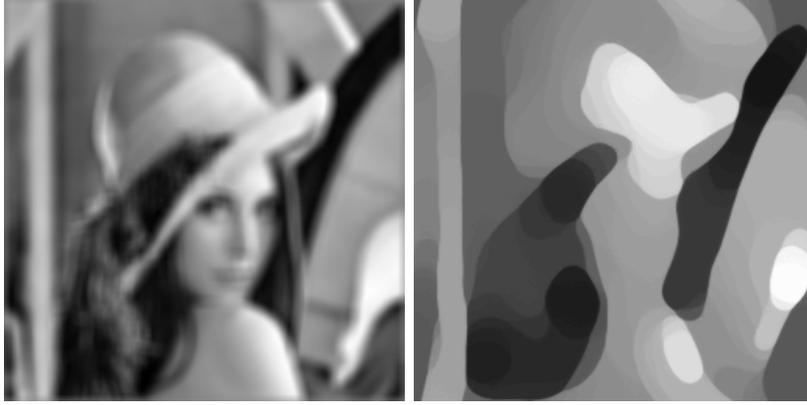


Figure 3.6: Left: “Lena” image blurred with a  $9 \times 9$  square blur. Right: Image estimate obtained for the first value of  $\lambda$ .

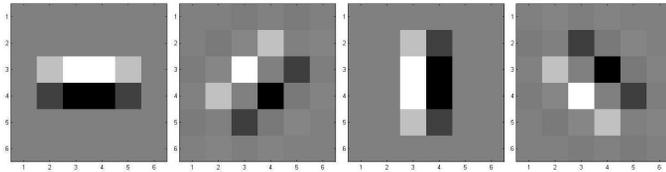


Figure 3.7: The set of edge detection filters, in the four orientations that were used. The leftmost filter is the base filter  $d_0$ .

filters of the set and by

$$g_{\theta}(x) = d_{\theta} * x \quad (3.10)$$

the filters’ outputs, the output of the edge detector is given by a combination of those outputs through an  $L_2$  norm,

$$f(x) = \sqrt{\sum_{\theta \in \Theta} g_{\theta}(x)^2}, \quad (3.11)$$

in which  $\Theta$  is the set of filter orientations under consideration. As an example of the detector’s operation, Fig. 3.8 shows the edges that were extracted, from the “Lena” image, by this edge detector, using the set of filters shown in Fig. 3.7.

### Image Prior

The prior that we use for images assumes that edges are sparse, and that edge intensities at different pixels are independent from one another (which obviously is a large simplification, but still leads to good results). The edge intensity at each pixel  $i$ , denoted  $f_i(x)$ , is assumed to follow a sparse prior with density

$$p[f_i(x)] \propto e^{-k[f_i(x)+\epsilon]^q}, \quad (3.12)$$



Figure 3.8: Edges of “Lena” computed using the proposed edge detector.

where  $k$  adjusts for the scale of edge intensities and  $q$  controls the prior’s sparsity;  $\epsilon$  is a small parameter which allows us to obtain finite lateral derivatives at  $f = 0$  (with  $0 < q < 1$ ), making the prior closer to actual observed distributions, and also making the optimization easier.

Assuming, for the noise  $n$  in (3.8), a Gaussian prior with zero mean and variance  $\sigma^2$ , the likelihood of the estimated pair (image + filter) is given by

$$p(x, h|y) \propto e^{-\frac{1}{2\sigma^2}\|y-h*x\|_2^2} \prod_i e^{-k[f_i(x)+\epsilon]^q}, \quad (3.13)$$

where  $i$  is an index running through all pixels. The log-likelihood is, apart from an irrelevant constant,

$$L(x, h|y) = -\frac{1}{2\sigma^2} \|y - h * x\|_2^2 - k \sum_i ([f_i(x) + \epsilon]^q). \quad (3.14)$$

Maximizing this likelihood is equivalent to minimizing the cost function

$$C(x, h) = \frac{1}{2} \|y - h * x\|_2^2 + \lambda \sum_i ([f_i(x) + \epsilon]^q), \quad (3.15)$$

where  $\lambda = k\sigma^2$ . This cost function is of the form given in (3.9). We can identify the data fidelity term,  $\frac{1}{2} \|y - h * x\|_2^2$ , and the regularizer,  $R[f(x)] = \sum_i [f_i(x) + \epsilon]^q$ . The regularizer  $R(f_i)$  is plotted in Fig. 3.9, for the parameters that were used in the experiments, and for typical values of the edge intensity.

This regularizer was chosen both because it favors sharp edges and because, for certain values of the parameters, the corresponding prior is rather close to actual observed distributions of the edges obtained from our edge extractor. A regularizer which favors sparse edges, such as this one, which is non-smooth and non-convex, typically favors piece-wise constant image estimates [165]. In our method this is quite visible in the first iterations (see Fig. 3.6), but is almost imperceptible in the final estimate, if there is no noise, because the final regularization is then very weak. When the image to be deblurred is noisy, the final regularization cannot

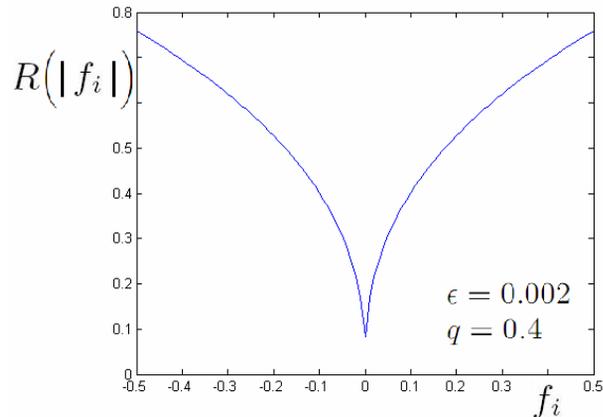


Figure 3.9: Plot of  $R(|f_i|)$ , computed for the values of the parameters used in the experimental tests.

be made so weak, to prevent the appearance of a strong amount of noise in the deblurred image, and that image will still retain some piece-wise constant character, as can be seen in the experimental results shown in Section 3.2.4. This is a compromise that has to be made, in our method and in several other ones, when one is simultaneously doing deblurring and denoising.

The cost function of (3.15) is similar to cost functions that have been used in other works on image deblurring (see [97, 99], for example). The well known total variation regularizer (in its discrete form) is a special case of our regularizer: it is obtained by using just two filters which compute horizontal and vertical differences, and by setting  $q = 1$  and  $\epsilon = 0$ . Despite these similarities with other methods, there are some important differences, in our approach, that are worth emphasizing:

- We use more elaborate edge detection filters than the simple horizontal and vertical differences used in many other works.
- We use  $q < 1$ , which makes the cost function non-convex, and therefore harder to optimize, but yields considerably better results.
- We use an optimization technique that leads to a good local minimum. As noted above, it is possible that the global minimum would not yield good deblurring results, but that is not the minimum that we seek in our method.

These three differences are crucial for the performance achieved by our method.

As was said above, we decrease  $\lambda$  during the optimization. Therefore, except for the last phase of the optimization,  $\lambda$  is not given by  $k\sigma^2$ . And in fact, even during that last phase,  $\lambda$  still is not, in general, given by that expression, because the noise  $n$ , besides allowing for possible noise in the blurred image, also allows for a mismatch between the estimated filter  $h$

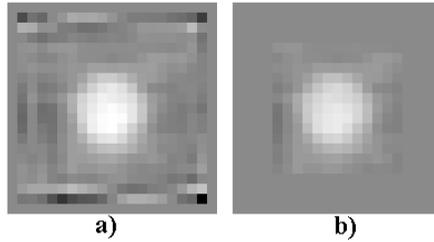


Figure 3.10: Filter estimate at an early stage of the method. a) With the safety zone. b) After discarding the safety zone.

and the true one. This mismatch leads to a difference between the reconstructed blurred image and the observed one, this difference being treated by the method as noise.

### Border effects

Since, in the first iterations of the method, the image estimates only contain the main image edges, the optimal filter estimates, at this stage, would not (even approximately) have a limited support. We constrain, in step 5 of the method, the filter estimate to have a limited support. This gives rise to undesired border effects in that estimate (see Fig. 3.10-a). These effects decrease in subsequent iterations, as the estimate of the deblurred image gets better. To avoid the influence of these effects, we use, in step 5 of the method, a “safety zone” with a width of a few pixels, around the desired filter support, and discard this zone, in each iteration, after the filter has been estimated (see Fig. 3.10-b).

A border effect of a different kind, this time relating to the estimated image, is due to the fact that, near the border of the image, a part of the filter’s support would fall outside the image. There is, therefore, a zone, adjacent to the image border, where estimation cannot be correctly performed. Estimation in this zone would typically lead to ringing artifacts parallel to the image borders. This border zone is not estimated, and is not included in the cost function of Eq. (3.15).

### Color and hyperspectral images

The method that we propose can also address color and hyperspectral images. A color image is a multichannel image which typically has three channels (red, green and blue). A hyperspectral image takes this concept much further, usually containing more than one hundred channels, which correspond to different frequency bands. In what follows we’ll speak of color images, but what is said can be extended, without change, to hyperspectral images.

To restore a color image, each of its channels should be restored. However, during the restoration process, one should take into account that the various channels should remain

aligned with one another. In the case of color images, even a small misalignment between the channels would lead to a significant visual degradation (color fringing along edges). Consequently, the color channels should be jointly processed, so that they maintain their alignment. A simple way to favor aligned solutions, is to apply the regularizer to the sum of the edge images from the three channels, instead of applying it separately to each channel. This corresponds to using the regularizer

$$R_f(x) = R[f(x)] \quad (3.16)$$

$$= R\left[\sum_c f^c(x)\right] \quad (3.17)$$

$$= \sum_i \left[\sum_c f_i^c(x) + \epsilon\right]^q \quad (3.18)$$

in which  $f^c$  is the edge image computed by applying  $f(\cdot)$  to the  $c^{th}$  color channel of  $x$ , and  $f_i^c(x)$  is the  $i$ th pixel of that image. The image

$$f(x) = \sum_c f^c(x) \quad (3.19)$$

is the edge image obtained, from the color image  $x$ , by separately applying the edge extractor  $f(\cdot)$  to each of the channels, and adding the results.

If we assume that all channels have suffered the same blurring degradation, we only need to estimate a single blurring filter. In this case, the cost function which is used to recover the color image is given by

$$C(x, h) = \frac{1}{2} \sum_c \|y_c - h * x_c\|_2^2 + \lambda \sum_i [f_i(x) + \epsilon]^q, \quad (3.20)$$

in which  $x_c$  is the  $c^{th}$  channel of the estimated image  $x$ ,  $y_c$  is the  $c^{th}$  channel of the degraded image  $y$  and  $f_i(x)$  is, as before, the  $i$ th pixel of the image  $f(x)$ . On the other hand, if we assume that different channels have suffered different blurs (which can happen, for example, if there is a significant amount of chromatic aberration from the lens that produced the image), then the cost function should be

$$C(x, h) = \frac{1}{2} \sum_c \|y_c - h_c * x_c\|_2^2 + \lambda \sum_i [f_i(x) + \epsilon]^q, \quad (3.21)$$

where  $h_c$  is the blur corresponding to the  $c^{th}$  channel.

## Multi-frame scenarios

The method can also be easily extended to address multi-frame scenarios, in which one has several frames, each with its own degradation, but all obtained from the same sharp scene. In this case, we can take advantage of the extra information that results from the existence of more than one blurred image of the same scene. Instead of using a single degradation model, the data connection term of the cost function must now take into account the degradation models of the various acquired frames. We index the frames with the subscript  $s$ , and assume that each acquired frame,  $y_s$ , was degraded by a different blurring operator,  $h_s$ , and a different additive noise,  $n_s$ :

$$y_s = h_s * x + n_s. \quad (3.22)$$

Assuming that all frames have Gaussian noises with the same variance, the multi-frame cost function is

$$C(x, h) = \frac{1}{2} \sum_s \|y_s - h_s * x\|_2^2 + \lambda R_f(x). \quad (3.23)$$

## Filter prior

The method was developed so that it would yield a good restoration performance without using any “strong” information on the blurring filter. Nevertheless, if prior information about the blurring filter is available, it can be used to advantage. If hard constraints on the blurring filter are known, they can be used in the filter optimization step (step 5 in Table 3.2). “Soft” constraints on the filter can be incorporated through the use of an additional regularizing term. If we assume a prior over the blurring filter

$$p(h) \propto e^{k_h R_h(h)}, \quad (3.24)$$

we are led to the cost function

$$C(x, h) = \frac{1}{2} \|y - h * x\|_2^2 + \lambda R[f(x)] + \lambda_h R_h(h), \quad (3.25)$$

in which  $\lambda_h$  and  $R_h(h)$  are, respectively, the regularizing parameter and the regularizing term of the blurring filter.  $R_h(h)$  can be a TV regularizer, as in [20, 97], for example, but other regularizers can also be used.

### 3.2.3 Quality measure

The measure that we used for evaluating the quality of the results of blind deblurring tests was the *increase in signal to noise ratio* (ISNR), similarly to what is commonly done in non-blind

deblurring. However, the computation of a meaningful ISNR in blind deblurring situations raises some special issues that we now address.

We start by recalling the basic concept of ISNR. Assume that  $x_0$  is an original image,  $y$  is a degraded version of that image and  $x$  is a recovered (enhanced) image, obtained from  $y$ . We start by defining the “signal” as image  $x_0$ , the “noise” of  $y$  as  $y - x_0$ , and the “noise” of  $x$  as  $x - x_0$ . The ISNR of the recovered image  $x$  relative to the degraded image  $y$  is, then, the difference between the SNR of  $x$  and the SNR of  $y$ . It can be computed, in dB, as

$$ISNR = 10 \log_{10} \frac{\sum_i (y^i - x_0^i)^2}{\sum_i (x^i - x_0^i)^2}, \quad (3.26)$$

where the superscript  $i$  indexes the images’ pixels, and the sums run through all pixels.

The special issues that arise in the computation of this measure in blind deblurring situations, are due to the following. The blind deblurring problem is strongly ill-posed. This means that non-regularized solutions have a large variability. There are two different kinds of variability that we need to distinguish here. One corresponds to changes in the *shape* of the estimated blurring filter’s PSF  $h$ , compensated by matching changes in the estimated image  $x$ . In this case, different estimated images will, in general, exhibit different amounts of residual blur and/or different artifacts (e.g. ringing), which affect their quality. These degradations should be taken into account by the quality measure. However, two forms of variability that are of a different kind are (1) affine transformations of the intensity scale of the filter, compensated by affine transformations of the estimated image, and (2) small translations of the blurring filter’s PSF, compensated by opposite translations of the estimated image. These degradations do not affect the quality of the deblurred image, and the restoration measure should be insensitive to them. As an example of the latter kind of variability, filter translations are visible in the positions of some of the estimated filters obtained in our tests, which are not fully centered (see Fig. 3.16 and Fig. 3.17).

The translation variability gets a bit more involved when we take into account that the images are processed in discrete form: There can be translations by a fractional number of pixels, which do not correspond to simple discrete translations of the discrete images, and involve interpolation between pixels.

To address these invariance issues, we have performed an image adjustment (spatial alignment and intensity rescaling) before comparing the images with the original sharp one. The estimated image was spatially aligned, and the pixel intensities were rescaled by an affine transformation, so as to minimize the image’s squared error relative to the original sharp image. As

a result, the noise energy of an image  $x$  relative to the sharp image  $x_0$  was given by

$$N(x) = \min_{a,b,\Delta_h,\Delta_v} \|ax_{\Delta_h,\Delta_v} + b - x_0\|_2^2, \quad (3.27)$$

in which  $x_{\Delta_h,\Delta_v}$  is the  $x$  image shifted by  $\Delta_h$  and  $\Delta_v$  pixels in the horizontal and vertical directions, respectively, and  $a$  and  $b$  are the parameters of the affine transformation.

The spatial alignment was performed with a resolution of 1/4 pixel, and with a maximum shift of 3 pixels in each direction. The estimated image  $x$  was saturated to the maximum and minimum values of the degraded image  $x_0$ , before alignment and rescaling. For the comparison to be fair, the alignment and rescaling were performed on both the deblurred image and the blurred one.

The ISNR of the recovered image  $x$  relative to the blurred image  $y$  was then computed as

$$ISNR = 10 \log_{10} \frac{N(y)}{N(x)}, \quad (3.28)$$

The sum on  $i$ , involved in the computation of  $N(\cdot)$  (see Eq. (3.27)) was restricted to the valid pixels. By “valid pixels” we mean all pixels of the image, except for the zone, adjacent to the image borders, where the estimation could not be correctly performed, as explained in Section 3.2.2, this zone being augmented by a width of 3 pixels to account for the maximal possible displacement due to the spatial alignment. The Matlab routines for image adjustment (alignment and rescaling) and for computing the ISNR are available at [http://www.lx.it.pt/~mscla/BID\\_QM.htm](http://www.lx.it.pt/~mscla/BID_QM.htm)

In order to be able to compare the reconstruction of the color images with the reconstruction of the corresponding grayscale ones, the ISNR of color images was computed on the luminance component  $I(x)$  through the expression used in the NTSC television standard, obtained from the image’s RGB channels  $x_r$ ,  $x_g$  and  $x_b$  through

$$I(x) = 0.2989x_r + 0.5870x_g + 0.1140x_b. \quad (3.29)$$

### 3.2.4 Experimental results

We tested the proposed method both on synthetic blurs and on actual blurred photos. We also performed comparisons with two other blind deblurring methods, published in the literature [20, 94].

Our method was implemented as outlined in Table 3.2, with the edge detection filters that are shown in Fig. 3.7. The PSF of the base filter that was used to generate these edge detection

filters (i.e., the PSF of the leftmost filter in Fig. 3.7) was

$$d_0 = \begin{bmatrix} 1 & 2 & 2 & 1 \\ -1 & -2 & -2 & -1, \end{bmatrix} / 12, \quad (3.30)$$

in which each matrix entry gives the value of a pixel of the PSF, and the arrangement of the matrix entries corresponds to the spatial arrangement of the corresponding pixels in the PSF. The other filters were obtained by rotating this base filter, with bicubic interpolation, by angles multiple of  $45^\circ$ . In the figures that we show ahead, the estimated image was first subjected to the affine transformation mentioned in Section 3.2.3, and was then saturated to the maximum and minimum values of the blurred image.

The blurred images were normalized so that black corresponded to  $-0.5$  and white (or maximum intensity, in the case of color channels of a color image) corresponded to  $0.5$ . Parameter  $\epsilon$  was set to  $0.002$ .

The sequence of values of  $\lambda$  was a geometric progression ( $\lambda_{n+1} = \lambda_n/r$ ), initialized at  $\lambda_1 = 2$ . The values that were used for  $r$  are given ahead for each specific case. For real-life photos, the iteration on  $\lambda$  was stopped on the basis of visual evaluation. In the synthetic experiments, we used the ISNR for deciding when to stop the iteration. The selection of the stopping point was rather important for noisy blurred images because, after the “optimal” point, the estimated image quickly became degraded with noise. For non-noisy blurred images, the method typically stopped progressing after a certain value of  $\lambda$ . After that value, the choice of the stopping point had almost no influence on the result.

All experiments were performed using the same sequence of values for parameter  $q$ :  $0.8, 0.8, 0.6, 0.6, 0.6, 0.6, 0.4, \dots, 0.4$ . The sequences of values of  $\lambda$  and  $q$  were experimentally found to be adequate for a wide variety of images and blurs, and can be used, without change, in most situations.

The support of the estimate of the blurring filter was limited to a square of size  $s \times s$  pixels, chosen to be slightly larger than the size of the actual blur (the value of  $s$  is given ahead for each case). We used a safety zone (see Section 3.2.2) with a width of three pixels around the support of the filter.

In most cases, the cost function was quadratic in  $h$ , and the optimization relative to  $h$  was performed by a relatively fast method (conjugate gradients, with 100 iterations for each value of  $\lambda$ ). In the cases in which this function was not quadratic (the cases in which a TV regularizer was used on the blurring filter and the cases in which we used a parametric model for the filter), gradient descent with adaptive step sizes [166] was used, because it can easily

deal with strongly non-quadratic functions and still is relatively fast. The optimization relative to the deblurred image  $x$  was also performed by gradient descent with adaptive step sizes (150 iterations for each value of  $\lambda$ ), because the cost function is a strongly non-quadratic, non-smooth function of  $x$ , and this method can also easily deal with non-smooth functions. The numbers of iterations mentioned above were experimentally chosen so as to achieve a good convergence of the corresponding optimizations. These numbers are not crucial: they could have been increased without negatively impacting the deblurring results, but such an increase would obviously also have increased the running time of the algorithm, with no significant advantage.

On an Intel Core 2 Duo system running at 2 GHz, programmed in Matlab and running on only one of the chip's processors, an iteration of the method, corresponding to one value of  $\lambda$ , took, for monochrome images of size  $256 \times 256$ , about 30 seconds, when conjugate gradients were used for the optimization relative to  $h$ . For color images, each iteration took about 70 seconds, also with conjugate gradient optimization of  $h$ .

The total deblurring time depended on the number of iterations in  $\lambda$ , which depended on the ratio  $r$  and on the stopping point of the  $\lambda$  sequence. Experiments with blurred monochrome photos (Section 3.2.4) were processed using  $r = 3$  and with around 10 values of  $\lambda$ , having taken about 5 minutes to be processed. Synthetic experiments were processed using a lower ratio of  $r = 1.5$  and approximately 55 and 23 iterations for non-noisy and noisy experiments, respectively, having taken about 28 and 12 minutes, respectively. Results with only slightly lower quality were obtained with a ratio of  $r = 3$  and with a somewhat higher stopping value for  $\lambda$  (see [9]), resulting in a much shorter processing time of 7.5 minutes for non-noisy images. For color images, all these times were multiplied by about 2.5.

### Synthetic degradations

In this section we first describe the main experiment, which was intended at showing that the proposed method can effectively deal, in a blind way, with a large variety of images and of blurs. After that, we describe additional tests that were performed to check some other aspects of the method's performance. All of these experiments were run with a ratio of  $r = 1.5$  in the sequence of values of  $\lambda$ . The iteration was run up to  $\lambda_{55} = 6.2 \times 10^{-10}$ , and the best stopping iteration was chosen based on the values of the ISNR measure, computed as described in Section 3.2.3.

The main experiment was performed with the five grayscale images shown in Fig. 3.11. Each image was blurred with each of the blurring filters shown in Fig. 3.12 (for a better

Blur	#1	#2	#3	#4	#5	#6	#7
Without noise	6.48	5.38	6.43	5.87	6.29	5.82	3.21
With noise	3.69	4.19	3.61	5.46	5.50	6.34	1.99

Table 3.3: Summary of the ISNR values obtained with our method, with no constraints and no regularization on the estimated filter. Each entry gives the average of the ISNRs obtained for the five test images, under the indicated conditions.

visualization, these filters, as well as the filter estimates to be presented further ahead, are shown with a band of zero-valued pixels around them). All filters were normalized to a DC gain of 1. The PSF of filter #1 is a uniform-intensity circle, and simulates an out-of-focus blur. Filter #2 simulates a linear, uniform motion blur. The PSF of filter #3 is a uniform-intensity square. Filter #4 was formed by choosing random pixel intensities with a uniform distribution in  $[-0.3, 0.7]$ , post-normalized to a DC gain of 1. Filter #5 simulates an irregular motion blur. Filter #6 corresponds to a circular motion blur, and was chosen because its frequency response has somewhat a non-lowpass character, and therefore is rather different from the most common blurs. Filter #7 is Gaussian, with a standard deviation of two pixels.

Filter #1 had a radius of  $l$  pixels, and filters #2, #3 and #4 had a size of  $l \times l$  pixels. We used different values of  $l$  for different cases:  $l = 9$  for “Lena”, “Barbara” and “Testpat1”, and  $l = 11$  for “Cameraman” and “Satellite”. For the “Lena”, “Barbara” and “Testpat1” images, the size of the estimated filters was set to  $s \times s$  pixels, with  $s = 15$ ; for “Cameraman” and “Satellite” we used  $s = 17$ , due to the larger size of the blurs. No constraints were imposed on the estimated filters. Each blurred image was used both without noise and with Gaussian i.i.d. noise at a *blurred signal to noise ratio* (BSNR) of 30 dB.

Figures 3.13-b to 3.15-b show a sample of the 70 blurred images used in this experiment, and the second rows of Figs. 3.13 to 3.15 show deblurred images that were obtained by the method. Table 3.3 gives a summary of the results, in terms of ISNR. Detailed results are given in Appendix 3.2.7. We can see that the method yielded, in almost all cases, a significant improvement in image quality. The blurring filters also were reasonably recovered, especially when there was no noise (see Figs. 3.16 and 3.17). As mentioned in Section 3.2.2, the images recovered in noisy situations had a slight piecewise-constant character.

The worst results corresponded to the Gaussian blur (#7). This has a simple explanation. Although, visually, the Gaussian filter doesn’t look worse than, say, the square or the circular ones, its frequency response decays much faster than those of the other filters. At the spatial frequency of  $f_{max} = 1/(2 \text{ pixels})$ , which is the maximum frequency allowed by the sampling

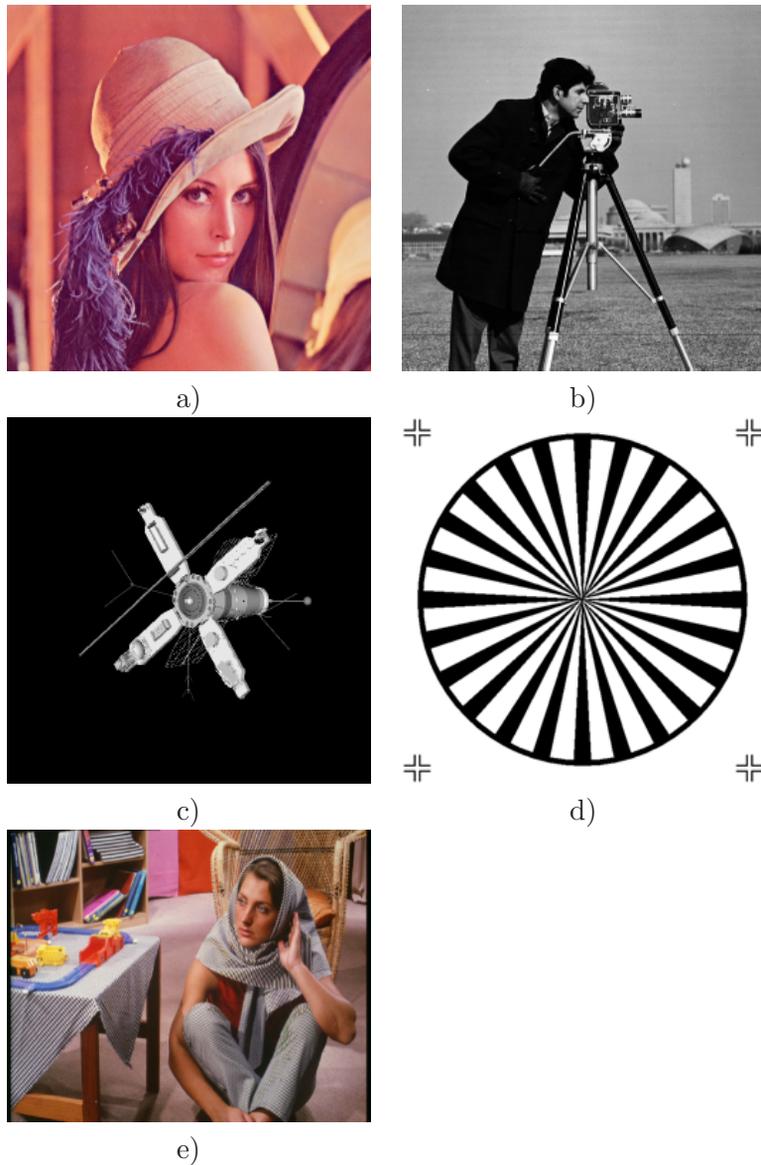


Figure 3.11: Set of images used for synthetic experiments. a) “Lena”. b) “Cameraman”. c) “Satellite”. d) “Testpat1”. e) “Barbara”.

theorem, the Gaussian filter presents an attenuation above 150 dB. At the frequency  $f_{max}/2$ , the attenuation is of more than 40 dB. This means that the filter eliminates essentially all the high frequency content from the original image, and there is no way to recover it (recall that, even without added noise, the blurred images do have noise due to rounding errors).

Our second test concerned the use of constraints on the blurring filter. In this test we used all blurring filters except #4 and #5, for which no simple constraints existed. For filter #1, we used as parametric model a uniform circle, with the diameter as parameter. Since the diameter could take any real, positive value, the filter’s PSF was obtained by computing the pixel intensities according to the fraction of each pixel that was covered by the circle with the

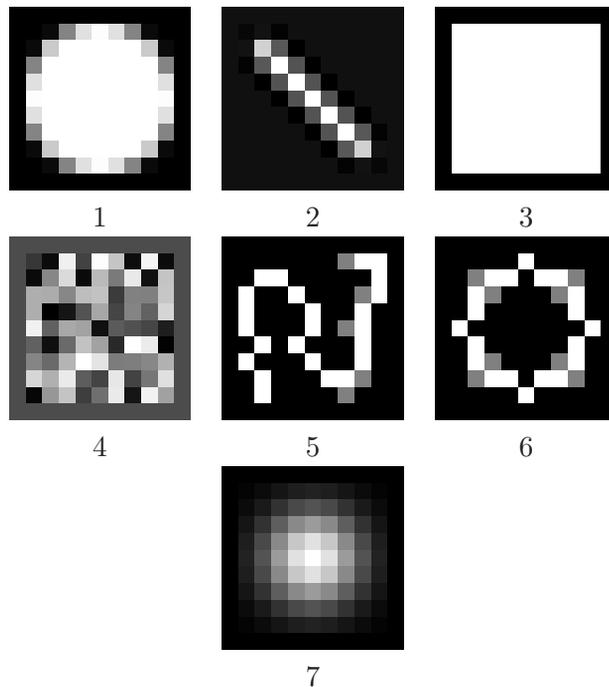


Figure 3.12: Set of blurring filters used in synthetic experiments. #1 - Out-of-focus blur. #2 - Linear motion blur. #3 - Uniform square. #4 - Random square. #5 - Nonlinear motion blur. #6 - Circular motion blur. #7 - Gaussian.

prescribed diameter; the gradient of the cost function relative to the diameter was computed taking this model into account. For filter #2, the constraint was symmetry relative to the central pixel. For filters #3, #6 and #7, the constraint was symmetry relative to the four axes that make angles multiple of  $45^\circ$  with the horizontal axis. The tests were run on all images of Fig. 3.11, without noise and with noise at 30 dB BSNR. Figures 3.13-d,f to 3.15-d,f show a sample of the noisy estimates (we don't show the noisy blurred images because, visually, they are almost indistinguishable from the non-noisy ones). Table 3.4 shows a summary of the ISNR values (the complete list of values is given in Appendix 3.2.7). It can be seen that, in most cases, the use of constraints improved the quality of the results. In some cases the improvement was quite impressive, and, on the other hand, in a very few cases, there was a very slight decrease in quality.

As mentioned above, we used a parametric model as constraint for filter #1. Therefore, in this case, the estimation of the blurring filter actually consisted of the estimation of its single parameter (the diameter). In the noiseless case, the estimated diameter values were of 8.84, 10.96, 10.96, 8.56 and 8.94 pixels, respectively, for the five images. These values compare well with the true diameter values of 9, 11, 11, 9 and 9 pixels, respectively.

A third test, performed only on the “Lena” image, concerned the use of a TV regularizer on the estimate of the blurring filter's PSF (without constraints). For this test, we used the

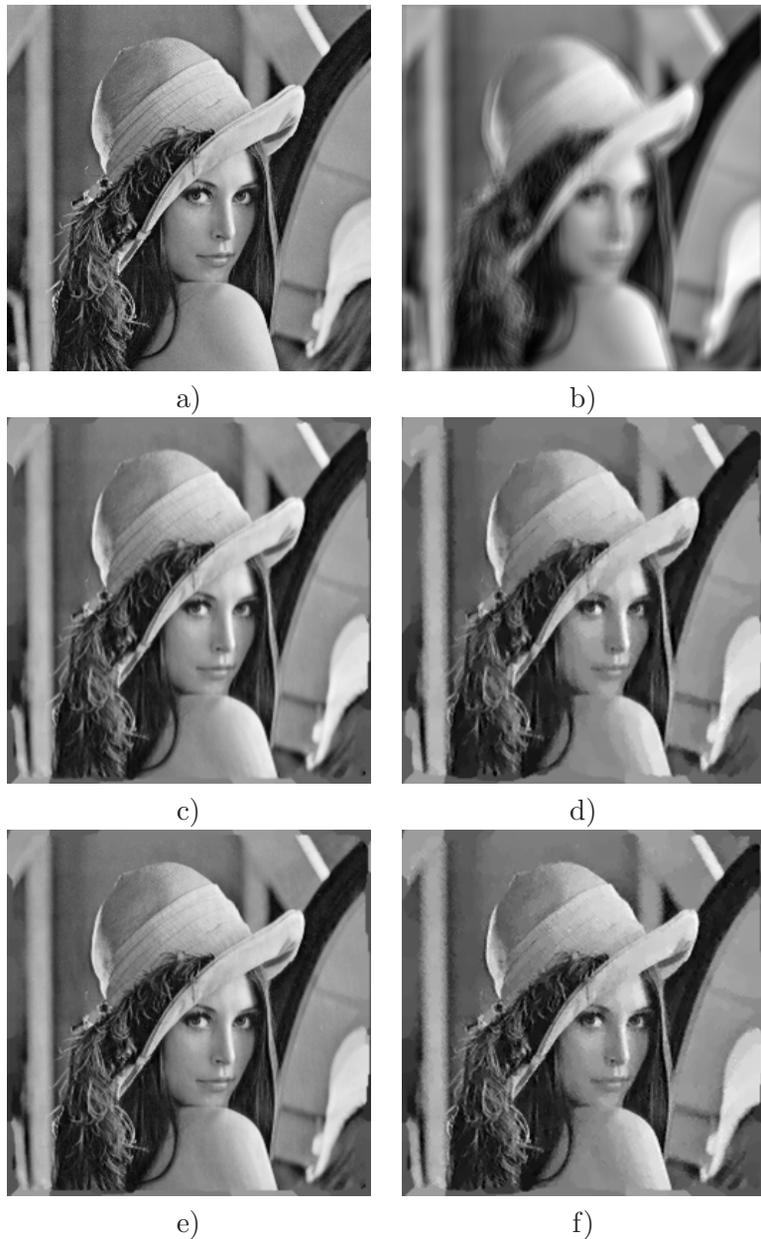


Figure 3.13: “Lena” image with blur #5. a) Sharp image. b) Blurred image. Next rows: Deblurred images. Left: Without noise. Right: With noise. Second row: Without filter constraints or regularization. Third row: Using TV regularization on the blurring filter.

cost function of (3.25), with a ratio  $\lambda_h/\lambda = 100$ . The ISNR values, given in Table 3.7 in Appendix 3.2.7, show that the use of the regularizer improved the SNR, in all cases, relative to the estimation without constraints. Although this regularizer yielded an improvement in the results, it was used here only as an example. Other regularizers may be more useful for specific situations.

The ISNR values attained in the tests described so far can be considered rather good, taking into account that the method under test is blind. In fact, these ISNR values are relatively close to the values attained by state of the art non-blind methods. For example, the recent work



Figure 3.14: “Barbara” image with blur #2. a) Sharp image. b) Blurred image. Next rows: Deblurred images. Left: Without noise. Right: With noise. Second row: Without filter constraints or regularization. Third row: With constraints.

[82] presents the ISNR values attained by several state of the art non-blind methods on the restoration of the “Lena” image blurred with a  $9 \times 9$  uniform blur, with 30dB BSNR. Those values range from 5.58dB, corresponding to the method from [102], to 6.1dB, attained by the method from [82]. Our method blindly attained, in the same problem, about 1.5 dB less (4.07dB without constraints, 4.47dB with constraints and 4.27 dB with TV regularization). The majority of the non-blind deblurring approaches have been tested with blurs produced with circular convolutions, while our method used linear convolutions. The circular convolutions introduce, in the original image, an artificial, long, straight, horizontal edge, corresponding to making the image top and bottom adjacent to each other, and also introduce a similar artificial

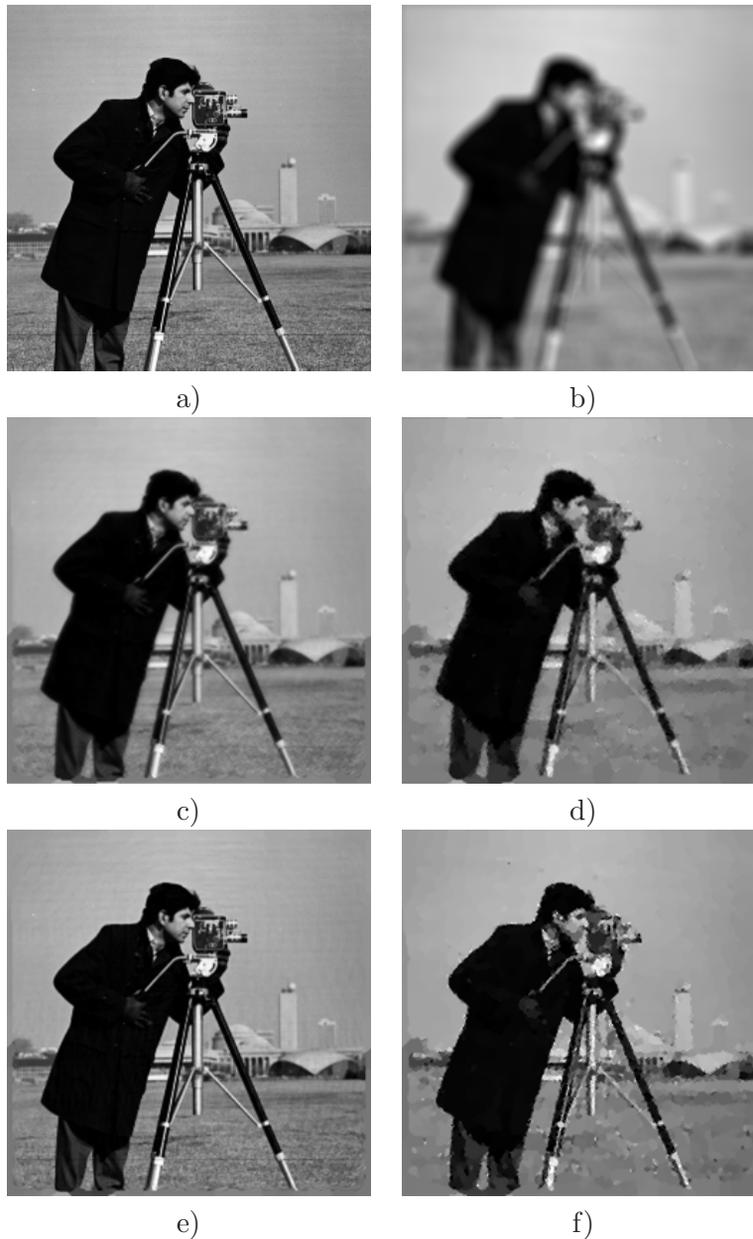


Figure 3.15: “Cameraman” image with blur #1. a) Sharp image. b) Blurred image. Next rows: Deblurred images. Left: Without noise. Right: With noise. Second row: Without filter constraints or regularization. Third row: With constraints.

vertical edge. The presence of these artificial edges would have helped our method to better estimate the blurring filter, slightly improving the deblurring results (evidence of this can be found in some results presented later, in Table 3.6, in which we find improvements between 0.03dB and 1.18dB, for our method, due to the use of circular convolutions).

We performed another test to check the method’s performance on color images. We used the cost function given by (3.20). For this test, we used the color “Lena” and “Barbara” images, with all the blurs described above, without constraints. Figures 3.18 and 3.19 show some results. The ISNR of the results is shown in Tables 3.7 and 3.10, in Appendix 3.2.7.

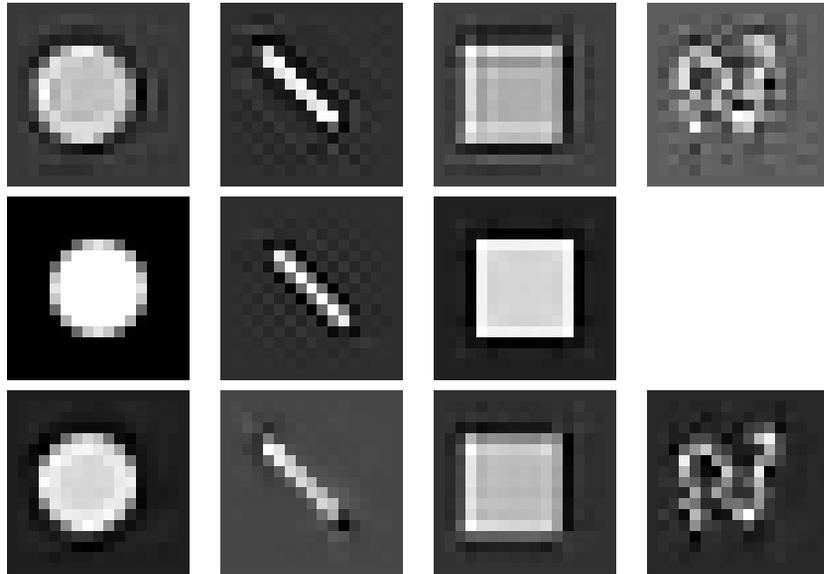


Figure 3.16: Filter estimates, for non-noisy experiments. First row: without restrictions. Second row: with restrictions. Third row: with TV regularization.

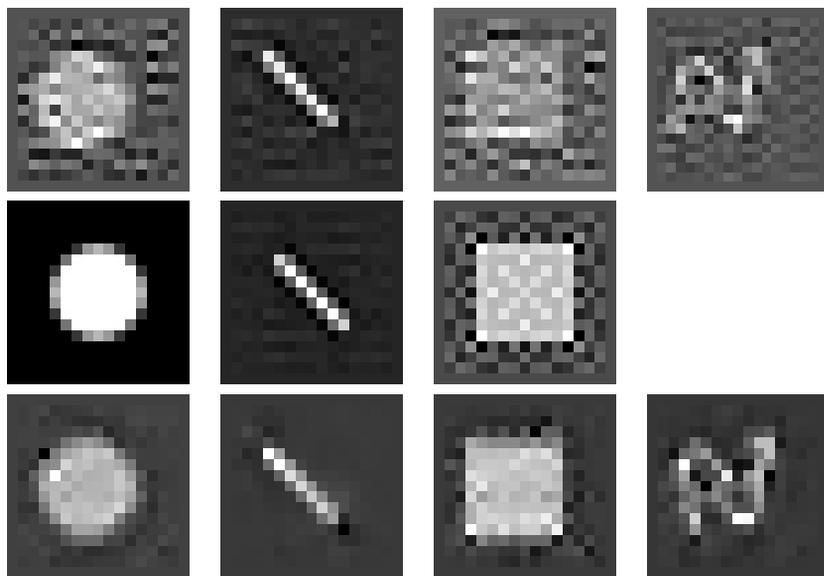


Figure 3.17: Filter estimates, for noisy experiments. First row: without restrictions. Second row: with restrictions. Third row: with TV regularization.

The results on color images were, in general, slightly better (and, in a few cases, significantly better) than those for grayscale images. This is not surprising, since the three color channels of color images contain more information about the blur than the single channel of grayscale images.

For testing the deblurring of multi-frame images, as described in Section 3.2.2, we used the “Lena” image with two blurred frames, both with motion blurs. Both blurs had a length of 11 pixels, and they had directions that made angles of  $45^\circ$  and  $135^\circ$ , respectively, with the

	Blur	#1	#2	#3	#6	#7
Without noise	Without Constraints	6.48	5.38	6.43	5.82	3.21
	With Constraints	<b>12.01</b>	<b>8.61</b>	<b>10.22</b>	<b>11.94</b>	<b>4.72</b>
With noise	Without Constraints	3.69	4.19	3.61	6.34	1.99
	With Constraints	<b>5.34</b>	<b>6.64</b>	<b>5.25</b>	<b>10.09</b>	<b>3.14</b>

Table 3.4: Comparison of the results obtained without and with constraints on the estimated filter. Each entry gives the average of the ISNRs obtained for the five tested images, under the indicated conditions. The best result for each case is shown in bold.



Figure 3.18: Deblurring of a color image. a) Sharp image. b) Image degraded with blur #2 and 30dB of noise. c) Image estimate without noise. d) Image estimate with noise.

horizontal axis. We used the noiseless blurred images, and also noisy images with BSNRs of 40 dB and 30 dB. To assess the advantage of using multiframe deblurring, we compared the images recovered in multiframe mode with the images recovered, in single-frame mode, from each of the two frames. Figure 3.20 shows the results for the 30dB noise level. Table 3.5 gives the ISNR values. We can see that the multiframe method was advantageous in situations with noise, but had no clear advantage in the situation without noise.



Figure 3.19: Deblurring of a color image. a) Sharp image. b) Image degraded with blur #5 and 30dB of noise. c) Image estimate without noise. d) Image estimate with noise.

Mode	Blur directions	Noise BSNR	SNR of the result
Single-frame	45°	-	15.50
Single-frame	135°		<b>16.22</b>
Multi-frame	45°,135°		16.09
Single-frame	45°	40dB	15.35
Single-frame	135°		15.14
Multi-frame	45°,135°		<b>15.88</b>
Single-frame	45°	30dB	14.12
Single-frame	135°		13.83
Multi-frame	45°,135°		<b>15.19</b>

Table 3.5: Multi-frame performance versus single-frame performance (“Lena” image with motion blurs of 11 pixels). The last column shows the SNR (in dB) of the deblurred images, relative to the original sharp one. The best results for each case are shown in bold.

### Blurred photos

Besides testing the method on synthetic degradations, we also applied it to real-life blurred photos. We used two different color scenes (Figs. 3.21-a and 3.22-a). The corresponding grayscale images were also processed (grayscale results are presented in Section 3.2.4).

We addressed two kinds of real-life degradations: the pictures in Fig. 3.21-d and Fig. 3.22-c were purposely taken with the camera wrongly focused, while in Fig. 3.21-e the camera was purposely rotated in the horizontal direction while the photo was being taken, to produce a motion blur. The photos of Fig. 3.21 were taken with a Canon S1 IS camera, and were coded

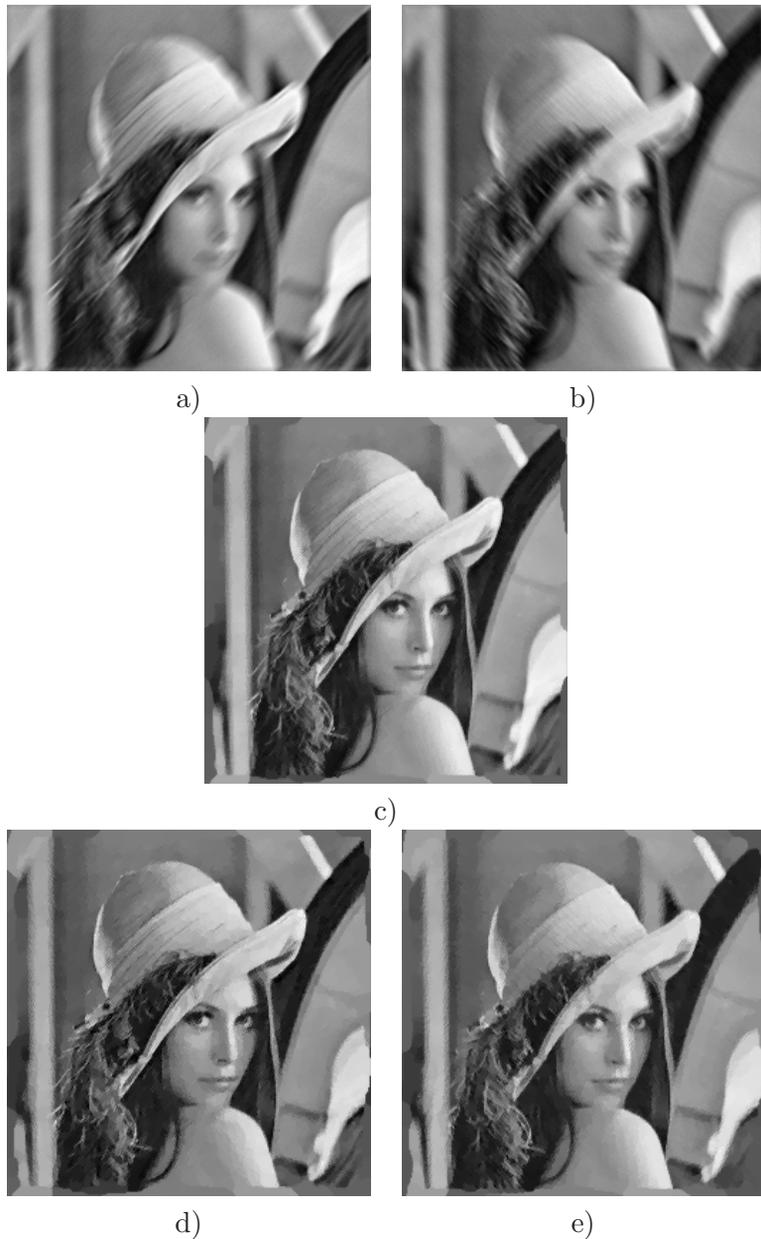


Figure 3.20: Multi-frame and single-frame estimates obtained with 30dB BSNR. First row: Degraded images. Second row: multi-frame image estimate. Third row: single-frame image estimates.

in JPEG format (this camera can't save images in RAW format). The photos of Fig. 3.22 were taken with a Panasonic DMC-FZ18 camera, and were coded in RAW format (i.e., using actual sensor data, after the demosaicing that interpolated colors among pixels).

The noise that was present in the photos was quite significant. The cameras that we used have small image sensors (about 1 cm diagonal), which yield images with a significant amount of noise (much larger than the 30 dB that we used in the synthetic experiments). In the images with full sensor resolution, the noise was too high to allow any significant improvement by means of the deblurring method. In order to simulate cameras with larger sensors, we reduced

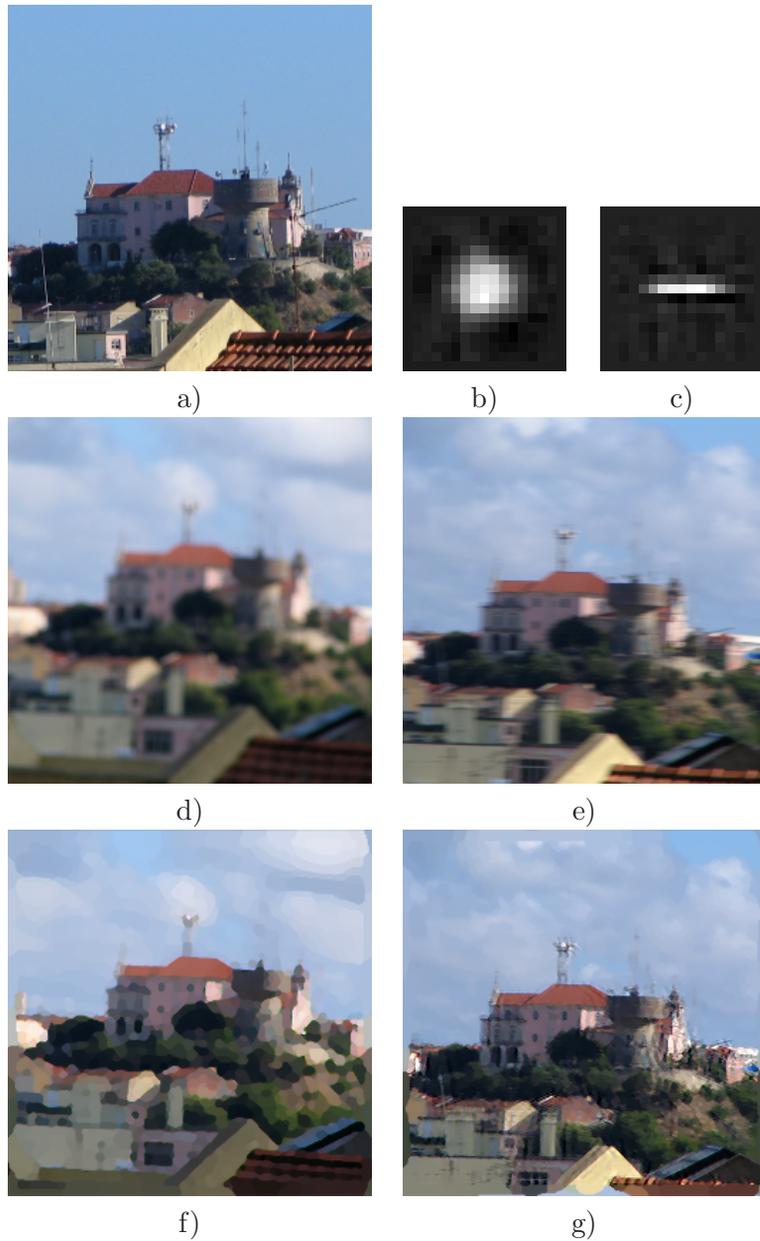


Figure 3.21: Results with real-life blurred photos. a) Sharp photo of the scene. b) and c) Filter estimates. Next rows: Left - out-of-focus blur. Right - motion blur. Second row: Blurred photos. Third row: Deblurred images.

the resolution of the camera images by averaging in squares of  $6 \times 6$  pixels, for the Canon camera, and of  $9 \times 9$  pixels, for the Panasonic camera. This reduced the noise to acceptable levels.

The size of the blur estimate was limited to a square of size  $15 \times 15$  pixels. We used a sequence of  $\lambda$  values with a ratio of  $r = 3$ , and truncated this sequence at  $\lambda_9 \approx 3.05 \times 10^{-4}$  for the experiments of Fig. 3.21 and at  $\lambda_{10} \approx 1.02 \times 10^{-4}$  for the experiment of Fig. 3.22.

All the recovered images were significantly sharper and had more visible details than the blurred ones, even though they had somewhat a “patchy” look, corresponding to somewhat a

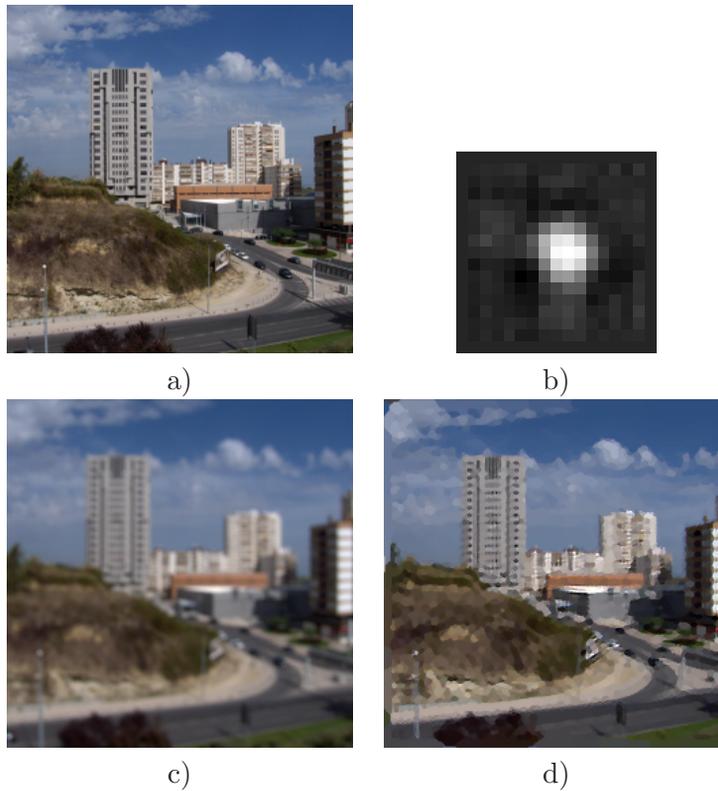


Figure 3.22: Results with an actual blurred photo. a) Sharp photo of the scene. b) Filter estimate. c) Blurred photo. d) Deblurred image.

piecewise-constant character. As had happened with the synthetic degradations, the restoration was slightly better for color photos than for monochrome ones (compare Fig. 3.21-f with Fig. 3.23-d, and Fig. 3.22-d with Fig. 3.24-d).

The results obtained with these photos were of lower visual quality than those obtained with the synthetic blurs. Two of the reasons for this probably were:

- The blurs that were present in the photos probably didn't exactly follow the model of Eq. (3.8). One of the main reasons may have been the presence of non-linearities in the image acquisition. It is known that image sensors may not be perfectly linear, due to the presence of anti-blooming circuits, for example. Furthermore, in the case of the Canon camera, for which we didn't have access to RAW data, we suspect that the camera also performed some nonlinear operations like denoising, sharpening and gamma compensation. In an actual application (for example, if deblurring is to be incorporated in the image processing performed in the camera itself), it should be possible to avoid, or to compensate for, these nonlinearities.
- The noise produced by image sensors is not Gaussian and (probably more important) its intensity is not independent from the image's intensity, contrary to the assumptions of

our method.

### Comparison with other methods

We compared our method with two other methods from the literature: APEX [94, 95] and the method from [20] (which we shall call YK method). These were the only two methods for which we were able to obtain implementations. APEX was simple enough for us to implement ourselves within useful time, and the authors of the YK method kindly provided a demo, coded in C++.

The APEX method [94, 95] is quite fast, but is limited to blurs which belong to the Levy family. This is a family with just two parameters, in which all functions have circular symmetry, and which encompasses the Gaussians. The method has two regularizing parameters ( $M$  and  $s$ ), whose values we have set to those recommended by the author ( $M = 500$  and  $s = 0.01$ ). The method has two further parameters (designated by  $A$  and  $t$ , respectively). For  $A$ , we used the values 2.00, 2.25, 2.50, ..., 7.75, 8.00, which cover the recommended interval. Parameter  $t$  can be varied between 1 and 0,  $t = 1$  corresponding to the blurred image, and  $t = 0$  to a “completely deblurred” one. We used the values 1, 0.75, 0.5, 0.25 and 0. For synthetic blurs, the ISNR values were computed for all combinations of values of  $A$  and  $t$ , and the best combination was selected. For real-life blurred photos, the best pair was chosen by visual inspection, since no ISNR values could be computed.

The YK method does not constrain the blur PSF, but assumes that it is piecewise smooth (and, from the comments made in [20], one can see that the method has some bias toward piecewise constant PSFs). The method has four parameters that must be manually chosen. We started by trying the values used in [20] but, with our blurred images, this produced results with very strong oscillatory artifacts. After several tests, we chose the following values, which seemed to produce the results with fewest artifacts: 1 and 2000 for the regularizing parameters of the image and of the PSF, respectively; 0.1 and 0.001 for the threshold parameters of the diffusion coefficients of the image and of the PSF, respectively. We should note that our tests were severely limited by the fact that the deblurring of each image, with this method, took about 12 hours, despite the fact that the method was coded in C++. Besides preventing us from doing a more extensive search of parameter values, this also prevented us from testing the method on noisy synthetic degradations.

The APEX method uses circular convolutions in the blurring model. The YK method uses convolutions computed from products of discrete cosine transforms (DCTs). Our method normally uses linear convolutions, but can easily be modified to use circular or DCT-based

convolutions. We tested all methods with degradations produced with linear convolutions, and also tested both our method and APEX with degradations produced with circular convolutions. Given the very poor results (to be seen ahead) obtained by the YK method with DCT-based convolutions, and the limited time that was available to us, we didn't consider it necessary to test our method with DCT-based convolutions.

For the comparison, we used both synthetic and real-life degradations. The synthetic degraded images were obtained from the grayscale “Lena” image, with blurs #3, #4, #5 and #7. Blur #7 is Gaussian, and therefore is within the family of blurs for which the APEX method is appropriate. Blur #3 is piecewise constant, and therefore appears to be appropriate for the YK method. Blur #5 may also be considered piecewise constant. Blur #7 is smooth, and therefore appears to be at least partially appropriate for that method, too. The real-life degraded images that we used were grayscale versions of two of the photos presented above, one with a motion blur and the other with an out-of-focus blur. The size of the estimated blurring filter was limited to  $15 \times 15$ , both in our method and in the YK one. APEX doesn't assume a limited size of the blurring filter.

Table 3.6 shows the ISNR values obtained with the synthetic degradations. We can see that, with linear convolutions, our method clearly surpassed the other two methods. APEX only yielded a significant improvement in the image quality for the Gaussian blur, as expected. When we used blurs produced with circular convolutions, which are the most appropriate ones for APEX, the results of our method improved slightly, in all cases but one. The results of APEX improved significantly in the case of the Gaussian blur (especially without noise), and improved only slightly for the other blurs. The APEX method only surpassed ours in the case of the Gaussian blur without noise. Even in that case, the advantage of APEX over our method was only slight, despite the fact that the Gaussian blur is within the class of blurs for which APEX was designed, and that APEX was only estimating two parameters, while our method was estimating 225.

The performance of the YK method was rather poor, which was somewhat a surprise to us. A possible explanation for the difference between these results and the ones presented in [20] is that, while the tests described in that reference involved the estimation of PSFs with up to 49 parameters, the tests performed by us involved PSFs with  $15 \times 15 = 225$  pixels. However, we should note that, with four parameters to be manually chosen in that method, it is hard to find a good set of values, and it may be questioned whether the method really should be considered blind. Still another explanation could be the fact that, as the algorithm's authors themselves say, the algorithm suffers quite severely from the problem of local minima.

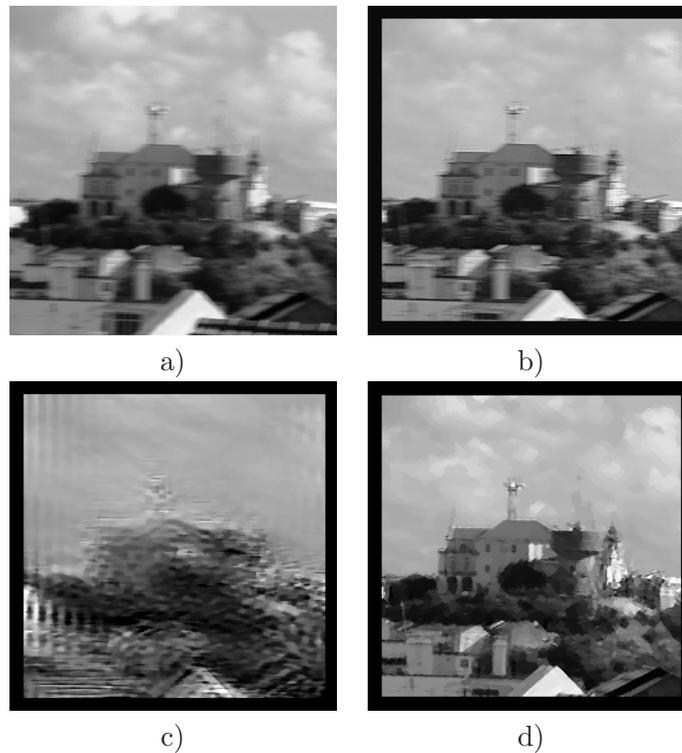


Figure 3.23: Deblurring of an actual photo with several methods. a) Blurred photo. b) Image deblurred with APEX. c) Image deblurred with the YK method. d) Image deblurred with our method.

Figures 3.23 and 3.24 show the results obtained with actual photos. The APEX method produced almost no improvement in the motion blur case, and produced a moderate improvement in the out-of-focus blur case. Both of these results are understandable: Motion blurs are well outside the family of blurs for which the method is appropriate, and the PSF of the out-of-focus blur that exists in the second photo seems not to be too far from a Gaussian (see Fig. 3.22-b), and probably is close to the family of blurs for which APEX is appropriate. Nevertheless, the result produced by our method was sharper, even in this case.

The results produced by the YK method show the kind of problem that affected many of the results obtained with that method: there were strong oscillatory artifacts, even with the parameter values that we used. We should note, however, that, for the test images sent by the method's authors, the method did yield results similar to the ones published by them. This gives us some confidence that the method was correctly applied. We have already speculated above about possible reasons for the poor results obtained in our tests with that method.

### Final comments

We stress that, although our method involves a few parameters, only one of them is crucial (and only for noisy blurs): the stopping point of the iteration. In fact, our tests have shown

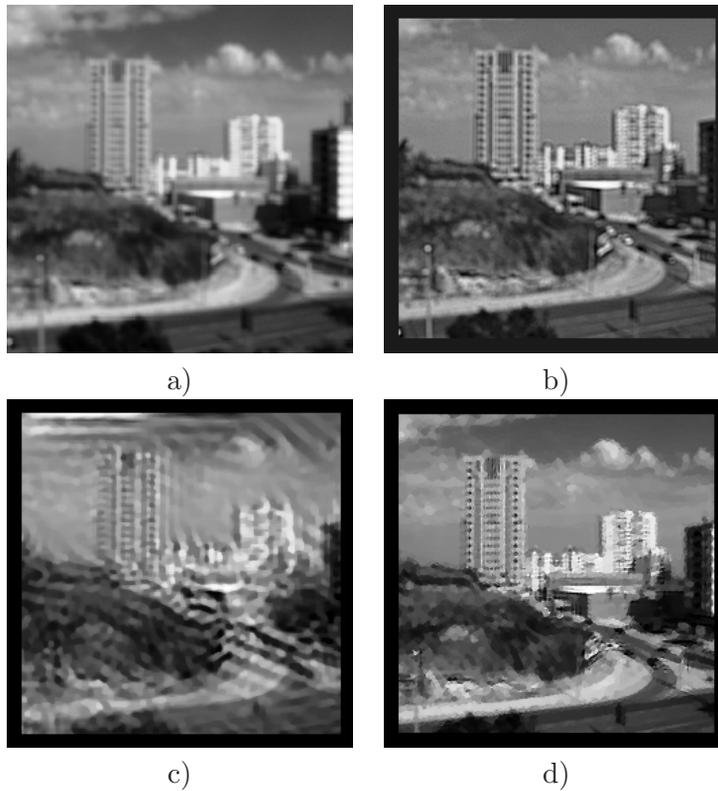


Figure 3.24: Deblurring of an actual photo with several methods. a) Blurred photo. b) Image deblurred with APEX. c) Image deblurred with the YK method. d) Image deblurred with our method.

that the same sequences of values of  $\lambda$  and of  $q$  yielded good results for a wide range of images and of blurs<sup>1</sup>. Therefore, these sequences can be fixed a priori, without knowing which image and which blur are to be handled. This being said, we should note that, by tuning these parameters, somewhat better results can be obtained than the ones that we have shown in this paper. In several practical applications, it may be quite possible to pre-tune such parameters. For example, in a digital camera, pre-tuned values can exist for different apertures, focal lengths, etc.

The choice of the stopping point of the iteration is not crucial for non-noisy images, as we said above. For noisy images, we don't have any good stopping criterion yet. The choice of the stopping point is very similar, in character, to the choice of the  $t$  value in the APEX method, and is a known difficult problem, even for non-blind methods, for which several solutions have been proposed (e.g. [110]). None of these solutions showed to be robust enough for our application.

---

<sup>1</sup>We have used two different values of the ratio  $r$  in different situations, but the smaller value can be used in all cases, with the only disadvantage of needing more iterations, and therefore taking a longer time, to reach the final value.

### 3.2.5 Conclusions

We have presented a method for blind image deblurring. The method differs from most other existing methods by only imposing weak restrictions on the blurring filter, being able to recover images which have suffered a wide range of degradations. Good estimates of both the image and the blurring operator are reached by initially considering the main image edges, and progressively handling smaller and/or fainter ones. The method uses an image prior that favors images with sparse edges, and which incorporates an edge detector that was specially developed for this application. The method can handle both unconstrained blurs and constrained or parametric ones, and it can deal with both single-frame and multi-frame scenarios.

Experimental tests showed good results on a variety of images, both grayscale and color, with a variety of synthetic blurs, without and with noise, with real-life blurs, and both in single and in multi-frame situations. The use of information on the blurring filter and/or of multi-frame data, when available, typically led to improvements in the quality of the results.

We have adapted the ISNR measure to the evaluation of the restoration performance of BID methods. The restoration quality of our method was visually and quantitatively better than those of the other methods with which it was compared.

So far, whenever the blurred image has noise, the processing has to be manually stopped, by choosing the iteration which yields the best compromise between image detail and noise or artifacts. An automatic stopping criterion will obviously be useful. This is a direction in which further research will be done.

The method can be extended in other directions: For example, (1) to address problems in which we aim at super-resolution, possibly combined with deblurring, and (2) to deblur images containing space-variant blurs (for example, a sharp scene containing one or more motion-blurred objects, or a scene containing objects at different distances from the camera, with different out-of-focus blurs). This latter extension has already shown useful results [143].

Finally, on a more theoretical level, but with possible practical implications, is the problem that we mentioned above, that the best deblurring solutions generally don't correspond to the global minimum of the cost function. This apparently means that a more appropriate cost function should exist. If it were found, it would probably lead to a better deblurring technique, both in terms of speed and of the quality of the results. This clearly is an important research direction.

### 3.2.6 Acknowledgment

We wish to acknowledge M. Kaveh and Y.-L.You for having provided the code of the BID method proposed in [20]. We also thank to R. J. Plemmons and D. Chen for having provided the satellite image that we have used. We wish to acknowledge very useful discussions with our colleagues J. Bioucas-Dias and M. Figueiredo. We are grateful for the reviewers' comments, which significantly contributed to improve the quality of this paper.

This work was partially supported by the Portuguese FCT and by the "Programa Operacional Sociedade do Conhecimento (POS-Conhecimento)", comparticipated by the FEDER European Community fund, under project POSC/EEA-CPS/61271/2004 and under the PhD fellowship SFRH/BD/23919/2005.

### 3.2.7 Appendix

#### Tables

Tables 3.7 to 3.11 show detailed ISNR values from several of the tests mentioned in this paper.

#### Gradients

Gradients of the method's cost function are explicitly given in this appendix. The cost function is

$$C(x, h) = \frac{1}{2} \|y - h * x\|_2^2 + \lambda \sum_i [f_i(x) + \epsilon]^q, \quad (3.31)$$

in which

$$f(x) = \sqrt{\sum_{\theta \in \Theta} (d_\theta * x)^2}. \quad (3.32)$$

Cost function (3.31) is quadratic on the filter  $h$ :

$$C(x, h) = \frac{1}{2} \|\text{vec}(y) - X \text{vec}(h)\|_2^2 + \lambda R(f(x)), \quad (3.33)$$

in which  $X$  is a matrix, corresponding to the linear operation of convolving an image with  $x$ , and  $\text{vec}()$  is the operation that vectorizes a matrix lexicographically. Considering (3.33), the gradient of (3.31) with respect to  $h$  is given by

$$\frac{\partial C(x, h)}{\partial \text{vec}(h)} = X^T [\text{vec}(y) - X \text{vec}(h)], \quad (3.34)$$

in which  $X^T$  is the transpose of  $X$ . Matrices  $X$  and  $X^T$  were not explicitly computed. Instead,

their corresponding operations were performed in the frequency domain:

$$X\text{vec}(w) = \text{ifft2}[\text{fft2}(x) .* \text{fft2}(w)], \quad (3.35)$$

$$X^T\text{vec}(w) = \text{ifft2}[\text{conj}(\text{fft2}(x)) .* \text{fft2}(w)], \quad (3.36)$$

in which  $\text{fft2}()$  is the 2D discrete Fourier transform and  $\text{ifft2}()$  is its inverse. The notations  $.*$  and  $\text{conj}()$  denote the point-wise product and complex conjugation operations, respectively.

Let  $H$  and  $D_\theta$  be the matrix operators corresponding to convolving an image with  $h$  and  $d_\theta$ , respectively. Cost function (3.31) can also be written as

$$\frac{1}{2} \|\text{vec}(y) - H\text{vec}(x)\|_2^2 + \lambda R[f(x)], \quad (3.37)$$

in which

$$R(f(x)) = \sum_i \left[ \left( \sqrt{\sum_{\theta \in \Theta} (D_\theta \text{vec}(x))^2} \right)_i + \epsilon \right]^q. \quad (3.38)$$

The derivative of (3.31) with respect to  $x$  is then given by

$$\frac{\partial C(x, h)}{\partial \text{vec}(x)} = H^T [\text{vec}(y) - H\text{vec}(x)] + \lambda \frac{\partial R[f(x)]}{\partial \text{vec}(x)}, \quad (3.39)$$

in which

$$\frac{\partial R[f(x)]}{\partial \text{vec}(x)} = \sum_{\theta \in \Theta} D_\theta^T \left[ \frac{D_\theta x}{\text{vec}[f(x)]} .* \text{vec} [(f(x) + \epsilon)^{q-1}] \right], \quad (3.40)$$

and  $./.$  is the point-wise division operator. Again, the products by  $H$ ,  $H^T$ ,  $D_\theta$  and  $D_\theta^T$  can be computed in the frequency domain, similarly to (3.36).

Lena $256 \times 256$		Linear convolutions				Circular convolutions			DCT-based conv.	
BSNR	Blur	Our method		With TV prior	APEX	YK	Our method		APEX	
		Without constraints	With constr.				Without constraints	With const.		
-	# 3	7.81	<b>9.55</b>	8.78	0.39	-6.89	8.74	<b>9.58</b>	0.52	0.43
	# 4	7.26	-	<b>9.82</b>	0.39	-1.39	<b>6.29</b>	-	0.56	-4.14
	# 5	8.87	-	<b>11.21</b>	0.38	-3.59	<b>10.05</b>	-	0.55	-0.27
	# 7	3.93	<b>3.99</b>	<b>3.99</b>	1.69	-7.99	4.03	4.12	<b>4.21</b>	-3.02
	# 3	4.07	<b>4.47</b>	4.27	0.60	-	4.26	<b>4.48</b>	0.74	-
	# 4	5.99	-	<b>6.74</b>	0.59	-	<b>6.03</b>	-	0.59	-
	# 5	6.76	-	<b>8.04</b>	0.53	-	<b>7.56</b>	-	0.69	-
30dB	# 7	2.23	<b>2.33</b>	2.29	1.77	-	2.40	<b>2.50</b>	2.29	-

Table 3.6: ISNR values (in dB) obtained for several methods. For each degradation, the best results are shown in bold.

Lena 256 × 256		Grayscale			Color
BSNR	Blur	Without const.	With const.	TV prior	Without constraints
-	# 1	7.87	<b>9.68</b>	9.66	10.75
	# 2	7.09	6.92	<b>7.83</b>	8.83
	# 3	7.81	<b>9.55</b>	8.78	9.70
	# 4	7.26	-	<b>9.82</b>	12.37
	# 5	8.87	-	<b>11.21</b>	10.00
	# 6	8.50	<b>11.51</b>	9.59	9.38
	# 7	3.93	<b>3.99</b>	<b>3.99</b>	4.62
30 dB	# 1	4.42	4.15	<b>4.45</b>	5.05
	# 2	5.56	5.33	<b>5.74</b>	6.15
	# 3	4.07	<b>4.47</b>	4.27	4.91
	# 4	5.99	-	<b>6.74</b>	6.85
	# 5	6.76	-	<b>8.04</b>	8.02
	# 6	7.35	<b>8.93</b>	7.36	8.24
	# 7	2.23	<b>2.33</b>	2.29	3.15

Table 3.7: ISNR (in dB) computed for experiments performed with the “Lena” image. For each degradation of the grayscale image, the best value is shown in bold.

Cameraman 256 × 256	No noise		BSNR at 30dB	
	Without constraints	With constraints	Without constraints	With constraints
# 1	6.32	<b>14.47</b>	4.27	<b>5.20</b>
# 2	4.87	<b>6.60</b>	4.15	<b>5.24</b>
# 3	5.51	<b>6.72</b>	4.07	<b>4.24</b>
# 4	<b>5.62</b>	-	<b>4.90</b>	-
# 5	<b>6.18</b>	-	<b>6.20</b>	-
# 6	8.06	<b>8.30</b>	6.64	<b>7.42</b>
# 7	2.72	<b>3.24</b>	1.81	<b>2.30</b>

Table 3.8: ISNR (in dB) computed for experiments performed with the “Cameraman” image. For each degradation, the best value is shown in bold.

Satellite 256 × 256	No noise		BSNR at 30dB	
	Without constraints	With constraints	Without constraints	With constraints
# 1	8.60	<b>13.60</b>	5.42	<b>6.34</b>
# 2	7.07	<b>8.37</b>	5.99	<b>7.38</b>
# 3	8.76	<b>10.24</b>	5.14	<b>5.33</b>
# 4	<b>7.76</b>	-	<b>8.17</b>	-
# 5	<b>8.72</b>	-	<b>7.08</b>	-
# 6	8.56	<b>12.23</b>	7.19	<b>10.92</b>
# 7	3.29	<b>3.44</b>	2.33	<b>2.75</b>

Table 3.9: ISNR (in dB) computed for experiments performed with the “Satellite” image. For each degradation, the best value is shown in bold.

Barbara 230 × 288		Grayscale		Color
BSNR	Blur	Without constraints	With constraints	Without constraints
-	# 1	<b>5.66</b>	4.31	5.83
	# 2	4.77	<b>5.68</b>	5.12
	# 3	6.22	<b>7.43</b>	6.19
	# 4	<b>4.56</b>	-	7.04
	# 5	<b>6.17</b>	-	6.87
	# 6	6.07	<b>13.43</b>	6.06
	# 7	2.99	<b>3.01</b>	3.02
30 dB	# 1	<b>3.26</b>	2.67	3.99
	# 2	3.37	<b>4.19</b>	4.26
	# 3	<b>3.73</b>	3.72	4.18
	# 4	<b>4.56</b>	-	4.87
	# 5	<b>4.83</b>	-	5.77
	# 6	5.18	<b>8.19</b>	5.13
	# 7	1.78	<b>1.81</b>	2.21

Table 3.10: ISNR (in dB) computed for experiments performed with the “Barbara” image. For each degradation of the grayscale image, the best value is shown in bold.

Testpat1 256 × 256	No noise		BSNR at 30dB	
	Without constraints	With constraints	Without constraints	With constraints
# 1	16.91	<b>17.99</b>	<b>8.43</b>	8.32
# 2	13.85	<b>15.49</b>	10.23	<b>11.05</b>
# 3	16.70	<b>17.14</b>	8.24	<b>8.50</b>
# 4	<b>15.98</b>	-	<b>14.59</b>	-
# 5	<b>14.11</b>	-	<b>13.63</b>	-
# 6	9.56	<b>14.22</b>	<b>18.05</b>	14.98
# 7	9.53	<b>9.94</b>	5.77	<b>6.51</b>

Table 3.11: ISNR (in dB) computed for experiments performed with the “Testpat1” image. For each degradation, the best value is shown in bold.

## Chapter 4

# Two-layer blind image deblurring

### Conference papers

Mariana S. C. Almeida and Luís B. Almeida, “Blind deblurring of foreground-background images,” in *IEEE International Conference on Image Processing - ICIP*, Cairo, Egypt, 2009, pp. 1301-1304.

Mariana S. C. Almeida and Luís B. Almeida, “Blind deblurring of two-layer images,” in *Portuguese Conf. on Pattern Recognition - RecPad*, Aveiro, Portugal, October 2009.



## 4.1 Blind deblurring of foreground-background images

### Abstract

This paper presents a method for deblurring an image consisting of two layers (a foreground layer and a background layer) which have suffered different, unknown blurs. This is a situation of practical interest. For example, it is common to find images in which we have a foreground object (e.g. a car) which has motion blur while the background is sharp (or vice-versa), or in which a foreground object and the background have different out-of-focus blurs.

We develop a model for this foreground + background degradation, and extend a previously introduced blind deblurring method to deal with this situation. As in the original blind deblurring method, the method presented here does not impose any strong constraints on the blurring filters. The method is almost completely blind, requiring, from the user, just a coarse indication of which are the foreground and background areas of the image.

The method has been tested with synthetic degradations and with real-life photos. We present some of the results. In all the experiments, the method was able to reasonably recover, from single degraded images: the complete deblurred image, the deblurred foreground and background images, and a mask providing the segmentation between foreground and background.

**keywords:** Blind image deconvolution, Space-variant blur, Image enhancement, Object segmentation.

### 4.1.1 Introduction

Image deblurring is an inverse problem whose aim is to recover images that have suffered blurring degradations. The blurring degradations can be shift-invariant (modeled by a convolution) or shift-variant (approximated by local convolutions). The present manuscript focuses on a special class of shift-variant degradations, in which two different regions of the image (that we call *foreground* and *background*, respectively) have different blurs. This is a situation often encountered in photography, when the scene is composed of an object that is closer to the camera, and a farther background. In this case, two typical scenarios are: (1) the closer object has motion blur and the background scene is static and in focus (or vice-versa); (2) the closer object and the farther background are both stationary, but have different focus blurs. These are two scenarios on which we make experimental tests in this paper, but it should be clear that the method to be presented is much more general, allowing the recovery from wide

classes of degradations both in the foreground and in the background. In the context of this paper, we shall call the images composed of foreground and background “two-layer images”, and the corresponding blurs “two-layer blurs”, while the ordinary images, with shift-invariant blurs, shall be called “single-layer”.

In single-layer blind image deconvolution (BID), both the image and the blur operator are unknown. This is an ill-posed problem, since there is an infinite number of pairs (original image + blurring filter) that could have led to the observed degraded image. Additionally, the blurring operator of BID typically is very ill-conditioned, making the estimate of the deblurred image very sensitive to noise and to mismatches between the estimated blurring filter and the true one. In shift-variant blind deblurring, the number of parameters to be estimated is even larger than in shift-invariant BID. In spite of the fact that the problem is rather challenging, a few shift-variant deblurring methods have been developed [20, 21, 37, 38, 126, 128–130, 160]. In [21, 37, 38, 126, 160], the problem’s difficulty is reduced by using multiple degraded images of the same scene. Results presented in [20] assume a highly restricted blur model (only out-of-focus), which is only tested on a simple synthetic image. [37, 126, 128, 129] address problems that are similar to those of the present work. However, [37, 126] use multi-frame scenarios. Methods of [128, 129] are either restricted to motion or to out-of-focus blurs, and do not correctly treat the regions’ boundaries. Furthermore, [37, 126, 128, 129] only address parameterized blurs, while the one that we present here handles more generic, non-parametric blurs. The method in [130] is restricted to motion blurs and estimates only the blur, without estimating the sharp image.

The approach that we propose in this paper evolved from a previously developed single-layer blind deblurring method [9, 120]. That method can satisfactorily deblur a wide range of degradations, without imposing any strong restrictions on the blurring filter. The flexibility of that method carries on to the method that we propose here, which can deal with rather unrestricted blurs in both the foreground and the background layers. The method is quasi-blind, requiring only a crude initial guess of the foreground and background zones, for region segmentation. The method’s performance was satisfactory when tested with synthetic degradations of natural images and with actual blurred photographs.

The structure of this paper is as follows: The single-layer deblurring method is briefly reviewed in Section 2. The two-layer method is then presented in Section 3. Section 4 reports experimental results.

### 4.1.2 Single-layer method

We consider the blurring model

$$y = h * x + n, \quad (4.1)$$

in which  $x$  is the original image,  $h$  is the PSF (point spread function) of the blurring filter and  $n$  is additive noise. The symbol  $*$  denotes mathematical convolution.

The deblurring method is based on two basic observations:

- In natural images, leading edges are sparse.
- In general, edges of a blurred image are less sparse than those of a sharp image, because they occupy a wider area (here, “edges” refers to continuously-valued variables that are related to the speed of spatial variation of the image’s intensity).

These two observations lead to the method’s central idea: using an objective function which favors images with sparse edges. The cost function that is used is of the form

$$C = \|y - \hat{y}\|_2^2 + \lambda_x R_f(x), \quad (4.2)$$

in which  $y$  is the observed blurred image,  $\hat{y}$  is the estimate of the blurred image, obtained through  $\hat{y} = h * x$ , where  $x$  is the estimated original image and  $h$  is the estimated blurring filter.  $R_f(x)$  is a regularizer which favors images with sparse edges. It uses an edge detector based on several directional edge-detecting filters (see Fig. 4.1). The output of the edge detector is given by

$$f(x) = \sqrt{\sum_{\theta \in \Theta} g_\theta^2}, \quad (4.3)$$

where  $g_\theta$  is the output of the edge detection filter with orientation  $\theta$  and  $\Theta$  is the set of orientations under consideration. The regularizer  $R_f(\cdot)$  operates on the edge intensity image  $f$  from Eq. (4.3):

$$R_f(x) = R[f(x)] = \sum_i (f_i + \epsilon)^q, \quad (4.4)$$

where  $i$  is an index running through all pixels,  $q$  controls the prior’s sparsity and  $\epsilon$  is a small parameter which yields finite lateral derivatives at  $f = 0$  (with  $0 < q < 1$ ), making the corresponding prior closer to actual distributions and also making the optimization easier.

Direct minimization of the objective function in (4.2), by itself, doesn’t generally lead to good deblurring results. The deblurring method uses a guiding procedure which essentially consists of starting with a large value of  $\lambda$  (heavy regularization) and progressively decreasing it as the estimate of the blurring filter becomes better. This guiding procedure allows the

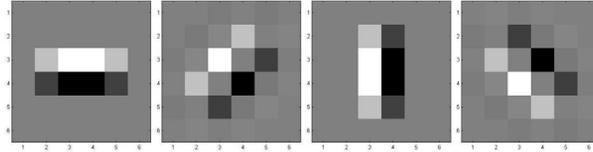


Figure 4.1: The set of edge detection filters, in the four orientations that were used.

method to be able to deblur a wide range of natural and artificial images with a wide range of blurs, without imposing any strict constraints on the blur PSF. References [9, 120] give more details, as well as experimental results.

### 4.1.3 Two-layer method

We now designate by  $x_1$  the background image, by  $x_2$  the foreground image, and by  $o$  an “opacity” mask, which describes the opacity of the foreground layer: If the foreground only contains opaque objects,  $o$  is a binary mask having ones in the pixels corresponding to the foreground objects and zeros in those corresponding to the background scene. If the foreground contains semi-transparent objects, then  $o$  is supposed to also contain that transparency information.

We assume that a sharp image would be formed according to

$$x_r = x_1 \cdot (1 - o) + x_2 \cdot o. \quad (4.5)$$

The operator  $\cdot$  denotes the Hadamard product. We model the blurred image by

$$y = (h_1 * x_1) \cdot (1 - \hat{o}) + (h_2 * x_2) \cdot \hat{o} + n, \quad (4.6)$$

in which  $h_1$  is the PSF of the blurring operator that has degraded the background image, and  $h_2$  is the PSF of the blurring operator that has degraded the foreground image.  $n$  represents additive noise, and  $\hat{o}$  is the blurred opacity mask, given by

$$\hat{o} = h_2 * o. \quad (4.7)$$

The estimation procedure is an extension of the single-layer one briefly presented in Section 2. All unknowns ( $x_1$ ,  $x_2$ ,  $h_1$ ,  $h_2$  and  $o$ ) are estimated by minimizing the cost function

$$C = \|y - \hat{y}\|_2^2 + \lambda_x [R_f(x_r) + R_f(x_1) + R_f(x_2)] + \lambda_o R_f(o), \quad (4.8)$$

in which  $\hat{y}$  is the estimate of the blurred image,

$$\hat{y} = (h_1 * x_1) \cdot (1 - \hat{o}) + (h_2 * x_2) \cdot \hat{o}, \quad (4.9)$$

and  $x_r$  is the estimated sharp image, which is constructed as in (4.5).  $\lambda_x$  is a regularizing

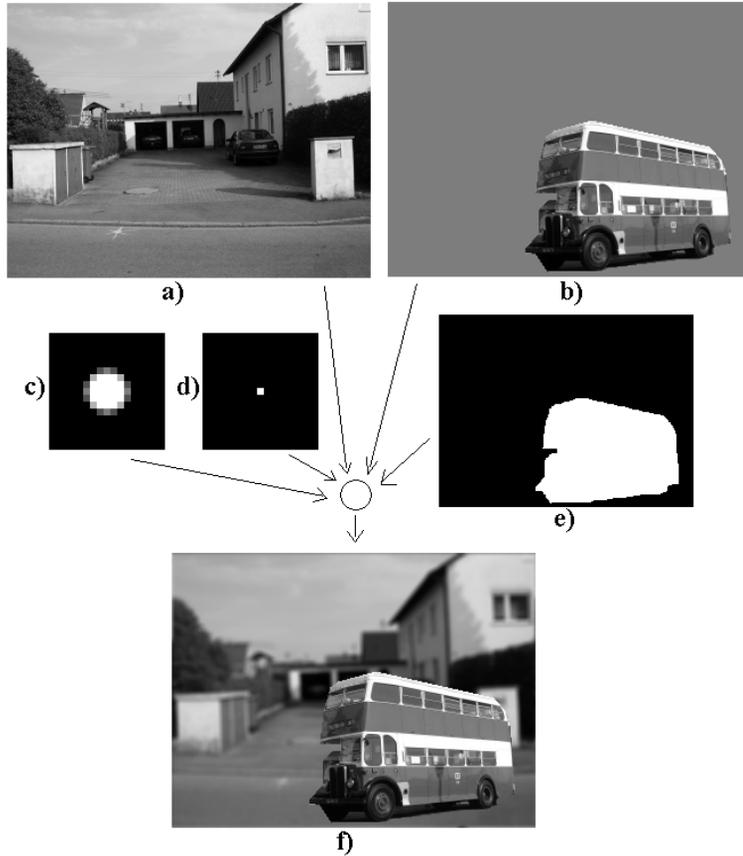


Figure 4.2: First synthetic degradation. a) Background image. b) Foreground object. c) Background blurring filter. d) Foreground blurring filter. e) Region segmentation. f) Blurred image.

parameter that controls the regularization applied to the estimated images (complete image, background and foreground), and  $\lambda_o$  is another regularizing parameter, controlling the regularization applied to the opacity mask  $o$ .

Like the single-layer method, the two-layer method only restricts the blurring filters to have a limited support. Among the infinite number of solutions, a suitable one is reached by starting with a large value for the regularizing parameters ( $\lambda_x$  and  $\lambda_o$ ) and progressively reducing them. Since we have two regularizing parameters ( $\lambda_x$  and  $\lambda_o$ ), it is convenient to fix a relationship between them. A good value for the ratio ( $\lambda_o/\lambda_x$ ) was empirically found to be 0.1.

For efficiency reasons, the images ( $x_1, x_2$ ), the filters ( $h_1, h_2$ ) and the segmentation ( $o$ ) are not simultaneously optimized. Table 4.1 outlines the method, for which a decreasing sequence of  $\lambda_x$  (and, consequently, of  $\lambda_o$ ) values and a non-increasing sequence of  $q$  values are assumed to have been previously chosen ( $q$  controls the regularizer's sparsity, as we'll see ahead).

During early iterations of the method,  $\lambda_x$  is large and only the main edges of the estimated image will survive, but they will be sharp [9, 120]. The sharp edges of the over-regularized

*Initialization:*

- 1 – Set  $h_1$  and  $h_2$  to the identity operator.
- 2 – Set  $x_1$  and  $x_2$  equal to  $y$ .
- 3 – Initialize  $o$  (as detailed in Section 4.1.3)
- 4 – Set  $\lambda_x$ ,  $\lambda_o$  and  $q$  to the initial values of their sequences.

*Optimization loop:*

- 5 – Find new estimates for  $x_1$  and  $x_2$ :  
 $(x_1, x_2) = \operatorname{argmin}_{x_1, x_2} C(x_1, x_2, h_1, h_2, o)$
- 6 – Find a new  $o$  estimate:  $o = \operatorname{argmin}_o C(x_1, x_2, h_1, h_2, o)$
- 7 – Find new estimates for  $h_1$  and  $h_2$ :  
 $(h_1, h_2) = \operatorname{argmin}_{h_1, h_2} C(x_1, x_2, h_1, h_2, o)$
- 8 – Set  $\lambda_x$ ,  $\lambda_o$  and  $q$  to the next values in sequence.
- 9 – If  $\lambda_x \geq \lambda_{x_{min}}$  go back to 5; otherwise stop.

Table 4.1: Method for deblurring a single image composed by two differently blurred areas.

images lead the estimates of the blurring operators to improve, and this, in iteration, allows smaller and fainter features to be progressively estimated, at the same time that the estimates of the blurring operators also improve. The gradual reduction of  $\lambda_x$  and  $\lambda_o$  results in a guiding technique which progressively considers smaller and fainter features.

### Prior information and preprocessing

The method needs to be initialized with a small amount of prior information, which is easily provided by a human. It consists of the indication of two compact regions in the image: one that is known to correspond to the background area, and another one that is known to correspond to the foreground area. Once these regions have been assigned, a preprocessing phase applies a segmentation algorithm [141] to find an initial estimate ( $o_0$ ) for the  $o$  mask (step 3 of Table 4.1). The method of [141] was implemented using horizontal and vertical image differences as features, with probability density functions that were estimated from the statistics of the assigned areas.

#### 4.1.4 Experiments

The proposed method was tested with synthetic and real-life degradations. The edge detector was implemented using the filters shown in Fig. 4.1, which were obtained from the basic filter

$$F_{\theta=0} = \begin{bmatrix} 1 & 2 & 2 & 1 \\ -1 & -2 & -2 & -1 \end{bmatrix} / 12 \quad (4.10)$$

by applying rotations by multiples of  $45^\circ$ . For all the experiments, we have used the same settings as in [9]: the images were normalized so that their intensity values would cover the interval  $[-0.5, 0.5]$ ;  $\epsilon$  was set to 0.002; the sequence of values for parameter  $q$  was: 0.8, 0.8, 0.6, 0.6, 0.6, 0.6, 0.4, ..., 0.4; the sequence for  $\lambda_x$  was geometric:  $\lambda_{x,n+1} = \lambda_{x,n}/r$  with  $\lambda_{x,0} = 2$ , and

with values of  $r$  that are given ahead. The sequence of values of  $\lambda_x$  was manually truncated at a value which led to the best compromise between image detail and noise or artifacts. We have manually chosen  $\lambda_o/\lambda_x = 0.1$ , as said above. The supports of the blurring filter estimates ( $h_1$  and  $h_2$ ) were limited to squares of size  $11 \times 11$  pixels each, and we used a safety margin of width 3 around these supports (see [9, 120] for details). All minimizations (steps 5, 6 and 7 of Table 4.1) were performed by gradient descent with adaptive step sizes.

In the first set of experiments, the method was tested with synthetic degradations. The sharp scene was constructed from two natural sharp images, corresponding to the background (Fig. 4.2-a) and foreground (Fig. 4.2-b). This sharp scene had  $240 \times 320$  pixels and was subjected to three different synthetic degradations: in the first degradation (Fig. 4.2-f) the foreground image was sharp and the background one had an out-of-focus blur; in the second degradation (Fig. 4.4-a) the foreground image had motion blur and the background was sharp; the third degradation (Fig. 4.5-a) combined the previous two degradations: the foreground image had motion blur and the background one had an out-of-focus blur.

For these experiments, we have set  $r = 2$  and the sequence of  $\lambda$  values was stopped at  $\lambda_{x,24} = 1.19 \times 10^{-7}$ . Fig. 4.6-a shows the manually chosen foreground and background regions that were used for this set of experiments. White corresponds to the manually chosen foreground region, black to the manually chosen background region, and gray to the unassigned region. The estimates obtained for the first degradation are shown in Fig. 4.3. For the other two degradations we only show the estimated sharp scene and blurring filters, due to lack of space (Figs. 4.4-b, c, d and 4.5-b, c, d). All these experiments led to sharper reconstructed images, in which more details were visible than in the corresponding blurred images. The estimated blurring filters and segmentation were also close to the true ones.

Another experiment was made with a real-life degradation corresponding to an actual photo obtained from the web (Fig. 4.7-a). The manually selected segments used for this experiment are shown in Fig. 4.6-b. We have set  $r = 1.5$  and the iteration was stopped at  $\lambda_{x,23} = 1.78 \times 10^{-4}$ . As we see from Fig. 4.7-b, the estimated scene was sharper than the original. In judging these results, one should take into account that photos normally involve some nonlinear processing, and thus do not exactly conform to our blurring model.

#### 4.1.5 Conclusions and future work

We have presented a method for deblurring shift-variant image degradations which often occur in photography. This kind of degradations often occurs when the photographed scene has a foreground object and a farther background. The approach presented in this paper is an

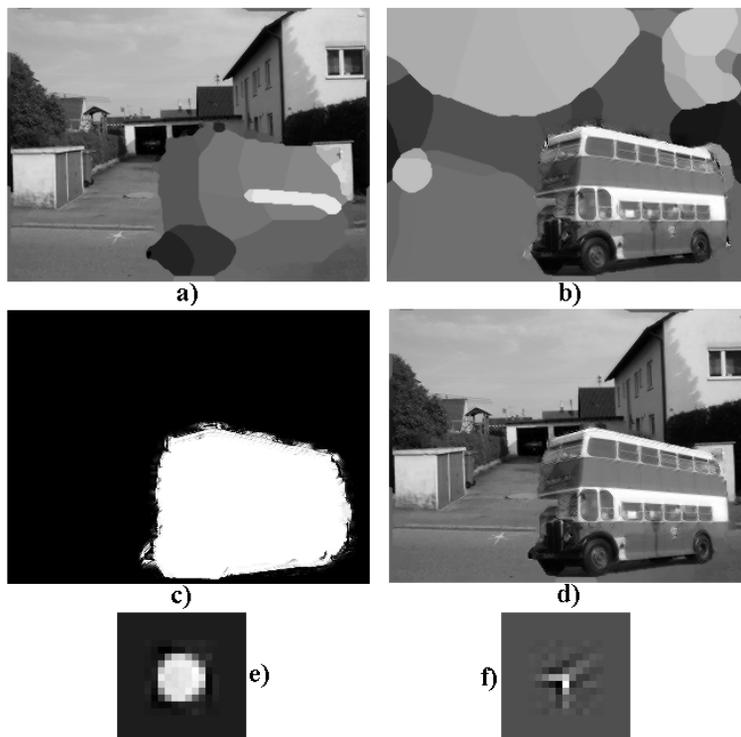


Figure 4.3: Results of the first synthetic experiment. a) Estimated background image, b) estimated foreground image, c) segmentation, d) estimated sharp scene, e) estimated background blur, f) estimated foreground blur.

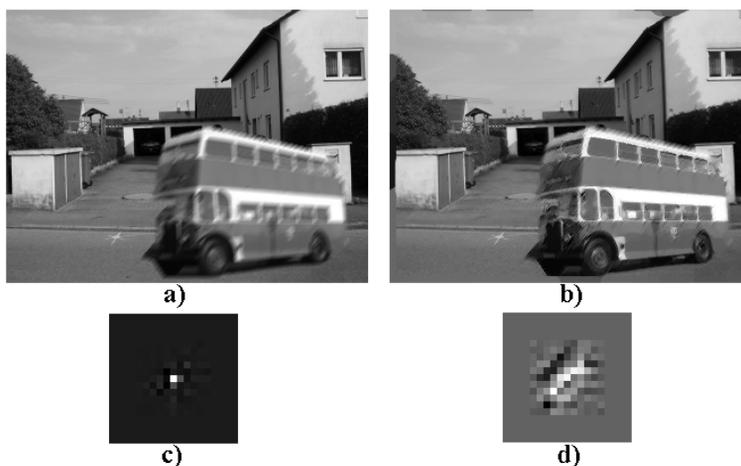


Figure 4.4: Second synthetic experiment. a) Blurred image. b) Estimated sharp scene. c) Estimated background blur. d) Estimated foreground blur.

extension of the shift-invariant BID method [9, 120]. As happens with that method, the one proposed here does not impose any strong restrictions on the blurring filter, and is able to deal with a wide range of blurs. The method's performance was tested on synthetic and real-life degradations.

In the future, we plan to extend the presented method in order to address more than two

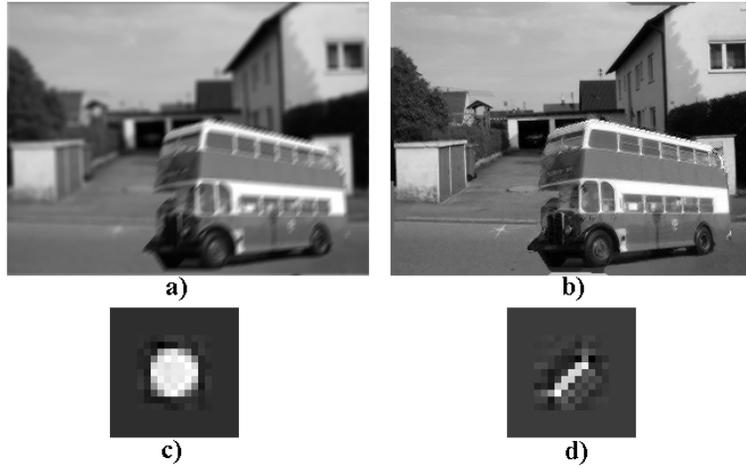


Figure 4.5: Third synthetic experiment. a) Blurred image. b) Estimated sharp scene. c) Estimated foreground blur. d) Estimated background blur.

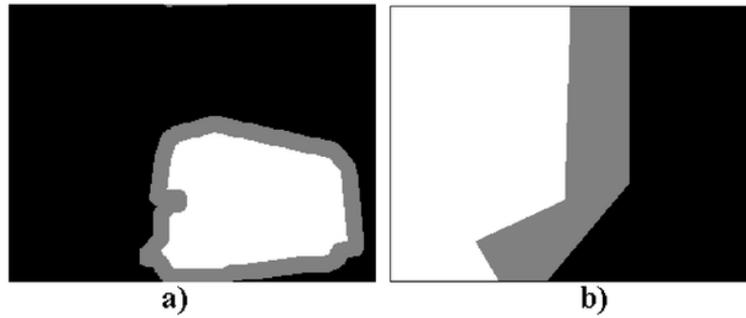


Figure 4.6: Manual supervision. a) Synthetic experiments. b) Photo. White: foreground. Black: Background. Gray: Unspecified.

blurred regions. We also plan to extend the originating BID method [9, 120] to address other shift-variant degradations of practical interest.

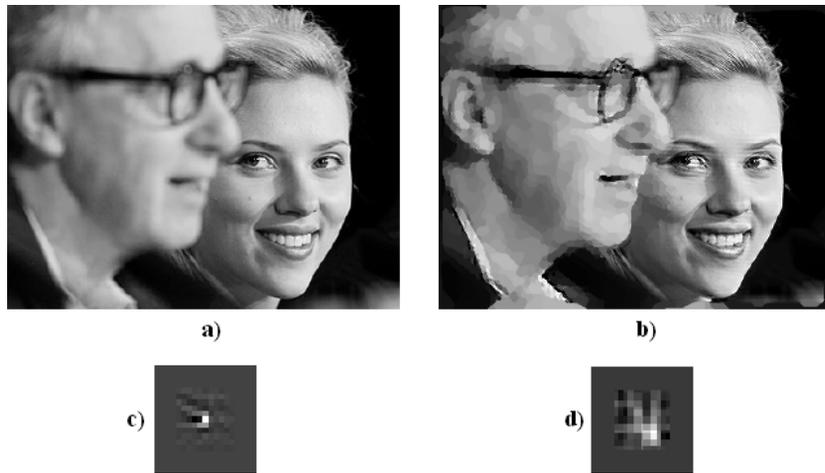


Figure 4.7: Deblurring a real photo. a) Blurred photo. b) Estimated sharp scene. c) Estimated background blur. d) Estimated foreground blur.

## 4.2 Blind deblurring of two-layer images

### Abstract

This work presents a method for deblurring an image consisting of two layers (a foreground layer and a background layer) which have suffered different, unknown blurs.

We extended a previously proposed blind deblurring method to the present foreground + background degradation. As in the original method, the approach presented here does not need to impose strong constraints on the blurring filters. The method is almost completely blind, requiring just a coarse indication of which are the foreground and background areas of the image.

Synthetic and real-life experiments show that the method is able to reasonably recover, from a single degraded image: the complete deblurred image, the deblurred foreground and background images, and a mask providing the segmentation between foreground and background.

### 4.2.1 Introduction

The present paper focuses on a special group of shift-variant degradations, in which two different regions of the image (that we call *foreground* and *background*, respectively) have different blurs. This is a common degradation in photographs that are composed by a closer object a farther background. Two typical photography scenarios which will be addressed are: (1) the closer object has motion blur and the background scene is static and in focus (or vice-versa); (2) the closer object and the farther background are both stationary, but have different focus blurs. In this paper, we shall call “two-layer images” the images composed by foreground and background, while we shall call “single-layer” the images which hav suffered a space-invariant blur.

Single-layer blind image deconvolution (BID) is an ill-posed inverse problem with and infinite number of solutions. In the two-layer blind deblurring problem the number of parameters under estimation is considerably larger than in the single-layer case, and the deconvolution problem gets even more challenging.

In this paper we present a method [120] which was evolved from a previously developed single-layer blind deblurring method [120]. That method was shown to overcome BID indeterminacies without the need to impose strong restrictions on the blurring filter [120]. The flexibility of this method carries on to the method that we propose here, which can deal with

rather unrestricted blurs in both the foreground and the background layers. The method only requires a crude initial guess of the foreground and background areas. Our approach is tested on a synthetic degradation and on an actual blurred photograph, with satisfactory results.

The manuscript is structured as follows: The two-layer method is briefly presented in Section 4.2.2. Section 4.2.3 reports experimental results and Section 4.2.4 concludes.

## 4.2.2 Two-layer method

We designate by  $x_1$  the background image, by  $x_2$  the foreground image, and by  $o$  an “opacity” mask, which describes the opacity of the foreground layer. Assuming that a sharp image would be formed according to

$$x_r = x_1 \cdot (1 - o) + x_2 \cdot o, \quad (4.11)$$

in which the operator  $\cdot$  denotes the Hadamard product, we model the degraded image by

$$y = (h_1 * x_1) \cdot (1 - \hat{o}) + (h_2 * x_2) \cdot \hat{o} + n, \quad (4.12)$$

in which  $h_1$  is the PSF of the blurring operator that has degraded the background image, and  $h_2$  is the PSF of the blurring operator that has degraded the foreground image.  $n$  represents additive noise, and  $\hat{o}$  is the blurred opacity mask ( $\hat{o} = h_2 * o$ ).

The estimation procedure is an extension of the single-layer method [120]. All unknowns ( $x_1$ ,  $x_2$ ,  $h_1$ ,  $h_2$  and  $o$ ) are estimated by minimizing the cost function

$$C = \|y - \hat{y}\|_2^2 + \lambda_x [R_f(x_r) + R_f(x_1) + R_f(x_2)] + \lambda_o R_f(o), \quad (4.13)$$

in which  $\hat{y}$  is the estimate of the blurred image, constructed according to the model (4.12), and  $x_r$  is the estimated sharp image (4.11).  $R_f$  is the regularizing function developed in [120].  $\lambda_x$  and  $\lambda_o$  are regularizing parameters that control the regularization applied to the estimated images (complete image, background and foreground) and to the opacity, respectively.

Like the single-layer method [120], this two-layer method only restricts the blurring filters to have a limited support. A suitable deblurring solution is reached by starting with a large value for the regularizing parameters ( $\lambda_x$  and  $\lambda_o$ ) and progressively reducing them.

The method needs to be initialized with a small amount of prior information, which is easily provided by a human. Two compact image regions must be indicated: one that corresponds to a background area, and another one corresponding to a foreground area. Once these regions have been assigned, a segmentation algorithm [141] is applied to find an initial estimate for the  $o$  mask.

More details about the single-layer and the two-layer method can be found in [120] and

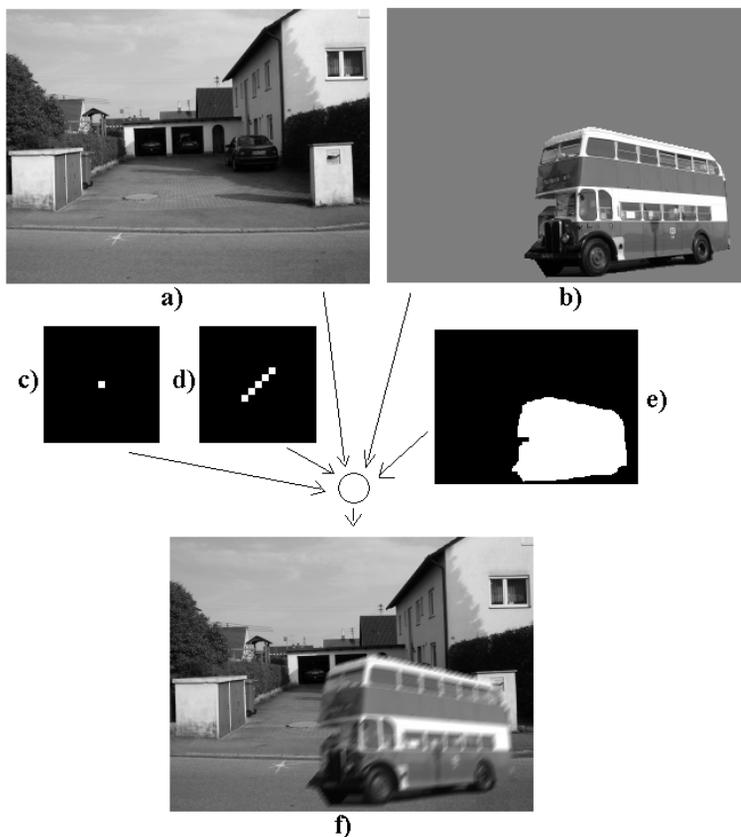


Figure 4.8: Synthetic degradation. a) Background image. b) Foreground object. c) Background blurring filter. d) Foreground blurring filter. e) Region segmentation. f) Blurred image.

[143], respectively.

### 4.2.3 Experiments

The proposed approach was tested, using the same parameters as in [143], on the synthetic degradation of Fig. 4.8 -f) and on the real blurred photo shown in Fig. 4.10 -a). For both experiments, the reconstructed image is sharper than the degraded one and the filter estimates are in accordance with the corresponding degradations (see Fig. 4.9 for the synthetic results and Fig. 4.10-b,c,d for the actual photo).

### 4.2.4 Conclusions

We have presented a method for deblurring two-layer image degradations which often occur in photographs composed by a closer foreground object and a farther background. The approach is an extension of the shift-invariant BID method presented in [120]. The method's performance was successfully tested on synthetic and real-life degradations.

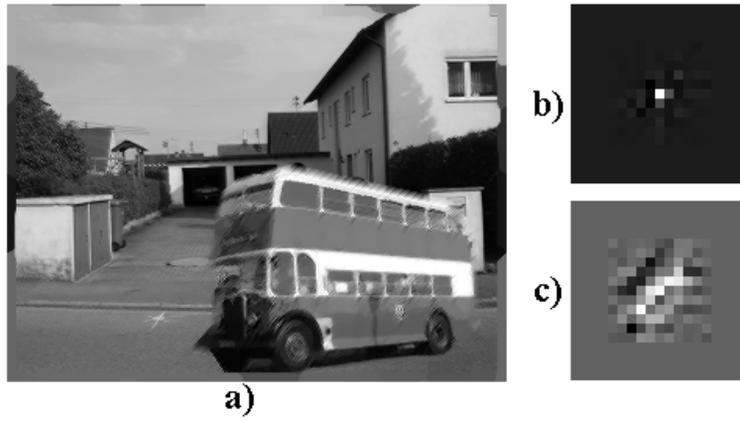


Figure 4.9: Synthetic experiment. a) Estimated sharp scene. b) Estimated background blur. c) Estimated foreground blur.

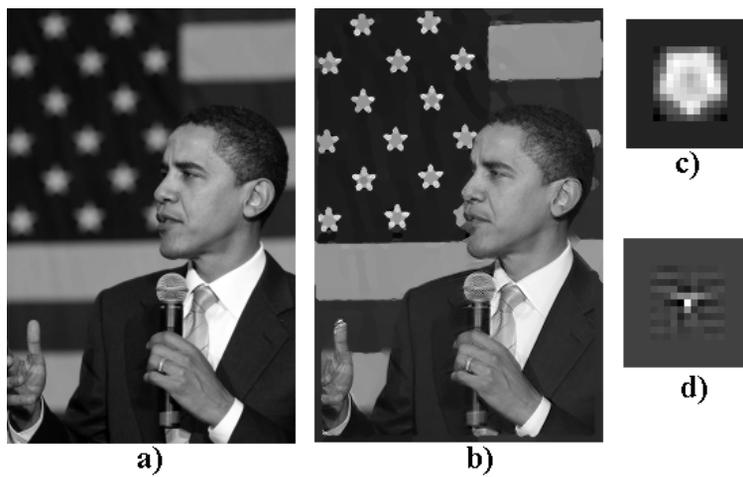


Figure 4.10: a) Blurred photo. b) Estimated sharp scene. c) Estimated background blur. d) Estimated foreground blur.

## Chapter 5

# Conclusions and future work

### 5.1 Conclusions

This dissertation focuses on two interesting nonlinear inverse problems in image processing and proposes new solutions to cope with their high levels of indetermination.

The first problem consists of inverting mixtures resulting from the show-through and bleed-through effects. Two efficient solutions for this separation problem are proposed in Chapter 2. The first one uses an ICA approach in which the show-through mixing process is physically modeled by a bi-affine function with four parameters. Due to the short number of parameters, this model manages to essentially eliminate the indeterminacy of nonlinear ICA, leading to separation results quantitatively better than those obtained with existing methods.

Based on simple proprieties of natural images and of the mixing process, a different separation method was also designed for both the show-through and the bleed-through processes. Since it does not directly resort to the independence criterion, this solution is not affected by the possible non independence of the source images. In contrast to previous approaches, the method does not assume the mixture to be space-invariant. It is, thus, suitable for mixtures with locally varying characteristics, such as those resulting from bleed-through or from wrinkled documents. This solution is non-iterative and achieves separation results visually competitive with those of other strategies.

Image deblurring is addressed on Chapters 3 and 4. Chapter 3 proposes a deblurring method for shift-invariant degradations, which is extended, in Chapter 4, to the shift-variant case of images consisting of two layers. In contrast to other existing methods, this deblurring technique only imposes weak restrictions on the blurring filter, being able to recover images which were subject to a wide range of degradations. The ill-posedness of blind image deblurring is overcome by initially considering the main image edges, and progressively handling smaller and/or fainter ones. Besides being originally a single-frame BID technique that does not

impose strong restrictions on the blurring filter, the method can deal with hard and soft filter constraints, and can be applied to multi-channel and multi-frame scenarios. The approach was tested on a wide set of images, both grayscale and color, with a variety of synthetic and real-life blurs, without and with noise, both in single and in multi-frame situations. The method reached good deblurring results. Compared to two other state-of-the-art methods, the proposed technique yielded better results and showed to be applicable to a much wider range of blurs.

The shift-invariant BID method was extended in order to address the shift-variant problem of images consisting of two layers (a foreground layer and a background layer) which may have suffered different blurring degradations. As in the original blind deblurring method, the shift-variant deblurring solution does not impose strong constraints on the blurring filters. Being almost completely blind, the method only requires, from the user, a coarse indication of which are the foreground and background areas of the image. The performance of the shift-variant deblurring method was successfully tested on synthetic and real-life degradations.

## 5.2 Future perspectives

This dissertation proposes solutions for separating show-through/bleed-through mixtures and for blindly restoring blurred images. Despite the good results that were obtained, each approach still deserves further research:

- The bi-affine model used in our first separation method has shown to fit the show-through process reasonably well. In spite of that, there is still room for model improvements: allowing bi-affine non-symmetric degradations (duplicating the number of the parameters of the model), using an extra parameter for a possible gamma correction, or, in a more complex scenario, accounting for light diffusion within the paper. More complex models can, however, become too flexible for an effective ICA-based approach.
- The two separation methods proposed in this thesis were designed for grayscale images. An interesting subject for future work is their extension to color images. Besides enabling the separation of color images themselves, an appropriate processing of multichannel images may lead to better separation results.
- So far, whenever the blurred image has noise, the BID method that was proposed involves a manual stop (by choosing the iteration that yields the best compromise between image detail and noise or artifacts). An automatic stopping criterion will obviously be useful.
- The BID method can still be extended in other directions. For example, (1) to address

problems in which we aim at super-resolution, possibly combined with deblurring, and (2) to deblur images containing space-variant blurs. Besides the two-layer case addressed in this thesis, there are other interesting shift-variant deblurring problems.

- On a more theoretical level, although with possible practical implications, is the issue that the best deblurring solution does not generally correspond to the global minimum of the cost function that was used. This apparently means that a more appropriate cost function should exist.
- Similarly to the shift-invariant deblurring method, the two-layer method described in this thesis is also extensible to multi-channel images and to multi-frame situations. These extensions have evident practical applications and would most probably lead to better deblurring results.
- The BID method proposed in this dissertation is relatively heavy in terms of computational cost. This drawback may become an issue for images with a large number of pixels, and clearly is an interesting topic for further research.
- Besides developing more efficient optimization techniques, we expect to improve the two-layer deblurring method through the use of techniques for matting estimation.



# Bibliography

- [1] A. Tonazzini, E. Salerno, and L. Bedini, “Fast correction of bleed-through distortion in grayscale documents by a blind source separation technique,” *International Journal of Document Analysis and Recognition - IJDAR*, vol. 10, no. 1, pp. 17–25, 2007.
- [2] G. Sharma, “Show-through cancellation in scans of duplex printed documents,” *IEEE Trans. on Image Processing*, vol. 10, no. 5, pp. 736–754, 2001.
- [3] A. Gupta, S. Kumar, R. Gupta, S. Chaudhury, and S. Joshi, “Enhancement of old manuscript images,” in *International Conference on Document Analysis and Recognition - ICDAR*, vol. 2, 2007, pp. 744–748.
- [4] X. Zhang, J. Lu, and T. Yahagi, “Blind separation methods for image show-through problem,” in *International Special Topic Conference on Information Technology Applications in Biomedicine*, 2007, pp. 255–258.
- [5] W. Boussellaa, A. Zahour, and A. Alimi, “A methodology for the separation of foreground/background in arabic historical manuscripts using hybrid methods,” *Journal of Universal Computer Science - JUCS*, no. 2, pp. 284–298, 2008.
- [6] M. Khan, H. Imtiaz, and M. Hasan, “Show-through correction in scanned images using joint histogram,” *Signal, Image and Video Processing*, June 2009.
- [7] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, “Removing camera shake from a single photograph,” *ACM Trans. on Graphics (TOG)*, vol. 25, no. 3, pp. 787–794, 2006.
- [8] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*. Bristol, Institute of Physics (IOP) Publishing, 1998.
- [9] M. S. C. Almeida and L. B. Almeida, “Blind deblurring of natural images,” in *IEEE Int. Conf. Acoustics, Speech, and Signal Processing - ICASSP*, 2008, pp. 1261–1264.
- [10] M. Bertero and P. Boccacci, “Image restoration methods for the large binocular telescope (LBT),” *Astronomy and Astrophysics Supplement Series*, vol. 147, pp. 323–333, 2000.
- [11] M. Mignotte and J. Meunier, “Three-dimensional blind deconvolution of SPECT images,” *IEEE Trans. on Biomedical Engineering*, vol. 47, no. 2, pp. 274–280, 2000.
- [12] S. Jefferies, K. Schulze, C. Matson, K. Stoltenberg, and E. K. Hege, “Blind deconvolution in optical diffusion tomography,” *Optics Express*, vol. 10, pp. 46–53, 2002.

- [13] M. M. Bronstein, A. M. Bronstein, M. Zibulevsky, and Y. Y. Zeevi, “Quasi-maximum likelihood blind deconvolution of images acquired through scattering media,” in *IEEE International Symposium on Biomedical Imaging - ISBI*, 2004, pp. 352–355.
- [14] H. Li, F. Qiao, O. R. Mawlawi, Y. Zheng, and R. X. Zhu, “Blind deblurring reconstruction technique with applications in PET imaging,” in *IEEE International Symposium on Biomedical Imaging - ISBI*, 2007, pp. 173–176.
- [15] J. Markham and J. Conchello, “Parametric blind deconvolution: a robust method for the simultaneous estimation of image and blur,” *Journal of the Optical Society of America A*, vol. 16, no. 10, pp. 2377–2391, 1999.
- [16] D. Adam and O. Michailovich, “Blind deconvolution of ultrasound sequences using non-parametric local polynomial estimates of the pulse,” *IEEE Trans. on Biomedical Engineering*, vol. 42, no. 2, pp. 118–131, 2002.
- [17] P. Pankajakshan, B. Zhang, L. Blanc-Féraud, Z. Kam, J.-C. Olivo-Marin, and J. Zerubia, “Blind deconvolution for diffraction-limited fluorescence microscopy,” in *IEEE International Symposium on Biomedical Imaging - ISBI*, 2008, pp. 740–743.
- [18] J.-P. P. Muller, *Digital Image Processing in Remote Sensing*. Taylor & Francis, 1988.
- [19] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising by sparse 3-d transform-domain collaborative filtering,” *IEEE Trans. on Image Processing*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
- [20] Y.-L. You and M. Kaveh, “Blind image restoration by anisotropic regularization,” *IEEE Trans. on Image Processing*, vol. 8, no. 3, pp. 396–407, March 1999.
- [21] M. Sorel and J. Flusser, “Space-variant restoration of images degraded by camera motion blur,” *IEEE Trans. on Image Processing*, vol. 17, no. 2, pp. 105–116, 2008.
- [22] B. Ophir and D. Malah, “Show-through cancellation in scanned images using blind source separation techniques,” in *IEEE International Conference on Image Processing - ICIP*, 2007, pp. 233–236.
- [23] H. Shen, L. Zhang, B. Huang, and P. Li, “A map approach for joint motion estimation, segmentation and super resolution,” *IEEE Trans. on Image Processing*, vol. 16, no. 2, pp. 479–490, 2007.
- [24] K. A. Patwardhan, G. Sapiro, and M. Bertalmío, “Video inpainting under constrained camera motion,” *IEEE Trans. on Image Processing*, vol. 16, no. 2, pp. 545–553, 2007.
- [25] Y. Tsaig and D. L. Donoho, “Compressed sensing,” *IEEE Trans. on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [26] K. T. Knox, “Show-through correction for two-sided documents,” *United States Patent 5646744*, 1997.
- [27] G. Sharma and K. T. Knox, “Show-through compensation apparatus and method,” *United States Patent 7145697*, December 2006.

- [28] Y. Huang and M. S. Brown, “User-assisted ink-bleed correction for handwritten documents,” in *ACM/IEEE-CS Joint Conference on Digital Libraries - JCDL*, 2008, pp. 263–271.
- [29] C. L. Tan, R. Cao, and P. Shen, “Restoration of archival documents using a wavelet technique,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 10, pp. 1399–1404, 2002.
- [30] D. Fadoua, F. L. Bourgeois, and H. Emptoz, “Restoring ink bleed-through degraded document images using a recursive unsupervised classification technique,” in *workshop on Document Analysis Systems - DAS*, 2006, pp. 38–49.
- [31] E. Dubois and A. Pathak, “Reduction of bleed-through in scanned manuscript documents,” in *Image Processing, Image Quality, Image Capture Systems Conference - PICS*, Montreal, Canada, 2001, pp. 177–180.
- [32] M. S. C. Almeida and L. B. Almeida, “Wavelet-based separation of nonlinear show-through and bleed-through image mixtures,” *Neurocomputing*, vol. 72, no. 1-3, pp. 57–70, December 2008.
- [33] A. S. Carasso, “Singular integral image deblurring method,” *United States Patent 7437012*, October 2008.
- [34] A. Falhi and P. Seitz, “Image deblurring and denoising system, device and method,” *World Patent 156329 (WO/2009/156329)*.
- [35] M. Ben-Ezra and S. K. Nayar, “Method for de-blurring images of moving objects,” *United States Patent 7440634*.
- [36] J. G. Nagy and D. P. O’Leary, “Restoring images degraded by spatially variant blur,” *SIAM Journal on Scientific Computing*, vol. 19, no. 4, pp. 1063–1082, 1998.
- [37] A. Kubota and K. Aizawa, “Reconstructing arbitrarily focused images from two differently focused images using linear filters,” *IEEE Trans. on Image Processing*, vol. 14, no. 11, pp. 1848–1859, Nov 2005.
- [38] M. Blume, D. Zikic, W. Wein, and N. Navab, “A new and general method for blind shift-variant deconvolution of biomedical images,” in *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, 2007, pp. 743–750.
- [39] Zeng and G. Lawrence, “Systems and methods for deblurring data corrupted by shift variant blurring,” *United States Patent Application 20080166063*, 2008.
- [40] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, “Understanding and evaluating blind deconvolution algorithms,” in *IEEE Conference on Computer Vision and Pattern Recognition - CVPR*, 2009, pp. 1964–1971.
- [41] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce, “Non-uniform deblurring for shaken images,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2010.

- [42] M. Šorel and F. Šroubek, “Space-variant deblurring using one blurred and one underexposed image,” in *IEEE International Conference on Image Processing - ICIP*, 2009, pp. 157–160.
- [43] P. Comon, “Independent component analysis – a new concept?” *Signal Processing*, vol. 36, pp. 287–314, 1994.
- [44] A. J. Bell and T. J. Sejnowski, “An information-maximization approach to blind separation and blind deconvolution,” *Neural Computation*, vol. 7, pp. 1129–1159, 1995. [Online]. Available: <ftp://ftp.cnl.salk.edu/pub/tony/bell.blind.ps>
- [45] A. Belouchrani, K. A. Meraim, J.-F. Cardoso, and E. Moulines, “A blind source separation technique based on second order statistics,” *IEEE Trans. on Signal Processing*, vol. 45, no. 2, pp. 434–444, 1997. [Online]. Available: <http://www.tsi.enst.fr/~cardoso/Papers.PDF/ieeesobi.pdf>
- [46] A. Hyvärinen and E. Oja, “A fast fixed-point algorithm for independent component analysis,” *Neural Computation*, vol. 9, no. 7, pp. 1483–1492, 1997.
- [47] A. Ziehe and K.-R. Müller, “TDSEP – An efficient algorithm for blind separation using time structure,” in *International Conference on Artificial Neural Networks - ICANN*, Skövde, Sweden, 1998, pp. 675–680. [Online]. Available: [http://wwwold.first.gmd.de/persons/Mueller.Klaus-Robert/ICANN\\_tdsep.ps.gz](http://wwwold.first.gmd.de/persons/Mueller.Klaus-Robert/ICANN_tdsep.ps.gz)
- [48] A. Hyvärinen and P. Pajunen, “Nonlinear independent component analysis: Existence and uniqueness results,” *Neural Networks*, vol. 12, no. 3, pp. 429–439, 1999. [Online]. Available: <http://www.cis.hut.fi/~aapo/ps/NN99.ps>
- [49] G. C. Marques and L. B. Almeida, “Separation of nonlinear mixtures using pattern repulsion,” in *Proc. Int. Worksh. Independent Component Analysis and Blind Source Separation - ICA*, Aussois, France, 1999, pp. 277–282. [Online]. Available: <http://www.lx.it.pt/~lbalmeida/papers/MarquesAlmeidaICA99.ps.zip>
- [50] G. Darmais, “Analyse générale des liaisons stochastiques,” *Rev. Inst. Internat. Stat.*, vol. 21, pp. 2–8, 1953.
- [51] A. Taleb and C. Jutten, “Source separation in post-nonlinear mixtures,” *IEEE Trans. on Signal Processing*, vol. 47, pp. 2807–2820, 1999.
- [52] H. Lappalainen and A. Honkela, “Bayesian nonlinear independent component analysis by multi-layer perceptrons,” in *Advances in Independent Component Analysis*, M. Girolami, Ed. Springer-Verlag, 2000, pp. 93–121. [Online]. Available: <http://www.cis.hut.fi/harri/ch7.ps.gz>
- [53] S. Harmeling, A. Ziehe, M. Kawanabe, B. Blankertz, and K. robert Müller, “Nonlinear blind source separation using kernel feature spaces,” in *Proc. Int. Worksh. Independent Component Analysis and Blind Source Separation - ICA*, S. Diego, California, U.S.A., 2001, pp. 102–107. [Online]. Available: [http://ica2001.ucsd.edu/index\\_files/pdfs/080-harmeling.pdf](http://ica2001.ucsd.edu/index_files/pdfs/080-harmeling.pdf)

- [54] D. Martinez and A. Bray, “Nonlinear blind source separation using kernels,” *IEEE Trans. on Neural Networks*, vol. 14, no. 1, pp. 228–236, 2003.
- [55] L. B. Almeida, “MISEP – Linear and nonlinear ICA based on mutual information,” *Journal of Machine Learning Research - JMLR*, vol. 4, pp. 1297–1318, 2003. [Online]. Available: <http://www.jmlr.org/papers/volume4/almeida03a/almeida03a.pdf>
- [56] T. Blaschke, T. Zito, and L. Wiskott, “Independent slow feature analysis and nonlinear blind source separation,” in *Proc. Int. Worksh. Independent Component Analysis and Blind Source Separation - ICA*, 2004, pp. 742–749. [Online]. Available: <http://itb.biologie.hu-berlin.de/~blaschke/publications/isfa.pdf>
- [57] L. B. Almeida, *Nonlinear Source Separation*. Morgan and Claypool, 2006.
- [58] K. Zhang and L. Chan, “Minimal nonlinear distortion principle for nonlinear independent component analysis,” *Journal of Machine Learning Research - JMLR*, vol. 9, pp. 2455–2487, 2008.
- [59] D. N. Levin, “Using signal invariants to perform nonlinear blind source separation,” in *Proc. Int. Worksh. Independent Component Analysis and Blind Source Separation - ICA*, Berlin, Heidelberg, 2009, pp. 58–65.
- [60] M. Haritopoulos, H. Yin, and N. M. Allinson, “Image denoising using self-organizing map-based nonlinear independent component analysis,” *Neural Networks*, vol. 15, no. 8-9, pp. 1085–1098, 2002.
- [61] H. Lappalainen, A. Honkela, X. Giannakopoulos, and J. Karhunen, “Nonlinear source separation using ensemble learning and MLP networks,” in *Proc. of the Symp. 2000 on Adaptive Sys. for Sig. Proc., Communications, and Control*, Lake Louise, Alberta, Canada, October 1-4 2000, pp. 187–192. [Online]. Available: <http://www.cis.hut.fi/projects/ica/bayes/papers/llouise.ps.gz>
- [62] S.-I. Lee and S. Batzoglou, “Application of independent component analysis to microarrays,” *Genome Biology*, vol. 4, no. 11, p. R76, 2003. [Online]. Available: <http://genomebiology.com/2003/4/11/R76>
- [63] A. Tonazzini, L. Bedini, and E. Salerno, “Independent component analysis for document restoration,” in *International Journal of Document Analysis and Recognition - IJDAR*, vol. 7, no. 1, 2004, pp. 17–27.
- [64] L. B. Almeida, “Separating a real-life nonlinear image mixture,” *Journal of Machine Learning Research - JMLR*, vol. 6, pp. 1199–1229, July 2005. [Online]. Available: <http://www.jmlr.org/papers/volume4/almeida03a/almeida03a.pdf>
- [65] A. Tonazzini, L. Bedini, and E. Salerno, “A markov model for blind image separation by a mean-field EM algorithm,” *IEEE Trans. on Image Processing*, vol. 15, no. 2, pp. 473–482, Feb. 2006.
- [66] M. S. C. Almeida, H. Valpola, and J. Särelä, “Separation of nonlinear image mixtures by denoising source separation,” in *Proc. Int. Worksh. Independent Component Analysis and Blind Source Separation - ICA*, J. Rosca, D. Erdogmus, J. Principe, and S. Haykin,

- Eds. Charleston, South Carolina, U.S.A.: Springer-Verlag, 2006, pp. 8–15. [Online]. Available: <http://www.cis.hut.fi/jaakkos/papers/nonlindss.pdf>
- [67] M. S. C. Almeida, “Extensions to denoising source separation,” Neural Networks Research Center of the Helsinki University of Technology and Instituto Superior Técnico, Tech. Rep., 2005, Graduation Thesis, Instituto Superior Técnico.
- [68] L. B. Almeida and M. Faria, “Separating a real-life nonlinear mixture of images,” in *Proc. Int. Worksh. Independent Component Analysis and Blind Source Separation - ICA*, ser. Lecture Notes in Artificial Intelligence, no. 3195. Springer-Verlag, 2004, pp. 729–736. [Online]. Available: <http://www.lx.it.pt/~lbalmeida/papers/AlmeidaICA04.pdf>
- [69] J. Särelä and H. Valpola, “Denoising source separation,” *Journal of Machine Learning Research - JMLR*, vol. 6, pp. 233–272, 2005.
- [70] B. Ophir and D. Malah, “Improved cross-talk cancellation in scanned images by adaptive decorrelation,” in *IEEE Convention of Electrical and Electronics Engineers in Israel, 2004. Proceedings*, Sept. 6-7 2004, pp. 388–391.
- [71] M. S. C. Almeida and L. B. Almeida, “Separating nonlinear image mixtures using a physical model trained with ICA,” in *IEEE Int. Worksh. on Machine Learning for Signal Processing - MLSP*, Maynooth, Ireland, 2006.
- [72] F. Merrih-Bayat, M. Babaie-Zadeh, and C. Jutten, “A nonlinear blind source separation solution for removing the show-through effect in the scanned documents,” in *European Signal Processing Conference - EUSIPCO*, Lausanne, Switzerland, 2008.
- [73] A. Tonazzini and I. Gerace, “Bayesian MRF-based blind source separation of convolutive mixtures of images,” in *European Signal Processing Conference - EUSIPCO*, 2005.
- [74] H. Nishida and T. Suzuki, “Correcting show-through effects on scanned color document images by multiscale analysis,” *Pattern Recognition*, vol. 36, pp. 2835–2847, December 2003.
- [75] M. S. C. Almeida and L. B. Almeida, “Separation of nonlinear show-through image mixtures using a physical model trained with ICA,” *Journal of Machine Learning Research - JMLR*, 2010 (submitted).
- [76] M. Faria, “Separação de misturas não lineares (Nonlinear source separation),” Tech. Rep., 2003, Graduation Thesis, Instituto Superior Técnico, in Portuguese.
- [77] M. S. C. Almeida and L. B. Almeida, “Wavelet based nonlinear separation of images,” in *IEEE Int. Worksh. on Machine Learning for Signal Processing - MLSP*, Maynooth, Ireland, 2006.
- [78] G. K. Chantas, N. P. Galatsanos, and A. Likas, “Bayesian restoration using a new nonstationary edge-preserving image prior,” *IEEE Trans. on Image Processing*, vol. 15, no. 10, pp. 2987–2997, 2006.
- [79] J. A. Guerrero-Colon and J. Portilla, “Deblurring-by-denoising using spatially adaptive gaussian scale mixtures in overcomplete pyramids,” in *IEEE International Conference on Image Processing - ICIP*, Atlanta, October 2006, pp. 625–628.

- [80] L. Bar, N. Sochen, and N. Kiryati, "Image deblurring in the presence of impulsive noise," *International Journal of Computer Vision*, vol. 70, no. 3, pp. 279–298, 12 2006.
- [81] V. Katkovnik, K. Egiazarian, and J. Astola, "A spatially adaptive nonparametric regression image deblurring," *IEEE Trans. on Image Processing*, vol. 14, pp. 1469–1478, 2005.
- [82] G. Chantas, N. Galatsanos, A. Likas, and M. Saunders, "Variational bayesian image restoration based on a product of t-distributions image prior," *IEEE Trans. on Image Processing*, vol. 17, pp. 1795–805, 2008.
- [83] G. K. Chantas, N. P. Galatsanos, R. Molina, and A. K. Katsaggelos, "Variational bayesian image restoration with a product of spatially weighted total variation image priors," *IEEE Trans. on Image Processing*, vol. 19, no. 2, pp. 351–362, 2010.
- [84] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Processing Magazine*, vol. 13, no. 3, pp. 43–64, May 1996.
- [85] P. Campisi and K. Egiazarian, *Blind image deconvolution: Theory and Applications*. CRC Press, 2007.
- [86] H. Yin and I. Hussain, "Blind source separation and genetic algorithm for image restoration," in *IEEE Int. Conf. Advances in space technologies - ICAST*, Islamabad, Pakistan, Sept. 2006, pp. 167–172.
- [87] F. Krahmer, Y. Lin, B. McAdoo, K. Ott, J. Wang, D. Widemannk, and B. Wohlberg, "Blind image deconvolution: motion blur estimation," University of Minnesota, Tech. Rep., 2006.
- [88] J. P. A. Oliveira, M. A. T. Figueiredo, and J. M. Bioucas-Dias, "Blind estimation of motion blur parameters for image deconvolution," in *Iberian Conference on Pattern Recognition and Image Analysis - IbPRIA*, Girona, Spain, June 2007.
- [89] M. S. Chang, S. W. Yun, and P. Park, "PSF search algorithm for dual-exposure type blurred image," *International Journal of Applied Science, Engineering and Technology*, vol. 4, no. 1, 2007.
- [90] F. Rooms, W. Philips, and J. Portilla, "Parametric PSF estimation via sparseness maximization in the wavelet domain," in *Proc. of SPIE, Wavelet Applications in Industrial Processing II*, vol. 5607, October 2004, pp. 26–33. [Online]. Available: [http://decsai.ugr.es/vip/files/conferences/vol5607pp26\\_33\\_2004.pdf](http://decsai.ugr.es/vip/files/conferences/vol5607pp26_33_2004.pdf)
- [91] A. S. Krylov. and A. V. Nasonov, "Adaptive image deblurring with ringing control," in *International Conference on Image and Graphics - ICIG*, 2009.
- [92] K.-H. Yap, L. Guan, and W. Liu, "A recursive soft-decision approach to blind image deconvolution," *IEEE Trans. on Signal Processing*, vol. 51, no. 2, February 2003.
- [93] L. Justen and R. Ramlau, "A non-iterative regularization approach to blind deconvolution," *Institute of Physics Publishing - Inverse Problems*, vol. 22, no. 3, pp. 771–800, 2006.

- [94] A. S. Carasso, “Direct blind deconvolution,” *SIAM Journal on Applied Mathematics (SIAP)*, vol. 61, pp. 1980–2007, 2001.
- [95] ———, “The APEX method in image sharpening and the use of low exponent lévy stable laws,” *SIAM Journal of Applied Mathematics*, vol. 63, no. 2, pp. 593–618, 2003.
- [96] Y.-L. You and M. Kaveh, “A regularization approach to joint blur identification and image restoration,” *IEEE Trans. on Image Processing*, vol. 5, no. 3, pp. 416–428, March 1996.
- [97] T. F. Chan and C.-K. Wong, “Total variation blind deconvolution,” *IEEE Trans. on Image Processing*, vol. 7, no. 3, pp. 370–375, March 1998.
- [98] A. C. Likas and N. P. Galatsanos, “A variational approach for bayesian blind image deconvolution,” *IEEE Trans. on Signal Processing*, vol. 52, pp. 2222–2233, 2004.
- [99] L. He, A. Marquina, and S. J. Osher, “Blind deconvolution using TV regularization and bregman iteration,” *International Journal of Imaging Systems and Technology*, vol. 15, pp. 74–83, 2005.
- [100] R. Kaftory and N. S. Y. Y. Zeevi, “Variational blind deconvolution of multi-channel images,” *International Journal of Imaging Systems and Technology, special issue on Blind Source Separation and Deconvolution in Imaging and Image Processing*, vol. 15, no. 1, pp. 55–63, 2005.
- [101] R. Molina, J. Mateos, and A. K. Katsaggelos, “Blind deconvolution using a variational approach to parameter, image, and blur estimation,” *IEEE Trans. on Image Processing*, vol. 15, pp. 3715–3727, 2006.
- [102] D. Babacan, R. Molina, and A. Katsaggelos, “Parameter estimation in total variation blind deconvolution,” in *European Signal Processing Conference - EUSIPCO*, August 2008.
- [103] N. Joshi, R. Szeliski, and D. J. Kriegman, “PSF estimation using sharp edge prediction,” in *IEEE Conference on Computer Vision and Pattern Recognition - CVPR*, June 2008, pp. 1–8.
- [104] J. H. Money and S. H. Kang, “Total variation minimizing blind deconvolution with shock filter reference,” *Image Vision Computing*, vol. 26, no. 2, pp. 302–314, 2008.
- [105] Q. Shan, J. Jia, and A. Agarwala, “High-quality motion deblurring from a single image,” *ACM Trans. on Graphics (TOG)*, vol. 27, no. 3, 2008.
- [106] J.-F. Cai, H. Ji, C. Liu, and Z. Shen, “Blind motion deblurring from a single image using sparse approximation,” in *IEEE Conference on Computer Vision and Pattern Recognition - CVPR*, 2009, pp. 104–111.
- [107] S. D. Babacan, R. Molina, and A. K. Katsaggelos, “Variational bayesian blind deconvolution using a total variation prior,” *IEEE Trans. on Image Processing*, vol. 18, no. 1, pp. 12–26, 2009.

- [108] D. G. Tzikas, A. C. Likas, and N. P. Galatsanos, “Variational bayesian sparse kernel-based blind image deconvolution with student’s-t priors,” *IEEE Trans. on Image Processing*, vol. 18, no. 4, pp. 753–764, 2009.
- [109] J. Jia, “Single image motion deblurring using transparency,” in *IEEE Conference on Computer Vision and Pattern Recognition - CVPR*, 2007.
- [110] P. C. Hansen, “The l-curve and its use in the numerical treatment of inverse problems,” in *Computational Inverse Problems in Electrophysiology, Advances in Computational Bio-engineering*. WIT Press, 2000, pp. 119–142.
- [111] J. P. Oliveira, J. M. Bioucas-Dias, and M. A. T. Figueiredo, “Adaptive total variation image deblurring: A majorization-minimization approach,” *Signal Processing*, vol. 89, no. 9, pp. 1683–1693, 2009.
- [112] G. Panci, P. Campisi, S. Colonnese, and G. Scarano, “Multichannel blind image deconvolution using the bussgang algorithm: spatial and multiresolution approaches,” *IEEE Trans. on Image Processing*, vol. 12, no. 11, pp. 1324–1337, 2003.
- [113] F. Sroubek and J. Flusser, “Multichannel blind deconvolution of spatially misaligned images,” *IEEE Trans. on Image Processing*, vol. 14, no. 7, pp. 874–883, 2005.
- [114] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum, “Blurred/non-blurred image alignment using sparseness prior,” in *Proceedings of the International Conference on Computer Vision - ICCV*. Blackwell Publishing Ltd, 2007.
- [115] V. Katkovnik, D. Paliy, K. Egiazarian, and J. Astola, “Frequency domain blind deconvolution in multiframe imaging using anisotropic spatially-adaptive denoising,” in *European Signal Processing Conference - EUSIPCO*, Florence, Italy, Sept. 4-8 2006.
- [116] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum, “Image deblurring with blurred/noisy image pairs,” *ACM Trans. on Graphics (TOG)*, vol. 26, no. 3, 2007.
- [117] M. M. Bronstein, A. M. Bronstein, M. Zibulevsky, and Y. Y. Zeevi, “Blind deconvolution of images using optimal sparse representations,” *IEEE Trans. on Image Processing*, vol. 14, no. 6, pp. 726–736, 2005.
- [118] D. Li, R. M. Mersereau, and S. J. Simske, “Blind image deconvolution through support vector regression,” *IEEE Trans. on Neural Networks*, vol. 18, no. 3, pp. 931–935, May 2007.
- [119] I. N. Aizenberg, D. Paliy, J. M. Zurada, and J. Astola, “Blur identification by multilayer neural network based on multivalued neurons,” *IEEE Trans. on Image Processing*, vol. 19, no. 8, pp. 248–270, 2008.
- [120] M. S. C. Almeida and L. B. Almeida, “Blind and semi-blind deblurring of natural images,” *IEEE Trans. on Image Processing*, vol. 19, no. 1, pp. 36–52, January 2010.
- [121] —, “Processo de focagem cega de imagens,” *Portuguese patent*, 2009.
- [122] L. Bar, N. A. Sochen, and N. Kiryati, “Restoration of images with piecewise space-variant blur,” in *International Conference on Scale Space and Variational Methods in Computer Vision - SSVM*. Springer, 2007, pp. 533–544.

- [123] M. Subbarao, Y. sik Kang, S. Dutta, and X. Tu, “Localized and computationally efficient approach to shift-variant image deblurring,” in *IEEE International Conference on Image Processing - ICIP*, 2008, pp. 657–660.
- [124] A. Kubota, K. Kodama, and K. Aizawa, “Registration and blur estimation methods for multiple differently focused images,” in *IEEE International Conference on Image Processing - ICIP*, 1999, pp. 447–451.
- [125] S. W. Hasinoff and K. N. Kutulakos, “A layer-based restoration framework for variable-aperture photography,” in *Proceedings of the International Conference on Computer Vision - ICCV*, 2007.
- [126] L. Bar, B. Berkels, M. Rumpf, and G. Sapiro, “A variational framework for simultaneous motion estimation and restoration of motion-blurred video,” in *Proceedings of the International Conference on Computer Vision - ICCV*, 2007.
- [127] S. Cho, Y. Matsushita, and S. Lee, “Removing non-uniform motion blur from images,” 2007.
- [128] J. Shin, S. Hwang, K. Kim, J. Kang, S. Lee, and J. K. Paik, “Blur identification and image restoration based on evolutionary multiple object segmentation for digital auto-focusing,” in *International Workshop Combinatorial Image Analysis - IWCIA*, Auckland, New Zealand, Dec. 1-3 2004, pp. 656–668.
- [129] A. Levin, “Blind motion deblurring using image statistics,” in *Advances in Neural Information Processing Systems - NIPS*, 2006, pp. 841–848.
- [130] S. Dai and Y. Wu, “Estimating space-variant motion blur without deblurring.” in *IEEE International Conference on Image Processing - ICIP*, San Diego, California, USA, pp. 661–664.
- [131] M. Sorel and F. Sroubek, “Space-variant deblurring using one blurred and one under-exposed image,” in *IEEE International Conference on Image Processing - ICIP*, Cairo, Egypt, 2009, pp. 157–160.
- [132] S. Dai and Y. Wu, “Motion from blur,” in *IEEE Conference on Computer Vision and Pattern Recognition - CVPR*, Alaska, USA, June 24-26 2008.
- [133] N. Joshi, R. Szeliski, and D. J. Kriegman, “PSF estimation using sharp edge prediction,” in *IEEE Conference on Computer Vision and Pattern Recognition - CVPR*, Alaska, USA, June 24-26 2008, pp. 1–8.
- [134] S. Dai and Y. Wu, “Removing partial blur in a single image,” in *IEEE Conference on Computer Vision and Pattern Recognition - CVPR*, 2009, pp. 2544–2551.
- [135] S. Bae and F. Durand, “Defocus magnification,” *Computer Graphics Forum*, vol. 26, no. 3, pp. 571–579, 2007.
- [136] R. Liu, Z. Li, and J. Jia, “Image partial blur detection and classification,” in *IEEE Conference on Computer Vision and Pattern Recognition - CVPR*, June 2008.

- [137] A. Levin, P. Sand, T. S. Cho, F. Durand, and W. T. Freeman, “Motion-invariant photography,” *ACM Trans. on Graphics (TOG)*, vol. 27, no. 3, 2008.
- [138] A. Levin, R. Fergus, F. Durand, and W. T. Freeman, “Image and depth from a conventional camera with a coded aperture,” *ACM Trans. on Graphics (TOG)*, vol. 26, no. 3, 2007.
- [139] M. Ben-Ezra and S. K. Nayar, “Motion-based motion deblurring,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 689–698, 2004.
- [140] Y.-W. Tai, H. Du, M. S. Brown, and S. Lin, “Image/video deblurring using a hybrid camera,” in *IEEE Conference on Computer Vision and Pattern Recognition - CVPR*, 2008.
- [141] M. A. T. Figueiredo, D. S. Cheng, and V. Murino, “Clustering under prior knowledge with application to image segmentation,” in *Advances in Neural Information Processing Systems - NIPS*, B. Schölkopf, J. Platt, and T. Hoffman, Eds. Cambridge, MA: MIT Press, 2007, pp. 401–408.
- [142] M. S. C. Almeida and L. B. Almeida, “Blind deblurring of two-layer images,” in *Portuguese Conf. on Pattern Recognition - RecPad*, Aveiro, Portugal, October 2009.
- [143] —, “Blind deblurring of foreground-background images,” in *IEEE International Conference on Image Processing - ICIP*, Cairo, Egypt, 2009, pp. 1301–1304.
- [144] A. Kraskov, H. Stögbauer, and P. Grassberger, “Estimating mutual information,” *Physical Review E*, vol. 69, p. 066138, 2004. [Online]. Available: <http://arxiv.org/pdf/cond-mat/0305641>
- [145] P. Moulin and J. Liu, “Analysis of multiresolution image denoising schemes using generalized-gaussian and complexity priors,” *IEEE Trans. on Information Theory*, no. 3, pp. 909–919, April 1999.
- [146] I. W. Selesnick, R. G. Baraniuk, and N. G. Kingsbury, “The dual-tree complex wavelet transform,” *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 123–151, November 2005.
- [147] G. P. Nason and B. W. Silverman, “The stationary wavelet transform and some statistical applications,” in *Wavelets and Statistics*, ser. Lecture Notes in Statistics. Springer-Verlag, 1995, pp. 281–300. [Online]. Available: [citeseer.ifi.unizh.ch/nason95stationary.html](http://citeseer.ifi.unizh.ch/nason95stationary.html)
- [148] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” *IEEE Trans. on Image Processing*, vol. 13, pp. 600–612, 2004.
- [149] H. R. Sheikh, A. C. Bovik, and G. de Veciana, “An information fidelity criterion for image quality assessment using natural scene statistics,” *IEEE Trans. on Image Processing*, vol. 14, no. 12, pp. 2117–2128, 2005.

- [150] E. P. Simoncelli and B. Olshausen, “Natural image statistics and neural representation,” *Annual Review of Neuroscience*, vol. 24, pp. 1193–1216, May 2001. [Online]. Available: <http://www.cns.nyu.edu/pub/eero/simoncelli01-reprint.pdf>
- [151] F. A. A. Kingdom, A. Hayes, and D. J. Field, “Sensitivity to contrast histogram differences in synthetic wavelet-textures,” *Vision Research*, vol. 41, no. 5, pp. 585–598, 2001.
- [152] Z. Wang and E. P. Simoncelli, “Reduced-reference image quality assessment using a wavelet-domain natural image statistic model,” in *Annual Symposium on Electronic Imaging*, Jan. 17-20 2005.
- [153] M. Miyahara, K. Kotani, and V. R. Algazi, “Objective picture quality scale (pqs) for image coding,” *IEEE Trans. on Communications*, vol. 46, no. 9, pp. 1215–1226, 1998.
- [154] R. Guidara, S. Hosseini, and Y. Deville, “Maximum likelihood blind image separation using nonsymmetrical half-plane markov random fields,” *IEEE Trans. on Image Processing*, vol. 18, no. 11, pp. 2435–2450, Nov. 2009.
- [155] L. Almeida, “Linear and nonlinear ICA based on mutual information – the MISEP method,” *Signal Processing*, vol. 84, no. 2, pp. 231–245, 2004. [Online]. Available: <http://www.lx.it.pt/~lbalmeida/papers/AlmeidaSigProc03.pdf>
- [156] K. Zhang and L. Chan, “Nonlinear independent component analysis with minimal nonlinear distortion,” in *ICML 07: Proceedings of the 24th international conference on Machine learning*, New York, NY, USA, 2007, pp. 1127–1134.
- [157] J. M. Bioucas-Dias, M. A. T. Figueiredo, and J. P. Oliveira, “Adaptive total-variation image deconvolution: A majorization-minimization approach,” in *European Signal Processing Conference - EUSIPCO*, Florence, Italy, September 2006.
- [158] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D*, vol. 60, pp. 259–268, 1992.
- [159] Y. P. Guo, H. P. Lee, and C. L. Teo, “Blind restoration of images degraded by space-variant blurs using iterative algorithms for both blur identification and image restoration,” *Image and Vision Computing*, vol. 15, no. 5, pp. 399–410, 1997.
- [160] M. Welk, D. Theis, and J. Weickert, “Variational deblurring of images with uncertain and spatially variant blurs,” in *DAGM Symposium for Pattern Recognition*, 2005, pp. 485–492.
- [161] L. G. Shapiro and G. C. Stockman, *Computer Vision*. Prentice-Hall, 2001, chap.5.
- [162] L. G. Roberts, “Machine perception on three-dimensional solids,” Massachusetts Inst. of Tech. Lexington Lab, Tech. Rep., 1963.
- [163] J. M. S. Prewitt, “Object enhancement and extraction,” *Picture Processing and Psychopictorics*, 1970.
- [164] L. S. Davis, “A survey of edge detection techniques,” *Computer Graphics and Image Processing*, vol. 4, pp. 248–270, 1973.

- [165] M. Nikolova, “Analysis of the recovery of edges in images and signals by minimizing non-convex regularized least-squares,” *SIAM Journal on Multiscale Modeling and Simulation*, vol. 4, no. 3, pp. 960–991, 2005.
- [166] L. B. Almeida, “Multilayer perceptrons,” in *Handbook of Neural Computation*, E. Fiesler and R. Beale, Eds., Institute of Physics. Oxford University Press, 1997. [Online]. Available: <http://www.lx.it.pt/~lbalmeida/papers/AlmeidaHNC.pdf>