Bleed-Through Document Image Restoration

A dissertation submitted to the University of Dublin
for the degree of Doctor of Philosophy

Róisín Rowley-Brooke
Trinity College Dublin, March 2014

SIGNAL PROCESSING AND MEDIA APPLICATIONS
DEPARTMENT OF ELECTRONIC AND ELECTRICAL ENGINEERING
TRINITY COLLEGE DUBLIN

To my family and friends

...dabit deus his quoque finem.

Abstract

Digitisation of original document sources for the purpose of conservation, detailed study, and facilitating access for a wider audience has been an increasing trend over recent years, particularly with constantly improving imaging technology available at ever decreasing costs. Many documents suffer from a wide variety of degradations that reduce their legibility and usefulness as sources. With the increase in digitisation has also come an increase in image processing based enhancement and restoration techniques. This thesis presents new approaches to automatic restoration of one particular type of degradation - bleed-through, which occurs when ink from one side of a page seeps through and interferes with the text on the other side, reducing legibility. Previous approaches to bleed-through restoration struggle in severe cases, when there is a significant overlap between bleed-through and foreground text intensities. There are also relatively few methods that focus on preserving the document appearance in the output.

The first contribution of this thesis is an approach to automatic bleed-through restoration where the observed degraded images are modelled as a per-pixel linear combination of the original clean images, and some proportion of the reverse side, however, bleed-through is explicitly constrained to non-foreground regions only. In order to preserve the document appearance, the model parameters, including clean image intensities themselves, are estimated in a Bayesian framework, with priors to ensure that restored regions are close in appearance to the background.

Though visual results show that this approach performs well on light to medium bleed-through examples, it does not perform as well on more severely degraded cases. Therefore a second approach for restoration is proposed where the bleed-through classification and restoration stages are decoupled to focus more explicitly on bleed-through identification. In this approach both sides of the page are classified jointly, resulting in a single label field which is then refined using rule-based connected component analysis. To maintain the goal of preservation of the document appearance, restoration is performed using exemplar based inpainting.

To overcome the issue of lack of access to high resolution image examples for testing, a small database of bleed-through example image pairs with manually generated ground truth foreground text masks, compiled for this work and made available for research purposes, is presented. A numerical evaluation methodology for bleed-through restoration based on the database is also described. Numerical and visual comparisons of the proposed restoration methods with existing techniques show that the second decoupled approach performs best overall.

In order to use information from both sides of the page in automatic bleed-through restoration approaches, they must first be spatially registered to a high degree of accuracy so that bleed-through on each side is aligned with its corresponding originating text. The final contribution of this thesis is an automatic bleed-through document registration approach that uses the page outline as the strongest matching feature between the two sides, rather than the bleed-through and text. A local grid point displacement-based refinement similar to existing approaches is then applied, however, the difference between both image intensity and gradient are used as the similarity metric. A numerical evaluation approach for bleed-through registration is also presented, and based on this the proposed method is shown to outperform previous works, and the global outline warp stage is shown to improve the performance of previous approaches.

Declaration

I hereby declare that this thesis has not been submitted as an exercise for a degree at this or any other University and that it is entirely my own work.

I agree that the Library may lend or copy this thesis upon request.

Signed,

Róisín Rowley-Brooke

March 14, 2014.

Acknowledgments

As the light at the end of this long tunnel becomes visible I would like to thank all of the many people who have helped keep me on track along the way, and a few in particular. Firstly, sincere thanks to my supervisor Dr. Anil Kokaram for introducing me to document restoration, and for his continued kindness, guidance, and encouragement over the past four years. Thank you to my supervisor Dr. François Pitié, to whom I am hugely indebted for his time, invaluable advice, good humour, and endless patience. Thank you also to Dr. Naomi Harte for her support and understanding.

To all the members of Sigmedia and the EEE Department over the past four years, and especially the occupants of Room 2.07, thanks for many great memories, and for making such a supportive and enjoyable environment. A big thank you to Finnian for the innumerable chats and laughs and coffee breaks that helped get me through the day, and also to Bernie for always having a kind word and a smile in the mornings. A grateful thank you to Dr. Gary Baugh for bailing me out on numerous occasions with code implementation, and to John Squires for his help with creating a manuscript display.

Thanks to Dr. Pádraig Ó Macháin for his assistance with access to the ISOS images, and for facilitating the database creation and website, and to Anne Marie O'Brien for managing updates and requests for access. Thanks also to Jeff Briedenbach for providing images from the Google Books collections, the Governors and Guardians of Marsh's Library, Dublin, for copyright permissions, and to Dr. Kerry Houston for trusting me with the St. Patrick's Cathedral organ MS.

This research was generously funded by the Irish Research Council EMBARK Initiative, and also part funded by Science Foundation Ireland and Google inc., to all of whom I am very grateful.

To my friends - especially all the Chapellers - thank you for providing welcome (often excessive!) distractions, along with much moral support.

A huge thank you to Mam, Dad, Naoise, Myles, and Cassa for their wonderful love and encouragement over the last four years and before, and to my nieces Mia, Róisín, and Caitlin, for often giving some much needed perspective!

Finally, to Darren, without whom this would have been impossible, for the unfailing love and support, for never tiring of 'this time it's different - you don't understand!', and for being the best part of the past eight years - thank you.

Contents

Contents **vii**

List of Acronyms **xi**

List of Manuscript Image Copyright Acknowledgements **xiii**

List of Manuscript Image Copyright Acknowledgements **xv**

1 Introduction **1**
- 1.1 Bleed-Through . . . 3
- 1.2 Thesis outline . . . 5
- 1.3 Contributions . . . 7
- 1.4 Publications . . . 8

2 A Review of Bleed-Through Restoration Techniques **11**
- 2.1 Intensity Information . . . 13
  - 2.1.1 Thresholding . . . 13
  - 2.1.2 Clustering . . . 15
  - 2.1.3 Blind Source Separation . . . 15
  - 2.1.4 Diffusion . . . 17
- 2.2 Spatial Information . . . 18
  - 2.2.1 Dual-Layer MRF . . . 18
  - 2.2.2 Wavelet Enhancement . . . 20
  - 2.2.3 Variational Model . . . 21
  - 2.2.4 Active Contours . . . 22
- 2.3 Summary . . . 24

3 Model-Based Approach **27**
- 3.1 The Case for a Linear Model . . . 28
- 3.2 Degradation Model . . . 30
- 3.3 Bayesian Framework . . . 30
  - 3.3.1 Likelihood . . . 31

3.3.2 Priors . . . 31
3.4 Solution . . . 33
3.4.1 Masks Estimate . . . 33
3.4.2 Mixing Parameters Estimate . . . 34
3.4.3 Clean Images Estimate . . . 34
3.4.4 Initialisation . . . 35
3.4.5 Algorithm . . . 37
3.5 Discussion . . . 37
3.6 Colour Restoration . . . 39
3.7 Summary . . . 41

4 Non-Parametric Approach 45
4.1 Preprocessing . . . 47
4.2 Classification . . . 48
4.2.1 Joint Histogram Segmentation . . . 51
4.2.2 Image Segmentation . . . 55
4.3 Restoration . . . 56
4.3.1 Clean Background . . . 56
4.3.2 Blending . . . 57
4.3.3 Colour Restoration . . . 57
4.4 Model Analysis . . . 64
4.5 Summary . . . 70

5 Evaluation 73
5.1 Bleed-Through Database . . . 74
5.1.1 Image Capture . . . 74
5.1.2 Crop Details . . . 74
5.1.3 Registration . . . 75
5.1.4 Ground Truth Creation . . . 76
5.1.5 Access . . . 76
5.2 Quantitative Analysis . . . 76
5.2.1 Binarisation Algorithm . . . 76
5.2.2 Bleed-Through Error Metrics . . . 77
5.2.3 DIBCO Evaluation Metrics . . . 78
5.3 Results . . . 80
5.3.1 Methods Selected for Comparison . . . 80
5.3.2 Visual Comparison . . . 89
5.3.3 Numerical Comparison . . . 90
5.3.4 Discussion . . . 94

5.4 Summary 95

6 Registration **97**
6.1 Initial Global Warp 99
6.2 Local Grid Warp 100
6.2.1 Data Term 101
6.2.2 Similarity Term 102
6.2.3 Optimisation 103
6.3 Quantitative Analysis 104
6.3.1 Local warp parameter variations 105
6.4 Results & Discussion 107
6.5 Summary 117

7 Conclusion **121**
7.1 Restoration 121
7.2 Registration 123
7.3 Future work 123
7.4 Final Remark 125

A Full Results for Chapter 5 **127**

B Full Results for Chapter 6 **133**

Bibliography **141**

List of Acronyms

BSS	Blind Source Separation
DCT	Discrete Cosine Transform
DIBCO	Document Image Binarisation Competition
DRD	Distance Reciprocal Distortion Metric
EM	Expectation Maximisation
ICA	Independent Component Analysis
ICM	Iterated Conditional Modes
ISOS	Irish Script On Screen Project
KNN	K-Nearest Neighbours
MAP	Maximum A Posteriori
MS	Manuscript
MRF	Markov Random Field
MSE	Mean Squared Error
OCR	Optical Character Recognition
OMR	Optical Music Recognition
PCA	Principal Component Analysis
PSNR	Peak Signal-to-Noise Ratio
QPBO	Quadratic Pseudo Boolean Optimisation
RP	Ranked Pairs Voting
SSD	Sum of Squared Differences
SVM	Support Vector Machine

List of Manuscript Image Copyright Acknowledgements

M01	With the kind permission of the Allan and Maria Myers Academic Centre, Newman College and St. Mary's College, University of Melbourne, Australia.	MS Brisleach Mhór Mhuighe Muirtheimhne	90r/v
M02	With the permission of the Governors and Guardians of Marsh's Library, Dublin.	Queen Elizabeth's Irish Primer	7r/v
M03			9r/v
M04	Courtesy of Google Books.		
M05			
M06			
M07			
M08			
M09	Courtesy of the National Library of Ireland.	MS G 18	147/148
M10			189/190
M11			273/274
M12			361/362
M13		MS G 311	265/266
M14	Courtesy of the James Hardiman Library Archives, National University of Ireland, Galway.	LS de hÍde 18	153/154
M15			155/156
M16	By permission of the Librarian, National University of Ireland Maynooth, from the collections of St Patrick's College, Maynooth.	MS M 86	13/14
M17			15/16
M18		MS R 68	80/81
M19	By permission of the President and Fellows of Corpus Christi College, Oxford.	MS.201	22r/v

List of Manuscript Image Copyright Acknowledgements

M20	By kind permission of Jesus College, Oxford.	MS.16	1r/v
M21	By permission of the Royal Irish Academy © RIA.	MS C iii 3	301r/v
M22			388r/v
M23			420r/v
M24	By kind permission of The Board of Trinity College Dublin.	MS 1333	9/10
M25		MS 1343	59/60
M26		MS 1435	147/148
M27		MS 1436	81/82
M28	With the kind permission of the Special Collections Librarian, University College Cork Library.	MS 96	547/548
M29	Images courtesy of ISOS; Original manuscript held in University College Dublin Library Special Collections.	Additional MS 14	384/385
M30			386/387
M31			726/727
M32	By kind permission of the UCD-OFM Partnership.	MS A 13	xxv/xxvi
M33		MS A 15	5/6
M34			37/38
M35		MS A 20	127r/v
M36		MS A 29	12r/v
M37			119r/v
M38			121r/v
M39		MS A 33	87/88

1

Introduction

The written word is one of the most important inventions in human history; the development of writing systems facilitated advances in civilization through the preservation of knowledge beyond individual memory and the oral tradition [26]. From the first inventories on clay tablets, through the development of literary traditions, complex legal systems, paper, and the printing press, there exists a vast wealth of information preserved in written form, representing a large part of modern cultural heritage. Original document sources are therefore often the first point of reference for many researchers, including historians, linguists, palaeographers, musicologists, and literary scholars to name a few.

One of the commonest problems encountered in the study of document sources, especially historical, is the reduction in legibility due to degradation. There are many different types and causes of degradation, with the severity affected by different factors such as the document materials, storage conditions, and even document content. Some of the degradations and causes of degradation that affect legibility include:

(i) ink fading due to excessive light exposure or water damage.

(ii) page 'ink corrosion' due to the acidic content of the ink used for writing.

(iii) ink bleed-through, often exacerbated by highly porous writing media, and damp storage conditions.

(iv) deliberate damage or ink effacing from censorship or conflicting opinions.

(v) earlier text occlusion as a results of reuse of the writing medium for more than one text - referred to as palimpsests[1].

(vi) acid migration from storage containers.

(vii) bacterial or fungal infection of the medium, often due to excessive storage humidity and temperature.

(viii) warping of the medium due to variations in storage temperature and humidity.

(ix) holes in the medium from pest damage, such as book beetles, silverfish, or mice.

(x) accidental damage, for example due to flood, fire, or over-handling of the document.

Physical restoration of degraded documents is an expensive and time consuming process, and can often result in unintended damage to the integrity of the original. For example, to remedy the fading of iron based inks, it was common since the 17th Century to add gallic acid to intensify the text. However, this restoration results in further degradation of the manuscripts over subsequent years as the added solution browns the surrounding page regions, sometimes rendering them illegible. In the 19th and early 20th Centuries other solutions were used for the same purpose, such as tea or tanning acid, ammonia, ammonium sulphide and potassium ferrocyanide, with similarly destructive properties [27, 36]. Another more famous example of damage done during restoration is the Book of Kells. For one of its several rebindings, in the early 19th Century, the entire book was heavily trimmed to make the edges more uniform, and they were subsequently gilded. However, this also resulted in the indiscriminate cropping of some of the rare illuminations that the book is most famous for [42]. The clear danger with physical restoration is that any mistakes made are generally irreversible. In recent years, due to the progress in digital capabilities amongst other reasons, it has been noted that the trend in restoration has moved towards a less invasive approach, with the focus moving towards preservation of original forms, and digitisation [36].

Digitisation of documents, and especially those that are degraded, has many benefits. As degradation is a continuous process, the digitisation allows a snapshot in time to be taken for future comparison and monitoring of the rate of decline. Also having a high quality copy reduces the amount of interaction with the original, thus preventing further damage from excessive handling. Though historical document imaging has been performed since the late 19th century [30], the copies created were generally limited to grayscale only, and could never provide a substitute for examining the physical document. With recent advances in digital imaging technology, the representations of documents have become more faithful to the original, and can even provide a

---

[1] from the Greek *πάλιν* - again, and *ψάω* - rub or wipe. Because of the relative expense of writing materials, the writing medium was often scraped clean of existing text that was deemed unnecessary, and reused for new work, often leaving small visible traces of the scraped text. It is often the underlying text that is of greater interest to researchers nowadays.

more detailed view than visible with the naked eye. The decrease in cost of sophisticated high resolution imaging equipment has also made document imaging a much more common practice. Furthermore, progress in online capabilities has made imaged documents previously accessible only to a privileged few available to a much wider audience, notable examples being the Early Manuscripts at Oxford University [4], and the Archimedes Palimpsest Project [1]. The availability of document images alongside modern scholarly editions now enriches the experience of casual readers as well as researchers, who can browse and access source material to gauge its usefulness before visiting and viewing the original, and can also refer back any number of times to the digital copy [27].

With the increase in prevalence of digitisation, has come an increase of interest in automatic document image analysis and restoration. Digital restoration of degraded documents has many advantages over physical. Firstly, in the digital domain, any number of changes may be made to the document appearance, whilst leaving the original intact - this allows for several different restoration approaches to be tested without risking damage to the physical document. Secondly, alongside high resolution imaging, digital analysis and restoration have the potential to reveal information that could not possibly be recovered using manual techniques. One of the more popular trends in recent times is the use of multispectral imaging to reveal hidden text information, especially in the work associated with the Archimedes Palimpsest [1], though this is not a recent innovation - in early 20th century the use of different artificial light filters, as well as ultra violet and infra red light sources, for photographing and recovering obscured or erased text, was developed [30]. However, the inclusion of image processing techniques and improvement in control over the range of wavelengths used for imaging has increased the range of degradations that can be restored. Even relatively simple image processing techniques such as contrast enhancement can be used to recover erased or faded text that is almost invisible to an impressive degree [2, 27].

1.1 Bleed-Through

This thesis is concerned with automatic restoration of a specific type of document degradation. Bleed-through is a form of textual interference that occurs when ink from one side of a page has seeped through and becomes visible on the other side, see Figure 1.1. This often results in reduction of legibility of the foreground text for those viewing and studying the document, and can also cause problems for automatic content based processing, such as binarisation, Optical Character Recognition (OCR), or layout analysis. Bleed-through restoration is therefore an important area of research, both in terms of improving the legibility of degraded images for individual researchers examining original sources, and also for larger scale digitisation projects, such as Google Books [5], that use automatic processing to make document collections searchable online. Bleed-through restoration is a challenging problem for a number of reasons. Firstly, in severe cases, there is a significant overlap between foreground, background, and bleed-through

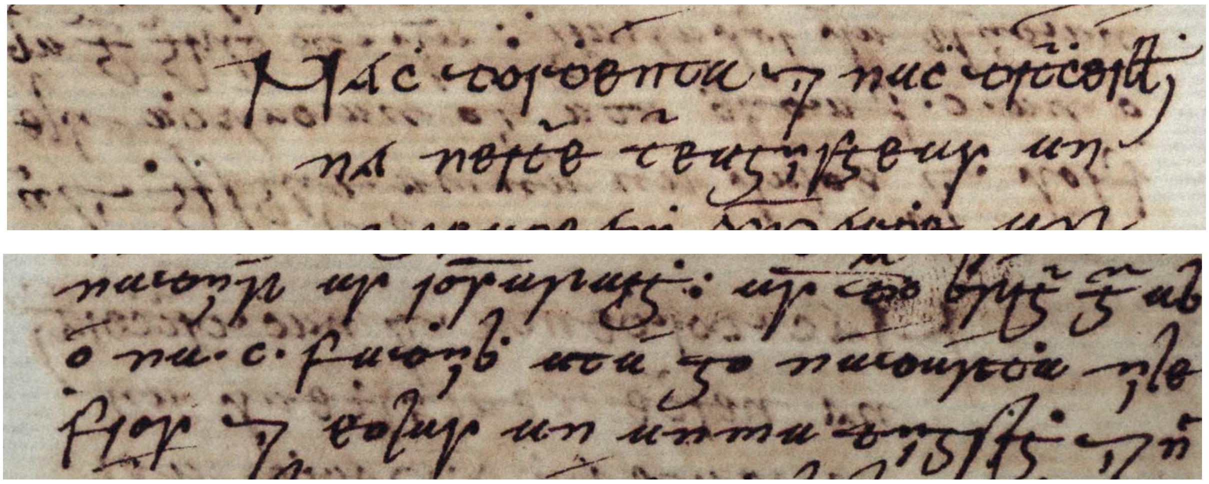


Figure 1.1: An example of a cropped recto (top) and verso image pair suffering from bleed-through degradation, taken from *Buaidh na Naomhchroiche*, a 17th century Irish translation of Girolamo Savonarola's *Triumphus Crucis*, fol. 121r/v. **M38**.

intensities. Secondly, in order to use information from both sides of the page to guide restoration, the two sides must first be spatially aligned, that is registered, to a high degree of accuracy, so that bleed-through degradation on one side of the page aligns with its corresponding originating text. Finally, as there is no one standard for document imaging, and as bleed-through can occur in a wide variety of different document types to a lesser or greater extent, the properties of bleed-through degradation can vary greatly between images.

Terminology

When referring to documents, especially historical, the following terminology is often used: *Folio* from the Latin *folium* ('leaf') has several meanings, one of which refers to a single sheet of paper or writing medium with two sides - *recto* and *verso*. The word folio can also refer to a sheet of paper that is folded to make two leaves of a book, a specific size of manuscript, or to a whole volume itself, made up of gathered pages [13]. The recto side of a page is the first side as when viewing a book, and the corresponding verso side is on the back of the recto. Document pages are often numbered therefore on the recto side of the folio only, and page referencing is in terms of recto and verso, for example 'pages 13v-15r'. In left-to-right languages, such as Latin and English, the recto will always be on the right hand side. However, in right-to-left languages, for example Arabic and Syriac, the recto will be on the left. In the case of rolled scrolls, recto refers to the inside, and verso to the outside. The derivation of the terms *recto* and *verso* originates from the process of making papyrus, where two layers of reeds stem strips were pressed together in water at right angles, then dried and polished with pumice to form a smooth writing surface [40]. In this case *recto* refers to the front side where the reeds run horizontally from the Latin *rectus* meaning straight, or ruled (this word has an alternative meaning 'right' hence *recto folio* - the page on the right - for Western books). *verso* is then the back side where the reeds run vertically, from *versus*, meaning turned. In document image processing, and more usually in blind bleed-through restoration, *recto* may often also refer to the side of the page being processed and restored, regardless of its location within the document, and *verso* is the source of the bleed-through degradation.

Current Issues

Previous work towards bleed-through restoration has struggled in more severely degraded examples, especially in regions where the foreground and bleed-through overlap, and in most cases, no attempt is made to preserve the inherent document characteristics in the output restored results. Access to high resolution, good quality bleed-through examples can be difficult unless connected with a specific library or digitisation project. Also, due to copyright issues, even if access is available, it can be difficult to publish results and make data available. As a result, there is a lack of available bleed-through databases and a standard numerical evaluation methodology for comparison of different bleed-through restoration approaches. Similar problems exist in the case of recto-verso registration of bleed-through degraded documents. In previous registration approaches, the accuracy of the registration itself is generally not considered, with the focus being on the performance of subsequent bleed-through restoration. The often required preprocessing of document images prior to registration has also not been addressed.

Objectives

Following from the issues described, the main aims of this thesis are as follows:

- To develop new bleed-through registration and restoration techniques that can cope with a wider range of degradations than previous approaches, whilst maintaining as much of the original document appearance as possible.
- To provide access to images used for testing, in order to facilitate comparisons with other works.
- To develop numerical evaluation approaches for both bleed-through restoration and registration results.

1.2 Thesis outline

The remainder of this thesis is organised as follows.

Chapter 2: Bleed-Through Removal: A Review

In this chapter the challenges associated with bleed-through removal are described, and a review of previous approaches proposed for bleed-through restoration is presented. Bleed-through restoration approaches generally fall into one of two groups: blind or non-blind, depending on whether they operate using one or both sides of the page. Different methods may be further classified depending on whether they use intensity information only for restoration, or intensity and spatial information, what type of output they generate, and the related motivation for restoration.

Chapter 3: Model-Based Approach

This chapter presents an approach for non-blind bleed-through restoration that seeks to preserve as much as possible of the character and content of the original document in the restored result, with legibility improved in bleed-through regions. A linear generative model for the degradation is used, where the observed degraded images are viewed as a per-pixel combination of the original clean images and some proportion of the corresponding reverse sides. The proposed model and solution are developed in a Bayesian dual-Markov Random Field (MRF) Maximum a Posteriori (MAP) framework, solving for 6 unknown parameters per pixel, including the clean image intensities themselves, and priors are included to ensure that the bleed-through restored regions are as close in appearance as possible to the surrounding background.

Chapter 4: Non-Parametric Approach

In the approach described in Chapter 3, the aim is to try and perform classification and restoration in the same step, focussing on the clean image intensity estimation, and it is heavily reliant on good initial estimates to obtain good results. To improve performance on more severe cases of bleed-through, this chapter presents a second framework for non-blind bleed-through restoration, where the classification and restoration processes are decoupled in order to focus explicitly on bleed-through region classification. Therefore the proposed framework contains several discrete stages including image preprocessing to remove local intensity variations, joint pixel region classification using a segmentation of the joint recto-verso intensity histogram followed by connected component analysis on the corresponding joint image labelling, and finally restoration using exemplar-based image inpainting to preserve the restored document character compared to the original.

Chapter 5: Evaluation

This chapter presents the details of a small manuscript bleed-through image database with ground truth foreground text masks and preregistered recto and verso sides, made available for research purposes. A numerical evaluation approach for bleed-through removal, based on the database, is then described, and the methods presented in Chapters 3 and 4 are compared visually and numerically against three previous non-blind approaches, using both the proposed evaluation and standard document binarisation evaluation metrics.

Chapter 6: Registration

This chapter presents a framework for bleed-through document recto-verso registration that consists of two distinct stages. Firstly, an initial global crop and alignment using binary page outline images is performed, designed both to remove extraneous image information from the registration, and to cope with examples where there is little or no bleed-through. Secondly, following previous works, a local grid-based warp is applied, however, using both the sum of squared differences (SSD) in image intensities and also gradient fields as a warp metric. An approach for numerical evaluation of bleed-through registration itself is also presented, and using this the results of the proposed approach is compared to previous methods.

Chapter 7: Conclusion

The final chapter summarises and assesses the contributions of this thesis, and outlines some directions for future work.

1.3 Contributions

The new work described in this thesis can be summarised as follows:

Chapter 3:

- An explicitly constrained double sided per-pixel linear model for bleed-through degradation, incorporating the spatial variation in bleed-through properties across a page.
- Incorporation of the linear model into a Bayesian framework for bleed-through restoration, where the clean image intensity itself is estimated in bleed-through regions.

Chapter 4:

- Joint intensity histogram domain classification of bleed-through degraded documents, where the two sides are classified simultaneously into four joint classes, rather than independently.
- Rule based connected label component analysis to refine bleed-through classification.
- Use of exemplar based image inpainting to preserve the original document characteristics in the restored result.

Chapter 5:

- A publicly available manuscript bleed-through document database with ground truth foreground text masks, and preregistered recto-verso pairs.
- A numerical evaluation approach for comparing the performance of different bleed-through restoration algorithms.

Chapter 6:

- Automatic cropping and global registration of recto-verso image pairs based on the page outline shape.
- Use of the difference between image gradient fields as well as intensity for a recto-verso registration matching metric.
- A numerical evaluation approach for comparing the performance of different bleed-through registration algorithms.

1.4 Publications

Portions of the work described in this thesis have appeared in the following publications:

Peer Reviewed Conference Papers

- Róisín Rowley-Brooke and Anil Kokaram. Bleed-Through Removal in Degraded Manuscripts. In *Proceedings of the Irish Signals and Systems Conference (ISSC '11)*, pages 255-260, Dublin, Ireland, June 2011.
- Róisín Rowley-Brooke and Anil Kokaram. Degraded Document Bleed-Through Removal. In *Proceedings of the Irish Machine Vision and Image Processing Conference (IMVIP '11)*, pages 70-75, Dublin, Ireland, September 2011.
- Róisín Rowley-Brooke and Anil Kokaram. Bleed-Through Removal in Degraded Documents. In *Proceedings of IS&T/SPIE Electronic Imaging Symposium, Document Recognition and Retrieval XIX (DRR '12)*, Burlinghame, CA, USA, January 2012.
- Róisín Rowley-Brooke, François Pitié, and Anil Kokaram. A Ground Truth Bleed-Through Document Image Database. In *Proceedings of the International Conference on Theory and Practice of Digital Libraries (TPDL '12)*, pages 185-196, Paphos, Cyprus, September 2012.
- Róisín Rowley-Brooke, François Pitié, and Anil Kokaram. A Non-Parametric Framework for Document Bleed-Through Removal. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '13)*, Portland, OR, USA, June 2013.
- Róisín Rowley-Brooke, François Pitié, Anil Kokaram. Nonrigid Recto-Verso Registration Using Page Outline Structure and Content Preserving Warps. In *Proceedings of the 2nd International Workshop on Historical Document Imaging and Processing (HIP '13)*, Washington DC, USA, August 2013.

Presentation

- Oral presentation of paper - Degraded Manuscript Restoration: A Case Study. In *Annual Conference of the Society for Musicology in Ireland (SMI'13)*, Maynooth, Ireland, June 2013.

2

A Review of Bleed-Through Restoration Techniques

The purpose of bleed-through restoration is to improve the legibility of affected documents, either for the benefit of human readers, or to assist further automatic processing, or both. As mentioned in the introduction to this thesis, previous approaches may be classified into one of two groups, blind or non-blind, depending on whether they operate with one or both sides of the document. Blind methods perform bleed-through restoration using one side of the page only, whereas non-blind methods use both sides of the page, the advantage being that there is more information available to work with. However, highly accurate registration of the two sides of the page is an essential pre-processing step to ensure that bleed-through on each side is aligned with its corresponding originating text - this is a non-trivial problem in itself (see Chapter 6). Furthermore, in some cases there may only be an image of one side of the page available.

In bleed-through restoration the main source of information used to generate the restored recto and verso images Y_r, Y_v, is the intensity of the observed images I_r, I_v, Most previous approaches perform restoration based on the grayscale intensity alone, though some do use full RGB colour. The intensity information alone however, is insufficient in severe bleed-through cases to distinguish between foreground and bleed-through regions, as there is often a significant overlap between the corresponding intensity profiles. Therefore some previous approaches also incorporate spatial information in order to model the relationship between neighbouring regions to improve restoration.

Previous approaches to restoration may also be grouped according to the final restored results produced. There are three types of output generated:

(i) Binary - where all background and bleed-through information is replaced with a uniform

Figure 2.1: Examples of the different bleed-through restoration outputs on a manuscript extract. Top row: Original image (left) and binary output. Bottom row: Pseudo-binary (left), and textured outputs. **M25**.

intensity, and foreground information is replaced with a contrasting uniform intensity.

(ii) Pseudo-Binary - where, again, all background and bleed-through information is replaced with a uniform intensity, but varying foreground intensities are preserved.

(iii) Textured - where varying foreground and background intensities are preserved, and bleed-through is replaced with an estimate of the background texture.

For modern printed documents, binary and pseudo-binary outputs will not affect the overall appearance for human readers. However, as can be seen in Figure 2.1 when these outputs are generated for historical documents on more textured writing media, though the legibility is improved, the character of the document is destroyed with the removal of background texture. A more visually pleasing output is one that preserves the texture of the page as much as possible. A further advantage of preserving background and foreground texture is that this allows for some ambiguity in classification of character edges, which in historical documents especially may not be sharply defined. The disadvantage of preserving the texture is that the complexity of the problem is increased by the need to estimate the background texture in bleed-through regions.

This chapter presents a brief review of previous work towards bleed-through restoration, examining the underlying intensity and spatial models used.

2.1 Intensity Information

As discussed above, the most common source of information is the intensity of the observed degraded images. Intuitively, it is clear that in most cases the background will correspond to the lightest regions of the document, the foreground will be the darkest, and the bleed-through can be assumed to be somewhere in between. Most approaches that use only intensity information exploit this, in general assuming that there is a clear distinction between the foreground and bleed-through intensity profiles.

2.1.1 Thresholding

Thresholding is perhaps the simplest method of image segmentation that splits the image into different regions based on one or more intensity threshold values. The most common use of thresholding in document processing is for binarisation, where the document is split into foreground and background. One of the most popular techniques that is often used as a benchmark for other methods, is that of Otsu [50], where the optimal threshold is defined so as to maximise the variance between foreground and background intensities, and to minimise the variance within. As mentioned previously, binarisation is an important process in document image analysis, and there have been numerous degraded document binarisation algorithms based on thresholding proposed that aim to cope with various degradations such as non-uniform intensity profiles, water and dirt stains, and poor contrast. However, most document binarisation approaches do not deal explicitly with bleed-through degradation - only those that do are discussed here. For further discussion and comparison of document binarisation techniques the reader is directed to [31, 48, 64, 66], and also the results of the annual IEEE Document Binarisation Contest(DIBCO) [23, 52–54], for more recent innovations.

The main drawback with use of thresholding for bleed-through removal, is that it assumes that, globally or locally, the foreground text is always darker than surrounding bleed-through degradations, and as a result will perform well on documents with light bleed-through only. In cases where there is a significant overlap between foreground and bleed-through intensities and text locations, thresholding alone produces poor results. A solution to this is proposed in [22], where the observation is made that, though some foreground text strokes may be fainter than bleed-through, they are very likely to be attached to foreground strokes that are darker. Therefore it is proposed that two intensity thresholds should be defined, either globally or locally; a high threshold, t_h, above which there will be negligible bleed-through, and a low threshold, t_l, below which there will be no foreground. The image is thresholded using t_h to remove all bleed-through and background, and then oversegmented foreground characters are regrown, with t_l as the limit for regrowth. To cope with regions where the recto text and bleed-through overlap, further empirically defined thresholds are applied in the regrowth stage on the gradient magnitude and direction, and maximum character size. This method produces some impressive results, however still falls short in the cases where some bleed-through text is darker than or as dark as the recto text, resulting in misclassification and regrowth of bleed-through interference. The approach is parameter intensive, as the threshold levels need to be tuned per image to obtain the best results.

Non-Blind Thresholding

Several methods have been proposed that seek to improve thresholding results by including information from the other side. Burgoyne et al. in [14] focus on music documents and propose extensions to the symmetric and non-symmetric Kullback-Leibler (KL) thresholding algorithms by adding in a second threshold level to distinguish between background and bleed-through, as well as background and foreground, based on the assumption that bleed-through will always be lighter than foreground text. They also propose a similar extension to the degraded document binarisation algorithm of Gatos et al. [24] (see Section 5.2 for more information), where a binarised version of the verso side is used to guide thresholding on the recto side, using foreground from the verso to indicate possible locations of bleed-through. The incorporation of registered verso information does not improve the results of the KL algorithm, but some minor improvements are shown in the Gatos method on historical documents with severe bleed-through, however it is concluded that these improvements are outweighed by the computational requirements of image registration.

Dubois and Pathak [21] use rule based comparison between recto and registered verso intensities to locate regions of bleed-through. In this approach a pixel is identified as bleed-through if the recto intensity is darker than some user defined threshold AND if the ratio between recto and verso intensities is less than a user defined ink attenuation coefficient. A similar comparison method is used in [75], where it is assumed that recto pixels that are darker than corresponding verso pixels are foreground text, and pixels that are lighter are bleed-through interference. These approaches both rely on the same global illumination properties for both sides of the page.

The thresholding method in [21], is modified in [20], where they identify four regions in each image to be segmented instead of locating bleed through only. The identified regions are foreground only, bleed-through only, background, and foreground bleed-through overlap. Firstly, the background threshold is set as 90% of the peak of the intensity histogram. Next foreground regions are located by examining minimum filtered versions of the recto and verso. The verso side is further brightened by an empirically chosen attenuation coefficient, and recto foreground set as regions where the filtered recto is darker than the filtered, brightened verso. The remaining pixels are identified as either bleed-through, or bleed-through and foreground, by thresholding the cross-correlation of the recto and verso intensities over a small window, as only bleed-through regions should have a high correlation between the two sides. The window size and correlation threshold are set empirically.

2.1.2 Clustering

Though clustering and thresholding techniques are often related, several restoration approaches focus explicitly on the use of clustering as the basis for bleed-through removal.

A blind recursive unsupervised classification approach for RGB document images is proposed in [19]. Firstly, PCA is applied to the input image to decorrelate and reduce the dimensions of the data. Then K-means ($K = 2$) is applied to the intensities of the reduced data and the two clusters back-projected to the original RGB colourspace in two new separate images. The resultant image with the darker intensity profile is then retained, and the lighter discarded. This process is iteratively applied until a visually acceptable result is obtained.

A non-blind clustering approach is proposed in [15] for restoration of early music documents. Firstly the background information is located and removed using the adaptive degraded document binarisation method of Sauvola and Pietikäinen [63], and the musical staff lines are also detected and removed using the approach proposed in [51]. A set of candidate bleed-through pixel locations are selected from the foreground layer where the recto intensity is lighter than the corresponding verso intensity. From these candidate pixels locations four features are examined: the recto intensity, verso intensity, the difference between intensities, and the correlation coefficient from a 9x9 region surrounding each candidate location on each side. Fuzzy c-means clustering [10] on the features is then used to label each candidate as bleed-through or foreground, and those labelled as bleed through removed.

Similarly to the thresholding approaches, these clustering methods assume that there is a distinction between foreground and bleed-through intensities and will perform poorly on images where this is not the case.

2.1.3 Blind Source Separation

There have been several works that treat bleed-through degradation as a blind source separation (BSS) problem [11,69–71,73,74]. A noiseless instantaneous linear mixing model for the observed intensity data generation process is adopted as follows.

$$I_i(x,y) = \sum_{j=1}^{N} a_{ij} Y_j(x,y), \quad i = 1, 2, \dots, M$$

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \vdots & \vdots \\ a_{M1} & \cdots & a_{MN} \end{bmatrix} \tag{2.1}$$

Where $I_i(x,y)$ is the ith observed intensity data, $Y_j(x,y)$ is the jth intensity source, and $\mathbf{A}$ the unknown mixing matrix. It is assumed that the number of observed and source signals is the same, such that $\mathbf{A}$ is an $N \times N$ square matrix.

Sources

The number of sources used in the model depends on the amount of available data - approaches for both blind and non-blind restoration have been developed. In [69, 74], a blind approach is proposed where the observations are the red green and blue colour channels of the observed degraded image, and the sources represent the foreground text, background, and bleed-through degradation. In [11, 70, 73], the linear model is extended to non-blind bleed-through removal, with the observed degraded grayscale recto and verso images used as the data observations, and the original clean images as the sources. This is then extended further in [71, 72] to cases where multispectral images of both sides are available. A system of mixing equations is generated with separate 2x2 mixing matrices for each image type, the observations and sources again being corresponding degraded and clean recto and verso images respectively.

Solution

For all the BSS based approaches the assumption is made that the extent of bleed-through is constant across the image. In the blind three source approach [69], Independent Component Analysis is used to estimate the mixing matrix, and the source containing the most foreground text visually chosen as the restored result. In [11], a non-blind approach, the results of a similar ICA stage are then binarised with the adaptive degraded document binarisation algorithm of Gatos et al. [24] to produce recto and verso foreground masks. The bleed-through regions are then replaced with estimated mean background intensity. In other non-blind models [71, 73] it is assumed that the foreground and bleed-through strengths are the same on both sides of the page, and therefore that the corresponding mixing matrix is symmetric and can be estimated using symmetric orthogonalisation on the observed data. This is extended to noisy mixtures in [70], where the problem is expressed in a dual MRF framework (see Section 2.2.1), however rather than using the graphical model directly for classification, the MRF framework uses smoothness priors to guide estimation of the inverse mixing matrix via an Expectation-Maximisation (EM) algorithm. There are also no intra-field priors in this case.

Convolution Model

To account for the possible non-instantaneous effect of ink spreading out horizontally as it bleeds-through the page, the framework in [70] is extended to noisy linear convolutive mixtures in [72] for multispectral views of one or both sides of the page, resulting in the degradation model

$$\begin{aligned} I_i(x,y) &= \sum_{j=1}^{N} a_{ij}\hat{Y}_{ij}(x,y) + n_i(x,y) \quad i = 1, 2, \ldots, M \\ \hat{Y}_{ij}(x,y) &= \mathbf{H}_{ij} Y_j(x,y), \end{aligned} \tag{2.2}$$

where $\hat{Y}_{ij}(x,y)$ is the degraded source $Y_j(x,y)$ in observation $I_i(x,y)$, and $\mathbf{H}_{ij}$ is the blur matrix that performs convolution between the corresponding source image $Y_j(x,y)$, and a blur kernel $\mathbf{B}_{ij}$. This model may be a more realistic representation of the degradation process, however in this approach it is again assumed that the extent of bleed-through is constant across the image. Also where non-blind restoration is performed, it is assumed again that the mixing matrix $\mathbf{A}$ is symmetric, and that the convolution matrices to generate the observed images are the same for both sides of the page.

As these approaches all rely on the assumption of uniform bleed-through across each page, the number of documents that they can successfully be applied to is limited to those with very light bleed-through.

2.1.4 Diffusion

The physical process of ink bleeding through the page may be considered to be diffusive. In [43], bleed-through degraded documents are viewed as dynamic processes whose change over time is directed by three diffusive components:

(i) The diffusion of the recto text as the ink spreads across and through the medium, and fades over time.

(ii) The diffusion of the background medium; this models effects such as dust becoming attached, the spread of mould or damp, and the increase in gaps in the medium.

(iii) The diffusion of the verso text as the ink bleeds through the page.

This model is expressed via the following equation:

$$\frac{\partial I_r}{\partial t} = \underbrace{\mathrm{DIFF}(I_r, Y_r, c_r)}_{\text{recto}} + \underbrace{\mathrm{DIFF}(I_r, I_{bg}, c_{bg})}_{\text{background}} + \underbrace{\mathrm{DIFF}(I_r, I_v, c_v)}_{\text{verso}} \tag{2.3}$$

where $I_r(t)$ is the observed image at artificial time t, Y_r is the clean recto image, I_{bg} the estimated background, and I_v the clean verso image. c_x is the diffusion coefficient for the corresponding source component I_x, and $\mathrm{DIFF}(I, Y, c)$ is a generalised diffusion operator, in order to maintain the flexibility of the overall model, that describes the diffusion from Y onto I with coefficient c.

Under the assumption that the ideal background is a constant intensity, this model creates good synthetic bleed-through documents and the following equation is proposed for restoration.

$$\frac{\partial I_r}{\partial t} = \mathrm{DIFF}(I_r, I_r, c_{i,r}) + \mathrm{DIFF}(I_r, I_{i,bg}, c_{i,bg}) - \mathrm{DIFF}(I_r, I_v, c_{i,v}), \tag{2.4}$$

where I_r and I_v are the recto and verso images (at time $t = 0$ they represent the observed degraded images). The subscript index i on the diffusion coefficients, $c_{i,x}$, is to differentiate between the degraded document creation and restoration processes as the verso diffusion rate has a different form for each case. Restoration of the degraded images is performed using finite-differences over 8 pixel neighbourhoods. This approach is very parameter intensive, with 5 empirically fixed parameters for the restoration process, and relies on the assumption of a uniform background. The model used also does not explicitly take into account regions of overlapping foreground and bleed-through - the synthetic images created do not ensure that the bleed-through is invisible through the foreground text.

2.2 Spatial Information

As has been discussed previously, intensity information alone is insufficient for identifying bleed-through in severely degraded cases, due to the overlap in intensity ranges between foreground, bleed-through and background regions. Some form of prior knowledge and spatial dependency needs to be introduced in order to improve the restoration. In what follows, an overview of previous work that incorporates both spatial and intensity information into the restoration process is presented

2.2.1 Dual-Layer MRF

Several methods have modelled the degraded images as a dual-layer MRF in order to capture the intrinsic spatial correspondence between pixels. In these cases the bleed-through restoration is formulated as a discrete labelling problem, where each pixel is assigned a label l from a predefined set of labels that varies between approaches. The optimal labelling then can be defined as one that minimises the following energy

$$E = E_d + \lambda E_s, \tag{2.5}$$

where E_d at a pixel site represents the cost associated with assigning label l to that pixel, and is based on the image intensity information. E_s is a smoothness cost associated with assigning different labels to neighbouring pixels, and the weight parameter λ balances the influence of the two energies.

Wolf [78] Blind Method

A blind dual MRF based method is proposed by Wolf in [78], where the intensities of the observed image foreground, bleed-through and background are modelled as Normal Distributions. Two binary label fields, L_r, L_v, are introduced to control which distribution is represented in different image regions, with $L = 0$ corresponding to background, and $L = 1$ corresponding to foreground. These label fields are used to ensure that bleed-through degradation is only visible in non-foreground regions, thus enforcing the observed fact that foreground text is opaque, that is, bleed-through strokes are not visible through the foreground text. With the inclusion of the two label fields, the optimisation now is to maximise the posterior probability: $p(L_r, L_v|I_r)$. The smoothness energy E_s, based on the two label fields, in this case describes the prior knowledge about the clean document, and is expressed as the sum of two Potts models, one corresponding to each label field. The motivation behind using the two fields is that the segmentation process is more suited to the document bleed-through generation model. Wolf argues that if a single label field were used, with a space of three or four labels (foreground/bleed-through/background/foreground&bleed-through overlap), the prior would be regularising a single field created by two processes, that is the bleed-through from the two sides. Instead the model uses two fields each chosen from a space of two labels (foreground/background) such that the prior regularises these two fields, each corresponding to a single bleed-through process. The two fields are assumed to be independent, and that allows for the factorisation of the prior probability. This assumption of independence of the recto and verso writing is only made on the prior in the absence of an observation. In other words the recto and verso text of the clean segmented document have no influence over each other, but in the presence of degradation due to bleed-through, there is interference and interaction between the two fields through the observation field, and this interaction is described in the data model.

The label fields are initialised using k-means clustering ($k = 3$) on the intensities, however the labels are not assigned based on the subsequent cluster intensity profiles. Instead the background cluster is chosen as the one having most pixels, and the foreground and bleed-through are assigned based on a connected component analysis of the remaining two; as foreground pixels cover bleed-through pixels, the foreground cluster is chosen as that which has the least connected components. The final labelling is obtained using graph cuts optimisation [37], and the final restored result is obtained by replacing bleed-through pixels with a local mean background intensity value - thus the document character in this technique is preserved.

This approach is shown to outperform other blind methods, however, it still relies on the initial label field estimates using intensity based clustering, and so does not perform well where there are large region of dark bleed-through. Furthermore, background replacement using local mean intensities can leave visible artefacts in highly textured image examples.

Huang et al. [34, 35] Non-Blind Method

A non-blind dual MRF approach is proposed in [34,35], where in this case the label for each pixel is taken from a set of three labels: foreground (fg), background (bg), or bleed-through (bl). The data cost is based on initial K-Nearest Neighbour (KNN) and Support Vector Machine (SVM) classification of the ratio of intensities between the two sides. Training data for the classifier is obtained in the form of user drawn coloured strokes in regions of each pixel class on both sides. Given the likelihood S_{fg}, S_{bg}, S_{bl} of each pixel to each class label $l \in \{fg, bg, bl\}$, obtained at initialisation, the data cost is defined as

$$\begin{aligned} E_d(l = fg) &= \frac{S_{bl} + S_{bg}}{2(S_{fg} + S_{bg} + S_{bl})} \\ E_d(l = bg) &= \frac{S_{fg} + S_{bg}}{2(S_{fg} + S_{bg} + S_{bl})} \\ E_d(l = bl) &= \frac{S_{bl} + S_{fg}}{2(S_{fg} + S_{bg} + S_{bl})} \end{aligned} \tag{2.6}$$

The smoothness energy, E_s, is defined with two components. An intra-field prior energy is used to ensure spatial smoothness in the classification of each layer, with penalties for assigning different labels to neighbouring pixels, based on the ratio of intensities between the two sides. An inter-field prior is also included to ensure that certain label combinations between layers cannot occur, such as bleed-through in the same location on both sides. The energy is minimised again using graph cuts [37], and areas classified as background or bleed-through are replaced with the mean background intensity value.

Due to the interesting incorporation of user assistance, this approach is used for comparison against the proposed methods in Chapter 5. The results presented show that this approach relies on a good user markup to produce good results, and tends to oversegment the foreground of one side in regions where foreground and bleed-through overlap.

2.2.2 Wavelet Enhancement

Wavelet transforms have been applied to many areas in image processing, including compression, enhancement, de-noising, texture analysis and segmentation, and feature detection. The use of wavelet transforms provides an alternative approach to modelling the spatial dependencies within images, as they perform decomposition into the different high and low frequency components. In bleed-through restoration they are useful for capturing the distinction between foreground and bleed-through regions, as the foreground is often more sharply defined, with a higher frequency profile than bleed-through, which is often blurred slightly due to the spread of ink across and through the page.

In [67] a non-blind bleed-through removal method is proposed that iteratively sharpens foreground text and smears bleed-through in the wavelet domain to improve the results of subsequent binarisation. In the case of the recto, binary enhancement and smearing features are defined based on an initial estimate of foreground text on the recto and verso sides respectively. This estimate is obtained by performing Canny edge detection on a scaled intensity difference image between the two sides, exploiting the fact that foreground text will be more sharply defined than bleed-through text. The edge detection is further improved by incorporating constraints on the stroke orientation - favouring slanted foreground strokes. The detected foreground edges are then used as loci to obtain text strokes of a predetermined width. These recto and verso stroke images are binarised using Niblack's document binarisation method [47] to obtain the enhancement and smearing features respectively. Iterative wavelet deconstruction and reconstruction are then applied to each image, scaling up the foreground text components and shrinking the bleed-through, guided by the enhancement and smearing features respectively. The enhanced images are then binarised, using the Canny edge detection, stroke location and thresholding approach used initially, to obtain the final output results.

A variation of this approach was proposed in [77] for blind bleed-through removal. In this case the assumption of slanted text is more heavily relied upon, as a directional wavelet transform is applied to the images, where the wavelet filters are convolved along directions of 45° and 135°, to capture foreground and bleed-through strokes in different components. The bleed-through is then reduced by iteratively enhancing one orientation component and smearing the other. The final processing stage used in [67] is then applied to obtain the binarised output.

For both these approaches there are again several parameters which must be tuned to obtain the best results per image, including the initial binarisation thresholds, the number of wavelet levels, and the wavelet enhancing and smearing factors. Also, regions of foreground overlapping on both sides of the page will both be enhanced and smeared, which could result in over segmentation of the foreground text, depending on the enhancing and smearing factors chosen.

2.2.3 Variational Model

In [44], a variational denoising model for bleed-through correction, using wavelet shrinkage, is proposed that includes models for both blind and non-blind restoration, incorporating the reverse diffusion approach described in Section 2.1.4. The non-blind method uses a function of the difference in intensity between the degraded recto and verso sides as an indicator of bleed-through and foreground text regions, and spatial smoothness is enforced in the wavelet domain. The variational model for each side consists of three terms:

(i) A ‘fidelity’ term ensures the restored image is close to the original in foreground regions. This corresponds to the data cost E_d in Equation 2.5.

(ii) A ‘reverse diffusion’ term ensures the restored image is close to a uniform target background in background and bleed-through regions. This is essentially an added prior on the clean document appearance.

(iii) A smoothness term, corresponding to the energy E_s in Equation 2.5, ensures that the restored image does not contain harsh cut-off at character edges, and that fine details are preserved.

The first two terms are weighted by a function of the intensity difference. The smoothness term is defined on the wavelet coefficients of the restored image, and weighted with a smoothing parameter λ chosen based on the estimated background of the degraded image. The model, for the recto side only, is expressed via minimisation of the energy functional

$$E(d_r) = \underbrace{||Y_r - I_r||^2_{w_0}}_{\text{fidelity}} + \underbrace{||Y_r - Y_b||^2_{w_{rev}}}_{\text{reverse diffusion}} + \underbrace{2\lambda||d_r||^p_{b^\alpha_p(l_p)}}_{\text{smoothness}}, \tag{2.7}$$

where d_r is the corresponding wavelet coefficient sequence of the clean recto image Y_r, I_r is the observed degraded recto image, and Y_b is the recto target background intensity. $||Y||_w$ represents the L_2 norm of Y weighted with function w. The reverse diffusion weight for the recto side w_{rev} is defined from the observed degraded images as follows

$$w_{rev}(I_r) = \frac{1}{2}\left(1 + \tanh\left(\frac{I_v - I_r - 2\sigma_{rev}}{\sigma_{rev}}\right)\right), \tag{2.8}$$

where σ_{rev} is a parameter controlling the degree of reverse diffusion. The fidelity weight w_0 in Equation 2.7 is defined as $1 - w_{rev}$. The solution of the model is obtained using hard wavelet shrinkage.

The variational model proposed for blind bleed-through removal is similar, the only difference being that the reverse diffusion weight function in Equation 2.7 is replaced with a background diffusion weight function based on the difference between the estimated background, I_b and the degraded image. This weight function (again for the recto side only) is expressed as follows:

$$w_{bkgd}(I_r) = \frac{1}{2}\left(1 + \tanh\left(\frac{I_b - I_r + 2\sigma_{bkgd}}{\sigma_{bkgd}}\right)\right), \tag{2.9}$$

where the parameter σ_{bkgd} represents the extent that gray values are affected by the background diffusion. The solution to this model is reconstructed using the inverse wavelet transform.

The non-blind approach in this case is also selected for comparison with the proposed methods in Chapter 5, as it focuses on the preservation of fine foreground features. The results presented show that this restoration approach is sensitive to the smoothness parameter chosen, with better results obtained in some cases when this is manually tuned rather than set based on the estimated background.

2.2.4 Active Contours

Active Contours is a shape based model driven segmentation approach, where the shape contour is explicitly modelled, rather than implicitly at segmentation boundaries. The principle of Active Contours is to search for a curve in the image where the weighted sum of internal energy and potential energy is minimum. The internal energy is defined by the curve itself to keep the model smooth during deformation, corresponding to E_s in Equation 2.5. The potential energy is computed from the image data to move the curve toward an object boundary or other desired feature within the image, corresponding to E_d in Equation 2.5. The well known Chan-Vese model [16] describes a method for region-based active contour segmentation using level set functions. This approach does not depend on sharp image gradients at object edges but rather obtains segmentation by minimising the intensity variance inside and outside the curves. This is useful in the case of degraded documents as character edges often fall off gradually due to the porosity of the writing medium or ink fading over time [29].

In [45] a framework for non-blind bleed-through removal based on the Chan-Vese method is described, where an active contour, $C(t)$, that approaches the segmented recto text as the artificial time $t \to \infty$, is located on the intersection between an evolving level-set function, or surface, $\phi(x, y, t)$, and the xy plane. The level set function evolution in this is controlled by three forces; a local thresholding force, F_T, based on the difference between the recto intensity and a selected threshold value, that raises the weight on pixels darker than the threshold, and reduces that of lighter pixels; a reverse diffusion force F_V, similar to Equation 2.8, again calculated based on the difference between intensities on the recto and verso sides; and a regularising force,

F_{REG}, that favours smooth boundaries:

$$\frac{\partial \phi}{\partial t} = \delta_{\phi}(F_T + \nu F_{REG} + \mu F_V), \tag{2.10}$$

where δ_{ϕ} controls the stability of the evolution as it limits change in the level set function to regions close to the zero set (xy plane). Parameters μ and ν control the importance of the diffusion force and smoothness of segmentation boundaries respectively, and as they must be estimated and tuned based on degradation levels and document type, this is another parameter intensive approach.

Another non-blind contour based method is proposed in [29], also based on the Chan-Vese model, that defines the segmenting curve, C, by minimising over the energy functional

$$E = E_1 + \lambda E_2, \tag{2.11}$$

where

$$E_1 = \mu|C| + \lambda_1 \int_{\phi_1} |I_r(x,y) - c_1|^2 dxdy + \lambda_2 \int_{\phi_2} |I_r(x,y) - c_2|^2 dxdy \tag{2.12}$$

is the original Chan-Vese functional with $|C|$ equal to the length of the curve, and c_1 and c_2 the mean intensities for the domains ϕ_1 and ϕ_2, that is inside the curve, corresponding to foreground text, and outside the curve respectively. μ,λ_1, and λ_2 are weighting parameters for each term, set to 1. The second energy term E_2 in Equation 2.11 is used to mitigate the under segmentation of the foreground text from the bleed-through where their intensity ranges overlap, and is defined based on the difference in intensity between the two sides as

$$E_2 = \int_{\phi_1} |I_r(x,y) - I_v(x,y) - c_3|^2 dxdy + \int_{\phi_2} |I_r(x,y) - I_v(x,y)|^2 dxdy, \tag{2.13}$$

where c_3 is the mean estimated foreground intensity inside the curve. This energy term is based on the assumption that the intensity of the recto foreground pixels is typically lower than corresponding bleed-through or background verso pixels, and so tends to over segment the foreground text in regions where there is foreground on both sides of the page. Therefore the combination of the two energies provides a balance between under segmentation where bleed-through and foreground intensities are not distinct, and over segmentation where there is foreground on both sides of the page. The parameter λ in Equation 2.11 balances the influence of the two energies, and is tuned manually for each image pair.

A post-processing step is proposed for severely degraded examples, where there is a significant overlap between bleed-through and foreground intensities. In these cases a high value for λ is necessary to remove all of the bleed-through, and this results in over segmentation in overlapping foreground regions. This stage attempts to complete broken characters using a modified Cahn-Hilliard inpainting model for binary images [8], with the darkest pixels in the background domain ϕ_2 identified as the target candidate completion domain.

As this approach is shown to outperform the non-blind wavelet enhancement method [67], and the user assisted dual-MRF approach [35], it is used for comparison with the proposed methods in Chapter 5. The results show that this method, with user tuned smoothness parameter, performs well on a range of different degradation types. However, the results still suffer when bleed-through is very dark - the stroke completion repairs bleed-through regions, and where there is significant foreground and bleed-through overlap - the gaps are too large for stroke completion to work.

2.3 Summary

This chapter has presented an overview of previous approaches to bleed-through removal, focussing on the underlying image processing used. The limitations of the methods described, fall into three categories: under segmentation of bleed-through due to overlapping bleed-through and foreground intensity ranges, over segmentation of regions where there is foreground on both sides of the page, and reliance on parameter tuning. Also, in many cases the assumptions made about the degraded documents limit the scope of the algorithms significantly, such as the BSS methods where uniform bleed-through is assumed across the images, and the wavelet enhancement approaches, where the text is assumed to be highly slanted.

User Interaction
Many of the described restoration methods claim to be fully automatic, however in reality there are often a number of parameters that need to be tuned on each image by a user before optimal results can be achieved. There has therefore been an increasing trend in approaches proposed that try to avoid parameter tuning and explicitly to make use of user interaction to improve performance. For example user assisted degraded document binarisation methods have been proposed in [18, 80]. Also the non-blind dual MRF based approach described in 2.2.1 incorporates user drawn strokes in regions of foreground, background, and bleed-through to use as training data for the classifiers. This approach is extended in [79] to a directed assistance framework, where the user is iteratively directed to regions of low confidence that may need further input for an improved classification result. In the contour based approach [29], the number of parameters required for tuning is limited to 1, and it is proposed to incorporate a slider that the user can employ to tune this parameter in real time to obtain the optimal output. The benefits of incorporation of user interaction will be discussed further in Chapter 5.

Restoration
As mentioned in the introduction to this chapter, there are three different types of restoration output generated. Barring a few exceptions, all of the techniques described produce binary or pseudo-binary output results. The approaches that preserve the background texture, replacing bleed-through regions only, all use an estimate of the local mean background intensity ([11, 20, 78]). This is suitable for documents that have a reasonably uniform background, however, in cases where the background medium is highly textured, this approach will leave visible artefacts. There is clearly a lack of emphasis on preserving the true character of the document in the output

results by trying to model the background texture as well as intensity. This problem will be addressed in the proposed approaches in Chapters 3 and 4.

3

Model-Based Approach

In Chapter 2, previous approaches to bleed-through removal were discussed, focussing on the use of intensity and spatial information available and the different output types generated. It was noted that bleed-through restoration approaches that produce binary or pseudo-binary outputs can destroy the character of manuscript document images, and also may not perform well at character edges where there is no clearly defined boundary - creating a harsh intensity cut-off effect. For viewing purposes it is preferable to leave as much of the document intact and unaltered as possible, so that the experience of studying the document *image* remains close to that of studying the *physical* document. Modelling the degraded images as the result of bleed-through degradation, and attempting to recover the original clean images allows for preservation of the document character, focussing on removing bleed-through regions only and preserving the background texture in those regions.

In this chapter therefore, an approach for bleed-through restoration is proposed that uses a linear generative model, where the observed degraded images are viewed as a per-pixel combination of the original clean images and some proportion of the corresponding reverse sides. Use of a strictly linear model however may not always accurately represent the bleed-through process, and therefore it is necessary to add constraints to limit the presence of bleed-through to certain locations only. The overlapping ink opacity constraint proposed in [78], where bleed-through can only occur in non-foreground regions, is enforced explicitly to ensure that the bleed-through process is accurately modelled. The proposed model and solution are developed in a Bayesian dual-MRF MAP framework, solving for 6 unknown parameters per recto-verso pixel (including the clean images themselves), and including both intra and inter-field priors to enforce spatial

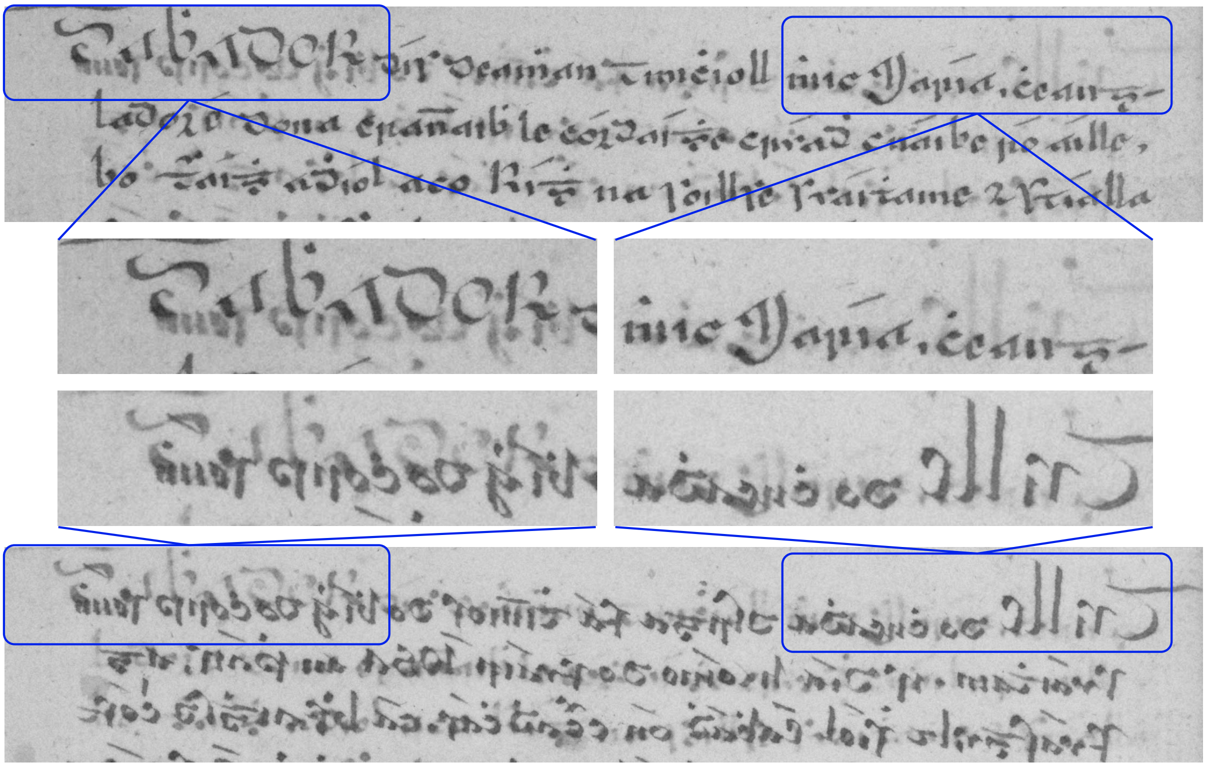


Figure 3.1: The strength of bleed-through cannot be assumed constant across a page: Recto and verso manuscript extracts, with cropped sections highlighting the variations in bleed-through strength. **M14**.

smoothness, and to ensure that the restored results appear realistic. It is assumed that registration has already been performed, and in the rest of this chapter ‘verso’ will refer to the flipped, registered verso side. The aim of this approach therefore is to produce clean results that remain as close to the original images as possible, but where legibility has been improved by the removal of bleed-through regions[1].

3.1 The Case for a Linear Model

As highlighted in Chapter 2, linear models have been used previously to model bleed-through degradation, though the physical process itself is nonlinear [21, 69, 70, 73]. In [73] the observed recto and verso images, $[I_r(x, y), I_v(x, y)]$ at pixel location (x, y), are treated as linear mixtures

---

[1] Parts of this chapter are based on the work published in [57, 58].

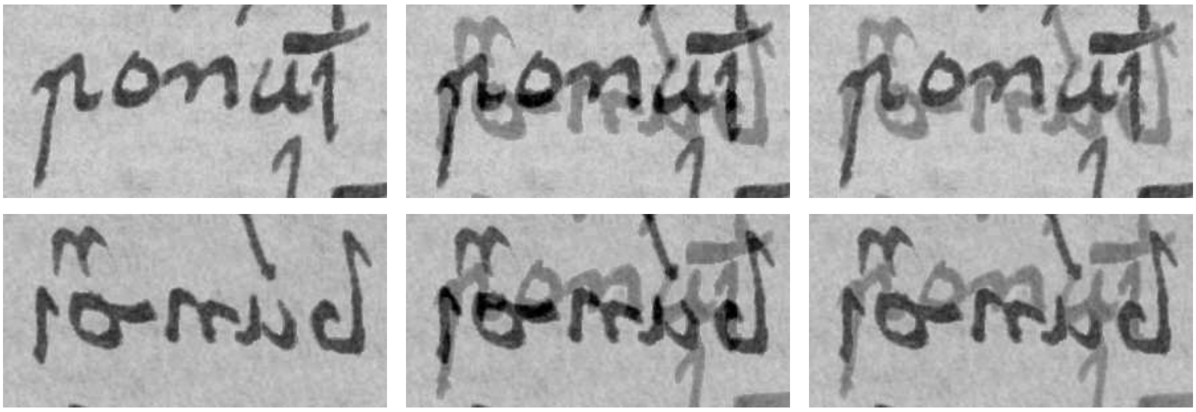


Figure 3.2: Comparison of linear and constrained linear model synthesised bleed-through outputs. From left to right: Synthesised clean recto (top) and verso sample images. Degraded images created using Equation 3.1, with $A_{11} = A_{22} = 1$, and $A_{12} = A_{21} = 0.34$. Degraded images created using a constrained linear model and variable mixing parameters.

---

of the target clean recto and verso source images, $[Y_r(x,y), Y_v(x,y)]$, as follows.

$$\begin{aligned} I_r(x,y) &= A_{11}Y_r(x,y) + A_{12}Y_v(x,y) \\ I_v(x,y) &= A_{21}Y_r(x,y) + A_{22}Y_v(x,y), \end{aligned} \tag{3.1}$$

where $\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ is the unknown mixing matrix between sources to be estimated, and A_{ij} is the proportion of one source in a given observation. This model is unrealistic in that it assumes that the degradation is uniform across the images, whereas this is often not the case, especially for manuscripts, as can be seen in Figure 3.1. For example, the variable pressure used when writing different characters, the size of character and amount/type of ink used, the potential variability in the porosity or thickness of hand made writing media, and damage to the medium due to water stains, are just a few factors that may affect the strength of bleed-through in local regions across a page. Furthermore, constant mixing parameters do not account for the fact that in regions where foreground and bleed-through overlap, the foreground is opaque and covers the bleed-through [78]. In these regions, restoration based on the model in Equation 3.1 would result in the foreground information being removed. It is more realistic therefore to have variable mixing parameters across the images, and to limit the presence of bleed-through to regions where there is no foreground text. As can be seen in Figure 3.2, a linear model modified by these assumptions produces more realistic synthetic images. The model used therefore simulates the physical degradation using a constrained per pixel linear model. In what follows, white pixels are considered to have intensity of 0, and black pixels 255, such that text regions have higher intensities than background regions.

3.2 Degradation Model

In the proposed model, an observed degraded recto pixel intensity $I_r(x,y)$, is a linear combination of the corresponding clean recto pixel $Y_r(x,y)$ and some proportion of the clean verso pixel $Y_v(x,y)$. The proportion of bleed-through is controlled by bleed-through parameter $\alpha_v(x,y)$, $(0 \leq \alpha_v < 1)$, and also the binary masks defined on both sides $M_r(x,y), M_v(x,y)$. Due to the symmetrical nature of the problem, the formation of the observed verso pixel $I_v(x,y)$ is similar. The model for each side of the page therefore is as follows.

$$\begin{aligned} I_r(x,y) &= Y_r(x,y) + M_r(x,y)(1 - M_v(x,y))\alpha_v(x,y)Y_v(x,y) + \epsilon_r(x,y) \\ I_v(x,y) &= Y_v(x,y) + M_v(x,y)(1 - M_r(x,y))\alpha_r(x,y)Y_r(x,y) + \epsilon_v(x,y) \end{aligned} \tag{3.2}$$

The binary mask terms, M_r, M_v are defined to be 0 where the corresponding image is foreground text, and 1 everywhere else. The combination of the mask terms explicitly ensures the opacity of foreground text limiting the presence of the clean verso image, Y_v in the observed recto image to regions where Y_r is background and Y_v is foreground. It is also assumed that the document image suffers from added noise due to physical degradation of the medium and the imaging process. This noise is represented in the terms ϵ_r, ϵ_v, and is assumed to be zero mean and Gaussian: $\mathcal{N}(0, \sigma^2_{\epsilon_r})$, $\mathcal{N}(0, \sigma^2_{\epsilon_v})$.

3.3 Bayesian Framework

The bleed-through removal method is formulated under a Bayesian Maximum *a posteriori* (MAP) framework. From the degradation model in Equation 3.2 it is clear that the unknown parameter vector for estimation contains six variables per pixel: $\boldsymbol{\theta} = [\alpha_v, \alpha_r, M_v, M_r, Y_r, Y_v]$. Applying Bayes' rule and discarding pixel coordinates (x,y) for clarity, the p.d.f. for the posterior probability of $\boldsymbol{\theta}$ given the observed data I_r, I_v, may be factorised as

$$p(\boldsymbol{\theta}|I_r, I_v, \tilde{M}_r, \tilde{M}_v, \tilde{\alpha}_r, \tilde{\alpha}_v) \propto p(I_r, I_v|\boldsymbol{\theta}, \tilde{M}_r, \tilde{M}_v, \tilde{\alpha}_r, \tilde{\alpha}_v)p(\boldsymbol{\theta}|\tilde{M}_r, \tilde{M}_v, \tilde{\alpha}_r, \tilde{\alpha}_v), \tag{3.3}$$

where $[\tilde{M}_r, \tilde{M}_v, \tilde{\alpha}_r, \tilde{\alpha}_v]$ represent the existing state of the estimated masks and bleed-through parameters in the neighbourhood of the pixel site currently being considered. In what follows, the case of the observed recto image only is considered as, again, due to the symmetrical nature of the model, the formulation is similar for the verso side.

3.3.1 Likelihood

Following the degradation model, the likelihood combines the influence of both the recto and verso sides to yield a joint Gaussian distribution

$$p(I_r, I_v | \boldsymbol{\theta}, \tilde{M}_r, \tilde{M}_v, \tilde{\alpha}_r, \tilde{\alpha}_v) \propto \exp - \left\{ \frac{1}{2\sigma^2_{\epsilon_r}} [I_r - Y_r - M_r(1 - M_v)\alpha_v Y_v]^2 \right. \\ \left. + \frac{1}{2\sigma^2_{\epsilon_v}} [I_v - Y_v - M_v(1 - M_r)\alpha_r Y_r]^2 \right\}. \tag{3.4}$$

Here it is assumed that the noise generating processes of the two sides are independent.

3.3.2 Priors

Prior models are defined for each of the unknown parameters to guide estimation as follows.

Masks

The mask variables M_r, M_v mark the estimated regions of foreground text, and should therefore be smooth in local patches. A Gibbs energy prior for spatial smoothness makes sense in this case. In regions where there is neither foreground text nor bleed-through degradation, there is no ambiguity and it is clear that these variables should be 1. It is possible to estimate roughly the regions of text, bleed-through and background (using K-means clustering on the degraded image, with K=3). Using this rough estimate it is sensible then to include a prior constraining the mask variables in the definite non-text, that is, background regions. Therefore the prior for the masks is

$$p(M_r | \tilde{M}) \propto \exp - \left\{ \sum_{i \in \mathcal{N}_{x,y}} (M_r - \tilde{M}_i)^2 \lambda_M + \beta_r (1 - M_r) \right\}. \tag{3.5}$$

Here $\tilde{M}$ represents the current state of M_r in the 8-connected neighbourhood $\mathcal{N}$ of the current site (x, y), and λ_M is a parameter to control the spatial smoothness (in essence an Ising model in this case due to the binary nature of the masks). In the second term, β_r represents a penalty for setting $M = 0$ in the regions of definite non-text, and is estimated in the initialisation step discussed in Section 3.4.4.

Mixing Parameters

Priors for the mixing parameters, α_v, α_r follow similar logic. Firstly, smoothness is encouraged with the Gibbs energy prior. Secondly, it is necessary when the mixing parameters are present to constrain them to values that would yield material that is close in intensity to the background regions in the clean image. Without this constraint there is nothing limiting the mixing parameters to create useful images since the smoothness of itself does not constrain the absolute value - see Figure 3.3. Finally to prevent the smoothness prior both from spreading the mixing parameters into foreground regions, and also from creating the impossible scenario of bleed-through on both sides in the same location, an intra-layer prior is used so that the mixing

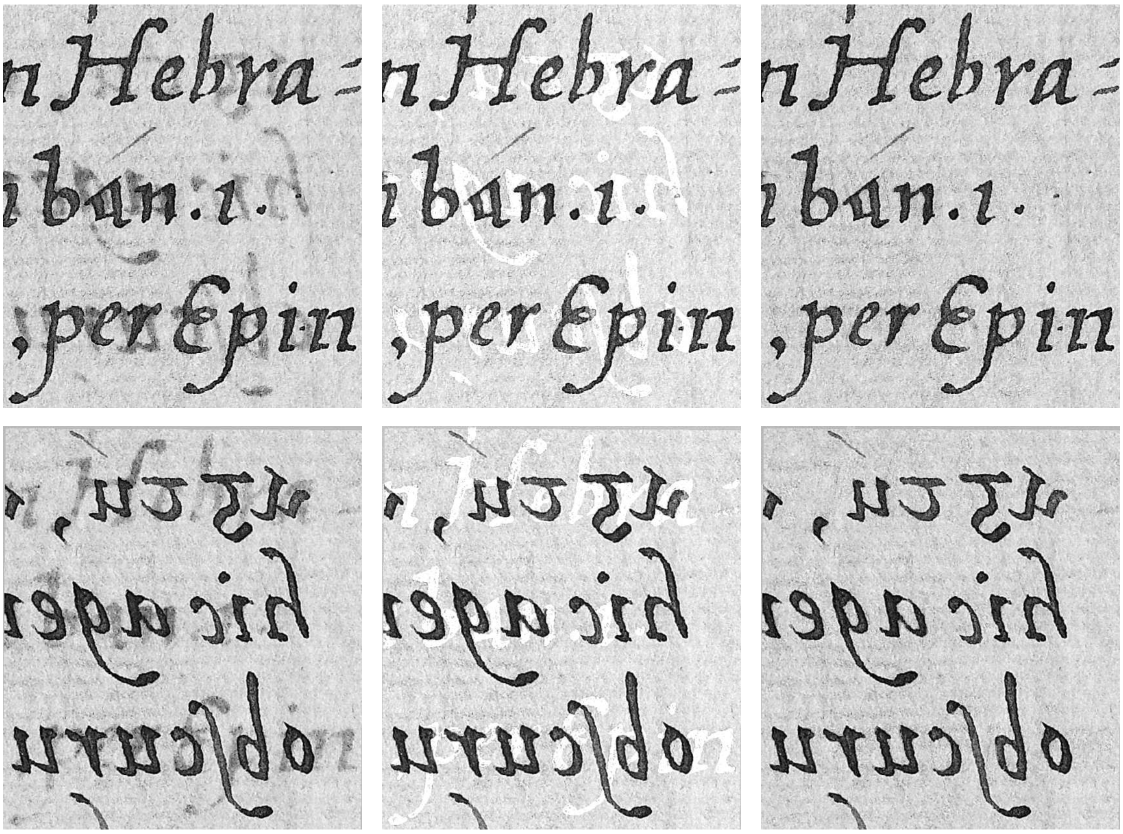


Figure 3.3: The constraining priors on the mixing parameters and clean image estimates ensure that bleed-through restored regions are close in intensity to the estimated local mean background intensity. Left to right: Original degraded recto (top) and verso sample images, restored without constraining priors, and restored with constraining priors. In the results without constraining priors there are visible artefacts in the bleed-through restored regions. **M02**.

parameters are constrained to the initial mask estimates, and against each other. Hence the prior model is

$$p(\alpha_v|\tilde{\alpha},\bar{\alpha}_v,\bar{M}_v) \propto \exp -\left\{ \sum_{i\in\mathcal{N}_{x,y}} (\alpha_v-\tilde{\alpha}_i)^2\lambda_\alpha + (\alpha_v-\bar{\alpha}_v)^2\lambda_c + (\bar{M}_v\alpha_r\alpha_v^2)\lambda_c \right\}, \tag{3.6}$$

where $\tilde{\alpha}$ are values of the α_v in the 8-connected neighbourhood of the current site (x,y), $\bar{\alpha}_v$ is a *rough* estimate of α_v, obtained during initialisation, $\bar{M}_v$ is the initial estimate for the verso mask, and λ_α, λ_c are parameters to control the degree of smoothness and constraint respectively. The smoothness parameter is much higher for the mixing parameters than for the masks so as to ensure that the final clean images retain as little trace of the bleed-through as possible, and to avoid harsh intensity cut-off at character edges; the purpose of the mixing parameters is to blend the bleed-through regions smoothly into the background. The final intra-layer term in Equation 3.6 adds a penalty for α_v defining bleed-through on the recto side in regions where the estimated verso mask is not foreground ($\bar{M}_v = 1$), and also in regions where α_r defines that there is bleed-through on the verso side, which must correspond with foreground text on the recto.

Clean Images

A prior is used for the clean image data to encourage, as for the mixing parameters, that the average brightness of the restored document in the bleed-through regions matches the average brightness of the background. This is again a simple Gaussian prior with parameters derived at initialisation.

$$p(Y_r|\bar{Y}_b) \propto \exp -\left\{ \frac{1}{2\sigma^2_{\epsilon_r}} M_r(1 - M_v)\alpha_v (Y_r - \bar{Y}_b)^2 \right\}. \tag{3.7}$$

Here $\bar{Y}_b$ is an estimate of the mean local background intensity of the clean recto image, and $\sigma^2_{\epsilon_r}$ is the variance of the background noise term as defined in Section 3.2. The multiplier $M_r(1 - M_v)\alpha_v$ ensures that this constraint is restricted to bleed-through restored regions only, and is proportional to the strength of bleed-through.

3.4 Solution

To solve for all the variables, a hybrid sampling scheme is applied that uses a combination of Iterated Conditional Modes (ICM) optimisation [9], and Gibbs sampling [25]. In this case the *mode* of the conditionals are selected for the masks and mixing parameters, while *samples* are drawn for the underlying clean images Y_r, Y_v. The solution for each of the variables is presented below for the case of the recto side only, as, again, the verso equations are similar.

3.4.1 Masks Estimate

As the masks are clearly present in both observations, the estimates for the masks are generated using the conditional

$$\begin{aligned} p(M_r|\alpha_r, \alpha_v, M_v, Y_r, Y_v, I_r, I_v, \tilde{M}) \propto \exp -\Bigg\{ & \frac{1}{2\sigma^2_{\epsilon_r}} (I_r - Y_r - M_r(1 - M_v)\alpha_v Y_v)^2 \\ & + \frac{1}{2\sigma^2_{\epsilon_v}} (I_v - Y_v - M_v(1 - M_r)\alpha_r Y_r)^2 + \sum_{i \in \mathcal{N}_{x,y}} (M_r - \tilde{M}_i)^2 \lambda_M + \beta_r(1 - M_r) \Bigg\}, \end{aligned} \tag{3.8}$$

where the terms are the same as for Equation 3.5. In this case the estimation is performed numerically since M_r is binary: both $M_r = 0$ and $M_r = 1$ are substituted in the expression above, and whichever yields the greater probability is selected.

3.4.2 Mixing Parameters Estimate

Each mixing parameter is present in only one of the observation terms, therefore the conditional probability at a site is

$$p(\alpha_v|M_v, M_r, Y_r, Y_v, I_r, I_v, \tilde{\alpha}, \bar{\alpha}_v, \alpha_r, \bar{M}_v) \propto \exp -\left\{ \frac{1}{2\sigma^2_{\epsilon_r}}(I_r - Y_r - M_r(1 - M_v)\alpha_v Y_v)^2 \right.$$
$$\left. + \sum_{i\in\mathcal{N}_{x,y}} (\alpha_v - \tilde{\alpha}_i)^2\lambda_\alpha + (\alpha_v - \bar{\alpha}_v)^2\lambda_c + (\bar{M}_v\alpha_r\alpha_v^2)\lambda_c \right\}, \quad (3.9)$$

where the terms are the same as in Equation 3.6. The estimate of α_v at each site is obtained analytically since the expression is quadratic in α_v. Hence

$$\hat{\alpha}_v = \frac{2\sigma^2_{\epsilon_r}\left[\sum_i \tilde{\alpha}_i\lambda_\alpha + \bar{\alpha}_v\lambda_c\right] + (I_r - Y_r)M_r(1 - M_v)Y_v}{2\sigma^2_{\epsilon_r}\left[\sum_i \lambda_\alpha + \lambda_c(1 + \bar{M}_v\alpha_r)\right] + (M_r(1 - M_v)Y_v)^2}. \quad (3.10)$$

3.4.3 Clean Images Estimate

The estimates for the clean images are generated with the conditional

$$p(Y_r|\alpha_r, \alpha_v, M_r, M_v, Y_v, I_r, I_v) \propto \exp -\left\{ \frac{1}{2\sigma^2_{\epsilon_r}}(I_r - Y_r - M_r(1 - M_v)\alpha_v Y_v)^2 \right.$$
$$\left. + \frac{1}{2\sigma^2_{\epsilon_v}}(I_v - Y_v - M_v(1 - M_r)\alpha_r Y_r)^2 + \frac{1}{2\sigma^2_{\epsilon_r}} M_r(1 - M_v)\alpha_v(Y_r - \bar{Y}_b)^2 \right\}. \quad (3.11)$$

Again, the terms here are the same as discussed for Equation 3.7. However, instead of maximising the conditional following the ICM process, a sample is drawn from this distribution within $\pm T$ standard deviations of the mean (T is set to 0.5 in what follows). This strategy has been employed by other authors working in video and audio restoration [49]. The reasoning behind this is that the mean tends to generate over smooth images, while using an unconstrained random draw is visibly chaotic - see Figure 3.4. Therefore by drawing samples within some distance of the mean, a textural component in the underlying signal is allowed for and the iterative process performs better. The required draw is therefore $Y_r \sim \mathcal{N}(\bar{Y}_r, \sigma^2_Y)$. By completing the square in the conditional above the mean and variance are extracted:

$$\bar{Y}_r = \frac{I_r + M_r(1 - M_v)\alpha_v(\bar{Y}_b - Y_v) + \sigma^2(I_v - Y_v)M_v(1 - M_r)\alpha_r}{1 + \sigma^2(M_v(1 - M_r)\alpha_r)^2 + M_r(1 - M_v)\alpha_v}, \quad (3.12)$$

where $\sigma^2 = \sigma^2_{\epsilon_r}/\sigma^2_{\epsilon_v}$

$$\sigma^2_Y = \frac{\sigma^2_{\epsilon_r}\sigma^2_{\epsilon_v}}{\sigma^2_{\epsilon_v}(1 + M_r(1 - M_v)\alpha_v) + (M_v(1 - M_r)\alpha_r)^2\sigma^2_{\epsilon_r}}. \quad (3.13)$$

The estimate for σ^2_Y depends on estimates for $\sigma^2_{\epsilon_r}, \sigma^2_{\epsilon_v}$, in the proposed method these estimates are obtained from the observed data using the initial mask estimates.

Figure 3.4: Drawing samples to estimate the clean images produces better results visually. Left to right: Original degraded recto (top) and verso sample images, clean image estimates (with known foreground masks) using the mean, unconstrained random samples, and constrained samples of the error distributions. **M19**.

A secondary estimation method for Y_r, Y_v, that is used every N iterations ($N = 10$ in this case), is to maximise conditionals that ignore the mask terms and priors, hence

$$p(Y_r|\alpha_r, \alpha_v, Y_v, I_r, I_v) \propto \exp - \left\{ \frac{1}{2\sigma_{\epsilon_r}^2}(I_r - Y_r - \alpha_v Y_v)^2 + \frac{1}{2\sigma_{\epsilon_v}^2}(I_v - Y_v - \alpha_r Y_r)^2 \right\}. \tag{3.14}$$

Due to the quadratic nature of Equation 3.14 in Y_r, an estimate of Y_r at a site is obtained analytically, giving

$$\hat{Y}_r = \frac{I_r - \alpha_v Y_v + \sigma^2 (I_v - Y_v)\alpha_r}{1 + \sigma^2 \alpha_r^2}, \tag{3.15}$$

where σ^2 is defined as above. Periodically reverting to this unconstrained linear model allows the heavily weighted smoothness priors on the mixing parameters to encourage reclassification in the masks of bleed-through regions misclassified as foreground. Figure 3.5 highlights the effectiveness of using the linear model to remove misclassified bleed-through regions showing sample extracts of clean image and mask estimates over 35 iterations. Furthermore using the secondary model for the final restored result ensures that resulting clean images appear smooth with minimal visible bleed-through artefacts remaining.

3.4.4 Initialisation

In this work the masks are initialised using K-means clustering on the intensity of the observed images, using 3 clusters. The lightest two clusters are considered to be background of the image

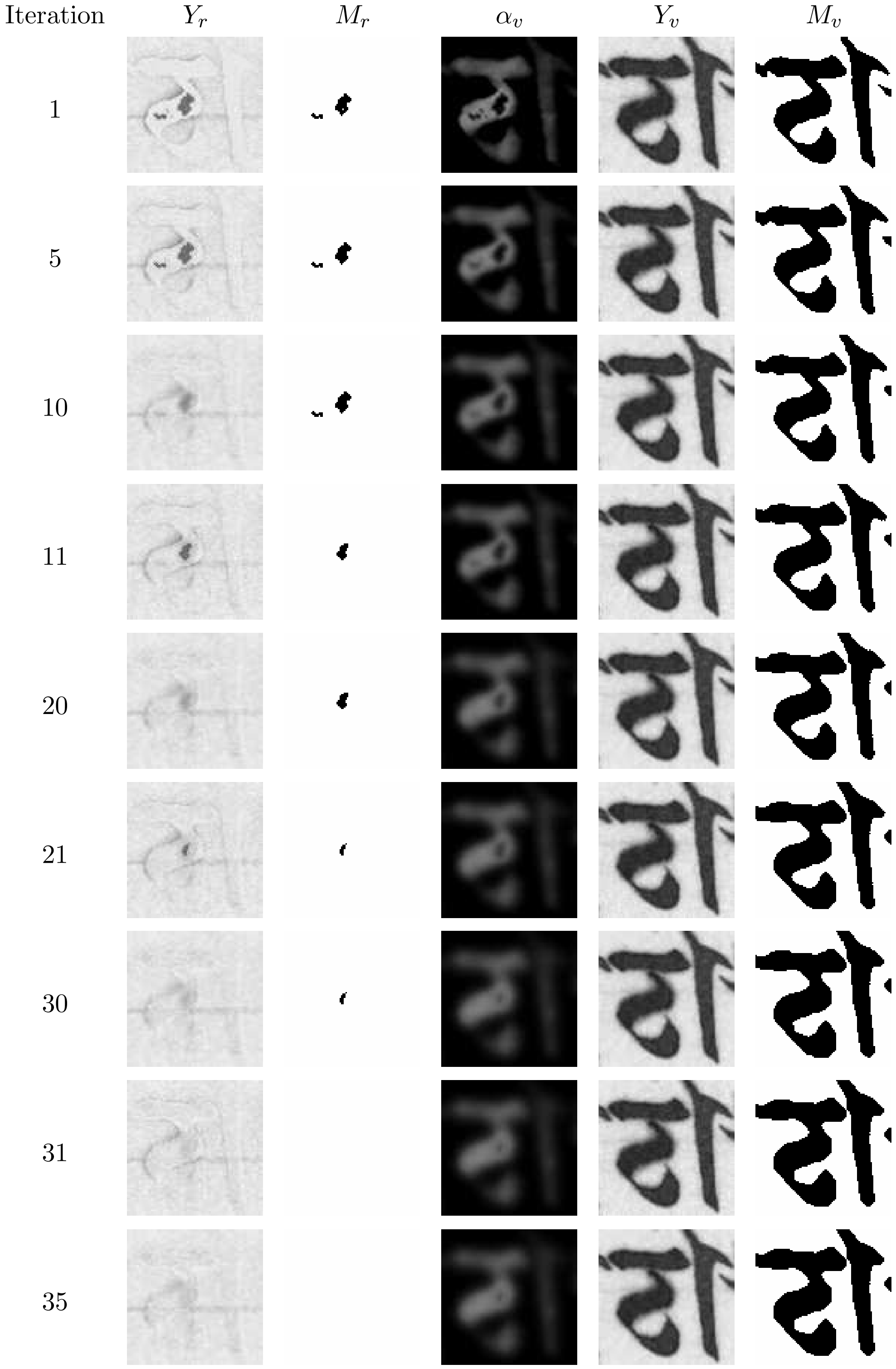


Figure 3.5: Sample extracts of clean image and mask estimates over 35 iterations. The unconstrained linear model, substituted every 10 iterations and for the final clean results, encourages reclassification of a dark bleed-through region on the recto side misclassified as foreground text, and that the final restored result has minimal residual bleed-through degradation. **M01**.

and the darkest cluster is then an estimate of the foreground text. This approach will create conservative initial estimates - ensuring that no foreground will be misclassified as bleed-through and removed erroneously. The mask penalties β_r, β_v are calculated in the same step by setting recto and verso background variables $B_r, B_v = 1$ in locations coincident with the brightest cluster where there is no text or bleed-through. The constraint variables β_r, β_v are then configured as $\beta_r = 100 \times B_r$, and $\beta_v = 100 \times B_v$. The masks and observed images are used to estimate the noise variance for both recto and verso sides, based on the variance of regions where $B_r = 0$ **and** $B_v = 0$. Also the estimates for local mean background intensity $\bar{Y_b}$ (for each side as appropriate) are generated using a Poisson diffusion process. The initial mixing parameters are then obtained from the mask estimates and the observed images as follows (for α_v).

$$\alpha_v = \begin{cases} \dfrac{Y_r - \bar{Y_b}}{Y_v}, & \text{if}(\bar{M_r} = 1) \wedge (\bar{M_v} = 0) \wedge (Y_r > \bar{Y_b}) \\ 0, & \text{else.} \end{cases} \tag{3.16}$$

In this case $(\bar{M_r} = 1) \wedge (\bar{M_v} = 0) \wedge (Y_r > \bar{Y_b})$ constitutes an initial estimate of recto bleed-through in regions where the initial recto mask is background, the verso mask is foreground, and the recto intensity is darker than the local mean background intensity. Finally, initial clean recto and verso estimates are obtained by substituting the relevant initial estimates into Equation 3.2.

3.4.5 Algorithm

The overall algorithm may be enumerated as follows:

1. Initialise all variables as described above.
2. Using a checker-board visitation pattern for sites, repeat until convergence:
   (a) Generate $\hat{M_v}$, $\hat{M_r}$, $\hat{\alpha_r}$, $\hat{\alpha_v}$ using the expressions above, (across all sites in separate image passes) updating in place.
   (b) If mod(iteration,10) $=0$, generate Y_r, Y_v without the masks as in Equation 3.15 Else, draw samples for Y_r, Y_v as described above in Section 3.4.

3.5 Discussion

Figures 3.6, 3.7, and 3.8 show sample results of the bleed-through removal approach on light, medium, and severe bleed-through example image pairs respectively. In all the results the character of the document is successfully preserved, and the algorithm removes the bleed-through on the light and medium cases. However, in the severe bleed-through example in figure 3.8, where there are large regions of bleed-through that overlap significantly in intensity with the foreground, the algorithm does not perform as well, with bleed-through artefacts remaining on the recto side. This highlights the dependence of the algorithm on a good initial estimate of

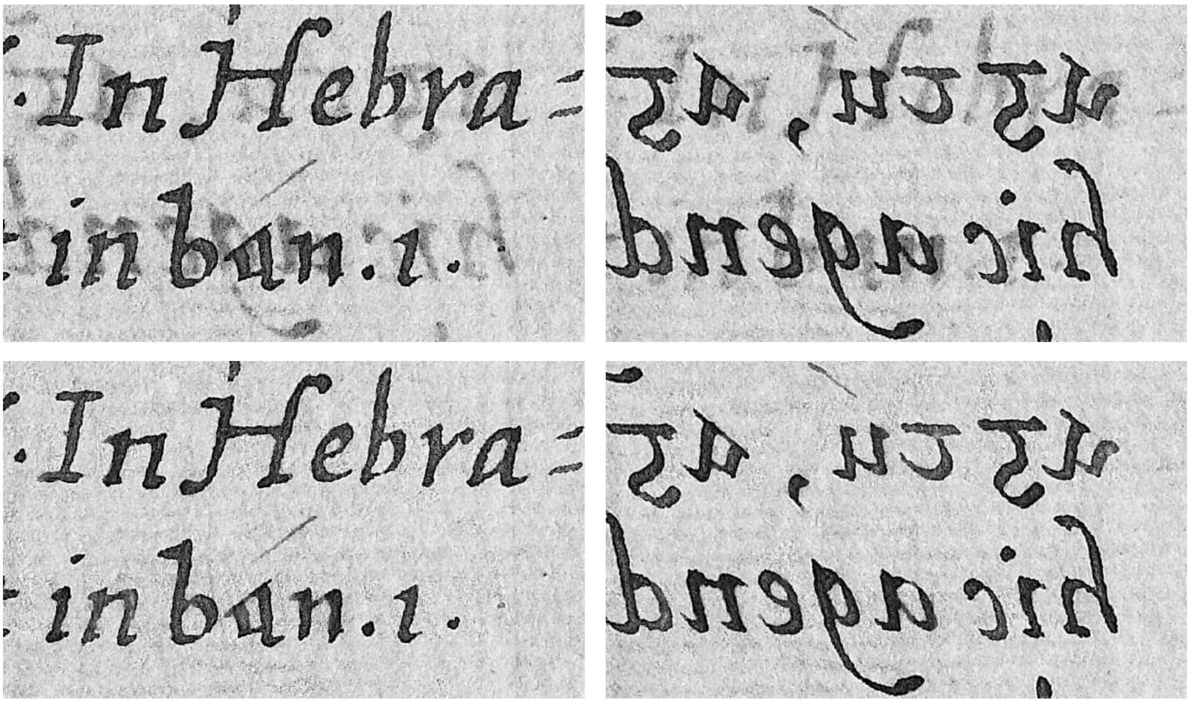


Figure 3.6: An example of restoration on a sample image pair with light bleed-through degradation. Top row: degraded recto (right) and verso samples. Bottom row: corresponding restored results, with the document character preserved and all of the bleed-through artefacts removed. **M02**.

the limiting masks, and the fact that the substitution of the unconstrained linear model to encourage reclassification of severe bleed-through regions misclassified as foreground only works well if the misclassified regions are small. A further limitation of this per-pixel approach is in cases where the bleed-through diffuses horizontally across the page, degrading a larger area than the originating text, an example of which can be seen in the image pair in Figure 3.9. In these cases the spread bleed-through pixels are not accounted for in the degradation model and will not be removed, resulting in a halo effect around restored bleed-through regions. The use of the unconstrained linear model for the final restored result can mitigate this effect when the spread is small, as the mixing parameters will blend those regions into the background. However, if the spread area is large, then the restored result will have artefacts remaining, as in Figure 3.9. A more complete evaluation and comparison of results with previous non-blind methods is presented in Chapter 5.

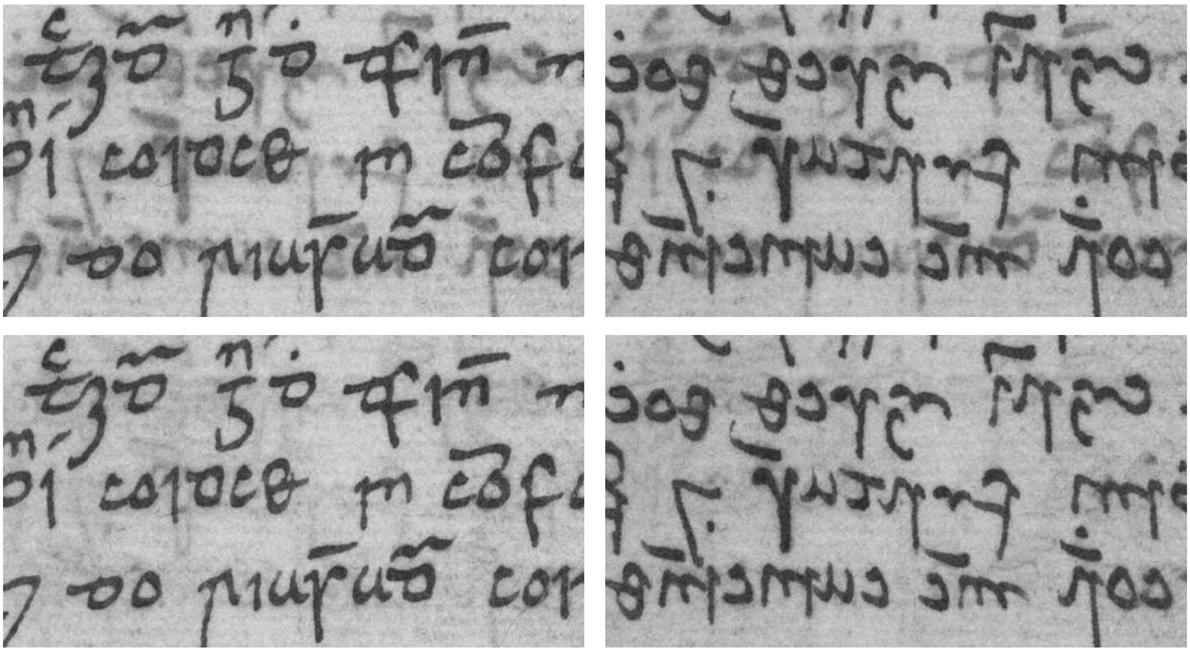


Figure 3.7: An example of restoration on a sample image pair with medium bleed-through degradation. Top row: degraded recto (right) and verso samples. Bottom row: corresponding restored results, with the document character preserved and most of the bleed-through artefacts removed. **M35**.

3.6 Colour Restoration

Where full colour images of the degraded pages are available, the results of the proposed method may be modified to produce clean colour results. This is achieved simply by converting the input RGB images to the Hue, Chroma, Luma (HCL) colourspace. The Luma channel is then replaced with the clean grayscale result from the proposed approach, and in the Hue and Chroma channels identified bleed-through regions are replaced with estimated local mean background values. An example of the restoration results applied to colour images can be seen in Figure 3.10. This simple approach works well in documents where the background colour is reasonably uniform. However in some illuminated manuscripts where background colouration is used to emphasise certain words or characters, applying this colour restoration method may interfere with such regions, negatively affecting the integrity of the clean document. An example of this effect can be seen in Figure 3.11. In these cases, restoration based on gayscale intensity alone is insufficient, and colour information needs to be taken into account.

Figure 3.8: An example of restoration on a sample image pair with severe bleed-through degradation. Top row: degraded recto (right) and verso samples. Bottom row: corresponding restored results. In this case the model does not remove a large proportion of the bleed-through on the recto side, due to the significant overlap in foreground and bleed-through intensities. However, the legibility of the restored result is not affected detrimentally. **M25**.

Figure 3.9: An example of restoration on a sample image pair where the bleed-through has spread horizontally across the page and has a larger area than its originating text. This effect is not taken into account by the per-pixel model, and visible bleed-through halo artefacts remain. **M16**.

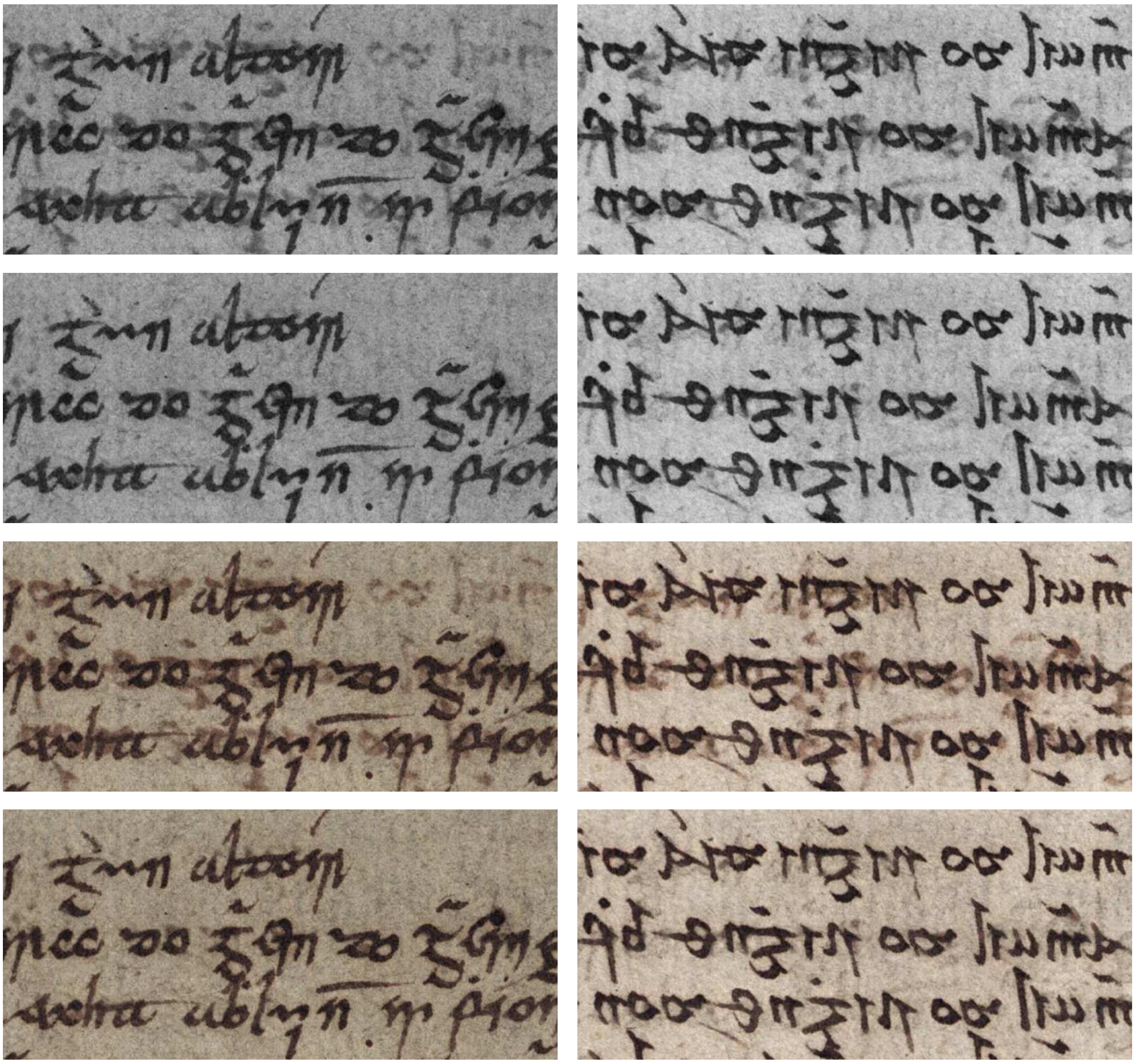


Figure 3.10: An example of colour restoration. Top to bottom: Original degraded recto (left) and verso sample images, restored results using the proposed approach, original degraded colour images, and clean colour images using the grayscale results. **M39**.

3.7 Summary

This chapter has presented a non-blind model-based approach to bleed-through restoration. The observed degraded images are modelled as a constrained linear combination of the original clean images, using binary foreground masks defined on both sides explicitly to limit the bleed-through to regions that are not foreground text, enforcing the opacity of foreground text over bleed-through. The intensity information on each side is used to define the masks initially, via K-means clustering, with the darkest of three clusters representing an estimate of foreground text. Spatial

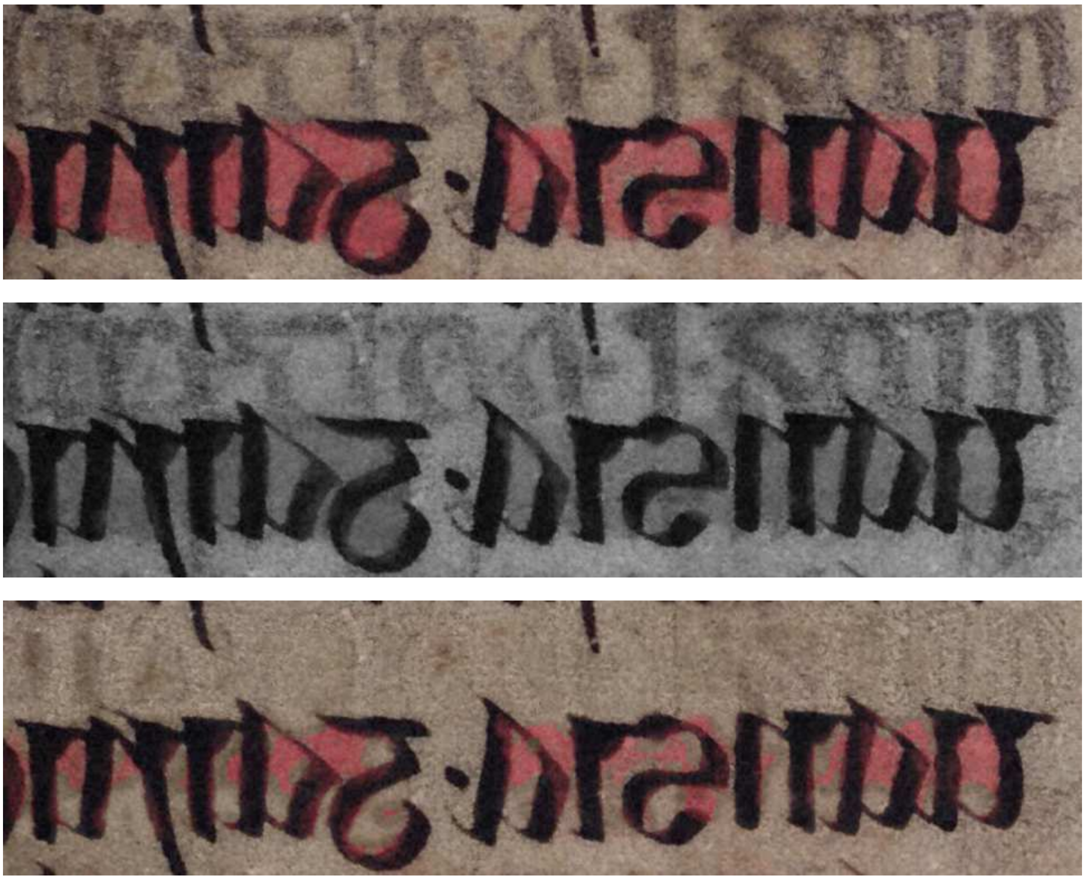


Figure 3.11: The colour restoration approach will not perform well on illuminated manuscript regions with background colouration. The colouration emphasis may not be translated to the grayscale image, and so is partially removed in the restored result. Top to bottom: Original degraded sample RGB image, grayscale image where colouration emphasis has been lost, and clean restored colour result with some colouration removed as bleed-through degradation. **M25**.

smoothness is enforced in a dual-layer framework, similar to [78] and [34], but instead of solving for a binary or ternary labelling, the model parameters including the clean image intensities are estimated using the degradation model. Intra-layer smoothness priors are used for the masks and mixing parameters on each side, and an inter-layer prior is also used for the mixing parameters to ensure that bleed-through cannot occur at the same location on both sides. Priors on the mixing parameters and clean images to ensure that restored regions are close to background intensity are also used. Estimates for the model parameters are obtained via a mixture of ICM optimisation and Gibbs sampling - drawing samples for the clean image estimates. A secondary linear model without limiting masks is substituted every 10 iterations for the clean image estimates to improve classification of bleed-through regions, and to ensure that the final estimates appear smooth with minimal visible bleed-through degradation remaining. The approach performs well on images with relatively light bleed-through, preserving the character of the documents. In more severe cases with large regions of relatively dark bleed-through, the method does not perform as well. However, in such cases the character of the documents is still maintained, and the document legibility is not affected detrimentally. In essence therefore this method is conservative, that is legibility is always improved in the output, and true foreground preserved, but at the cost of also preserving dark bleed-through in severe cases. Restoration results can also be extended to colour images using a simple colourspace transformation, as long as the document background does not include regions with added colouration that is not distinguishable in the grayscale domain.

4

Non-Parametric Approach

The previous chapter presented a conservative bleed-through removal method that produces visually pleasing results, but performs poorly in severely degraded cases due to the reliance on good initialisations. In order to improve performance on the severe cases it is necessary to decouple the bleed-through identification and restoration processes that are performed jointly in the previous approach, and to focus on explicit bleed-through classification rather than estimation of the clean image intensity. A number of factors can be identified to further improve performance as follows.

Firstly it is observed that the intensity distributions of foreground, background, and bleed-through regions may not be uniform across a single image due to non-uniform lighting during imaging, which is especially prevalent in photographed rather than scanned pages, or degradation of the background medium for example due to water stains or page warping. Classification based on global intensity properties will be detrimentally affected by these variations, as the intensity distributions of each class will overlap more significantly. Classification performance may be improved therefore either by examining local regions independently, or by removing any variations in the global intensity properties as a preprocessing step.

The initialisation used in the previous method is based on the intensity information of recto and verso sides independently, and the clean image intensities for each side are subsequently estimated using the intensity information from both sides. As the output relies on a good initialisation, it is observed that use of intensity information independently is a limiting factor in classification performance where foreground and bleed-through intensity distributions overlap significantly. An improvement in severe case classification can therefore be obtained by exam-

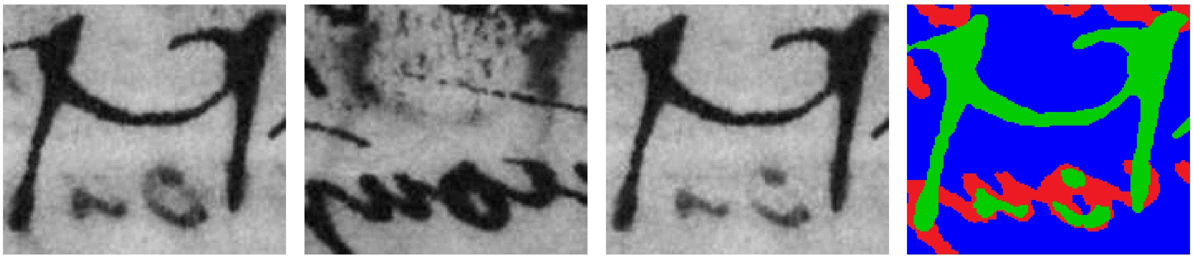


Figure 4.1: Highlighting the usefulness of a connected component based refinement to classification. Left to right: Degraded recto extract, corresponding degraded verso extract, clean recto output using the method from Chapter 3, and generated output labelling using the foreground mask estimates, where green represents foreground, blue background, and red bleed-through. The bleed-through regions that are not removed in the restored result are connected to correctly identified regions. **M38**.

ining intensity information from both sides of the page jointly rather than independently. Use of a joint intensity feature was first proposed in [33], where it is used to guide labelling of recto and verso images separately, using a dual-MRF framework with two observation fields and one label field.

A final observation is that the spatial information of a per-pixel based approach is not powerful enough to reclassify dark bleed-through misclassified as foreground initially, as the neighbourhood of foreground and bleed-through pixels will often be similar. However, examining classification labelling output in Figure 4.1, generated using the final foreground mask estimates from the approach in Chapter 3, it is clear that misclassified bleed-through regions are connected to correctly classified regions, shown in red. This suggests that rather than a pixel based refinement, a larger connected component based refinement will be more powerful in identifying bleed-through misclassified as foreground.

Based on these observations, this chapter presents a second framework for non-blind document bleed-through removal. This approach is designed to cope with a wider range of bleed-through degradations than that proposed in Chapter 3, whilst maintaining the ultimate goal of preserving as much of the original document appearance as possible. The proposed framework includes image preprocessing to remove local intensity variations, pixel region classification based on a segmentation of the joint recto-verso intensity histogram and connected component analysis on the subsequent image labelling. Finally restoration of the degraded regions is performed using exemplar-based image inpainting[1]. Again, in this chapter it is assumed that the registration stage has already been performed, and in what follows 'verso' will refer to the flipped, registered verso side.

---

[1] Sections of this chapter are based on the work published in [60].

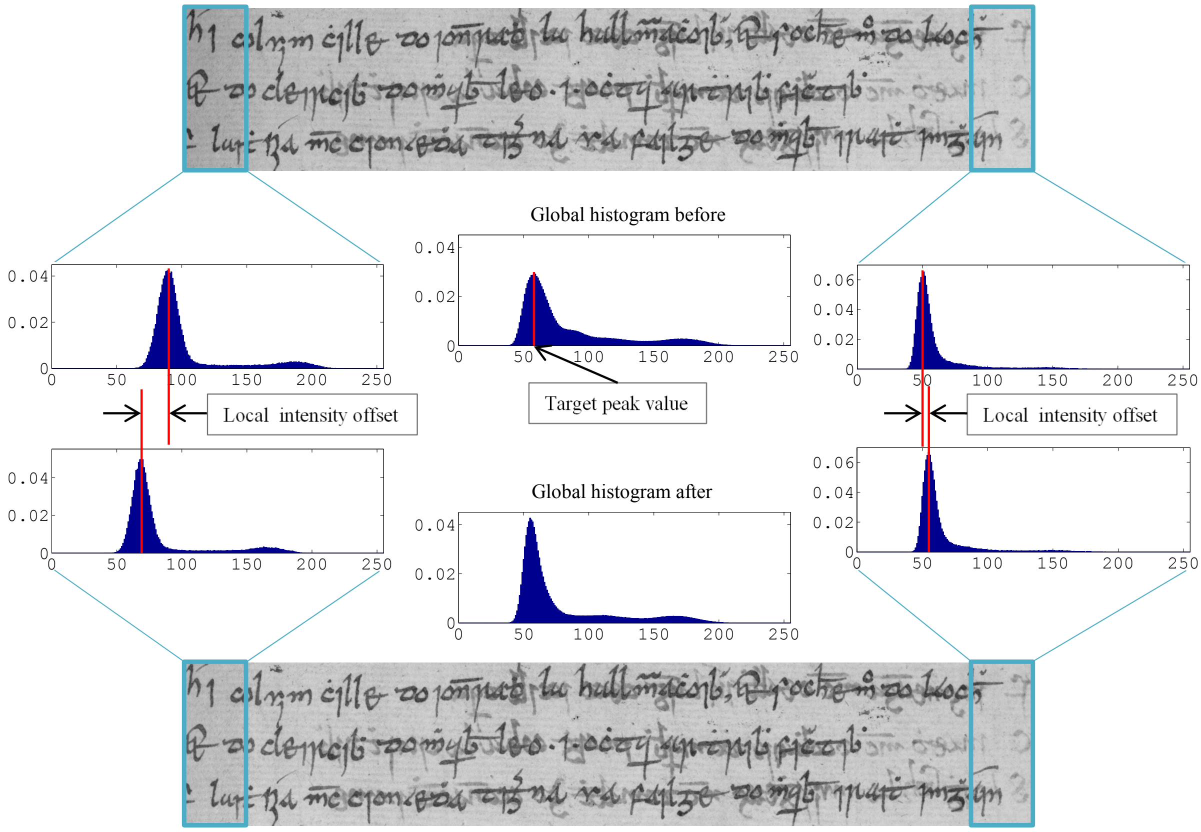


Figure 4.2: An illustration of the preprocessing stage on a sample recto image, **M21**.

4.1 Preprocessing

As classification is performed using joint global intensity properties, this approach is sensitive to local variations in intensity profile over the document image, for example due to reduced lighting close to the binding, or water stains. Therefore prior to classification it is necessary to compensate for any such variations. Since many document imaging projects perform little or no image enhancement it cannot be assumed that the resultant images have uniform global intensity properties. Any trends are removed by exploiting firstly the fact that the only constant in most document images is the background medium; the presence and strength of any text or other foreground information may change across a page, and secondly that the largest proportion of pixels corresponds to background. Therefore the recto and verso images are adjusted separately by applying local intensity offsets such that the peaks of the local intensity histograms, corresponding roughly to mean local background intensities, are aligned. This is performed by examining intensity histograms of overlapping blocks in the original image. An intensity value for each block is stored that corresponds to the maximum peak in the block intensity histogram. The mean of these intensity values is set as the target mean background intensity, and an offset

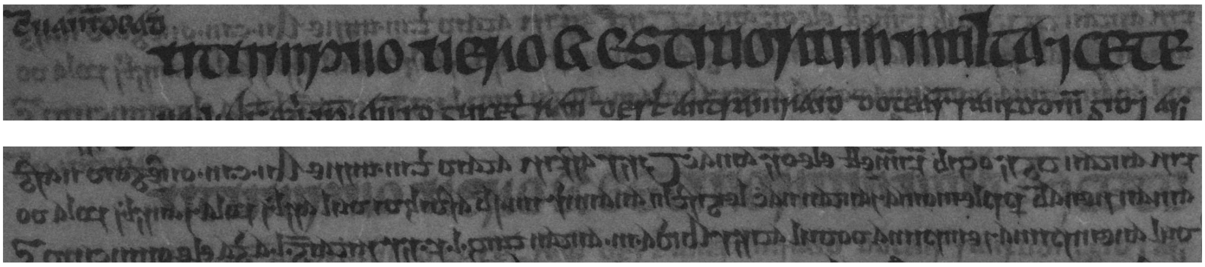


Figure 4.3: An example of a severely degraded document recto-verso pair sample with significant overlap between foreground, background, and bleed-through intensity distributions, and a relatively large proportion of overlapping foreground text on both sides of the page. Top: recto sample. Bottom: corresponding registered verso sample. **M24**.

matrix is defined from the difference between the target and stored local intensity values. The block size is set empirically to 200 × 200 pixels with an overlap of 50. An illustration of this stage is shown in Figure 4.2.

4.2 Classification

This proposed method aims to create a joint labelling of recto and verso images, from a set of four 'pair' labels: background on both sides, *bgbg*, recto foreground and verso bleed-through, *fgbl*, recto bleed-through and verso foreground, *blfg*, or foreground on both sides, *fgfg*. As discussed in Chapter 2, the degree of intensity overlap between different labels can be reduced significantly by combining intensity information from both sides of the page. In previous approaches, the ratio and difference of intensities have been used [15,21,29,35,44,45,67]. Figure 4.4 shows a comparison of both 1D and joint ground truth label distributions for the sample image pair in Figure 4.3, using the recto intensity, intensity difference, and intensity ratio. For illustration purposes, the sample image pair in Figure 4.3 used to obtain these distributions is severely degraded, with a significant overlap between foreground, background, and bleed-through intensity distributions, and a relatively large proportion of *fgfg* regions. As can be seen in Figure 4.4, the difference and ratio measures do not distinguish between foreground on both sides of the page (*fgfg*), and background on both sides (*bgbg*), which may result in important foreground information being removed from one or both sides. Use of difference and ratio measures therefore requires further information to preserve *fgfg* regions. There is also still a significant overlap between bleed-through and foreground label distributions when using these features. Figure 4.5 shows contour plots of joint intensity histograms of recto-verso pairwise intensities from the sample in Figure 4.3 with 1D and joint ground truth labels. It is clear that using pairwise intensities can reduce the label overlap regions and preserve the distinction between *fgfg* and *bgbg* labels. Thus in this approach the recto and verso images I_r, I_v are treated as a joint image I_p, and each

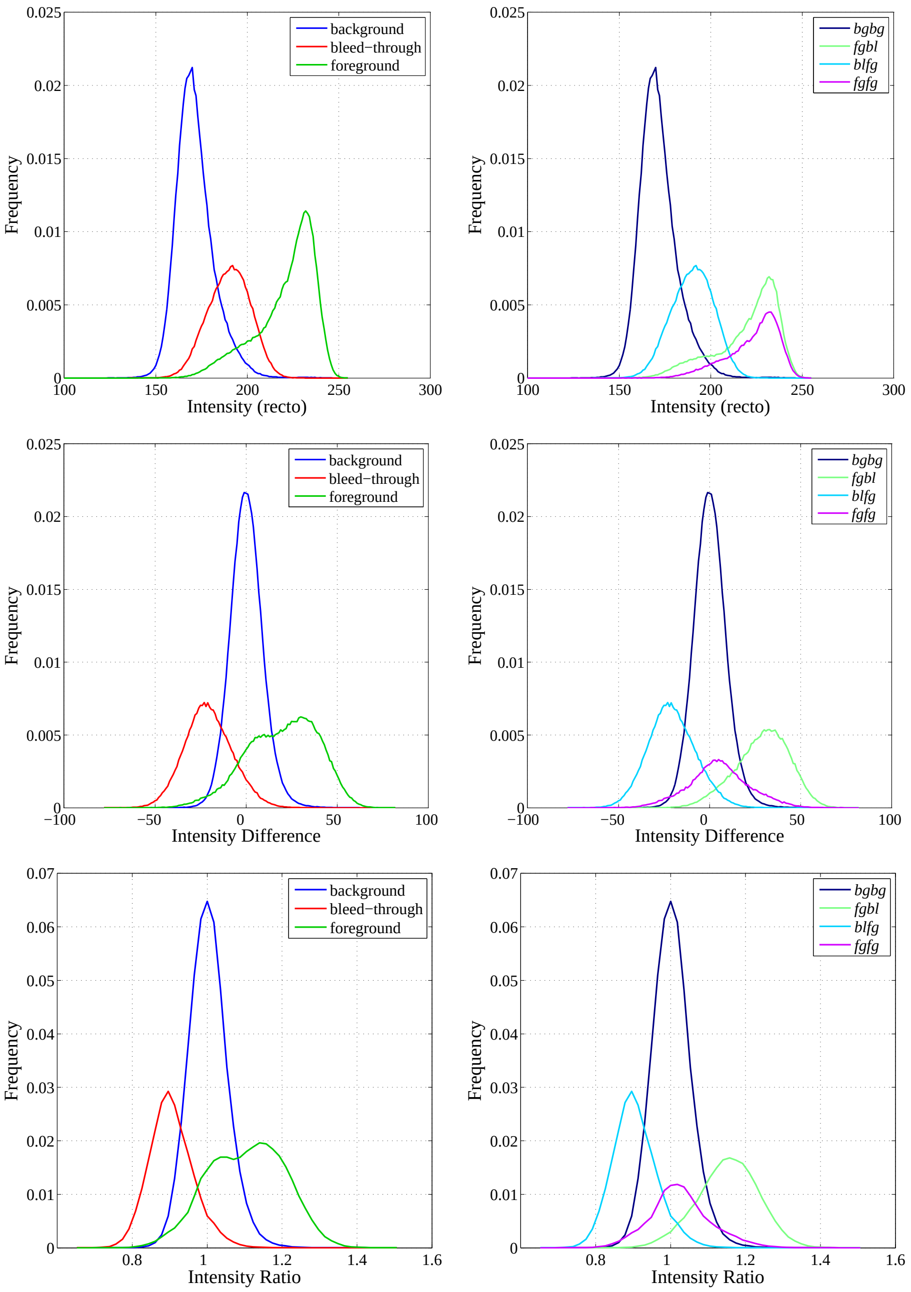


Figure 4.4: There is a significant overlap between 1D and joint label distributions in a severe bleed-through example (shown in Figure 4.3) when using single intensity features. The left column shows histograms of ground truth 1D labels for the recto side, the right column similar histograms for 2D joint labels of both sides. Top to bottom: recto intensity (I_r), intensity difference ($I_r - I_v$), intensity ratio ($\frac{I_r}{I_v}$)

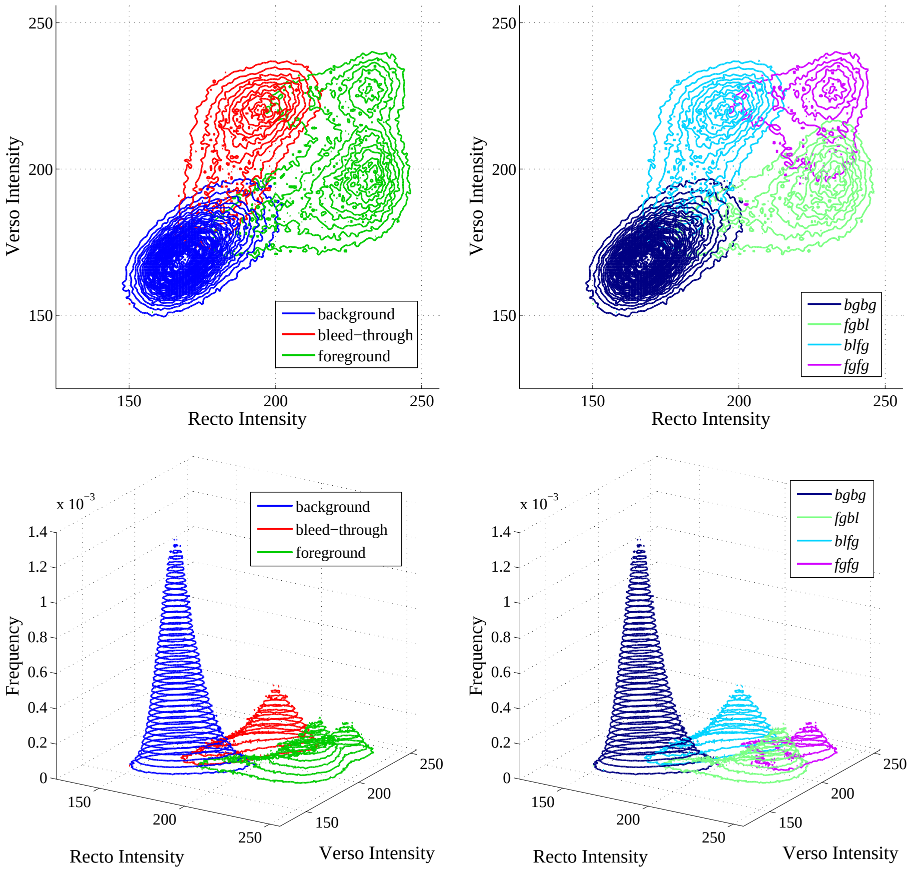


Figure 4.5: The overlap between 1D and 2D label distributions in a severe bleed-through example can be reduced using the joint intensity feature. Top row: contour plots of joint intensity histograms of recto-verso intensities, $([I_r, I_v])$, from the sample in Figure 4.3, highlighted with 1D (left) and 2D (right) ground truth labels. Bottom row: 3D view of the contour plots.

---

pixel pair $I_r(x, y), I_v(x, y)$ is treated as a single pixel $I_p(x, y)$ with intensity pair $\mathbf{i}$ in the range $[0, 255] \times [0, 255]$, where, as in Chapter 3, 0 corresponds to white, and 255 to black. Further processing can then be performed in the spatial domain to correct any misclassifications due to the remaining overlap regions.

There are therefore two stages to classification, firstly a joint histogram of intensity pairs is segmented into four regions corresponding to the four labels using a MRF-based clustering approach. This histogram labelling is then used as a look up table to obtain an initial image labelling. Secondly, a set of rules governing connected label components in the image labelling is applied to produce the final label field for the recto-verso image I_p.

4.2.1 Joint Histogram Segmentation

Examining the joint intensity histograms in Figure 4.5, it is clear from the large peak in the points with lighter intensity that the largest proportion of pixels in I_p will correspond to regions where both recto and verso are background ($bgbg$). There are then three smaller peaks corresponding to the three remaining labels. In this approach, labels are assigned based on the distance from these maxima. However, it is necessary to ensure that the histogram segmentation maps to an image segmentation that is spatially smooth. So the labelling is formulated as a MRF framework with a spatial smoothness prior based in the recto-verso image domain rather than the joint histogram domain. The energy E corresponding to a particular labelling l is defined as

$$E(l) = \sum_{\mathbf{i} \in \mathcal{I}} \left[\beta_{\mathbf{i}} U_{\mathbf{i}}(l_{\mathbf{i}}) + \alpha \gamma_{\mathbf{i}} \sum_{\mathbf{j} \in \mathcal{N}_{\mathbf{i}}} V(l_{\mathbf{i}}, l_{\mathbf{j}}) \right], \tag{4.1}$$

where $l_{\mathbf{i}} \in \{bgbg, fgbl, blfg, fgfg\}$ is the label at point $\mathbf{i}$ in the joint histogram h, and $\mathcal{I}$ corresponds to the set of non-zero histogram entries. The remaining terms in Equation. 4.1 are discussed below.

Unary Terms

The unary energy $U_{\mathbf{i}}(l_{\mathbf{i}})$ represents the cost of a point $\mathbf{i}$ in the joint histogram being assigned label $l_{\mathbf{i}}$. In order to include the frequency distribution of intensity pairs in the classification of each label, $U_{\mathbf{i}}(l_{\mathbf{i}})$ is defined as the mahalanobis distance between point $\mathbf{i}$ and the centre of the label cluster corresponding to $l_{\mathbf{i}}$.

Binary connections

The neighbourhood structure $\mathcal{N}_{\mathbf{i}}$ of a point $\mathbf{i}$ is defined in the image domain as

$$\mathcal{N}_{\mathbf{i}} = \left\{ \mathbf{j} \,|\, \mathbf{j} = I_p(x', y'), \mathbf{i} = I_p(x, y), (x'y') \in \mathcal{N}_{x,y} \right\}. \tag{4.2}$$

So each instance of an intensity pair $\mathbf{i}$ is located in the recto-verso image I_p, then the corresponding points in the joint histogram of the 4-connect neighbours in I_p of these instances are added to the neighbourhood of $\mathbf{i}$. This means that the size of the neighbourhood for each intensity pair depends on the frequency of that point in the joint histogram.

Binary Terms

The pairwise energy $V(l_{\mathbf{i}}, l_{\mathbf{j}})$ represents the cost of neighbouring points in the histogram being assigned labels $l_{\mathbf{i}}$ and $l_{\mathbf{j}}$ respectively. In order to obtain estimates for these energy terms, ground truth joint image label fields were generated for a database of 25 preregistered recto-verso image pairs with manually generated ground truth foreground text masks. This database is presented with further details in Chapter 5. Examining a subset of the ground truth label fields, the pairwise energies are set empirically as the negative log likelihood of co-occurrence of labels in that subset (see Table 4.1).

Table 4.1: Likelihood of co-occurrence of labels

$l_{\mathbf{i}}$ \ $l_{\mathbf{j}}$	*bgbg*	*fgbl*	*blfg*	*fgfg*
bgbg	0.66	0.00065	0.0069	0.00013
fgbl	0.0065	0.13	0.0001	0.0022
blfg	0.0069	0.0001	0.13	0.0021
fgfg	0.00013	0.0022	0.0021	0.046

Model Variations

The parameters $\beta_{\mathbf{i}}$ and $\gamma_{\mathbf{i}}$ in Equation 4.1 represent the weighting applied to binary and unary energies assigned to each intensity pair value. Three different variations of the general model in Equation 4.1 are examined in this work:

1: $[\beta_{\mathbf{i}} = h(\mathbf{i}), \gamma_{\mathbf{i}} = 1]$. In this case the unary energy term for each intensity pair is weighted by the corresponding histogram frequency, so each unary term depends on the number of instances of its intensity in the recto-verso image. This model is therefore similar to the standard dual-MRF used in [33,35], however classification is performed jointly, and the constraint that each instance of an intensity pair must have the same label is explicitly enforced in this case as the classification is performed in the joint histogram rather than the image domain.

2: $[\beta_{\mathbf{i}} = 1, \gamma_{\mathbf{i}} = \frac{1}{h(\mathbf{i})}]$. In this model it is the binary energy that is weighted by the inverse of the histogram frequency, and the unary energy for each intensity pair has equal weight in the overall energy. Essentially this means that the energy contribution for each intensity pair becomes the corresponding unary energy plus an average of the pairwise energy for the 4-connect neighbourhoods of each instance of that intensity in the recto-verso image domain.

3: $[\beta_{\mathbf{i}} = 1, \gamma_{\mathbf{i}} = 1]$. In this case neither of the model terms is weighted by the histogram frequency, so the unary energy for each intensity is given equal weight in the overall energy, and the pairwise energy amounts to a sum over all the 4-connect neighbourhoods of each instance in the recto-verso image.

Smoothness Weight

A frequency independent smoothness weight, α, is applied to $V(l_{\mathbf{i}}, l_{\mathbf{j}})$ to balance the global influence of the binary and unary energies on the histogram label field. An unique value for each model variation is chosen empirically by examining the output histogram label fields generated using different smoothness weight values. Figure 4.6 shows the percentage histogram error for each model on each image over a range of smoothness weight values, and the chosen minimum error weights. Unsurprisingly the optimal range for Model 3 is much lower than Models 1 or 2, as no frequency dependent weighting is applied for Model 3, and therefore the pairwise energy

Table 4.2: Average proportions of each label as percentages of joint histogram intensity pairs and joint image pixels.

	bgbg	*fgbl*	*blfg*	*fgfg*
Histogram area (%)	17.98	33.08	32.31	16.63
Image area (%)	65.48	14.16	14.95	5.39

is much higher than the unary energy due to the large neighbourhood. It is also interesting to note that for three of the images in the database the corresponding minimum error weight is extremely small compared to the average, which would indicate that none of the models accurately captures the image properties in these cases. This could be due to the fact that the images contain a higher proportion of *fgfg* than the average, or a higher text to background ratio, and so the pairwise energy terms defined in Table 4.1 may not accurately represent the particular image statistics.

Initialisation

In most bleed-through cases, the *bgbg* label accounts for the largest proportion of pixels in the joint image, but as can be seen in Table 4.2, has a relatively low range of distinct intensities compared to the *fgbl* and *blfg* labels. The *fgfg* label corresponds to the smallest proportion of pixels in both the image and histogram label fields. The initialisation of the joint histogram labelling is therefore designed to avoid over-classification of intensities as *bgbg* and under-classification of *fgfg*. Initial estimates for the cluster centres are obtained as follows:

(i) Find the peak of the joint histogram and set as *bgbg* centre.

(ii) Set *fgfg* centre as $[\max(I_r), \max(I_v)]$.

(iii) Obtain quadrants of joint histogram corresponding roughly to *fgbl* and *blfg* using centre line defined by *bgbg* and *fgfg* centres.

(iv) Set peaks of quadrants as *fgbl* and *blfg* cluster centres.

(v) Update *fgfg* centre to $[fgbl(1), blfg(2)]$.

(vi) Cluster intensity pairs using the minimum Euclidean distance to the estimated centres.

The clusters and centres are then updated once using the Mahalanobis distance for *bgbg* and Euclidean distance for the remaining three labels. The use of the different distances initially ensures that the strong *bgbg* distribution is not over-classified, and that the relatively weak *fgfg* distribution is favoured in overlap regions.

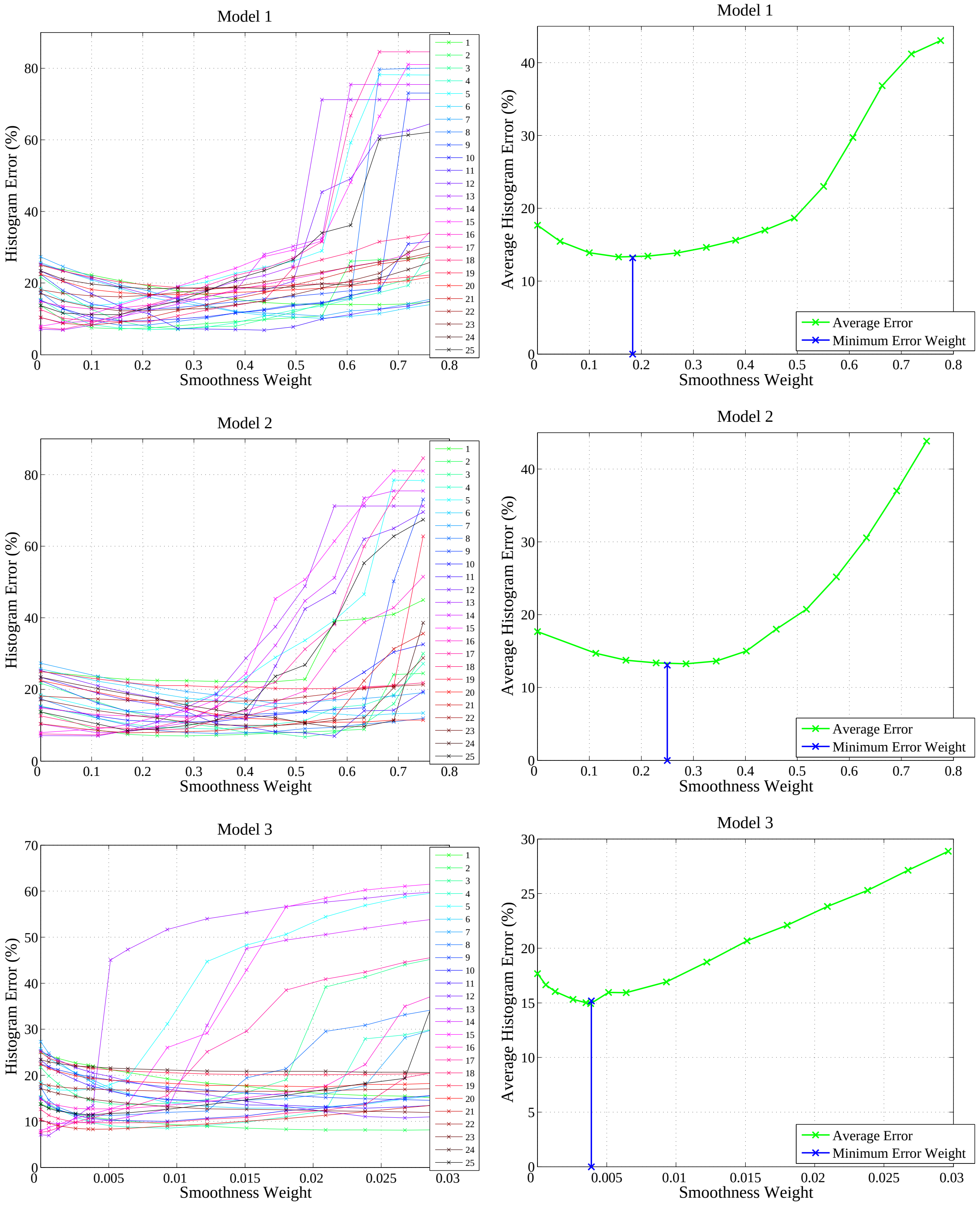


Figure 4.6: Percentage Histogram labelling error for ranges of different smoothness weights for each model variation over the database (left), and average errors (right) with the empirically chosen smoothness weights shown blue.

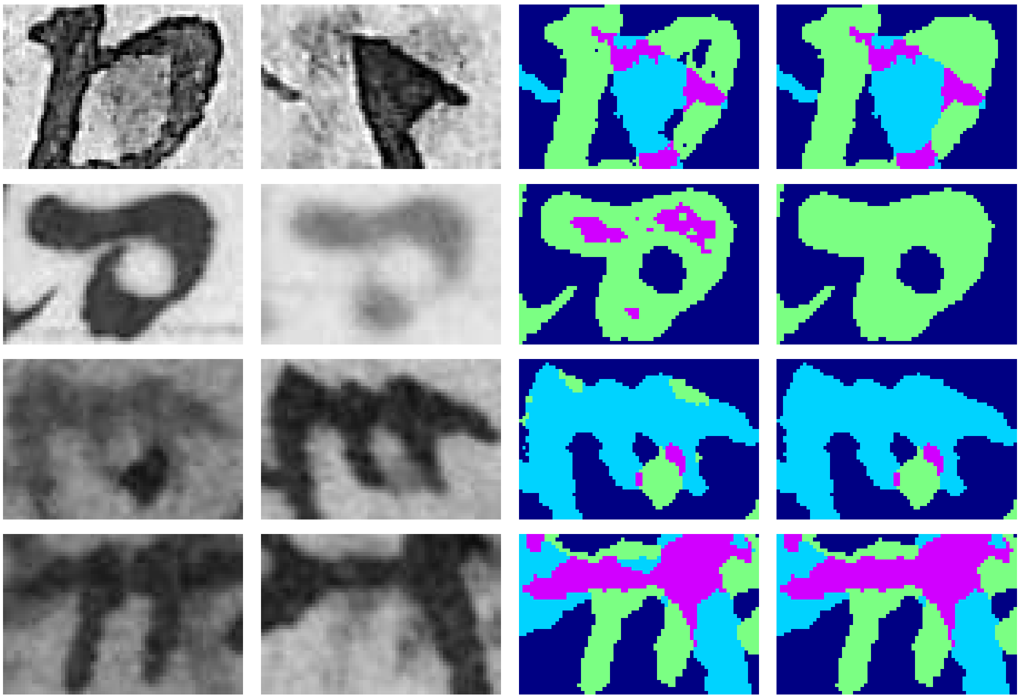


Figure 4.7: This image illustrates the implementation of the connected component boundary rules. Left to right: recto extract, verso extract, image labelling before rules applied, and image labelling after rules applied. Row 1: Misclassified *bgbg* components (dark blue) are corrected, **M02**. Row 2: *fgfg* components (pink) are replaced with *fgbl* (green), **M01**. Row 3: *fgbl* components (green) connected to *blfg* (light blue), but not *fgfg* (pink) are replaced with *blfg*, **M39**. Row 4: A *blfg* component is connected to *fgfg*, but not *bgbg* so is replaced with *fgfg*, **M39**.

Optimisation

Quadratic Pseudo-Boolean Optimization (QPBO) [56] is used to minimise the energy in Equation 4.1, repeatedly iterating through each of the four label options in the order [*fgfg*,*blfg*,*fgbl*,*bgbg*] until convergence. Different label permutations were tested and found to have little impact on the subsequent results.

4.2.2 Image Segmentation

Following colour segmentation, the image labelling is initialised by using the histogram labelling as a look up table for pixels in the recto-verso image I_p. A subset of pixels will inevitably be misclassified due to the overlapping nature of the histogram label boundaries, however as the pairwise energy used in the histogram segmentation is derived from neighbourhoods in the image domain, spatial smoothness has already been enforced, which increases the likelihood that misclassified pixels will be grouped together and connected to correctly classified groups.

Therefore a full per-pixel analysis is not performed on the image labelling, and instead connected components of each label are examined, and rules governing permitted neighbouring components iteratively applied to correct misclassifications until convergence. The rules for each label are as follows.

bgbg: This label covers the greatest proportion of the image, and so connected components will mostly be larger than the average character size. Smaller components correspond either to valid within character spaces, such as in ‘a’ and ‘o’, or to misclassifications. To avoid relabelling valid within character spaces, only the connected components that are less than 10% of the average character component size are analysed. Presumed to be mislabelled these components are relabelled with the label corresponding to the largest proportion of their neighbours.

fgfg: Conversely, this label covers the smallest proportion of the image, and as very dark bleed-through can often be mislabelled as *fgfg*, no assumptions can be made about the size of components and all are examined. The outer edges of components with this label must contain both *fgbl* and *blfg* labels, as overlapping text regions will originate from text alone on both sides. If this is not the case the component is relabelled *fgbl* or *blfg* according to which is present in the outer edge, or as *bgbg* if neither is present.

fgbl: For this label, again only components less than 10% of the average character size are examined. The outer edges of these components must contain either *fgfg* and *bgbg*, or *bgbg* only. If the outer edge of a component contains *fgfg*, but not *bgbg* also, then the component is relabelled as *fgfg*. If the outer edge of a component contains the label *blfg* , but not *fgfg*, then it is relabelled *blfg*.

blfg: Components labelled *blfg* are processed in exactly the same way as *fgbl*, with the two labels interchanged. Example implementations of these rules are illustrated in Figure 4.7.

4.3 Restoration

As for the approach in Chapter 3, the aim of this method is to preserve as much of the document as possible. The background texture is therefore again preserved to ensure that the experience of studying the document *image* remains close to that of studying the *physical* document. The restored recto and verso images $Y_r(x,y), Y_v(x,y)$ are obtained by replacing identified bleed-through regions, that is, where $l_\mathbf{i}$ = *fgbl* for $Y_r(x,y)$, and $l_\mathbf{i}$ = *blfg* for $Y_v(x,y)$, with background texture from clean background images $B_r(x,y), B_v(x,y)$.

4.3.1 Clean Background

The background images are generated using the exemplar based inpainting algorithm of Criminisi et al. [17], that uses small sampled patches of nearby texture to fill image holes. The images $B_r(x,y), B_v(x,y)$ for recto and verso sides are generated using regions labelled as *bgbg* as the texture source, and inpainting all other label regions. Problems may be encountered with this

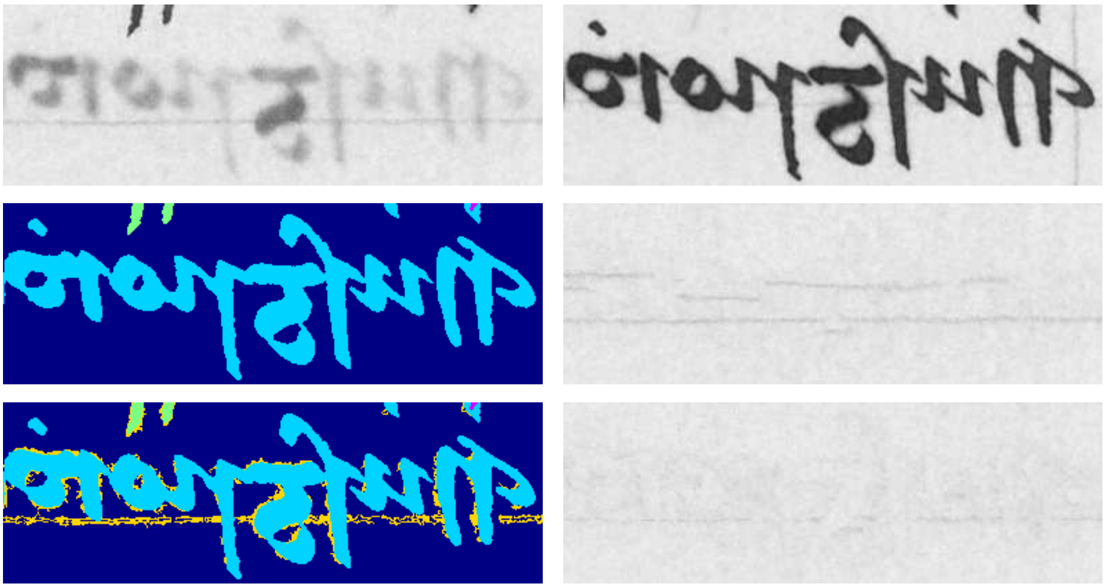


Figure 4.8: An example extract of an image where the gradient needs to be examined to improve the background plate. Top row: left - degraded recto extract with faint ruled lines, right - corresponding verso. Second row: left - image labelling (dark blue=texture source), right - visible artefacts in the recto background plate. Bottom row: left - labelling with 10% of source gradients removed (yellow), right - the improved background plate. **M01**.

approach in regions where faint foreground information might not have been identified during classification (b_3 errors as defined in Section 4.4), with the result that foreground patterns are replicated in the background images. To mitigate this the gradients of the regions labelled as *bgbg* are examined and the highest 10% of gradients removed from the inpainting source (see Figure 4.8). The advantage of using exemplar-based inpainting rather than texture synthesis is that no assumptions need to be made about the document background intensity distribution.

4.3.2 Blending

Using a per-pixel replacement of bleed-through pixels with corresponding clean background pixels creates restored images with visible seams at the edges of replaced regions. These artefacts are prevented by identifying the seam locations as the inner and outer edges of components labelled *fgbl* or *blfg*, and then performing a weighted sum of the original and background images across these edges to produce a smooth transition between original and replaced regions in the final restored results. The effect of this stage is illustrated in Figure 4.9.

4.3.3 Colour Restoration

Similarly to the method presented in Chapter 3, where full colour images of the degraded pages are available, the results of the proposed method may be reused to produce clean colour results,

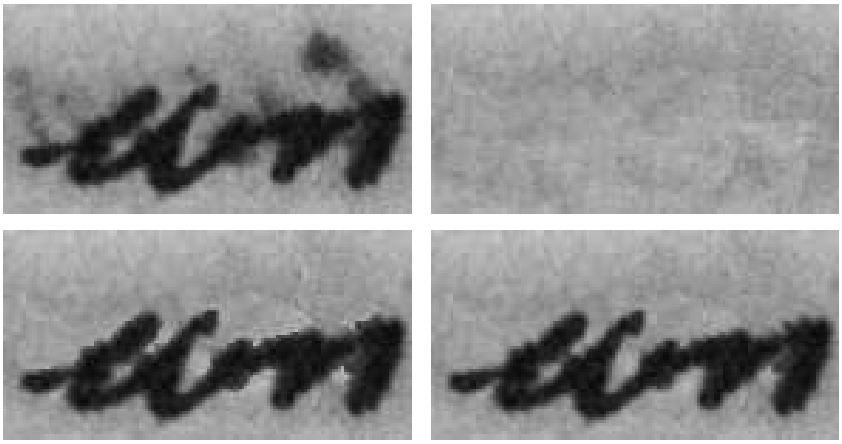


Figure 4.9: An example of blending the background image with the degraded image in bleed-through boundary regions. Top row: left - original degraded image sample, right - corresponding clean background. Bottom row: left - restored image with no blending, right - restored image with blending. **M37**.

using the colour images to generate the clean background images. As identified bleed-through regions are replaced with texture generated from background samples in this approach, rather than estimated clean intensity combined with estimated mean background values for Hue and Chroma channels, the colour restoration here is more robust to background colouration, as can be seen in Figure 4.10

Further colour image results for the proposed Model 2 on images from outside the database are shown in Figures 4.11-4.14. As these images, sourced from the Google Book Scanning Project [5], are full page examples that have different characteristics from those contained in the database, the restoration results show that this approach is potentially robust to a wider range of document types than those on which it was developed. The examples contained in Figures 4.12 and 4.13, as well as bleed-through, also suffer from the related ‘press-on’ degradation. This occurs when ink from the opposite page, rather than the reverse side of the same page, has migrated to its corresponding facing page over time, and interferes similarly with the foreground text, though in these examples the press-on is not severe. As this degradation is not included in the restoration model, it is treated as faint foreground text and left untouched in the restored result. Figure 4.14 shows an example where a modern library stamp imprinted on the blank verso side has bled-through and damaged the foreground text on the recto side. The proposed approach successfully removed the stamp interference in the restored result, however a small section of foreground text in the bottom left hand corner of the recto has also been removed due to the relatively low proportion of *blfg* in this example.

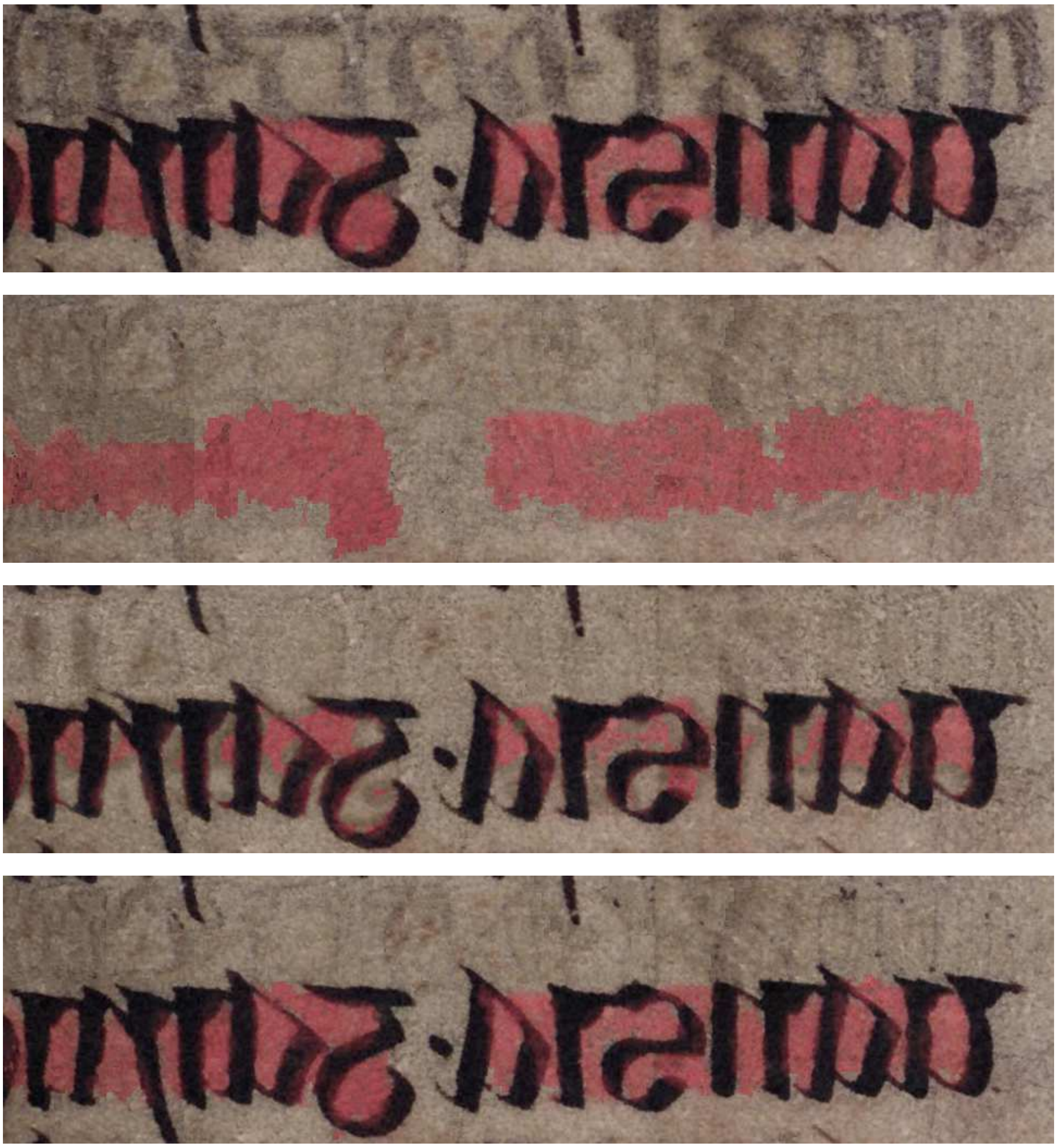
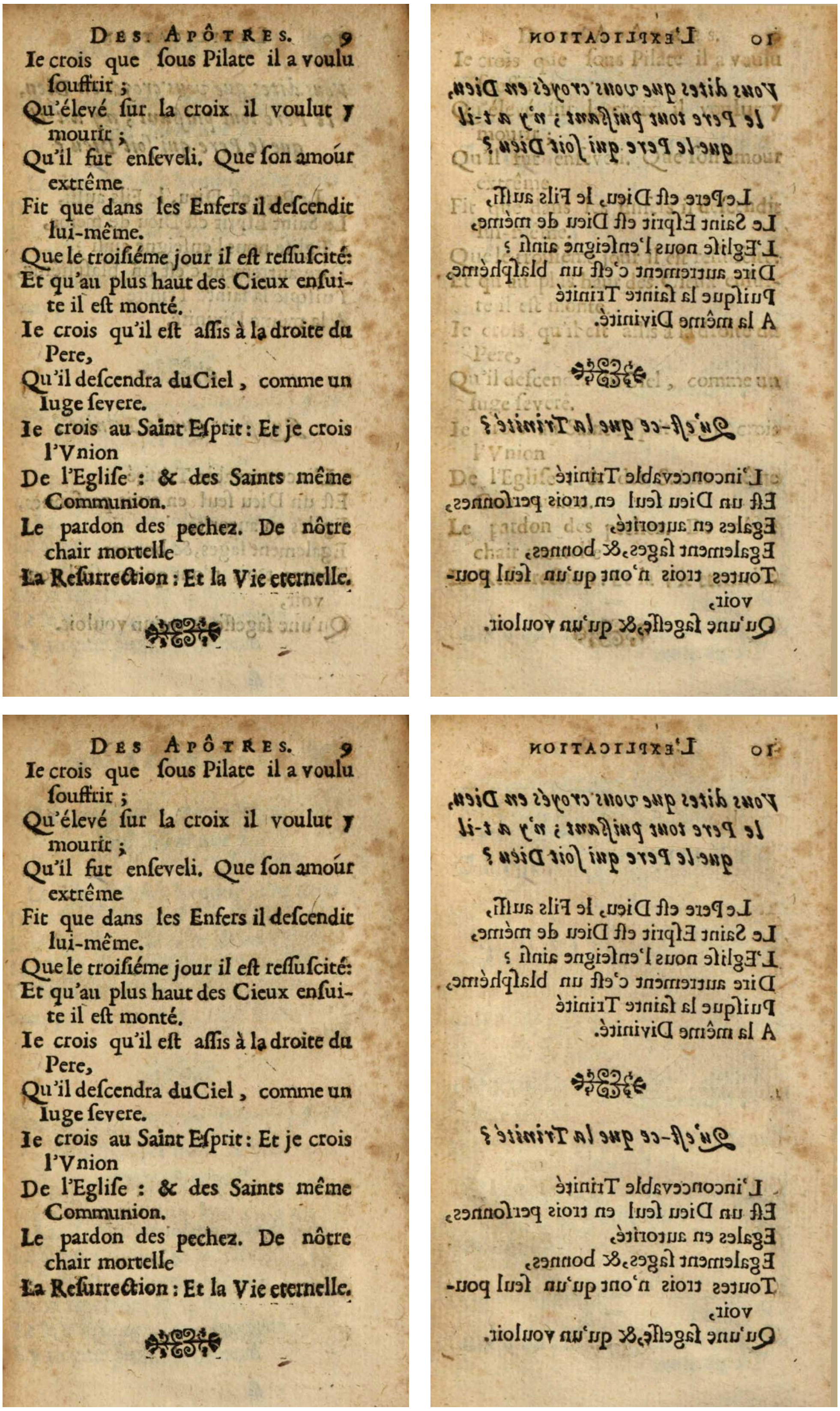


Figure 4.10: The colour restoration approach is more robust to background colouration than the method in Chapter 4. Top to bottom: Original degraded sample RGB image, generated clean background image, restored colour result from the method presented in Chapter 3 with some colouration removed as bleed-through degradation, and clean restored colour result from the method proposed in this Chapter. **M25**.



DES APÔTRES. 9

Ie crois que ſous Pilate il a voulu ſouffrir ;
Qu'élevé ſur la croix il voulut y mourir ;
Qu'il fut enſeveli. Que ſon amour extrême
Fit que dans les Enfers il deſcendit lui-même.
Que le troiſiéme jour il eſt reſſuſcité:
Et qu'au plus haut des Cieux enſuite il eſt monté.
Ie crois qu'il eſt aſſis à la droite du Pere,
Qu'il deſcendra du Ciel, comme un Iuge ſevere.
Ie crois au Saint Eſprit: Et je crois l'Vnion
De l'Egliſe : & des Saints même Communion.
Le pardon des pechez. De nôtre chair mortelle
La Reſurrection : Et la Vie eternelle.

10 L'EXPLICATION

Vous dites que vous croyés en Dieu, le Pere tout puiſſant ; n'y a-t-il que le Pere qui ſoit Dieu ?

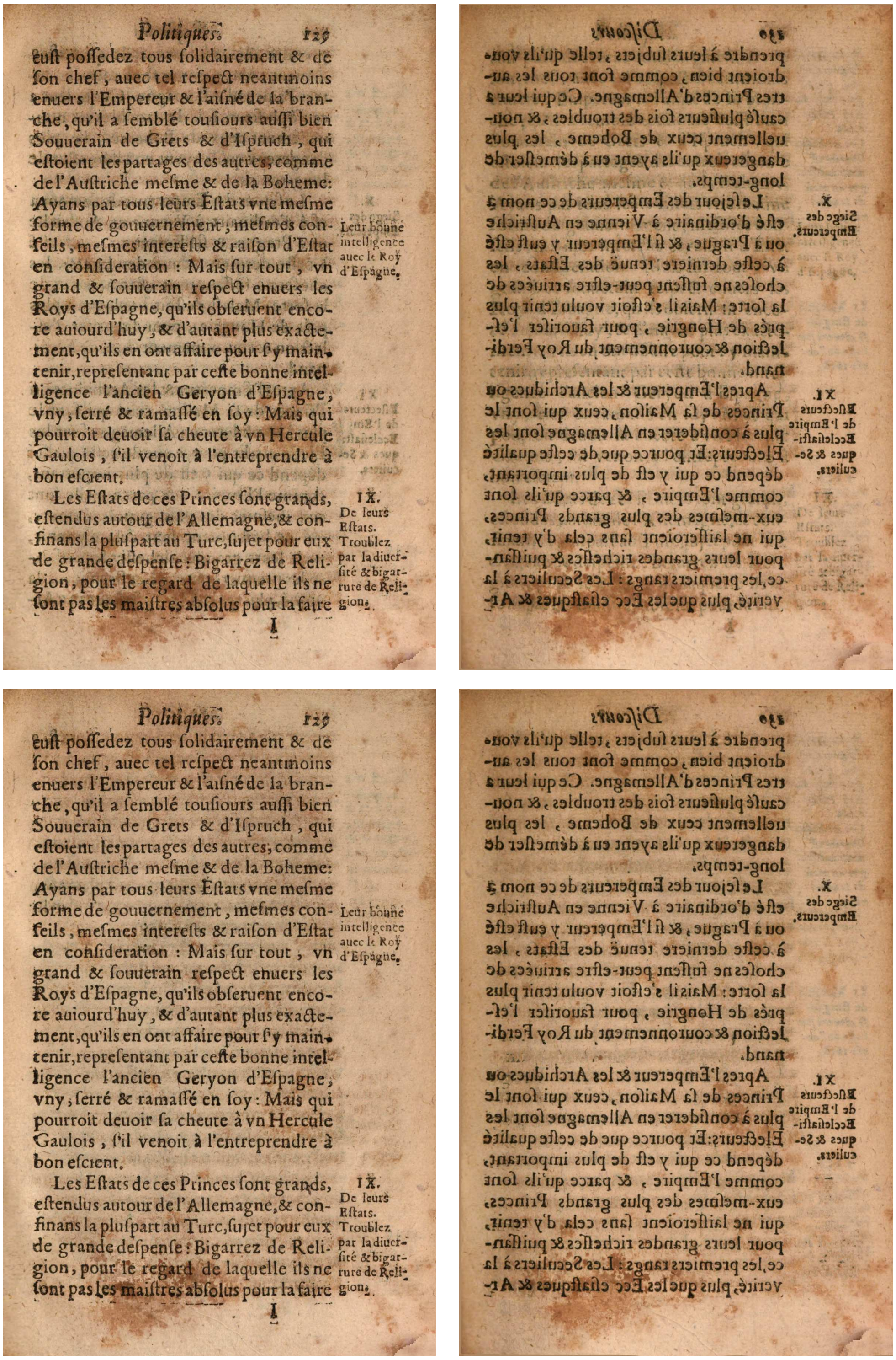
Le Pere eſt Dieu, le Fils auſſi,
Le Saint Eſprit eſt Dieu de même,
L'Egliſe nous l'enſeigne ainſi ;
Dire autrement c'eſt un blaſphême,
Puiſque la ſainte Trinité
A la même Divinité.

Qu'eſt-ce que la Trinité ?

L'inconcevable Trinité
Eſt un Dieu ſeul en trois perſonnes,
Egales en autorité,
Egalement ſages, & bonnes,
Toutes trois n'ont qu'un ſeul pouvoir,
Qu'une ſageſſe, & qu'un vouloir.

DES APÔTRES. 9
Ie crois que ſous Pilate il a voulu ſouffrir ;
Qu'élevé ſur la croix il voulut y mourir ;
Qu'il fut enſeveli. Que ſon amour extrême
Fit que dans les Enfers il deſcendit lui-même.
Que le troiſiéme jour il eſt reſſuſcité:
Et qu'au plus haut des Cieux enſuite il eſt monté.
Ie crois qu'il eſt aſſis à la droite du Pere,
Qu'il deſcendra du Ciel, comme un Iuge ſevere.
Ie crois au Saint Eſprit: Et je crois l'Vnion
De l'Egliſe : & des Saints même Communion.
Le pardon des pechez. De nôtre chair mortelle
La Reſurrection : Et la Vie eternelle.
10 L'EXPLICATION
Vous dites que vous croyés en Dieu, le Pere tout puiſſant ; n'y a-t-il que le Pere qui ſoit Dieu ?
Le Pere eſt Dieu, le Fils auſſi,
Le Saint Eſprit eſt Dieu de même,
L'Egliſe nous l'enſeigne ainſi ;
Dire autrement c'eſt un blaſphême,
Puiſque la ſainte Trinité
A la même Divinité.
Qu'eſt-ce que la Trinité ?
L'inconcevable Trinité
Eſt un Dieu ſeul en trois perſonnes,
Egales en autorité,
Egalement ſages, & bonnes,
Toutes trois n'ont qu'un ſeul pouvoir,
Qu'une ſageſſe, & qu'un vouloir.

Figure 4.11: Colour bleed-through restoration results on a sample full document image from outside the database, **M04**. Top row: degraded recto (left) and verso images. Bottom row: corresponding restored recto (left) and verso images.



Politiques. 129

euſt poſſedez tous ſolidairement & de ſon chef, auec tel reſpect neantmoins enuers l'Empereur & l'aiſné de la branche, qu'il a ſemblé touſiours auſſi bien Souuerain de Grets & d'Iſpruch, qui eſtoient les partages des autres, comme de l'Auſtriche meſme & de la Boheme: Ayans par tous leurs Eſtats vne meſme forme de gouuernement, meſmes conſeils, meſmes intereſts & raiſon d'Eſtat en conſideration : Mais ſur tout, vn grand & ſouuerain reſpect enuers les Roys d'Eſpagne, qu'ils obſeruent encore auiourd'huy, & d'autant plus exactement, qu'ils en ont affaire pour ſ'y maintenir, repreſentant par ceſte bonne intelligence l'ancien Geryon d'Eſpagne, vny, ſerré & ramaſſé en ſoy : Mais qui pourroit deuoir ſa cheute à vn Hercule Gaulois, ſ'il venoit à l'entreprendre à bon eſcient.

Leur bonne intelligence auec le Roy d'Eſpagne.

IX.
De leurs Eſtats.
Troublez par la diuerſité & bigarrure de Religion.

Les Eſtats de ces Princes ſont grands, eſtendus autour de l'Allemagne, & confinans la pluſpart au Turc, ſujet pour eux de grande deſpenſe : Bigarrez de Religion, pour le regard de laquelle ils ne ſont pas les maiſtres abſolus pour la faire

I

Politiques. 129

euſt poſſedez tous ſolidairement & de ſon chef, auec tel reſpect neantmoins enuers l'Empereur & l'aiſné de la branche, qu'il a ſemblé touſiours auſſi bien Souuerain de Grets & d'Iſpruch, qui eſtoient les partages des autres, comme de l'Auſtriche meſme & de la Boheme: Ayans par tous leurs Eſtats vne meſme forme de gouuernement, meſmes conſeils, meſmes intereſts & raiſon d'Eſtat en conſideration : Mais ſur tout, vn grand & ſouuerain reſpect enuers les Roys d'Eſpagne, qu'ils obſeruent encore auiourd'huy, & d'autant plus exactement, qu'ils en ont affaire pour ſ'y maintenir, repreſentant par ceſte bonne intelligence l'ancien Geryon d'Eſpagne, vny, ſerré & ramaſſé en ſoy : Mais qui pourroit deuoir ſa cheute à vn Hercule Gaulois, ſ'il venoit à l'entreprendre à bon eſcient.

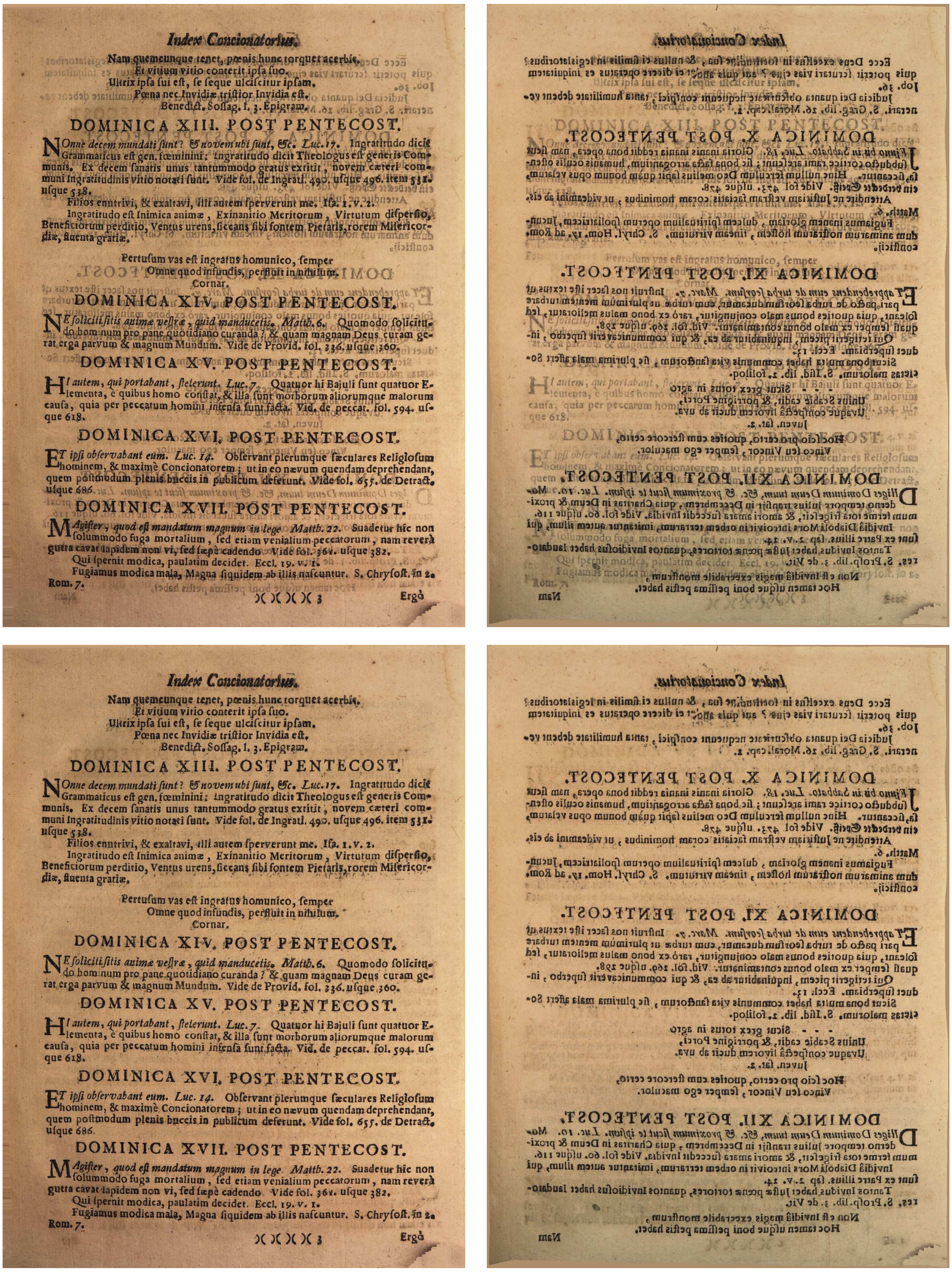
Leur bonne intelligence auec le Roy d'Eſpagne.

IX.
De leurs Eſtats.
Troublez par la diuerſité & bigarrure de Religion.

Les Eſtats de ces Princes ſont grands, eſtendus autour de l'Allemagne, & confinans la pluſpart au Turc, ſujet pour eux de grande deſpenſe : Bigarrez de Religion, pour le regard de laquelle ils ne ſont pas les maiſtres abſolus pour la faire

I

Figure 4.12: Colour bleed-through restoration results on a sample full document image from outside the database, **M05**. This example contains 'press-on' from the pages opposite the recto and verso sides, which is not removed in the restored result. Top row: degraded recto (left) and verso images. Bottom row: corresponding restored recto (left) and verso images.



Index Concionatorius.

Nam quemcunque tenet, pœnis hunc torquet acerbis.
Et vitium vitio conterit ipſa ſuo.
Ultrix ipſa ſui eſt, ſe ſeque ulciſcitur ipſam.
Pœna nec Invidiæ triſtior Invidia eſt.
Benedict. Soſſag. l. 3. Epigram.

DOMINICA XIII. POST PENTECOST.

N*Onne decem mundati ſunt? & novem ubi ſunt, &c. Luc. 17.* Ingratitudo dicit Grammaticus eſt gen. fœminini: Ingratitudo dicit Theologus eſt generis Communis. Ex decem ſanatis unus tantummodo gratus extitit, novem cæteri communi Ingratitudinis vitio notati ſunt. Vide fol. de Ingrati. 490. uſque 496. item 531. uſque 538.

Filios enutrivi, & exaltavi, illi autem ſperverunt me. Iſa. 1. v. 2.

Ingratitudo eſt Inimica animæ, Exinanitio Meritorum, Virtutum diſperſio, Beneficiorum perditio, Ventus urens, ſiccans ſibi fontem Pietatis, rorem Miſericordiæ, fluenta gratiæ.

Pertuſum vas eſt ingratus homunico, ſemper
Omne quod infundis, perfluit in nihilum.
Cornar.

DOMINICA XIV. POST PENTECOST.

N*E ſoliciti ſitis animæ veſtræ, quid manducetis. Matth. 6.* Quomodo ſolicitudo hominum pro pane quotidiano curanda? & quam magnam Deus curam gerat erga parvum & magnum Mundum. Vide de Provid. fol. 336. uſque 360.

DOMINICA XV. POST PENTECOST.

H*I autem, qui portabant, ſteterunt. Luc. 7.* Quatuor hi Bajuli ſunt quatuor Elementa, è quibus homo conſtat, & illa ſunt morborum aliorumque malorum cauſa, quia per peccatum homini infenſa ſunt facta. Vid. de peccat. fol. 594. uſque 618.

DOMINICA XVI. POST PENTECOST.

E*T ipſi obſervabant eum. Luc. 14.* Obſervant plerumque ſæculares Religioſum hominem, & maximè Concionatorem; ut in eo nævum quendam deprehendant, quem poſtmodum plenis buccis in publicum deferunt. Vide fol. 655. de Detract. uſque 686.

DOMINICA XVII. POST PENTECOST.

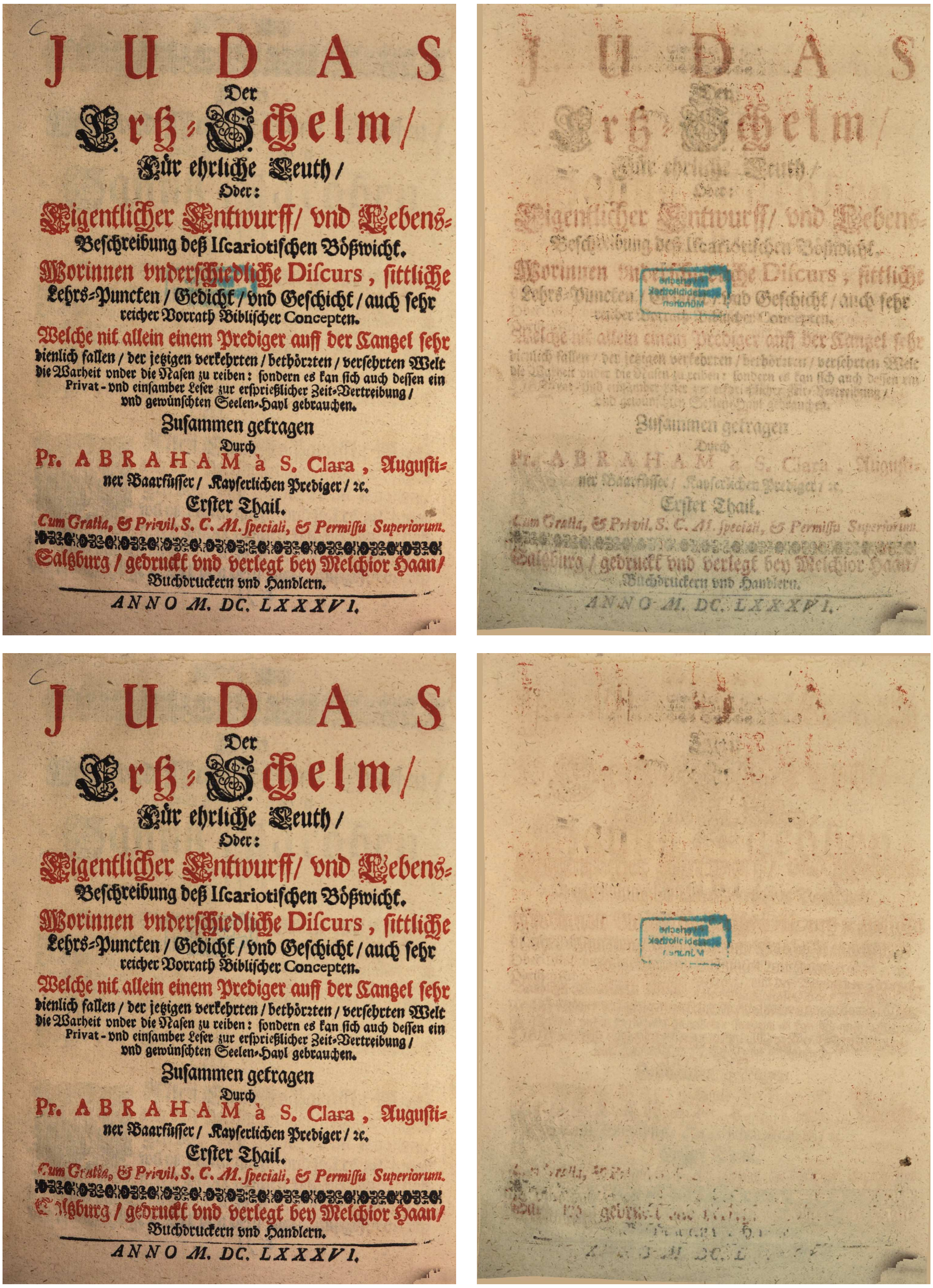
M*Agiſter, quod eſt mandatum magnum in lege. Matth. 22.* Suadetur hic non ſolummodo fuga mortalium, ſed etiam venialium peccatorum, nam reverà gutta cavat lapidem non vi, ſed ſæpè cadendo. Vide fol. 361. uſque 382.

Qui ſpernit modica, paulatim decidet. Eccl. 19. v. 1.

Fugiamus modica mala, Magna ſiquidem ab illis naſcuntur. S. Chryſoſt. in 2. Rom. 7.

)( )( )( )(3 Ergò

Figure 4.13: Colour bleed-through restoration results on a sample full document image from outside the database, **M06**. This example contains ‘press-on’ from the pages opposite the recto and verso sides, which is not removed in the restored result. Top row: degraded recto (left) and verso images. Bottom row: corresponding restored recto (left) and verso images.



JUDAS
Der
Ertz-Schelm/
Für ehrliche Leuth/
Oder:
Eigentlicher Entwurff/ vnd Lebens-Beſchreibung deß Iſcariotiſchen Bößwicht.
Worinnen vnderſchiedliche Diſcurs, ſittliche Lehrs-Puncten/ Gedicht/ vnd Geſchicht/ auch ſehr reicher Vorrath Bibliſcher Concepten.
Welche nit allein einem Prediger auff der Cantzel ſehr dienlich fallen/ der jetzigen verkehrten/ bethörrten/ verſehrten Welt die Warheit vnder die Naſen zu reiben: ſondern es kan ſich auch deſſen ein Privat- vnd einſamber Leſer zur erſprießlicher Zeit-Vertreibung/ vnd gewünſchten Seelen-Hayl gebrauchen.
Zuſammen getragen
Durch
Pr. ABRAHAM à S. Clara, Auguſtiner Baarfüſſer/ Kayſerlichen Prediger/ ꝛc.
Erſter Thail.
Cum Gratia, & Privil. S. C. M. ſpeciali, & Permiſſu Superiorum.
Saltzburg/ gedruckt vnd verlegt bey Melchior Haan/ Buchdruckern vnd Handlern.
ANNO M. DC. LXXXVI.

Figure 4.14: Colour bleed-through restoration results on a sample full document image from outside the database, **M07**. In this example, the bleed-through of the stamp from the verso to the recto is successfully removed. Top row: degraded recto (left) and verso images. Bottom row: corresponding restored recto (left) and verso images.

4.4 Model Analysis

In this section, a comparative analysis of variations in the histogram segmentation stage, presented in Section 4.2.1, is performed. Since in previous works performance of bleed-through removal is generally evaluated in terms of binary or pseudo-binary results, a set of evaluation metrics specific to this 4-class classification approach is proposed and examined for each variation. For a more complete discussion on previous numerical evaluation approaches, and a comparison of the proposed bleed-through removal approaches with other methods see Chapter 5.

There are two main classification error types associated with bleed-through removal that affect the legibility of the document. Firstly true foreground information may misclassified and removed from the image (referred to here as b_1), and secondly bleed-through may be misclassified and left in the image (b_2). Of the two, b_1 errors will affect the integrity of the document more seriously as the restored result will be reduced in legibility compared with the *original* degraded image, whereas b_2 errors result in reduced legibility compared with the *optimal* result. There are two further error types, b_3 and b_4, associated with this particular approach, as can be seen in the four label confusion matrix in Table 4.3. b_3 errors represent foreground information being left in the background texture source, thus potentially introducing artefacts into the restored result, and b_4 errors represent true background regions being replaced with synthesised background texture - see Section 4.3 for further information. However, as b_3 and b_4 errors are not concerned explicitly with legibility they are not examined in detail for model evaluation purposes.

These errors alone do not give a clear indication of the classification performance with respect to each label individually. Therefore per-label F-scores are also examined:

$$F_\beta = \frac{(1+\beta^2)\cdot \text{precision} \cdot \text{recall}}{(\beta^2 \cdot \text{precision}) + \text{recall}}, \tag{4.3}$$

where precision $= \dfrac{TP}{TP+FP}$, recall $= \dfrac{TP}{TP+FN}$, and TP, FP, FN are the True Positive, False Positive, and False Negative values respectively. The standard F_1-score is used to evaluate the classification performance for *fgbl* and *blfg*, but as discussed in Section 4.2.1, it is preferable to

Table 4.3: Confusion matrix for the 4 class joint histogram labelling, highlighting the different error types.

		Ground Truth			
		bgbg	*fgbl*	*blfg*	*fgfg*
Predicted	*bgbg*	TP	b_2/b_3	b_2/b_3	b_3
	fgbl	b_4	TP	b_1/b_2	b_1
	blfg	b_4	b_1/b_2	TP	b_1
	fgfg	b_4	b_2	b_2	TP

Table 4.4: Comparison of average errors from different initialisation variations.

		bgbg $F_{0.5}$	*fgbl* F_1	*blfg* F_1	*fgfg* F_2	$F_{1\mathrm{M}}$	b_1	b_2
Histogram Labelling	No update	69.13	82.87	83.47	77.03	80.61	3.87	15.20
	K-means	76.08	76.82	75.46	73.26	75.64	4.56	19.89
	M	86.62	**87.54**	**88.12**	**79.95**	**84.63**	3.24	**9.09**
	E+M	**86.67**	86.77	87.58	79.88	83.92	**2.99**	10.06
Refined Image Labelling	No update	91.52	80.15	79.45	61.46	78.15	2.24	7.77
	K-means	93.13	78.73	78.28	60.00	79.94	2.29	7.85
	M	93.01	84.36	81.71	62.88	83.32	2.27	**6.35**
	E+M	**93.02**	**84.44**	**81.90**	**64.00**	**83.51**	**2.17**	**6.35**

avoid over-classification of *bgbg* and under-classification of *fgfg* so the $F_{0.5}$-score is used for *bgbg* to weight precision higher than recall, and the F_2-score is used for *fgfg* to weight recall higher than precision. An overall per image classification score is also evaluated using the macro-average of F_1-scores for each label:

$$F_{1\mathrm{M}} = \frac{2 \cdot \mathrm{precision}_{\mathrm{M}} \cdot \mathrm{recall}_{\mathrm{M}}}{\mathrm{precision}_{\mathrm{M}} + \mathrm{recall}_{\mathrm{M}}}, \tag{4.4}$$

where

$$\mathrm{precision}_{\mathrm{M}} = \frac{\sum_{i=1}^{4} \frac{TP_i}{TP_i + FP_i}}{4}, \qquad \mathrm{recall}_{\mathrm{M}} = \frac{\sum_{i=1}^{4} \frac{TP_i}{TP_i + FN_i}}{4}. \tag{4.5}$$

The macro-average is used as it gives equal weight to each class rather than micro-average where the performance of large classes tend to dominate the score [41, 65].

Initialisation Comparison

Four variations of the initial cluster and centres update described in Section 4.2.1 were tested to obtain the best input to the optimisation stages: 'No update' - the centres and clusters were left as the original estimates; 'K-means' - the initial centre estimates were used as centroid starting locations for K-means clustering; 'M' - the centres and clusters were updated once using the minimum Mahalanobis distance to the initial centres for all labels, and the updated centres set as the means of the updated clusters; 'E+M' - the centres and clusters were updated once as described in Section 4.2.1 firstly using the minimum Mahalanobis distance to the initial centres to obtain the *bgbg* label cluster, and then using the minimum Euclidean distance to the 3 initial centres for the remaining labels, to obtain the corresponding clusters, and the updated centres again set as the mean of the updated clusters. Table 4.4 shows average performance of the four variations over all the measures described, evaluated compared to minimum error histogram labellings and ground truth image labellings. It is clear that the 'E+M' approach performs best overall, and especially in terms of *fgfg* F_2 and b_1 results, where the improvement over the other variations is relatively large.

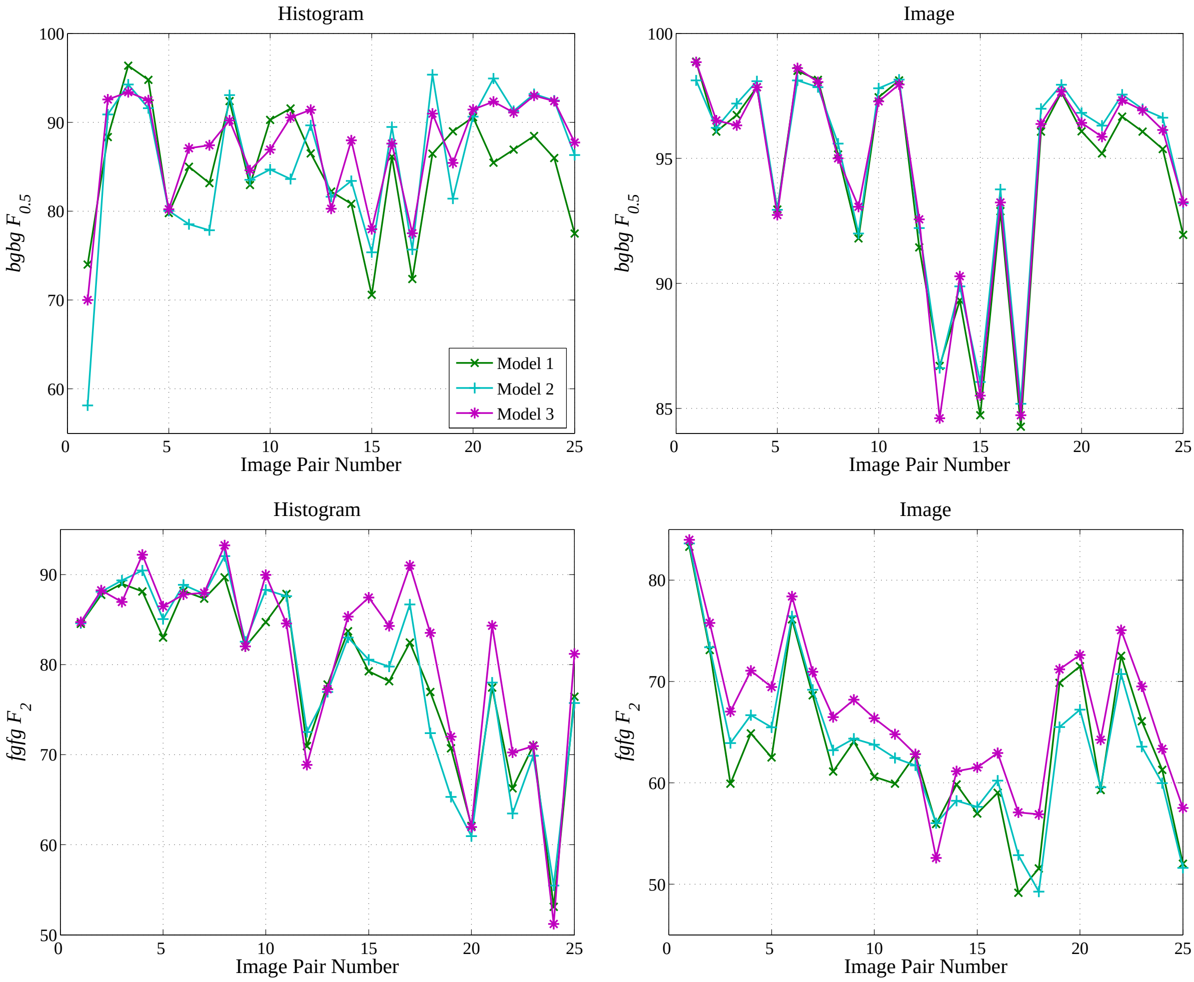


Figure 4.15: *bgbg* $F_{0.5}$ scores (top), and *fgfg* F_2 scores (bottom) for the three model variations on both histogram (left) and corresponding image (right) labellings.

Model Comparison

To evaluate the performance of the three model variations described in Section 4.2.1, again both the histogram labelling and final image labelling results are examined. However, as can be seen in Table 4.4, the histogram scores do not necessarily correlate with the image scores. This would indicate that the minimum error histogram labelling generated using the ground truth image labelling is not necessarily the optimal input to the image labelling refinement stage. Therefore while both sets of results are examined, the image labelling results are used to define the overall performance of each model.

Figure 4.15 shows *bgbg* $F_{0.5}$ scores (top), and *fgfg* F_2 scores (bottom) for the three model variations on both histogram (left) and corresponding image (right) labellings. It is clear that the performance varies widely depending on the input image pair, and that for *bgbg* $F_{0.5}$ scores, each

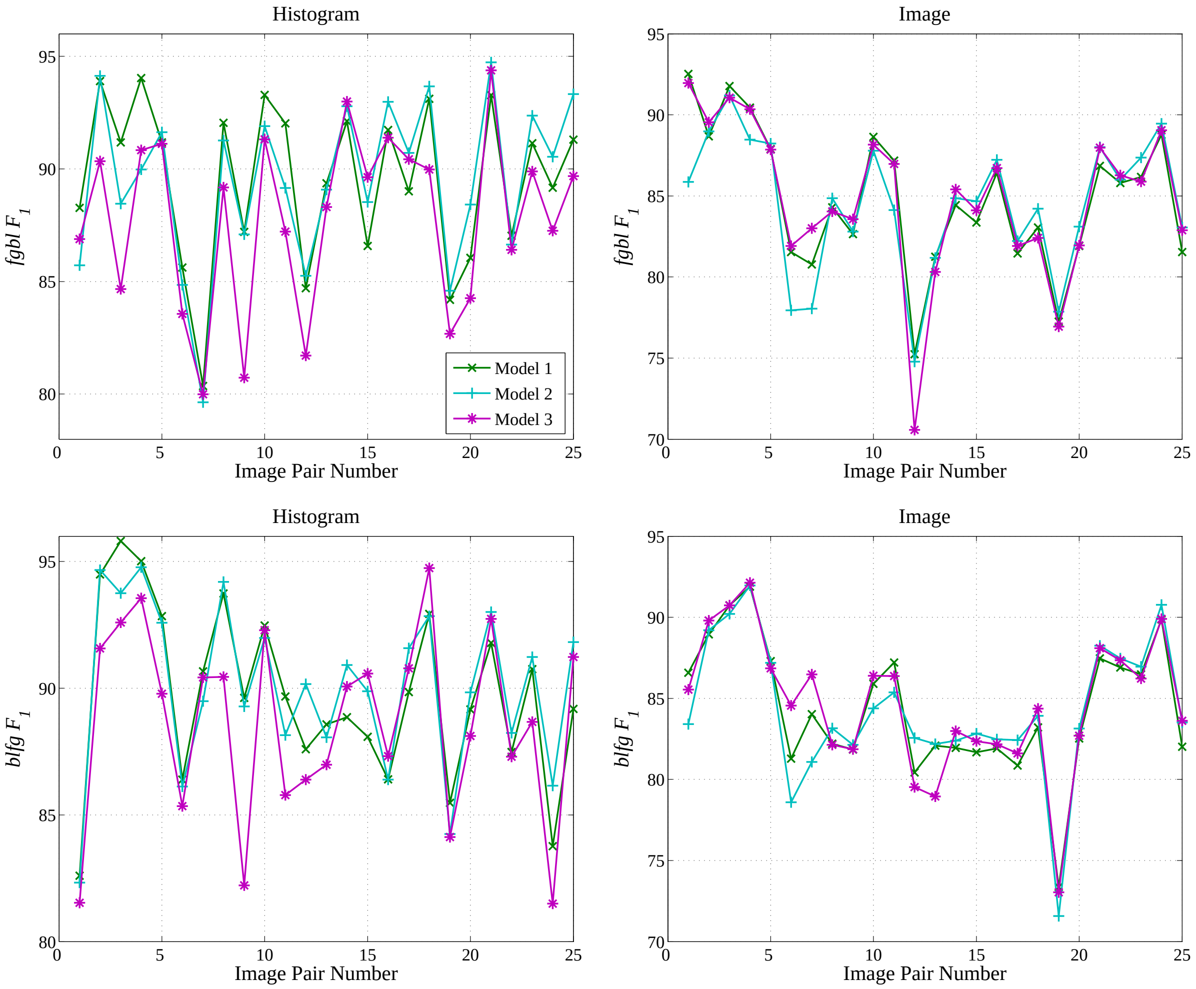


Figure 4.16: *fgbl* (top), and *blfg* (bottom) F_1 scores for the three model variations on both histogram (left) and corresponding image (right) labellings.

method performs better on different images. For *fgfg* F_2 scores Model 3 performs significantly better overall for both the histogram and image labellings.

Figure 4.16 shows the *fgbl* (top), and *blfg* (bottom) F_1 scores for the three model variations on both histogram (left) and corresponding image (right) labellings. Again, the performance of each model varies widely depending on the input image pair. Model 3 performs worse over both labels in the histogram labelling, this is likely related to the high performance in the *fgfg* histogram results, as the overlap regions are segmented in favour of *fgfg* rather than *fgbl* or *blfg*. However this poor performance is mitigated in most of the image labelling results, where the performance of each model is similar overall.

Figure 4.17 shows $F_{1\mathrm{M}}$ scores (top) and b_1 vs b_2 error distributions for the three model variations, again on both histogram (left) and image (right) labellings. As macro averaged F_1

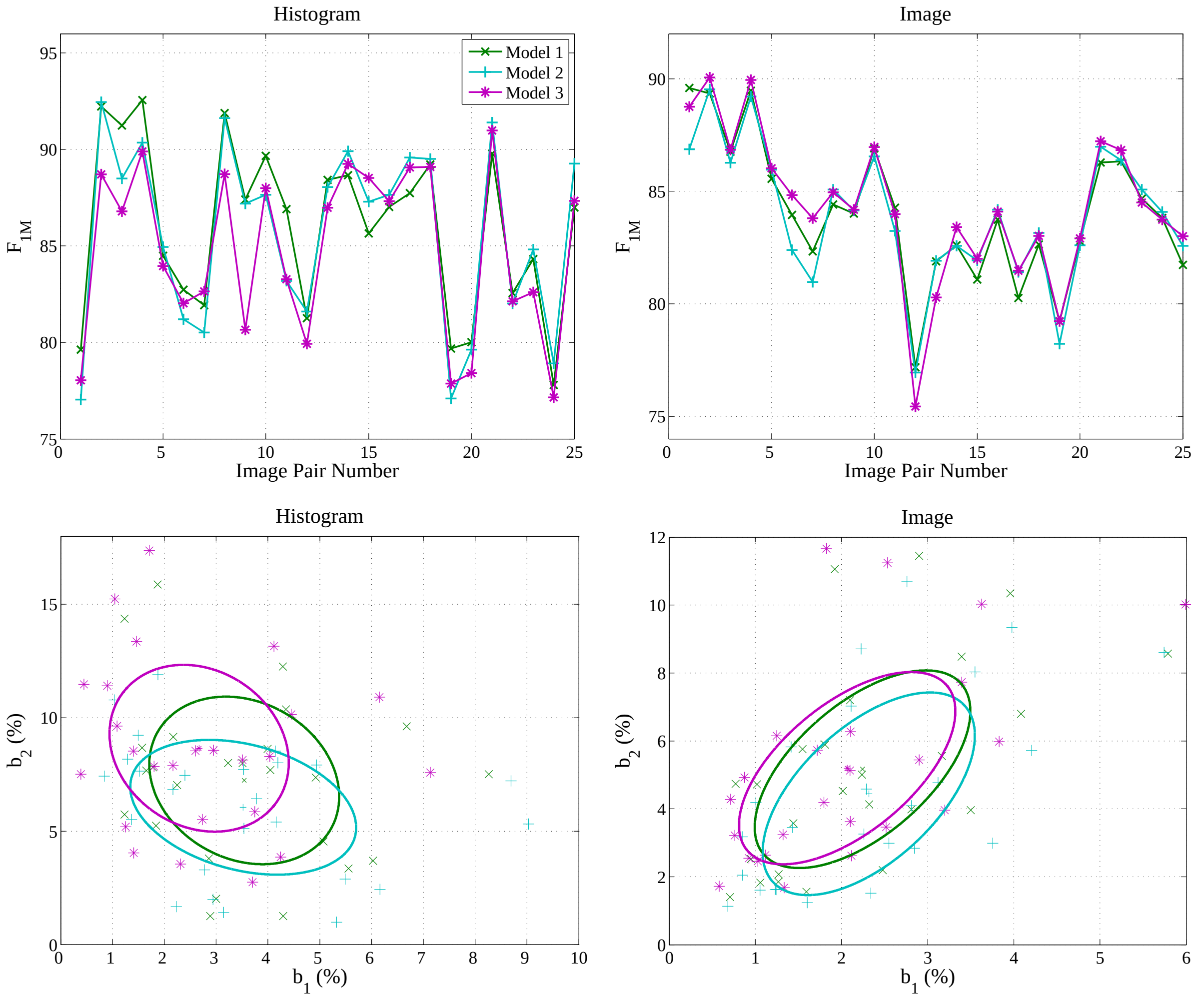
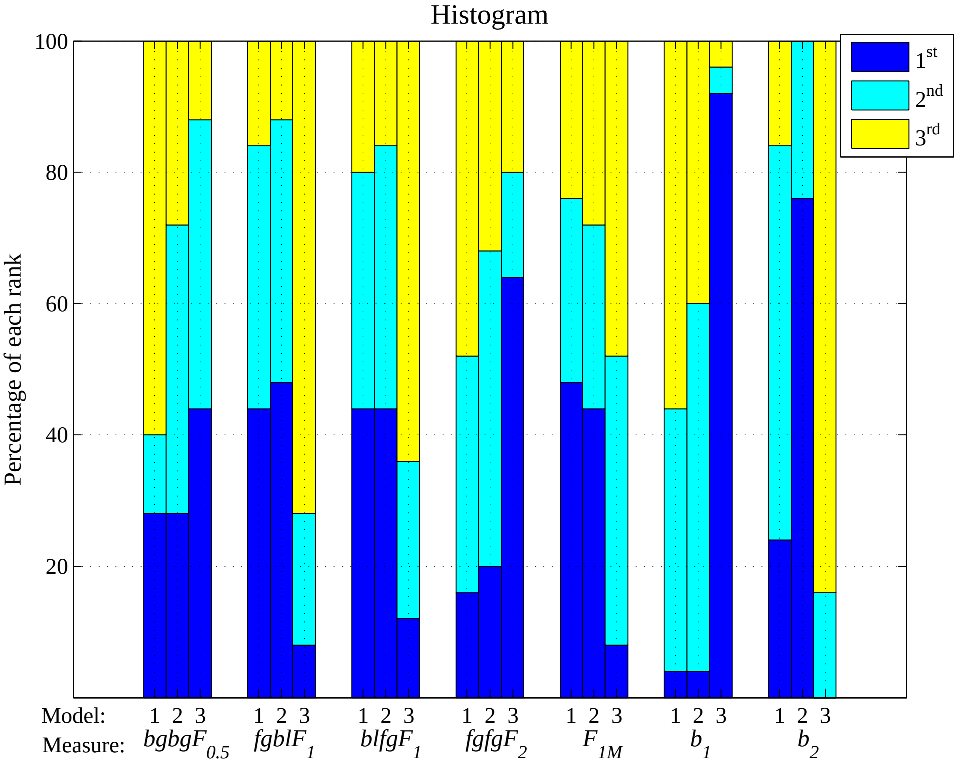


Figure 4.17: $F_{1\mathrm{M}}$ scores (top) and b_1 vs b_2 error distributions for the three model variations on both histogram (left) and image (right) labellings.

scores gives equal weight to each label, even though the *fgfg* may represent a small proportion of the image, classification performance on this label is given equal weight, therefore the measure favours conservative classification. The b_1 vs b_2 distributions highlight the performance of each model with respect to foreground preservation and bleed-through removal explicitly - the target for perfect restoration is $[0, 0]$. Both the histogram and image results indicate that Model 3 performs best in terms of foreground preservation, but Model 2 performs best in terms of bleed-through removal.

As the results vary widely per image pair, the performance of each model is further evaluated by ranking the results measures for each image pair. Figure 4.18 shows the percentage of each rank obtained for each error measure in both the histogram (top) and image (bottom) domains. The ranking results correlate with observations already made - for the image labelling, Model 3



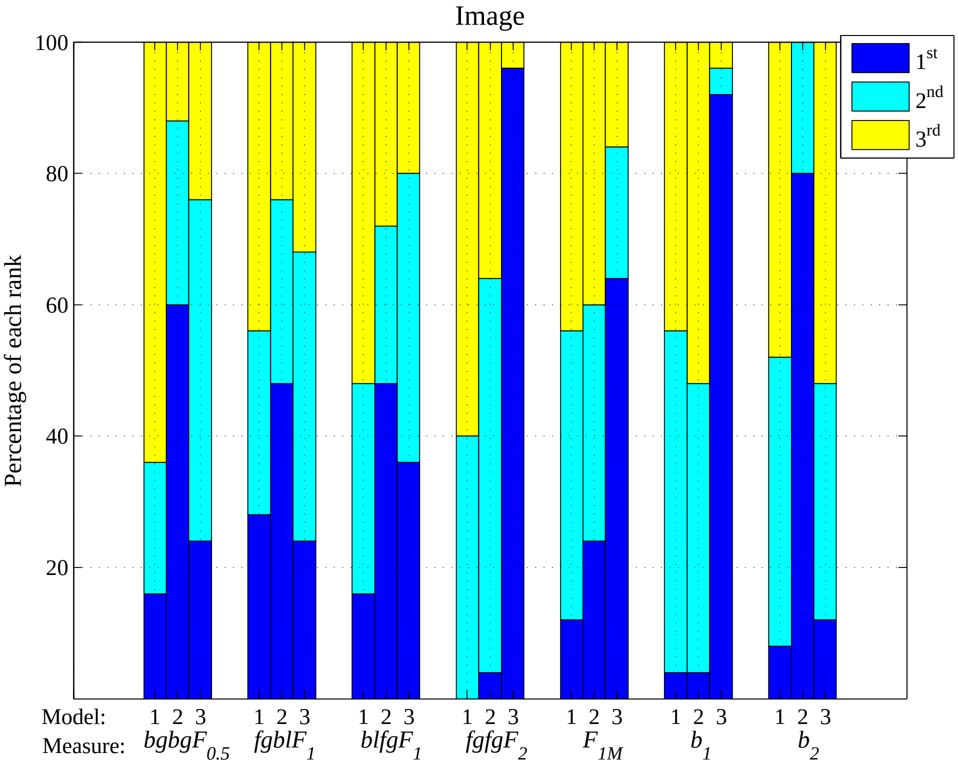


Figure 4.18: Illustrations of the ranking of the performance of each model for different error measures. Top: errors in histogram labelling, bottom: errors in final image labelling.

Table 4.5: Comparison of average errors from different model variations.

		$\beta_{\mathbf{i}}$	$\gamma_{\mathbf{i}}$	*bgbg* $F_{0.5}$	*fgbl* F_1	*blfg* F_1	*fgfg* F_2	$F_{1\mathrm{M}}$	F-Rank	b_1	b_2	b-Rank
Histogram Labelling	1	$h(\mathbf{i})$	1	85.10	**89.51**	89.73	79.14	**85.99**	247	3.54	7.24	111
	2	1	$\frac{1}{h(\mathbf{i})}$	85.48	89.50	**90.07**	79.42	85.66	**233**	3.52	**6.05**	**90**
	3	1	1	**87.30**	87.79	88.64	**81.35**	84.78	270	**2.67**	8.65	99
Refined Image Labelling	1	$h(\mathbf{i})$	1	94.24	84.79	84.34	63.28	84.03	298	2.24	5.17	120
	2	1	$\frac{1}{h(\mathbf{i})}$	**94.73**	84.49	84.28	63.47	83.93	239	2.32	**4.44**	92
	3	1	1	94.53	**84.83**	**84.63**	**66.84**	**84.30**	**213**	**2.07**	5.20	**88**

is ranked 1^{st} in 94% of the image pairs for *fgfg* F_2, and 92% for b_1, and Model 2 is ranked 1^{st} in 80% of the image pairs for b_2, and 60% for *bgbg* $F_{0.5}$.

Finally Table 4.5 shows the average error measure values for each model, and also the total accumulated rankings for each of the F-scores (F-rank), and b errors (b-rank). The F-scores and b errors are ranked separately, as they are essentially different representations of the same measures. The average F-score image results are very similar for each model, except in the case of the *fgfg* label, where Model 3 has a significantly higher score. The average b-error values again emphasise the foreground preservation strength of Model 3 and bleed-through removal performance of Model 2. The total accumulated F-Score ranks and b-error ranks indicate that, for the selected smoothness weights, Model 3 performs best overall. However, as mentioned previously, both the $F_{1\mathrm{M}}$-score and *fgfg* F_2-score favour conservative results. Therefore the F-Score rank, whilst representing a measure of balanced label performance, does not necessarily indicate the best Model in terms of visual results. In addition, as the b-rank difference between Models 2 and 3 is minimal, and as these results are based on an empirically fixed smoothness weight for each model for all image pairs, they are not definitive and further analysis is necessary. Finally, as can be seen in Figure 4.6, the performance of all the model variations on some of the test cases is highly sensitive to the smoothness weight chosen and this would indicate that a better performance overall may be achieved by selecting a separate weight for each image pair. A better choice of pairwise energy terms could also reduce the sensitivity of the model results to the smoothness weight selected.

4.5 Summary

This chapter has presented a non-blind non-parametric framework for bleed-through removal involving three stages: preprocessing, classification, and restoration. In the preprocessing stage any intensity trends in the input images are removed using local offsets derived from the local intensity histograms. In the classification stage the observed degraded images are treated jointly as a single two dimensional image. Classification then proceeds in two steps, firstly by performing a joint intensity histogram segmentation, then by applying rules to label connected components in the corresponding image labelling. Three model variations are examined for the histogram segmentation. Finally, in the restoration stage, clean background images for both recto and verso sides are generated using exemplar-based image inpainting. A blend between the clean background images and original degraded images is then performed in identified bleed-through regions to create the final restored results. A numerical comparison approach for evaluating the performance of the model variations is then presented, and results indicate that the three variations perform similarly, with the unweighted variation performing slightly better overall, especially in terms of foreground preservation. However, the performance of this approach is sensitive to the choice of smoothness weight in some image-pair examples; an individual weight for each image pair may improve the output significantly. Further analysis is presented in Chapter 5, where visual comparisons with the approach from Chapter 3 show that all three proposed model variations perform significantly better on a wide range of degradations, whilst maintaining the document character.

5

Evaluation

There are two main challenges that face researchers working in the area of bleed-through restoration. Firstly, it can be difficult to obtain access to high resolution and quality degraded image examples unless connected with a specific library or digitisation project. Even when access to images is available, it is often very difficult to publish results based on those images due to copyright issues. Secondly, as for all document restoration techniques, problems arise when trying to analyse results quantitatively, as there is no actual ground truth available. This problem may be overcome either by creating synthetic degraded images with known ground truth, [22, 35], or by creating synthetic ground truth data for given real degraded images, [15]. Alternatively, performance may be evaluated by quantifying how the restoration affects a secondary step, such as the performance of an Optical Character Recognition (OCR) system on the document image, [73, 78] (in which case ground truth information about the document content is required), or visually counting bleed-through errors at a word level, [29, 35, 67, 76]. A further issue with quantitative evaluations for performance comparison is the different formats of the results of different methods, as described in Chapter 2; providing a fair comparison between binary or pseudo-binary and textured results is a challenging problem. It is proposed that a fair quantitative comparison between methods can only be achieved if they are converted to the same format then compared to a ground truth that is also of the same format, and that the simplest way of achieving this is to binarise all the results and compare them to a binary ground truth.

The first part of this chapter describes the details of a small bleed-through image database with manually generated foreground text masks, made available for research purposes. Since converting manuscript images to a suitable format for use in testing bleed-through restoration algorithms can be a time consuming process, it is hoped that this database, where all necessary processing has been performed, will prove to be a useful resource.

The second part of the chapter describes the proposed evaluation approach, and compares the performance of the bleed-through removal algorithms presented in Chapters 3 and 4 with three non-blind removal approaches[1].

5.1 Bleed-Through Database

A manuscript bleed-through document image database has been compiled, consisting of 25 registered recto/verso image pairs, taken as crops from larger manuscript images with varied degrees of bleed-through degradation. The manuscripts selected range in date from 15^{th} to 19^{th} Century, and were either written on vellum or paper. The average crop size is roughly 570×2390 pixels. All images contained in the database are taken from the collections of the Irish Script On Screen Project (ISOS) [6]. ISOS is a project of the School of Celtic Studies, Dublin Institute for Advanced Studies, Dublin, Ireland, and is funded by the Dublin Institute for Advanced Studies [3]. The object of ISOS is to create digital images of manuscripts written in Irish, and to make these images accessible as an electronic resource for researchers. There are no such bleed-through datasets with ground truth freely available for research purposes in the public domain at the time of writing.

5.1.1 Image Capture

Image capture was performed by the individual libraries hosting the documents. Each manuscript page was scanned at 600dpi, and also photographed. The images used for the database were the photographs, taken using a 5×4 format viewing camera with a Phase One P45 digital back. Both camera and manuscript were positioned on a specially adapted book-cradle. Each image was then processed in Adobe Photoshop to crop to an optimum canvas size and superimpose a text header and footer to distinguish individual pages. A ruler was also placed alongside each image to indicate scale. Digital enhancement was not performed at this stage.

5.1.2 Crop Details

Some pre-processing was necessary to crop out binding, ruler markers, and digital labels that could influence the performance of intensity based algorithms. Also, as the full high resolution manuscript images are very large in size, it is not practical to use them for testing - smaller sections are preferable. For the database, crops were taken from the larger images such that they would contain a sentence or phrase of text on both the recto and verso sides. The reason for this was to allow for the possibility of restoration evaluation using legibility improvement as a metric. All the images were converted to grayscale and saved in tif format. File names

---

[1] Parts of this chapter are based on the work published in [59, 60].

in the database follow the format 'lib'.'MS'.'fol'.tif. 'lib' represents the library from which the manuscript contained in the image originates and can be one of eight labels:

(i) AC - The Allan and Maria Myers Academic Centre, University of Melbourne, Australia.

(ii) FH - The Benjamin Iveagh Library, Farmleigh House, Ireland.

(iii) NLI - The National Library of Ireland.

(iv) NUIG - The James Hardiman Library, National University of Ireland, Galway.

(v) NUIM - The Russell Library, National University of Ireland, Maynooth.

(vi) RIA - The Royal Irish Academy Library.

(vii) TCD - Trinity College Dublin Library.

(viii) UCD - University College Dublin Library.

'MS' refers to the manuscript number, and 'fol' refers to either the page number, or the folio number followed by 'r' or 'v' to denote the recto or verso side. For example, 'RIA.MSCiii3.301v.tif' is the crop of the verso side of folio 301 taken from Manuscript C iii 3 in the collections of the Royal Irish Academy Library. The ground truth images are labelled as for the degraded images, but with appended 'gt.tif' to differentiate between the two.

5.1.3 Registration

As mentioned in Chapter 2, to perform non-blind bleed-through restoration the recto and verso sides must first be registered so that the bleed-through interference on each side is aligned with its corresponding originating text from the opposite side. The registration method used for the database involved three stages. Firstly a set of corresponding control points on both recto and verso images were manually selected. These points indicated locations of the same textual features on each side. A global affine warp model for the verso side of each image pair was then derived from a least squares fit to the displacements between these locations. Secondly, this affine model was used as an initialisation to the affine warp based bleed-through registration method of Dubois et al. [21]. Finally, local adjustments were made to the registration manually, using the 'gridwarp' function in NUKE[2], that defines a Bezier grid over the source image and allows the user to reposition the corners of squares in the grid, applying a nonrigid warping to the local image region correspondingly. Some difficulties were encountered in registration of images where the crops contained text close to the manuscript binding. In these regions the page deformation is highly nonlinear and an affine model is unsuitable. This problem was overcome by using manual registration only, with a very fine grid over the whole image.

---

[2] The Foundry's node-based compositor, `http://www.thefoundry.co.uk/products/nuke`

5.1.4 Ground Truth Creation

The ground truth foreground images were created manually, by drawing around the outline of foreground text on both recto and registered verso sides. These outline layers were then extracted from the images and filled in to create binary foreground text images, with black representing text, and white representing background. In handwritten documents, the edges of characters can often be blurred or gradually fade into the background due to ink absorption by the medium, or due to the angle and pressure of the writing instrument. This makes marking the precise location of the boundary between text and background a very subjective decision. For all the images it was decided that the edge of characters would be defined where the last traces of ink were visible when viewed in detail, as it was considered preferable to preserve as much of the foreground text shape as possible. This results in foreground mask text that often appears larger than the image text, however this is taken into account in the bleed-through removal performance evaluation proposed in Section 5.2.

5.1.5 Access

The database has been made freely available online, and is hosted by ISOS at

`http://www.isos.dias.ie/libraries/Sigmedia/english/index.html`

along with a visual comparison of bleed-through removal results. However, as the copyright on the image crops in the database rests with the individual institutions holding the original documents, it is necessary to obtain a username and password from ISOS in order to download the database.

5.2 Quantitative Analysis

As mentioned in the introduction to this chapter, it is difficult to perform objective evaluation of bleed-through performance in the absence of ground truth, and to perform comparison between results of different methods that have different output formats. The quantitative analysis approach used here is to binarise all the results and then compare the binary images with the manually generated ground truth masks.

5.2.1 Binarisation Algorithm

The binarisation method selected to obtain the results for comparison with the ground truth is the adaptive degraded document image binarisation algorithm of Gatos et al. [24]. This method involves several distinct stages. Firstly the image is preprocessed using a low-pass Wiener filter to reduce noise, smooth the background texture, and enhance the contrast between foreground and background regions. Then a rough estimate of the foreground is obtained using the adaptive thresholding method of Sauvola and Pietikäinen [63], and from this a background estimate is generated using interpolation of neighbouring identified background intensities. The image is then thresholded by comparing the distance between the estimated background with the original image, and a threshold image defined based on the background intensity. Finally a post-processing step involving successive application of shrink and swell filtering is performed in order to improve the quality of text regions and preserve stroke connectivity.

The use of binarisation on the clean image results is likely to favour the pseudo-binary and binary over the textured results, especially in images where there are many faint text features, the main text body is close in intensity to the background, or where the background medium is highly textured. However, the binarisation output to a certain extent mirrors the impact of different output formats on the visual appearance of the results. Faint foreground features that are successfully preserved are more ‘noticeable’ visually when contrasted against a plain background rather than textured, and the binarisation algorithm is more likely to classify them correctly as foreground. Conversely, bleed-through artefacts left in error are also more noticeable visually in contrast to a plain background, and are less likely to be missed by the binarisation stage than if they are left against a textured background. However, in order to obtain a fair comparison, the binarisation technique used here was selected in order to mitigate partially the potential advantage to the outputs of certain methods, as it is designed to be robust to varying background intensities and therefore is capable of obtaining good estimates of foreground text from a noisy background.

5.2.2 Bleed-Through Error Metrics

As discussed in Chapter 4, there are two errors associated with bleed-through removal that affect the legibility of the restored results - removing true foreground information, and failing to remove bleed-through. However, as the comparison in this case is between binary per image label fields and not 4-class joint label fields, the bleed-through error cannot be examined explicitly and the results for recto and verso sides are evaluated individually. Three per pixel probability error measures are used to compare the performance of each method: *FgError*, the probability that a pixel in the foreground text was classified as background, or false negative rate, *BgError*, the probability that either a background or bleed-through pixel was classified as foreground - false positive rate, and *TotError* the probability that any pixel in the image was misclassified - total error rate. These are defined for each image output as

$$\mathit{FgError} = \frac{FN}{FN + TP} \tag{5.1}$$

$$\mathit{BgError} = \frac{FP}{FP + TN} \tag{5.2}$$

$$\mathit{TotError} = \frac{FP + FN}{TP + TN + FP + FN}, \tag{5.3}$$

where TP, FP, TN, FN are the True Positive, False Positive, True Negative, and False Negative values respectively. In order to account for the ambiguity of the character edges in the ground truth, pixels on the inner and outer edges of characters are removed from the error calculations using morphological dilation and erosion.

Comparisons of error values between different image examples can be misleading, as these values depend on the relative size of the background and foreground regions in each image with respect to the text character size. For example an image that is mostly background, with small text, is likely to have a much smaller *BgError* value than an image with large text characters covering most of the image and proportionally less background. Therefore, while examining average error metrics per image can give a general indication of relative performance, it is more useful to rank the performance of each method on each image separately, via the three error metrics (least probability of error to greatest). Ranked Pairs Voting (RP) [68] is used to obtain these ranks as this approach is based on pairwise comparisons between individual method performances.

5.2.3 DIBCO Evaluation Metrics

In addition to the proposed metrics, standard binarisation evaluation metrics are also used to evaluate the binarised results of each method tested. The DIBCO series [23, 52–54] is an annual competition for comparing recent advances in document image binarisation algorithms using established objective evaluation performance measures.

In each competition, four measures are used to evaluate the binarised output of each approach: (i) F_1-score; (ii) Pseudo F_1-score; (iii) Peak signal-to-noise ratio (PSNR); (iv) Distance Reciprocal Distortion Metric (DRD).

F_1-score

The F_1-score is the same as that used in Chapter 4, repeated here for clarity

$$F_1 = \frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}, \tag{5.4}$$

where $\text{precision} = \frac{TP}{TP + FP}$, $\text{recall} = \frac{TP}{TP + FN}$, and TP, FP, FN are again the True Positive, False Positive, and False Negative values respectively.

Pseudo F_1-score

This measure, proposed in [48], differs from the standard F_1-score in the recall term used. A pseudo-recall (recall_{ps}) is calculated based on a skeletonized binary ground truth image. The motivation for this is that each character will have an unique silhouette that can be summarised with its skeleton. A perfect recall can therefore be achieved in cases where the foreground regions of binarised results include each skeleton component. The skeletonized ground truth is defined as

$$GT_s(x,y) = \begin{cases} 1, & \text{foreground.} \\ 0, & \text{background.} \end{cases}, \tag{5.5}$$

then the pseudo recall is calculated as follows

$$\text{recall}_{ps} = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} GT_s(x,y) \cdot B(x,y)}{\sum_{x=1}^{M} \sum_{y=1}^{N} GT_s(x,y)}, \tag{5.6}$$

and the pseudo F_1-score

$$F_{1ps} = \frac{2 \cdot \text{precision} \cdot \text{recall}_{ps}}{\text{precision} + \text{recall}_{ps}} \tag{5.7}$$

Peak Signal-to-Noise Ratio

The PSNR is a measure of the similarity between two images:

$$PSNR = 10 \log \Big(\frac{P^2}{MSE} \Big), \tag{5.8}$$

where P is the maximum intensity difference between images ($P = 1$ in this case), and

$$MSE = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} (GT(x,y) - B(x,y))}{M \cdot N}, \tag{5.9}$$

with $M \cdot N$ the number of pixels in each image.

Distance Reciprocal Distortion Metric

This measure was proposed in [39] as an alternative to the PSNR, which does not match well with subjective visual binarisation performance evaluation, due to its point based nature. The DRD is designed to correlate with human visual perception of document images, and calculates the distortion of each misclassified pixel $B(x_k, y_k)$, in a 5×5 window surrounding the pixel, as the sum of ground truth pixels in that window that differ from the misclassified pixel value:

$$DRD_k = \sum_{i=-2}^{2} \sum_{j=-2}^{2} |GT(x_k + i, y_k + j) - B(x_k, y_k)| \cdot W_n(3 - i, 3 - j), \tag{5.10}$$

where W_n is a normalised weight matrix, defined in [39] as

$$W(i,j) = \begin{cases} 0, & i = 3 \, \text{and} \, j = 3. \\ \frac{1}{\sqrt{(i-3)^2 + (j-3)^2}}, & \text{otherwise.} \end{cases}, \tag{5.11}$$

$$W_n(i,j) = \frac{W(i,j)}{\sum_{i=1}^{5} \sum_{j=1}^{5} W(i,j)}. \tag{5.12}$$

The total DRD for all S misclassified pixels is then

$$DRD = \frac{\sum_{k=1}^{S} DRD_k}{NUBN}, \tag{5.13}$$

where *NUBN* is the number of non-uniform 8×8 blocks in the ground truth image.

5.3 Results

5.3.1 Methods Selected for Comparison

Three non-blind techniques that use both spatial and intensity information were chosen for comparison with the approaches proposed in Chapters 3 and 4. These methods were selected as in previous works they have been shown to outperform other approaches discussed in Chapter 2. A brief summary of each, along with implementation details for these and the proposed methods, is contained in what follows.

Wavelet
The wavelet based bleed-through removal approach of Moghaddam and Cheriet in [44], referred to as Method 1, uses a function of the *difference* in intensity between the degraded recto and verso sides as an indicator of bleed-through and foreground text regions, with spatial smoothness enforced in the wavelet domain. This approach was implemented as described in [44]. However, the automatically generated results on some image pairs in the database were not optimal, with large regions of background and bleed-through remaining; the automatically estimated smoothing weight in these instances was too low. For these images therefore, the value of smoothing weight was chosen manually to produce the best result visually.

User Assisted
The method proposed by Huang et al. in [34] and [35], referred to as Method 2 aims to classify pixels from each side individually as foreground, background, or bleed-through based on the *ratio* of intensities between the recto and verso sides, initialised based on user input training data, and spatial smoothness is enforced in a dual-layer MRF framework. This approach was implemented as described in [35]. The per image user markup consisted of $9 - 12$ strokes drawn on both recto and verso sides that attempted to capture any variations in foreground, background and bleed-through profiles. For some image pairs the background and bleed-through were overclassified on recto and verso sides, with foreground underclassified therefore reducing legibility significantly. In these cases the user markup was repeated with a greater number of strokes highlighting the foreground regions.

Active Contour
The approach proposed by Hanasusanto et al. in [29], referred to as Method 3, uses a modified Chan-Vese active contour model [16], with an additional energy functional based on the *difference* in intensity between the degraded recto and verso sides, whose influence is controlled by a difference weight parameter. A post-processing stage for broken stroke completion is also included. The implementation was kindly provided by the authors. In this approach, the difference weight parameter is manually selected and critical to achieving good results. The weight was chosen for each image individually based on the initial segmentation performance, and as

there is often a balance to be struck between foreground preservation and bleed-through removal, two sets of outputs were generated. For one set, referred to as Method 3.i, a conservative difference weight was selected so as to preserve as much foreground as possible whilst still removing the most bleed-through. For the second set of results, referred to as Method 3.ii, a higher difference weight was selected so as to remove as much bleed-through as possible, inevitably resulting in a larger amount of broken foreground strokes than for the conservative weight. The post-processing stroke completion stage was then applied to both sets.

Model-Based

For the implementation of the approach in Chapter 3, referred to as Model 4, the optimisation was performed over 35 iterations for each image pair, with the alternative linear model for the clean images substituted every 10 iterations and in the final iteration. The various parameters were set empirically as: $\lambda_M = 1$, $\lambda_\alpha = 55$, $\lambda_c = 5$.

Non-Parametric

Finally, this approach was implemented as described in Chapter 4, with fixed smoothness weights for each model variation, referred to as Method 5.i, Method 5.ii, and Method 5.iii respectively. As the performance on some images was sensitive to the weight chosen, a further set of results was generated for the most conservative Model 3, using manually selected smoothness weights, referred to as Method 5.iv.

Methods

For clarity and ease of reference, the method names and their corresponding references are listed here as follows:

Method 1 - Wavelet based approach with some manually tuned results [44].
Method 2 - User assisted approach [35].
Method 3.i - Contour based method [29], with conservative manually selected difference weight.
Method 3.ii - Contour based method [29], with high manually selective difference weight.
Method 4 - Proposed model-based approach from Chapter 3.
Method 5.i - Proposed approach from Chapter 4, Model 1 (frequency weighted Unary Energy).
Method 5.ii - Proposed approach from Chapter 4, Model 2 (frequency weighted Pairwise Energy).
Method 5.iii - Proposed approach from Chapter 4, Model 3 (no frequency weighting).
Method 5.iv - Proposed approach from Chapter 4, Model 3, with manually selected smoothness weight.

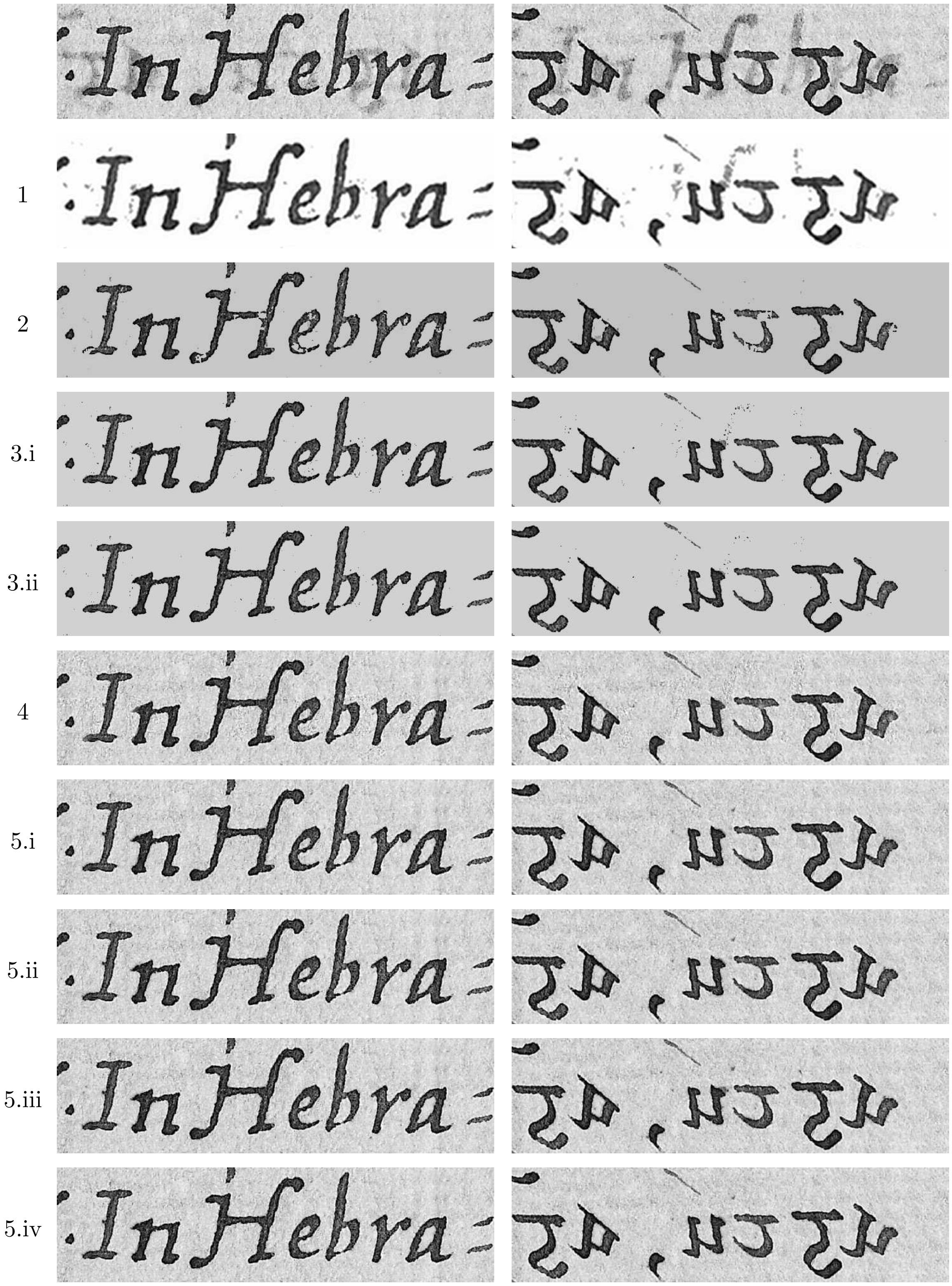


Figure 5.1: Sample cropped restoration results from all methods on an image pair from the database with light bleed-through. Left column: Original degraded recto and clean results. Right column: corresponding degraded verso and clean results. **M02**.

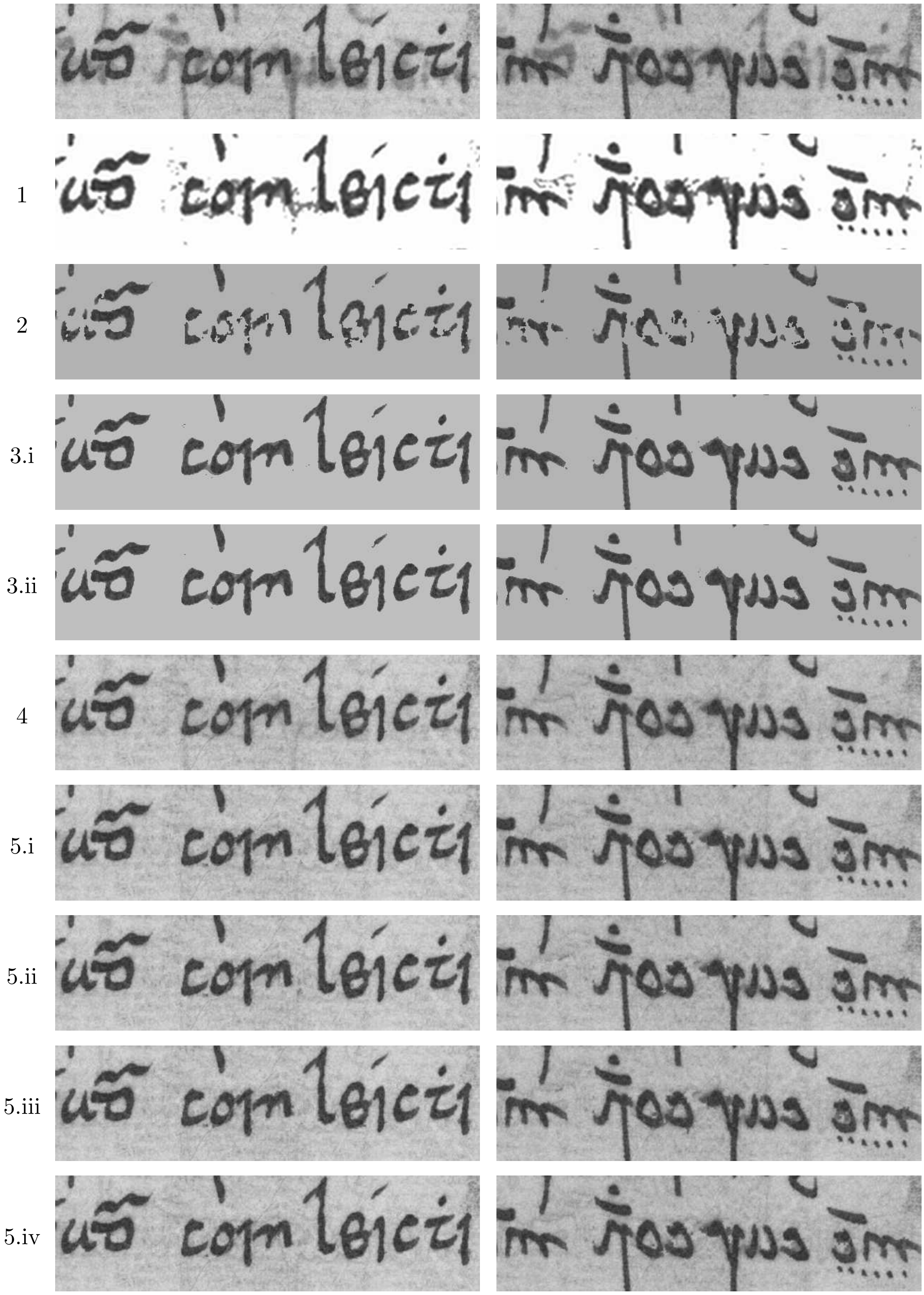


Figure 5.2: Sample cropped restoration results from all methods on an image pair from the database. Left column: Original degraded recto and clean results. Right column: corresponding degraded verso and clean results. **M35**.

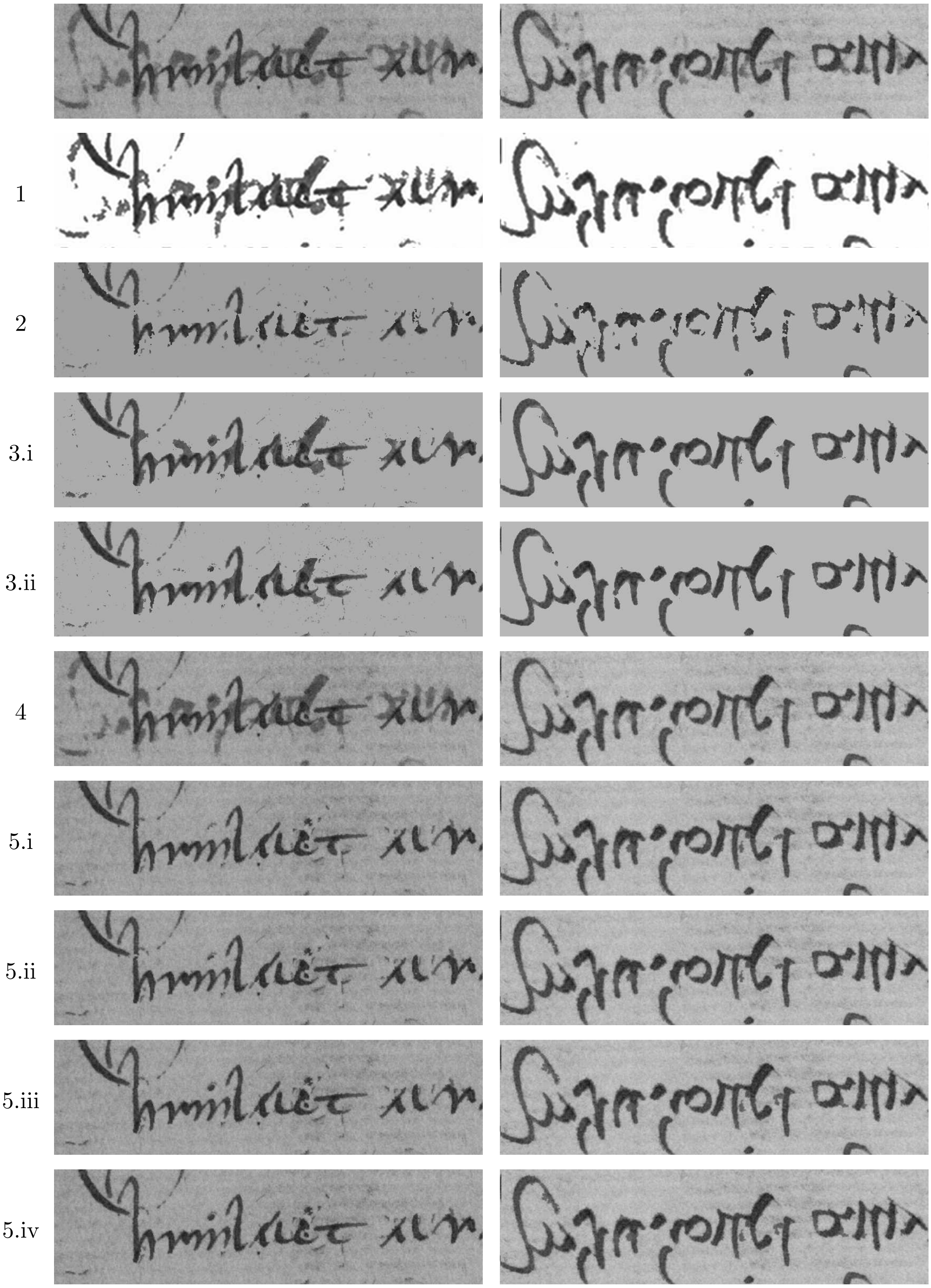


Figure 5.3: Sample cropped restoration results from all methods on an image pair from the database. Left column: Original degraded recto and clean results. Right column: corresponding degraded verso and clean results. **M32**.

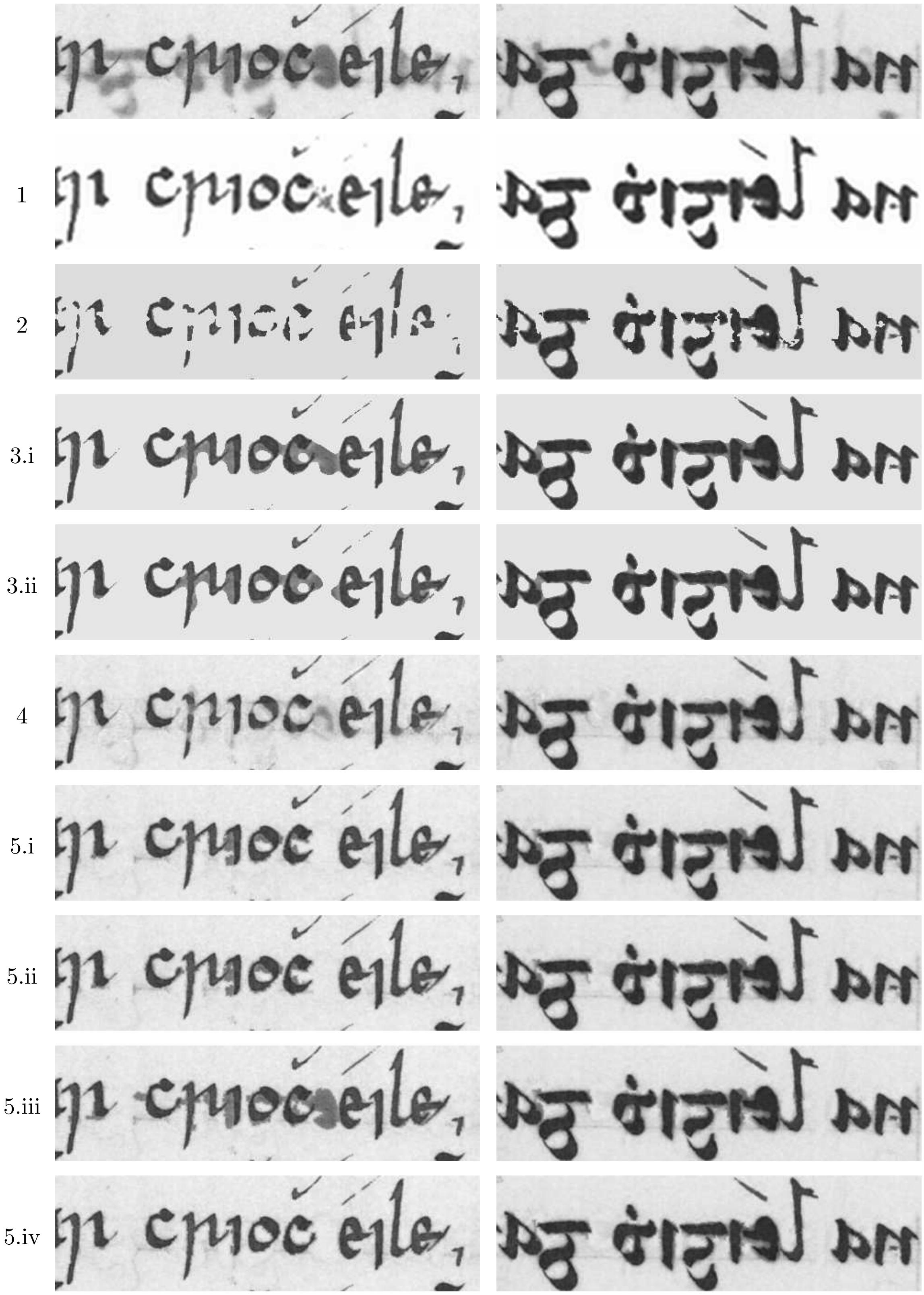


Figure 5.4: Sample cropped restoration results from all methods on an image pair from the database. Left column: Original degraded recto and clean results. Right column: corresponding degraded verso and clean results. **M01**.

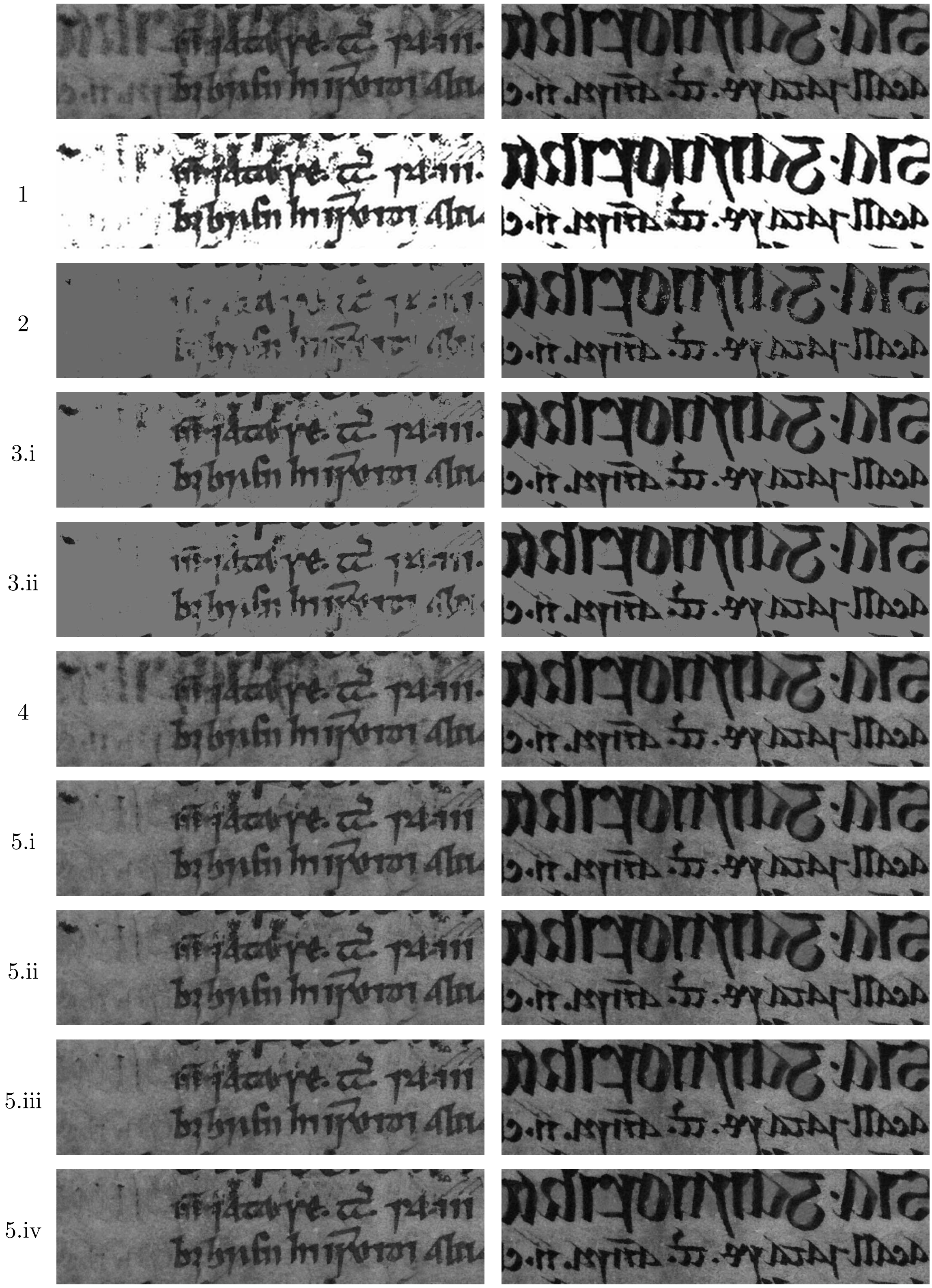


Figure 5.5: Sample cropped restoration results from all methods on an image pair from the database. Left column: Original degraded recto and clean results. Right column: corresponding degraded verso and clean results. **M25**.

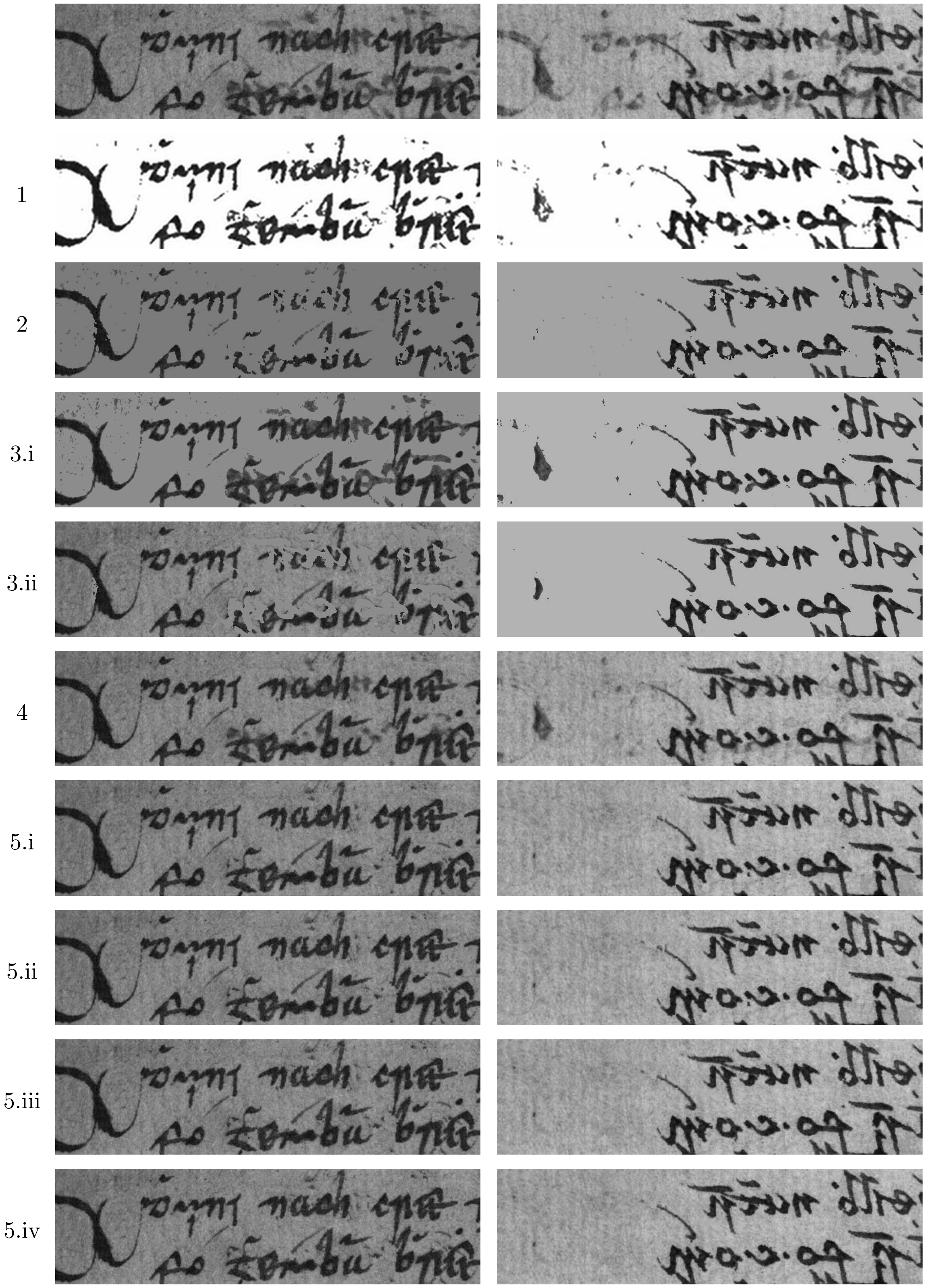


Figure 5.6: Sample cropped restoration results from all methods on an image pair from the database. Left column: Original degraded recto and clean results. Right column: corresponding degraded verso and clean results. **M39**.

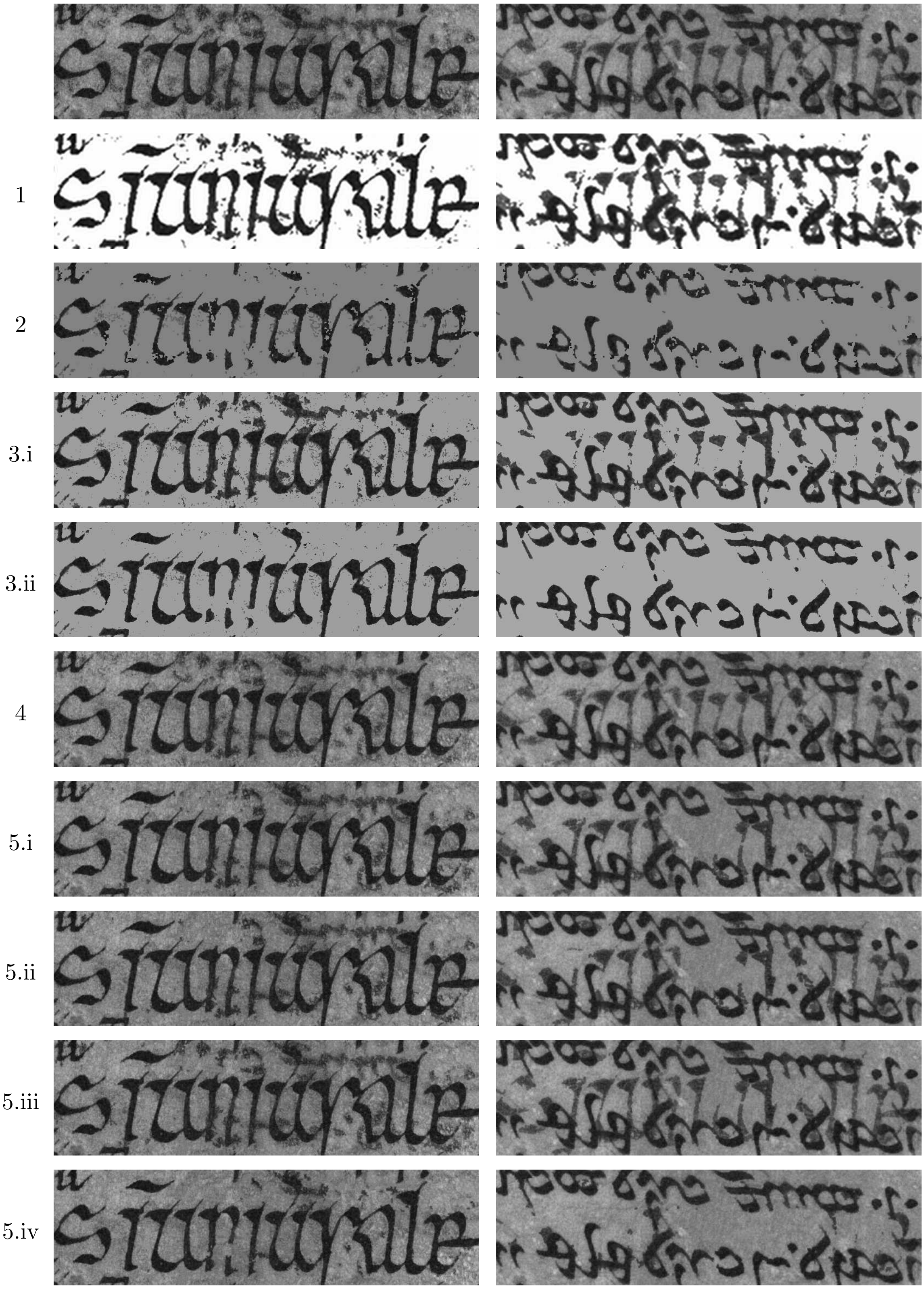


Figure 5.7: Sample cropped restoration results from all methods on an image pair from the database. Left column: Original degraded recto and clean results. Right column: corresponding degraded verso and clean results. **M27**.

5.3.2 Visual Comparison

Figures 5.1-5.7 show a visual comparison of a selection of extracts of results from each method tested. Visual comparison results on all full database images can be found in the accompanying CD.

It is clear from the figures that the wavelet based Method 1, similarly to the proposed Method 4, is generally conservative - preserving the foreground in all examples and not performing as well in more severe bleed-through cases. However, as mentioned earlier since the output background texture is not preserved, bleed-through artefacts left by Method 1 become more noticeable in the restored results due to the increase in contrast between background and bleed-through intensities, thus reducing legibility compared to the original degraded images. This is especially noticeable in the more severe bleed-through examples in Figures 5.3, 5.5, 5.6, and 5.7. In comparison, the artefacts left by Method 4 appear unchanged from the original.

The user assisted Method 2 generally removes most bleed-through successfully, however also removes some sections of true foreground information in each example, especially in regions of overlapping foreground on both sides. This is undesirable as it clearly reduces the legibility of the restored result in each case. This method also classifies some background regions as foreground in the examples with more highly textured background in Figures 5.3, 5.5-5.7.

The active contour Methods 3.i and 3.ii perform well at preserving the foreground text in most examples due to the broken stroke completion stage, however as can be seen in Figures 5.3 and 5.4 this results in some bleed-through artefacts also being completed. As expected, the conservative difference weight used in Method 3.i results in less bleed-through removed in each example, however in the severe cases shown in Figures 5.5 and 5.7, the foreground is better preserved. In Figure 5.6, the recto results for both Method 3.i and 3.ii are not optimal. For Method 3.i the difference weight was set to 0 in order to preserve the foreground, which resulted in most of the bleed-through being preserved too, and in Method 3.ii, the background was segmented as foreground, and all foreground strokes in overlapping regions removed as bleed-through. This is most likely due to the fact that in this particular example the recto side is much darker than the verso side, and this negates any assumptions made about the difference between intensities in foreground, background and bleed-through regions. This method therefore may not cope well with image examples where the global intensity properties *between* sides is significantly different.

Comparing the performance of the two proposed approaches from Chapters 3 and 4, there is little difference between the outputs of Method 4 and Methods 5.i-5.iv on the relatively light and medium bleed-through examples in Figures 5.1 and 5.2 respectively. However as expected, Methods 5.i-5.iv perform significantly better on the more severe cases. It is clear from all the examples shown that both proposed approaches successfully preserve the document character in each case.

Examining the different variations of the approach in Chapter 4, the performance of all the Methods 5.i-5.iv is similar in Figures 5.1-5.3, and 5.6, with only minor variations in the outputs generated. The results for these cases appear less sensitive to the choice of smoothness parameter. Figures 5.4 and 5.7 on the other hand show cases where the outputs of each model variation are not similar. In Figure 5.4, Methods 5.i and 5.ii, generate near identical results, only with 5.ii preserving the faint foreground accent in the recto image. Method 5.iii performs worse than all the compared methods in this case, most noticeably in the recto image, with a large proportion of the bleed-through left. The manually tuned Method 5.iv however performs best in this case at removing bleed-through, though the recto accent is broken.

Figure 5.7 shows a sample extract where the fixed smoothing weight is insufficient for all Methods 5.i, 5.ii, and 5.iii. with large sections of bleed-through left in both sides. Again, the manually tuned Method 5.iv has a significantly better output in this case, however there is some bleed-through left in both sides where there is overlapping foreground. The active contour Method 3.ii performs better at removing bleed-through in this example, with only slightly more broken strokes. The main strength of the Methods 5.i-5.iv is apparent in Figure 5.6, where a large region of dark bleed-through left by most of the other methods is removed, while the foreground text, including faint features are preserved.

It is clear from visual comparison of results that the most difficult regions to classify accurately for all approaches are those where there is overlapping foreground on both sides combined with severe bleed-through degradation. In these regions, there is often little or no distinction between foreground and bleed-through intensities, and the character edges are often almost indistinguishable.

5.3.3 Numerical Comparison

The error metrics described in Section 5.2, were evaluated for the results of each method on both recto and verso sides of each image pair in the database and the results are discussed in what follows. Figure 5.8, shows the *BgError* plotted against *FgError* for all images and each method, with trend ellipses centred at the mean *BgError* and *FgError* values. This figure highlights the performance of each method in terms of obtaining a balance between foreground preservation and bleed-through (and background in this case) removal, where the optimal restoration target is at $[0, 0]$. The user assisted Method 2 clearly stands out as having the lowest *BgError* results, with highest *FgError* overall. This is due to the overclassification of background and bleed-through, resulting in broken true foreground strokes. The wavelet Method 1 also stands out as performing relatively poorly in terms of *BgError* results, highlighting the visual observation that bleed-through and background artefacts left by this approach are much more noticeable than in the original due to the increased contrast with the background. The two active contour variations again compare as expected, with the conservative variation, Method 3.i (dark green), having lower *FgError* and higher *BgError* overall, and the approach with stronger difference weight, Method 3.ii (pink), conversely performing much better in terms of

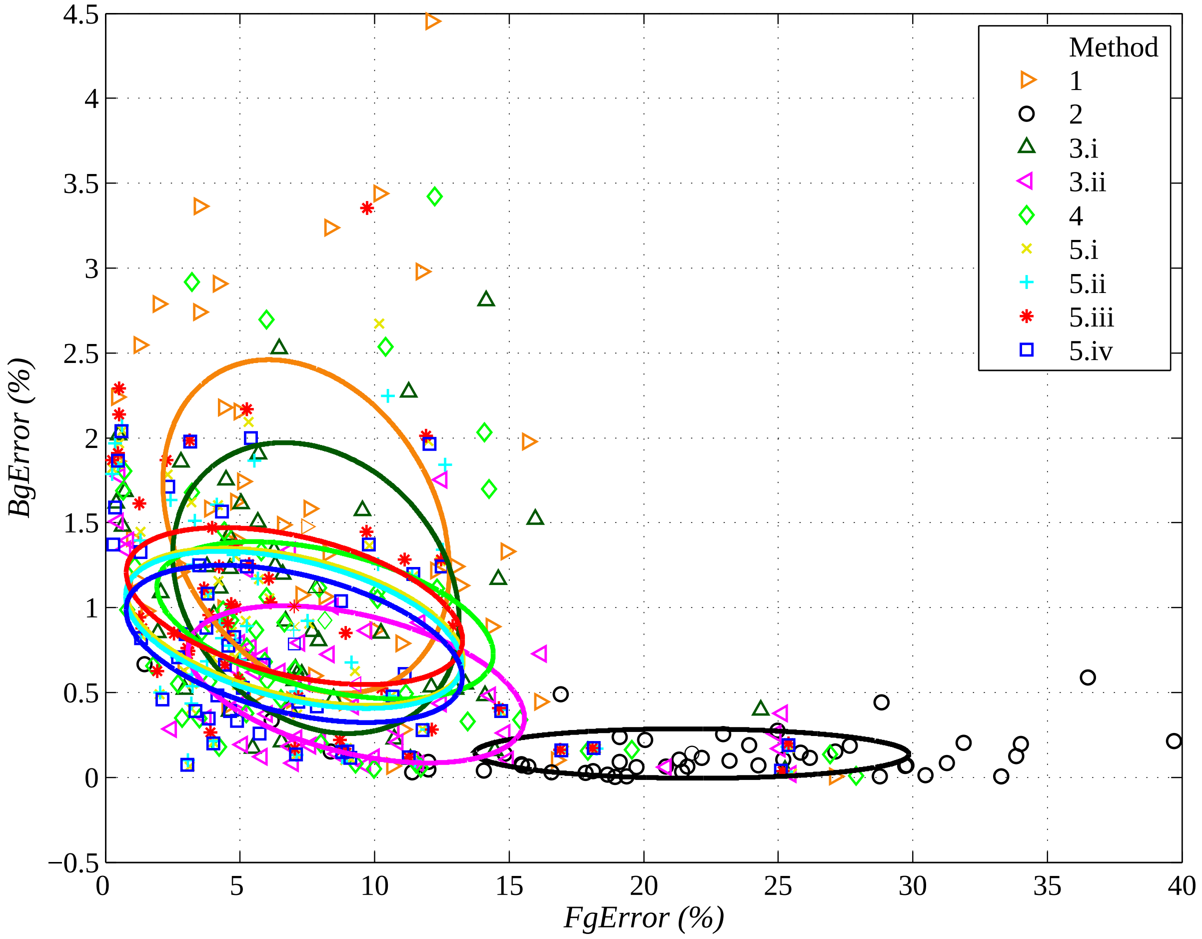


Figure 5.8: *BgError* plotted against *FgError* for all 50 images restored with each method tested, with trend ellipses centred at the mean *BgError* and *FgError* values.

BgError and worse in terms of *FgError* . The proposed methods, due to the similarity in output format, have similar trend performance overall, with lower *FgError* on average than previous works. Method 4 (light green) from Chapter 3 however, has more instances of high *BgError* than the methods from Chapter 4, highlighting the relatively poor performance on images with dark bleed-through. A clearer differentiation between the three models from Chapter 4 can be seen in this evaluation than in Chapter 4; where the average *FgError* is similar for all three, the models that include frequency weighting, Methods 5.i (light blue) and 5.ii (yellow) have better bleed-through removal capability than the model with no weighting, Method 5.iii (red). However, the un-weighted model performance is increased significantly with the manually tuned smoothness weight, Method 5.iv (dark blue), performing best overall.

It is clear from Figure 5.8 that the performance of all methods varies significantly between image examples due to the different properties of each image pair. Therefore rather than ex-

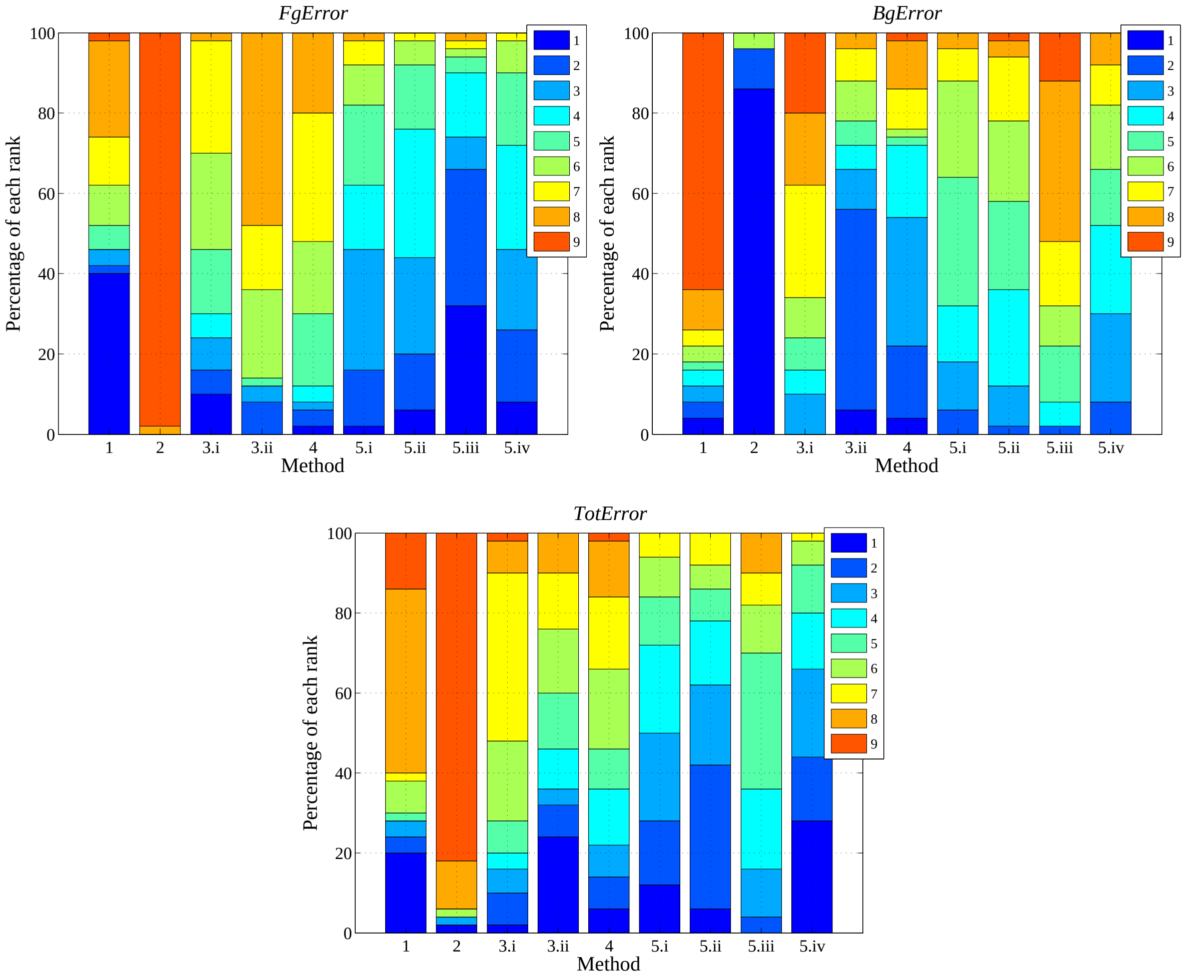


Figure 5.9: Illustration of the ranking of the performance of each method for the bleed-through error measures. Top left: *FgError*, right: *BgError*. Bottom: *TotError*.

amining per image results for all metrics to compare methods, the ranking performance of each method for each metric is examined. For completeness full results for each metric are presented in Appendix A. Figures 5.9, and 5.10 show illustrations of the ranking performance of each method, with the percentage of each rank highlighted. The rank performances for *FgError* and *BgError* shown in Figure 5.9 correlate with the observations already made. However, the relative ranking performance of the methods from Chapter 4 emphasises the disparity between ranking and mean error values, as the mean *FgError* for these methods is very similar, whereas Method 5.iii performs significantly better in the rank results. The *TotError* ranks again confirm the observations made about balancing between bleed-through removal and foreground preservation, however as the background region generally constitutes a much larger proportion of the image than the foreground, this metric is influenced more by *BgError* than *FgError* results. It

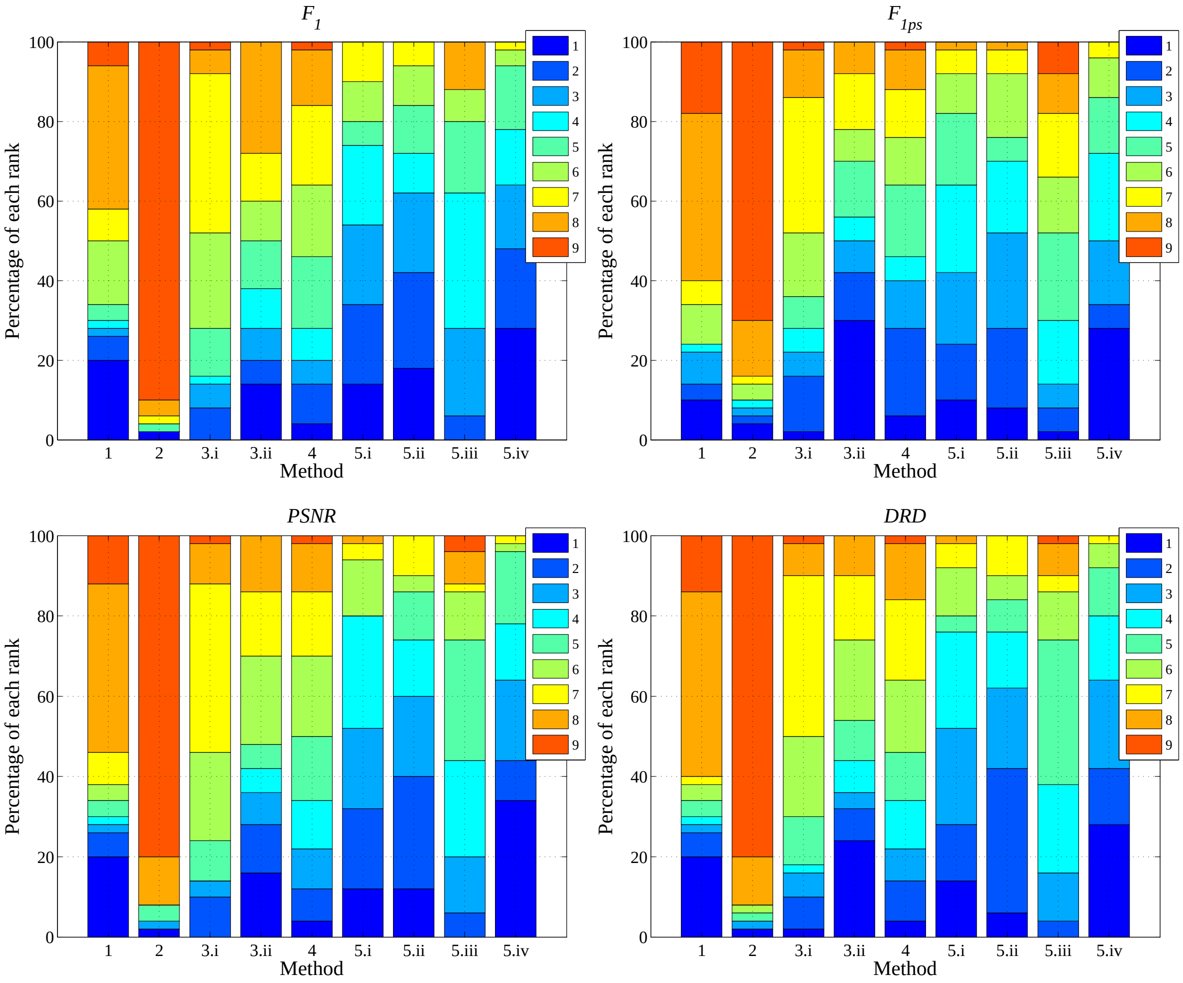


Figure 5.10: Illustration of the ranking of the performance of each method for the DIBCO error measures. Top left: F_1-Score, right: F_{1ps}-Score. Bottom left: $PSNR$, right: DRD

is clear that the manually tuned proposed Method 5.iv performs best, with Methods 5.ii and 5.i 2nd and 3rd respectively. Though the active contour Method 3.ii is ranked 1st in more cases than 5.i and 5.ii, it has comparatively more lower ranked results also. The method proposed in Chapter 3 outperforms Methods 1, 2, and 3.i, however is again limited by poor performance on severe cases.

The ranking performance for the DIBCO metrics, shown in Figure 5.10, again correlate with observations already made, with all methods performing similarly for each metric. It is interesting to note the improvement in ranking for the active contour methods between the F_1-score and F_{1ps}-score results and between the $PSNR$ and DRD results. As the F_{1ps} and DRD take into account stroke completeness and human visual perception of error, this improvement shows the effect of the broken stroked completion stage on the final results appearance.

Table 5.1: Average errors, RP ranks, and cumulative rank scores for each method tested over all 50 images.

	1	2	3.i	3.ii	4	5.i	5.ii	5.iii	5.iv
FgError	7.46	21.79	7.82	9.32	8.16	7.05	**6.99**	7.02	7.01
RP *Fg*-Rank	5	9	6	8	7	4	3	1	2
BgError	1.48	**0.14**	1.12	0.55	0.92	0.89	0.87	1.01	0.79
RP *Bg*-Rank	9	1	8	2	3	5	6	7	4
TotError	2.44	3.95	2.32	2.18	2.20	1.98	1.96	2.08	**1.90**
RP *Tot*-Rank	8	9	7	4	6	3	2	5	1
Rank Score	902	941	884	720	778	629	610	725	**561**
F_1	87.59	81.23	87.90	88.03	88.11	88.91	88.98	88.63	**89.14**
RP F_1-Rank	8	9	7	6	5	3	2	4	1
F_{1ps}	92.78	90.11	93.40	93.97	93.64	93.98	94.00	93.67	**94.19**
RP F_{1ps}-Rank	8	9	7	2	5	4	3	6	1
$PSNR$	13.50	12.31	13.66	13.84	13.81	14.06	14.08	13.92	**14.16**
RP $PSNR$-Rank	8	9	7	6	5	3	2	4	1
DRD	9.15	12.87	8.66	8.23	8.46	7.80	7.75	8.09	**7.58**
RP DRD-Rank	8	9	7	5	6	3	2	4	1
DIBCO Rank Score	1199	1680	1172	910	1027	727	691	991	**603**

All the results discussed are summarised in Table 5.1, where average error values, per error RP ranks, and Rank Scores for each method are shown. It is clear from the combination of all the metrics that proposed manually adjusted Method 5.iv performs best overall, followed closely by Method 5.ii. The approach proposed in Chapter 3, Method 4, is ranked 6th overall.

5.3.4 Discussion

Aside from bleed-through removal, subjectively it is easy to say that the proposed methods produce visually ‘nicer’ results than previous approaches. However, an important factor to be taken into account when comparing outputs visually is the use that the bleed-through removal was intended for. Though the main goal of the proposed approaches is the preservation of document character whilst improving legibility, this may not be important for approaches where the main goal is accurate binarisation input for OCR, handwriting recognition or some other processing stage. Therefore, whilst visual comparisons can give an idea as to performance, it is very necessary to include a more objective analysis. A clear advantage of classification based approaches is that the choice of output type is not fixed; binary and pseudo-binary outputs may be generated directly from the image label fields, and textured outputs may be generated using the restoration approach described in Section 4.3.

From the results presented, in general the visual and numerical comparisons show that the approach proposed in Chapter 4 performs best overall in terms of balancing bleed-through removal and foreground preservation. The main drawback of this approach is that there is no method for splitting component shapes in cases where they may contain both true foreground and bleed-through. A direction for future work would be to incorporate contour based or gradient information into the connected component analysis, so that component shapes could be modified in cases where they contain an obvious character edge. The connected component rules could also be modified to include this information. A further drawback of this method is that, like Method 4, it has no way of dealing with bleed-through that spreads horizontally across the page.

The significant improvement that was achieved in the results for some of the image pairs when using manually tuned smoothness weights in both the proposed approach from Chapter 4, the wavelet based Method 1, and also the high performance of the manually tuned Method 3.ii, suggests that full automation with robustness to the wide variety of different bleed-through document types, image resolutions, and degradation strengths may not be feasible, and that incorporation of user assistance, as stated in [29, 35], is in fact the best way to achieve consistently optimal results. A further observation that can be made from examining results over the whole database is that, as severity of degradation increases, there is a point at which, even with manually tuned parameters, no method will successfully remove all bleed-through without further contextual content-based information.

A user interactive approach is most suited to the needs of individual researchers seeking to restore a small number of documents to improve legibility, or for presentation purposes. However, for large digitisation projects, any user interaction must necessarily be kept to a minimum. The high performance of the proposed Methods 5, coupled with visual results presented in Chapter 4 on test images from outside of the database, show that this approach, with improved automatic selection of smoothness weights, has the potential to improve the appearance of bleed-through degraded documents, even if not to an optimal extent, and could easily be incorporated as part of a larger framework of document analysis to improve the performance of subsequent stages. To further test the robustness of the proposed approaches, the ground truth database would need to be expanded to include a greater variety of different document types.

A key performance metric not addressed in detail in this work is that of computational complexity, as the implementations for each approach were not optimised for speed.

5.4 Summary

This chapter has presented the details of a database of 25 registered recto-verso images with manually generated ground truth foreground masks. An objective evaluation approach has also been proposed that seeks to provide a fair comparison between results of different methods, taking into account their different formats. Using the database and proposed evaluation, along with standard binarisation evaluation metrics, the performance of the two proposed restoration approaches, detailed in Chapters 3 and 4, have been compared against 3 previous non-blind approaches. The results showed that the method proposed in Chapter 4 performed best overall, and that parameters manually selected per image can improve performance significantly over automatically selected or fixed, leading to the conclusion that for any approach to be robust to a wide variety of input degradations and formats, some form of user input is necessary.

6

Registration

The previous chapters have presented and evaluated two new approaches for non-blind bleed-through restoration. However in both cases it was assumed that registration of recto and verso sides had already been performed. Accurate registration of document recto and verso sides with bleed-through degradation is essential for accurate automatic non-blind bleed-through restoration - even relatively small misalignments will result in bleed-through artefacts remaining in the restored result. Registration of document pages is a non-trivial task for a number of reasons. Firstly the intensity profiles of corresponding ink and bleed-through regions are very different; it may be difficult to find correspondences between faint bleed-through and its originating foreground text. Secondly, bleed-through may only occur sparsely across a page, leading to the chance of foreground text from both sides becoming warped to match. Furthermore, page topography can cause problems for accurate registration, especially in manuscripts and printed books with tight binding where the degree of curvature towards the spine can be very different on recto and verso sides, or in documents on vellum or paper that have become warped due to ageing or poor storage conditions.

There exist a number of previous approaches to recto-verso registration. In one of the earliest methods, [21], the parameters of a global affine warp are optimised using the sum of squared differences (SSD) in intensity between the two sides as the similarity metric. In [55] a global method for printed documents is proposed that considers rotation and translation only, using a layout analysis of verso foreground ink and recto bleed-through. The rotation angle is selected based on an estimation of the bleed-through profile only, and then the vertical and horizontal translations are selected using both the estimated ink and bleed-through profiles. These methods all perform global alignment only, but in many cases this is not sufficient to obtain the fine level registration accuracy needed for bleed-through removal.

Previous local registration methods all consist of two warp stages: a coarse global registration followed by local region alignment. The approach proposed in [75] performs global alignment considering translational deformation only, under the assumption that the two images are of the same spatial resolution. The translation is obtained by locating corresponding regions in areas of heavy bleed-through on both sides, using Normalised Cross-Correlation (NCC) based block matching over a fixed search window. Secondly, a finer alignment is performed by dividing the recto and verso into small fixed blocks and using the same block matching over smaller windows, in this case considering both translational and rotational deformations. In regions of warped block overlap or divergence, radial basis function (RBF) interpolation [28] of the image intensity is used. A similar local approach is used in [35], except that local translational deformation only is considered, and the displacements for each block are interpolated over the whole image using thin plate spline interpolation to obtain the final warp. These approaches both rely on the images being of the same spatial resolution, which is often not the case, especially for large volumes digitised with a fixed imaging rig. They also require the input images to be aligned enough to within the initial global search window to perform accurate registration.

Another similar local registration method is proposed in [76], using a multiresolution approach. In this case the global warp stage considers full affine transformations, and the local refinement is based on a coarse grid point deformation, rather than block based, using hierarchichal levels of B-spline warps [62] to obtain the full displacement field. The uniform spline grid is defined based on detected corresponding control points on the recto and verso sides located based on the assumption of slanted text, and to prevent overlapping points in the output. The registration error metric used is the residual complexity [46], defined as the compression complexity of the difference in intensity between the registered images, and evaluated as a function of the Discrete Cosine Transform (DCT) of this difference, or residual image. A smoothness constraint on the warp is also used to prevent excessive distortion in the registered result. As this approach relies explicitly on locating correspondences between slanted bleed-through and foreground text, the presence of strong bleed-through across a page is required for accurate alignment. Furthermore all of these previous approaches do not take into account the preprocessing necessary for registration if the document images contain sections of the opposite page and binding.

This Chapter presents a registration framework that follows a similar approach to the previous local warp methods; it consists of two distinct warp stages[1]. However, the initial warp stage is based on the fact that the page outline will be the same for both sides irrespective of the strength or presence of bleed-through. The images of two sides are automatically cropped to remove any binding or opposite page information, and then the outlines are registered using a global similarity transformation. In the second stage a local grid-point based warp, similar

---

[1] Parts of this chapter are based on the work published in [61].

to [76] is applied, however, rather than using the residual complexity which can be computationally intensive especially for large images, the sums of the squared difference in image intensities and also gradient fields are used as the error metric. To prevent excessive distortion of the warped image, a 'content-preserving' smoothness term [38] is included in the warp estimation. A numerical evaluation approach for bleed-through registration performance is also presented, where the automatically generated warp displacement field results are compared against ground truth displacement fields generated from manually registered images. The error for each image pair is calculated in terms of the smallest bleed-through feature width so as to be able to compare registration performance on images of different resolution, text size, and strength of bleed-through degradation. Comparisons of the proposed approach with existing methods shows improved registration performance.

6.1 Initial Global Warp

As it is difficult to calculate *a priori* how much bleed-through a page contains, and how strong the degradation is, it is difficult to know if there is enough correspondence between the recto and verso sides to produce a satisfactory registration result. Therefore for the initial global warp the outline shape of the page itself and not the information on each side is used, as even on a page with no bleed-through, the outline will be the same on the recto and verso sides. At present, it is assumed that the image region surrounding the document is reasonably uniform and dark in intensity, and that an interleaf has been placed behind the imaged page during image capture. As the methods for document imaging vary widely between different projects, the use of an interleaf is not standard - it can add time to the imaging process, and can risk causing damage to more fragile documents. However, use of interleaves is not uncommon, and in some projects the documents are even unbound to facilitate imaging, which results in the same effect.

To isolate the outline shape, both recto and verso sides are binarised globally using Otsu's thresholding method [50], and then large scale morphological opening and closing is performed to the resultant binary images to remove all remaining text features, leaving the shape of the pages only. An illustration of this stage is shown in Figure 6.1. In cases where the surrounding image region is lighter than the writing medium, a simple inversion of the binarisation process will achieve the same result.

These binary images are then automatically cropped on the left hand side to remove any of the opposite page present, as this will be a different page in the recto and verso images and also a different width thus affecting both the outline global, and subsequent local registration. The crop width is obtained by observing that where the two pages meet the binding, there is a sharp change in the outline edge gradient. So by performing Harris corner detection on the top and bottom left quadrants of the binary images these points are located, and the crop width is set as the minimum of the horizontal coordinates of the two (see Figure 6.2). A simple adjustment of the search window for corner detection can be made for cases where the full two

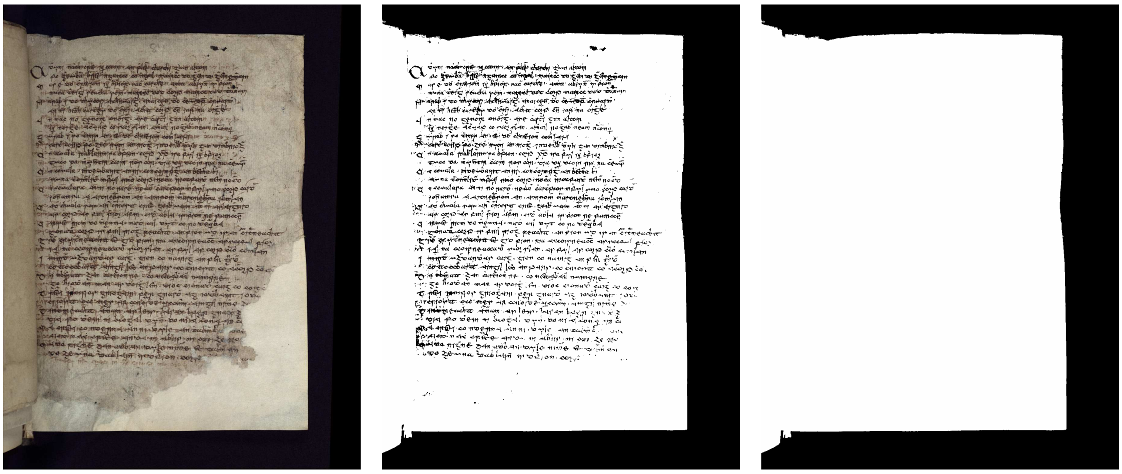


Figure 6.1: Illustration of the page outline extractions. Left to right: Manuscript page, binarised, and after morphological opening and closing. **M39**.

pages are imaged together as there will still be a change in the outline edge gradient because of the binding.

Following the binding crop stage, an initial global registration is performed between the two binarised pages. To prevent any loss of resolution, the reference image $J(x, y)$ is set as the side with the larger area of the two binary images, and the warped image $I(x, y)$ as the smaller. The global warp $\mathbf{W}_g(x, y; \mathbf{p}_g)$ is a similarity transformation and is defined as

$$\mathbf{W}_g(x, y; \mathbf{p}_g) = \begin{bmatrix} p_1 & -p_2 \\ p_2 & p_1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} p_3 \\ p_4 \end{bmatrix} \tag{6.1}$$

Uniform scaling is enforced as the visible width of the page may not be the same on both sides due to the curvature of the binding, and this could lead to undesirable stretching of the text; the height of the page is the most accurate measure of scale between the two images (see Figures 6.11 and 6.12). The similarity metric used for the global warp is the SSD between the two binarised images, and the warp parameters $\mathbf{p}_g$ are optimized using multi-resolution gradient descent optimisation, with three resolution levels, coarse to fine.

6.2 Local Grid Warp

Once the outer page edges are roughly aligned, only small adjustments should be necessary to align accurately the bleed-through regions with their originating text. As the nature of the warp required can vary widely across a single page, the warp used is a grid-point based deformation, similarly to [76], and is obtained by minimising a function of two weighted energy terms: a data term for each grid point displacement, and a similarity transform term to ensure that text and bleed-through regions are not adversely distorted.

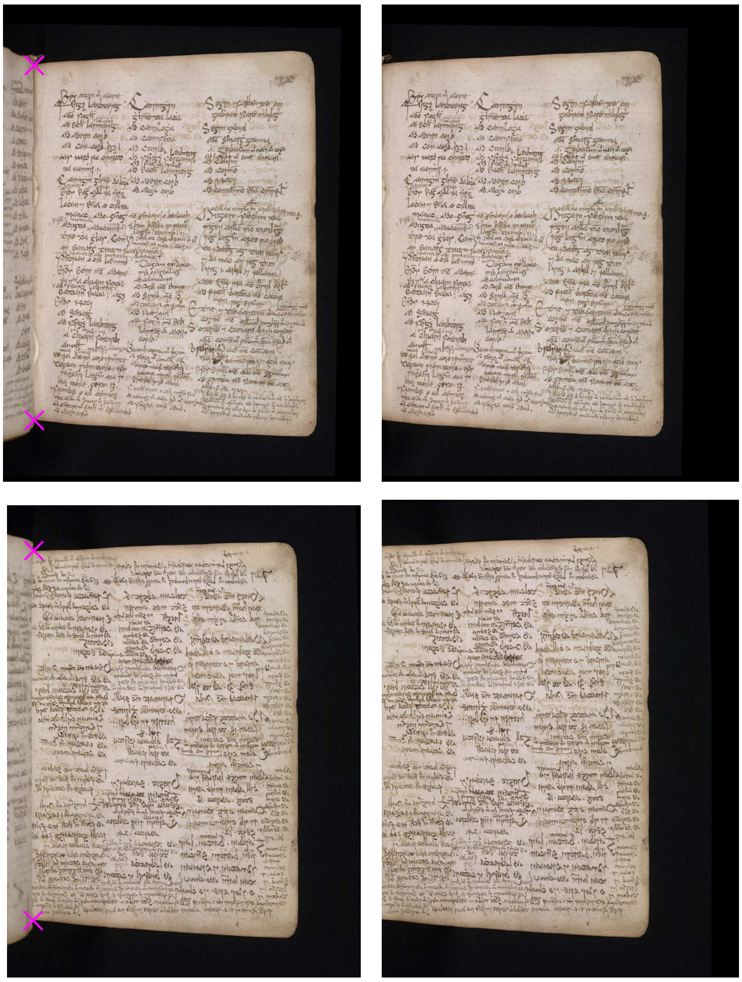


Figure 6.2: Illustration of the automatic crop stage. From left to right: Recto (top) and flipped verso images with located binding corners, and corresponding cropped images. **M31**.

6.2.1 Data Term

The sum of squared intensity difference is again used as the error metric between the two images. However, as the foreground and corresponding bleed-through intensity profiles are often very different, use of intensity information alone may lead to foreground text on one side erroneously aligning with foreground text on the other. Thus the difference in the gradient fields of the two images is also used, and the full data energy term may be expressed as

$$\begin{aligned} E_d = \sum_{x,y} \Big\{ & (I(x+u, y+v) - J(x,y))^2 \\ & + \lambda \big[(I_x(x+u, y+v) - J_x(x,y))^2 \\ & \qquad + (I_y(x+u, y+v) - J_y(x,y))^2 \big] \Big\}, \end{aligned} \tag{6.2}$$

where $I(x,y)$ is the globally warped warped image, λ is a weight parameter to control the influence of the gradient difference term, and the displacement fields $u(x,y), v(x,y)$ are controlled by a smaller number of displacement estimates lying on a coarse rectilinear grid over the image:

$$u(x,y) = \sum_{i=1}^{n_x}\sum_{j=1}^{n_y} w_{ij}(x,y)u_{ij} \tag{6.3}$$

$$v(x,y) = \sum_{i=1}^{n_x}\sum_{j=1}^{n_y} w_{ij}(x,y)v_{ij}, \tag{6.4}$$

where u_{ij}, v_{ij} are the horizontal and vertical displacements respectively of the set of $n_x \times n_y$ grid points (x_i, y_j), and w_{ij} are bilinear weights controlling the influence of the grid displacements on pixel displacements. The advantage of choosing bilinear interpolation is that the local control is smaller than for B-Spline, thus reducing the computational load. However the trade-off is that the resultant interpolated warp fields are not as smooth.

6.2.2 Similarity Term

To ensure that the document text and bleed-through regions do not become adversely distorted during registration, a constraint on the warp parameters is included. Content-preserving warps [38] are employed for this purpose. In this approach each grid cell is split into two triangles and a similarity term measures the deviation of each output triangle from a similarity transformation of its corresponding input triangle. Each point of a triangle may be represented in a local coordinate system defined by the vector between the other two points, and the 90 degree rotation of that vector, see Figure 6.3. For example, one point $P_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$ of a triangle $\{P_0, P_1, P_2\}$ may be expressed as follows.

$$P_0 = P_1 + p(P_2 - P_1) + qR_{90}(P_2 - P_1), \quad R_{90} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \tag{6.5}$$

where (p,q) are the known local coordinates. If the corresponding warped output triangle $\{P'_0, P'_1, P'_2\}$ has not undergone a similarity transformation, P'_0 will not correspond with the warped location calculated with P'_1, P'_2, and (p,q). Therefore the smoothness energy for this point is derived from the distance between P'_0 and its desired location under a similarity transformation:

$$E_s(P'_0) = \left\| P'_0 - \big(P'_1 + p(P'_2 - P'_1) + qR_{90}(P'_2 - P'_1)\big) \right\|^2 \tag{6.6}$$

The energy for the warped output triangle is therefore the sum of the energy of each point:

$$E_s(\{P'_0, P'_1, P'_2\}) = w_s\big(E_s(P'_0) + E_s(P'_1) + E_s(P'_2)\big), \tag{6.7}$$

where w_s is a saliency weight for each triangle set to the intensity variance of the corresponding enclosing grid cell. The variance is used to ensure that the similarity constraint is high in regions

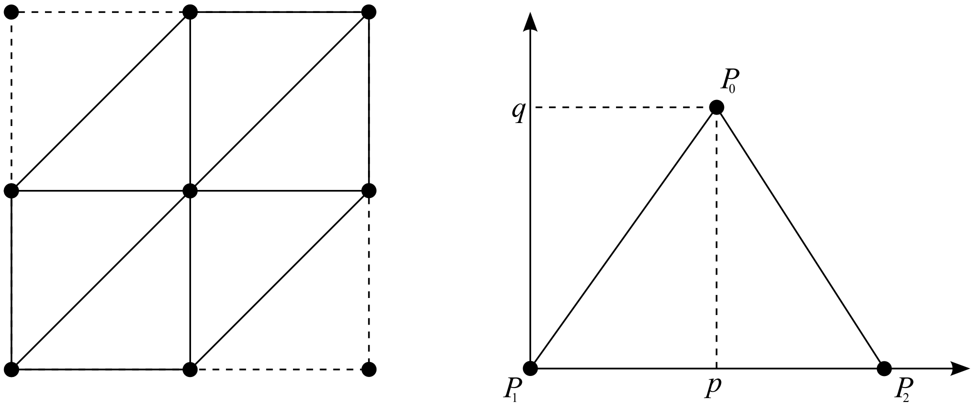


Figure 6.3: Left: Six triangle neighbourhood structure of a grid point. Right: each point in a triangle may be represented in a local coordinate system defined by the vector between the other two points, and the 90 degree rotation of that vector, with known local coordinates p, q.

with foreground text or bleed-through, where non-similar warps are most undesirable, and low in background regions where registration accuracy is not so crucial for accurate bleed-through removal. In this work there are six triangles connected to each point in the grid, excluding boundaries (see Figure 6.3), and therefore similarity energy for a single point is obtained by summing Equation 6.7 for each triangle, and the total similarity energy E_s of a warped grid is the sum of that energy for each point.

6.2.3 Optimisation

The full energy equation for the grid warp is defined as follows:

$$E = E_d + \alpha E_s, \tag{6.8}$$

with parameter α controlling the influence of the similarity term, set to 1 in what follows. By letting $\mathbf{W}(x, y; \mathbf{p})$ represent the grid warp, where $\mathbf{p} = [u_{11}, ...u_{n_x n_y}, v_{11}, ...v_{n_x n_y}]^\intercal$, and assuming that an estimate for $\mathbf{p}$ is known, Equation 6.8 can be minimised iteratively with respect to increments in the warp parameters, $\hat{\mathbf{p}}$, until convergence:

$$\begin{aligned} E_d = \sum_{x,y} \Big\{ & (I(\mathbf{W}(x,y;\mathbf{p}+\hat{\mathbf{p}})) - J(x,y))^2 \\ & + \lambda \big[(I_x(\mathbf{W}(x,y;\mathbf{p}+\hat{\mathbf{p}})) - J_x(x,y))^2 \\ & \quad + (I_y(\mathbf{W}(x,y;\mathbf{p}+\hat{\mathbf{p}})) - J_y(x,y))^2 \big] \Big\} \end{aligned} \tag{6.9}$$

To facilitate minimisation, this expression is linearised using a 1st order Taylor Series expansion of the warped image intensity and gradient fields [7]: $I(\mathbf{W}(x,y;\mathbf{p}+\hat{\mathbf{p}}))$,$I_x(\mathbf{W}(x,y;\mathbf{p}+\hat{\mathbf{p}}))$, and

$I_y(\mathbf{W}(x,y;\mathbf{p}+\hat{\mathbf{p}}))$:

$$\begin{aligned} E_d = \sum_{x,y} \Big\{ & (I(\mathbf{W}(x,y;\mathbf{p})) + \nabla I \frac{\partial \mathbf{W}}{\partial \mathbf{p}} \hat{\mathbf{p}} - J(x,y))^2 + \\ & \lambda \big[(I_x(\mathbf{W}(x,y;\mathbf{p})) + \nabla I_x \frac{\partial \mathbf{W}}{\partial \mathbf{p}} \hat{\mathbf{p}} - J_x(x,y))^2 \\ & + \big(I_y(\mathbf{W}(x,y;\mathbf{p})) + \nabla I_y \frac{\partial \mathbf{W}}{\partial \mathbf{p}} \hat{\mathbf{p}} - J_y(x,y))^2 \big] \Big\} \end{aligned} \tag{6.10}$$

Substituting Equation 6.10 into Equation 6.8, the resulting energy expression is quadratic in $\hat{\mathbf{p}}$, and is iteratively minimised by solving a sparse linear system using the method of least squares. To improve the computation time, the optimisation is again performed at three resolution levels, coarse to fine.

6.3 Quantitative Analysis

Similarly to the case of bleed-through removal, objective evaluation of registration accuracy is difficult in the absence of ground truth. In previous approaches the accuracy of the registration stage itself is generally not considered, but instead the focus of performance evaluation is on subsequent bleed-through restoration, [21, 35, 75, 76]. In [55] manual rigid registration on a number of bleed-through images is performed and the parameters of the manual warp compared with the automatically generated parameters - a similar approach is proposed here. 24 full manuscript image pairs from 15 different manuscripts on paper and vellum, varying in the amount of warping required and severity of bleed-through, have been manually registered using the same Bezier grid-based warp as used for the final registration stage of the bleed-through database in Chapter 5. By convention the verso side of each example is flipped, however as described in Section 6.1, to prevent any loss of resolution, the side with the smaller page area was registered to the larger. The images range in size from approximately 1800 × 1900 to 5000 × 8000 pixels, with average resolution of 300dpi. All examples barring one were imaged with interleaves (see Figure 6.13), and most of the images were again taken from the collections of Irish Script On Screen Project [6], with two exceptions. Firstly, in order to include an example of vellum pages with interleaves, one image pair was taken from the Early Manuscripts at Oxford University [4], and secondly, the example image pair without interleaves was obtained from the Google Books collections [5], however without any processing applied for online display purposes. Following registration, for each registered image pair, a mask of the text and bleed-through regions was created, and a ground truth two dimensional displacement field extracted from the masked region. Performance evaluation is based on the error between this displacement field and a similar field generated from the automatic registration method.

As bleed-through document images can vary widely in terms of resolution, text size, and bleed-through strength, comparing the mean displacement errors between images alone does not give any idea as to how much any subsequent bleed-through restoration method will be affected.

A misalignment of a few pixels will affect restoration of low resolution images, where characters may only be a few pixels wide, much more severely than high resolution images with much larger characters. This also means that taking average displacement error values between images, will not necessarily provide an accurate representation of registration performance. Therefore the registration error is evaluated in terms of text feature size. In the test images, the smallest features generating bleed-through were generally between 5 and 15 pixels across the narrowest point (referred to in what follows as 'x'); a misalignment of more a quarter of this width would be visibly noticeable, and more than half would cause serious problems for accurate bleed-through restoration. Therefore the proposed measures of accuracy are the percentage of pixel displacement errors less than half the width of the smallest bleed-through feature, $0.5x$ error, the percentage less than a quarter of the width, $0.25x$ error, and also the mean displacement error. Each metric is further ranked, similarly to the bleed-through metrics in Chapter 5, using Ranked Pairs (RP) voting [68].

6.3.1 Local warp parameter variations

A number of variations in the local warp parameters were tested in order to examine the effect that changes in each parameter has on the registration results. Firstly, the number of grid points was set to be the same for both horizontal and vertical axes, $n_x = n_y = n$, and evenly spaced over the page region in the image only, defined from the bounding box of the binarised outline from the initial global warp. The numbers of grid points examined were 10, 15, 20, 25, 30, and 50. Secondly the gradient difference weight, λ, was varied over 0.1, 1, 10, and 100. Finally, three variations of the content preserving energy were tested. Though the content preserving weight, α was fixed at 1, the influence of the weight selection for each triangle, w_s was examined, with variations chosen as (i) the variance of the enclosing grid cell, as described above in section 6.2; (ii) fixed at 0.5; (iii) fixed at 1. The latter two cases essentially represent a variation in the content preserving parameter.

Figure 6.4 shows average displacement error results over all the images tested for increasing gradient difference weights, λ, with each colour representing a different number of grid points n, and each shape representing a different triangle weight, w_s, variant. Full per image results are shown in Appendix B for completeness. There is a clear discrepancy between the $0.5x$ and $0.25x$ error and the mean displacement error results; As λ increases from 0.1-10 in general the $0.5x$ and $0.25x$ error results improve, however the mean error worsens. This highlights the problem with comparing mean error values over all images between variations, as the results for $\lambda = 10$ are skewed by poor performance on one image. In general there is little difference between the results where w_s is set as the variance ($\sigma^2, \times$ markers), and $w_s = 0.5$ (circle markers), with the $w_s = 0.5$ variations performing slightly better when $\lambda = 0.1$, and slightly worse when $\lambda = 100$. The results for $w_s = 1$ variations (triangle markers) are significantly worse than the others when $\lambda = 0.1, 1, 10$, but better when $\lambda = 100$. This is likely due to the fact that when λ is high, this

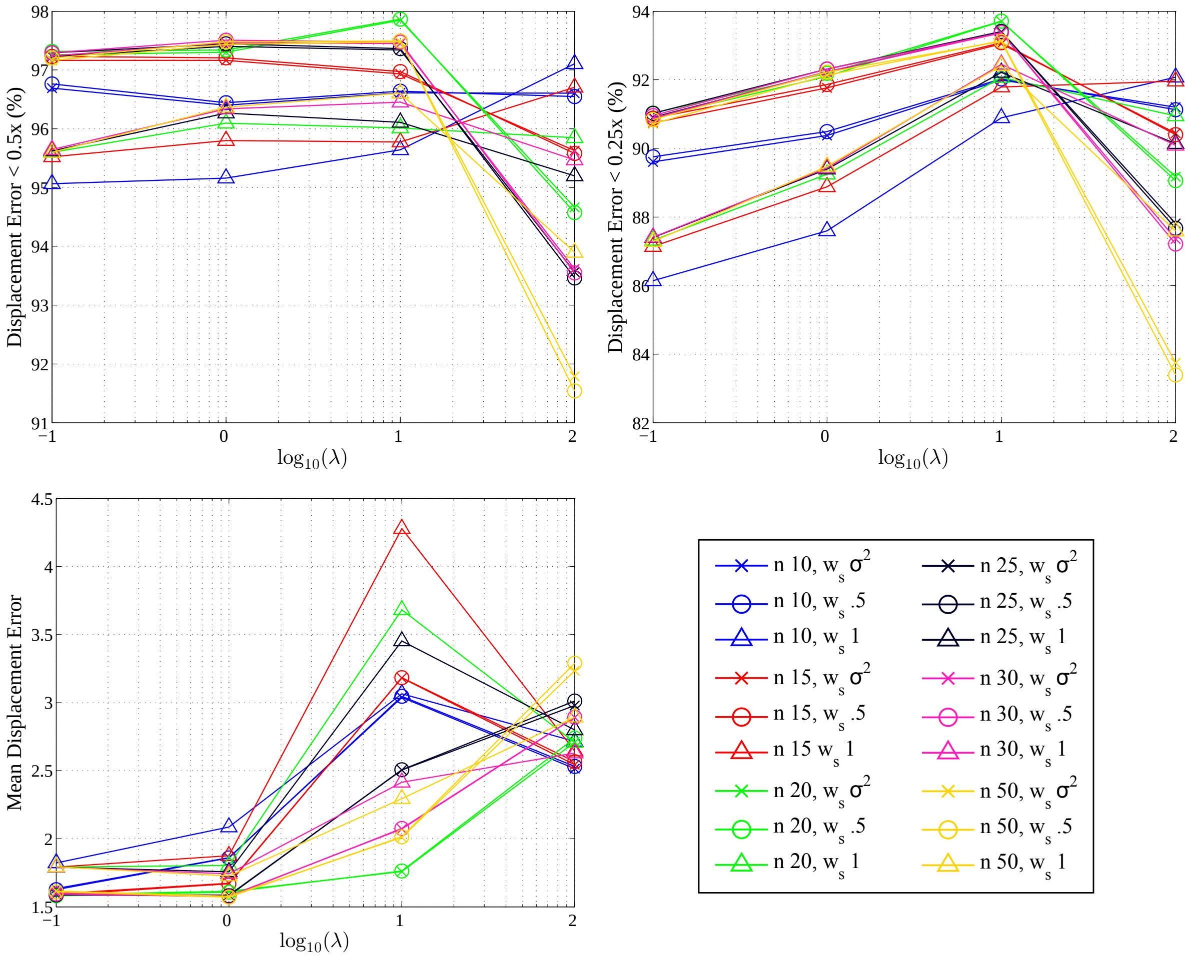


Figure 6.4: The effect of increasing λ with different fixed grid size (colour variations) and triangle weights (shape variations). Average errors over all 24 registered results. Top left: $0.5x$ error, right: $0.25x$ error, bottom right: mean displacement error.

disproportionately weights the data energy, and a higher w_s balances this out to a certain extent.

Figure 6.5 shows similar results for increasing number of grid points, n, with each colour representing a different gradient weight, λ, and each shape again representing a different triangle weight, w_s, variant. For $\lambda = 0.1, 1$ and all w_s variations the results improve with increasing n up to 20, with only minor variations for higher values. The performance for $\lambda = 10$, however, with $w_s = \sigma^2$, or fixed at 0.5 has a clear peak at $n = 20$ for the $0.5x$ and $0.25x$ errors, and a minimum for the mean error values. The performance for $\lambda = 100$ decreases with increasing n. Again the discrepancy between $0.5x$ and $0.25x$ and mean errors is clear. Though this analysis is by no means exhaustive, using these results and ranking the performance of each variation per image, the following parameters were empirically selected for comparison with previous approaches: $\lambda = 10, n = 20, w_s = \sigma^2$

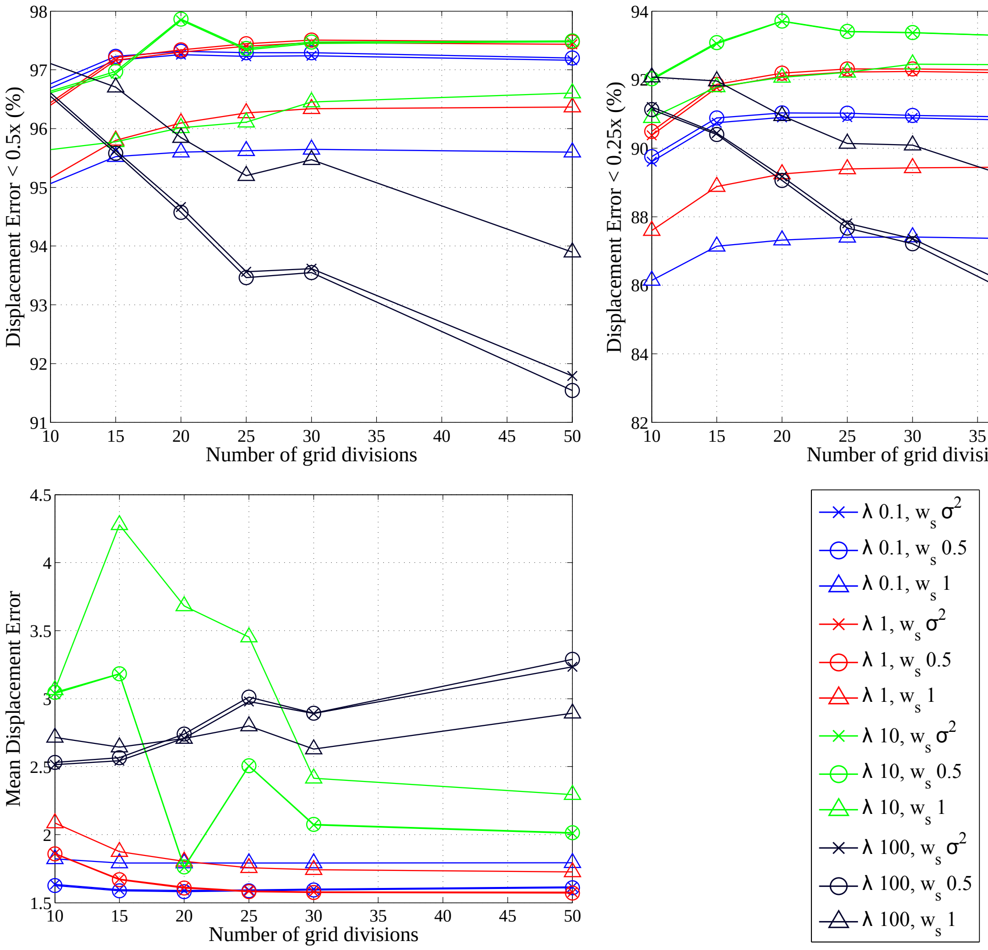


Figure 6.5: The effect of increasing n with different fixed λ (colour variations) and triangle weights (shape variations). Average errors over all 24 registered results. Top left: $0.5x$ error, right: $0.25x$ error, bottom right: mean displacement error.

6.4 Results & Discussion

To compare the performance of the proposed approach, the global affine method with SSD metric, in [21], and also the non-rigid method, with residual complexity metric in [76], were implemented. For the non-rigid method implementation [76], the initial results were poor due to mismatches or too few points in the corresponding control point selection. This is due to the fact that correspondences are based on the assumption of slanted text, which most of the test image pairs do not contain, and information from the binding on the opposite page also adversely affecting point selection. So instead of defining the uniform B-Spline grid from corresponding control points, it was generated automatically based on the image dimensions, with the resolution of the control mesh increasing with each image resolution level. This approach was also tested

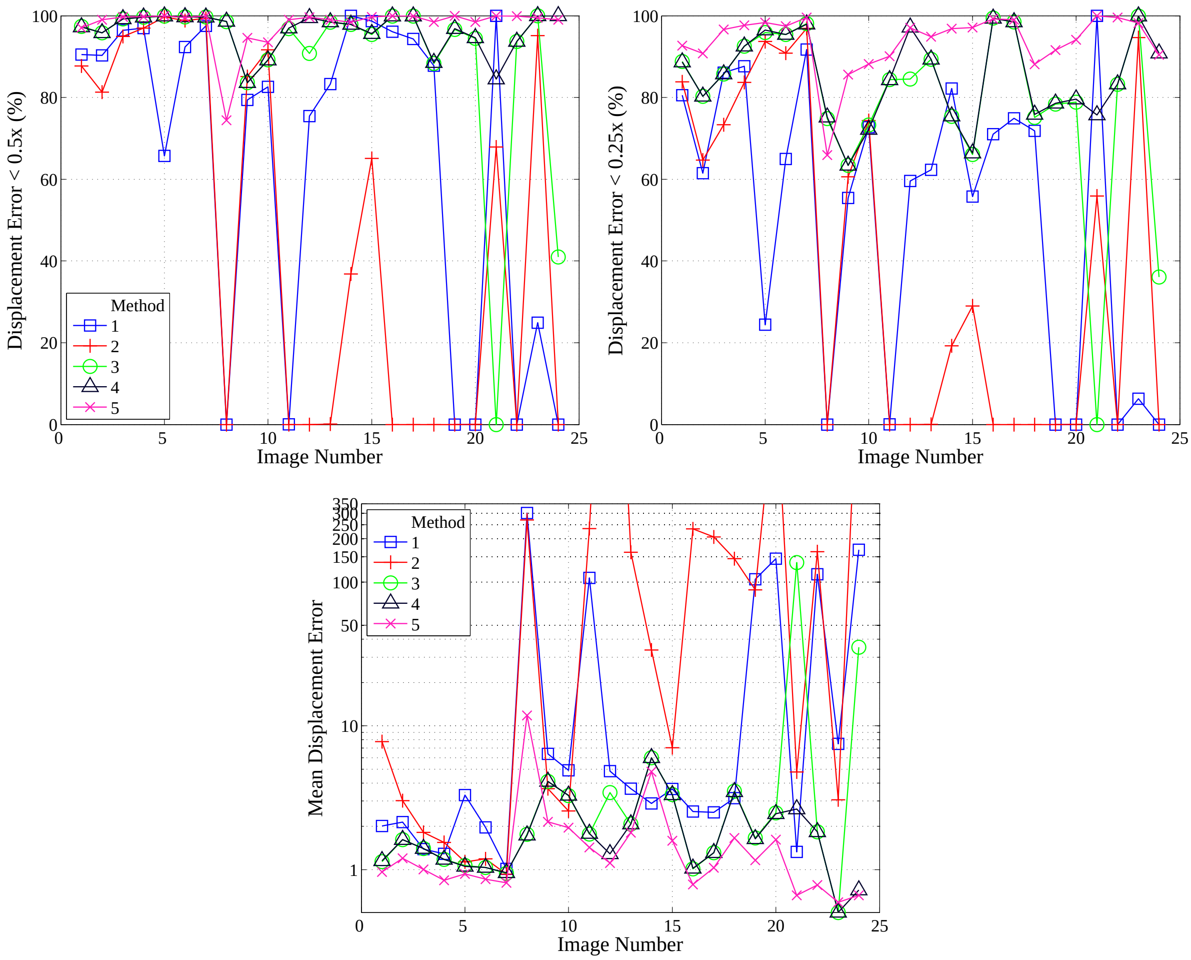


Figure 6.6: Comparison of the performance of each registration method on all 24 images. Top left: 0.5x error, right: 0.25x error, bottom: mean displacement error. There is a clear improvement in the performance of the nonrigid Method 2 ([76]) after the initial crop, Method 3 and outline warp, Method 4. The proposed Method 5 (pink) performs best in most cases.

on the images after the proposed automatic crop and outline alignment were performed. The comparison methods are each referred to as follows:
Method 1 - Global affine SSD-based approach [21].
Method 2 - Local nonrigid residual complexity based approach [76].
Method 3 - Method 2, following the proposed automatic crop stage.
Method 4 - Method 2, following the proposed automatic crop and global alignment stages.
Method 5 - The proposed approach.

Numerical Comparison

Figure 6.6 shows the 0.5x errors, 0.25x errors, and mean displacement errors for each method on

Table 6.1: Summary of displacement error with respect to ground truth

Method		1	2	3	4	5
	Mean	41.47	74.89	9.09	1.98	1.75
	Rank	4	5	3	2	1
$< 0.5x$	Min	0	0	0	83.62	74.41
	Max	100	99.71	100	100	100
	Mean	64.67	45.86	89.84	96.20	97.84
	SD	41.14	45.07	22.60	4.91	5.26
	Rank	4	5	3	2	1
$< 0.25x$	Min	0	0	0	63.37	65.91
	Max	100	96.92	99.87	99.87	100
	Mean	50.37	38.36	79.04	85.08	93.70
	SD	35.63	40.21	22.10	10.87	7.27
	Rank	4	5	3	2	1

each image tested, and these results are summarised in Table 6.1 with RP ranks for each metric shown. As expected, the the performance of the global SSD based Method 1 [21] (blue) on the full images is highly varied depending on the amount of local warping required, though it is clear that for some images the global affine model is sufficient. The nonrigid residual complexity based Method 2 [76] (red) in most cases performs worse than the global affine Method 1, and fails in two examples. As the initial warp used for Method 2 is also affine, this suggests that the use of residual complexity as an error metric for the initial alignment stage in some cases may be detrimental to the performance of the subsequent local refinement. The results for Methods 3 and 4 (green and black) on the cropped, and cropped+outline warped images respectively show a significant improvement over Method 2, highlighting the performance of the proposed initial crop and outline warp. The proposed Method 5 (pink) performs best in most cases, this is most noticeably shown in the $0.25x$ error results - Method 5 is on average 8.6% higher than Method 4. The summary of results in Table 6.1 again show that the proposed Method 5 performs best overall, and that there is a large improvement in the performance of the previous nonrigid approach following the proposed initial processing and warp stages.

Visual Comparison

Figures 6.7-6.10 show a visual comparison of sample results of each method, with both output registered images and displacement errors shown for each example. Again, full results may be found in the accompanying CD. For many of the images tested, the largest registration errors for each method were near the binding, due to the page curvature. Figures 6.7 and 6.8 show the results for an example image pair where this is the case. As the uncropped warped image contains textual information from the opposite page, but the reference image opposite page is blank, this causes a large registration error for Method 2 near the binding. Methods 3, 4, and 5 perform better, as the opposite page information has been mostly removed. However, there are still large regions of high error close to the binding. As a similar result was achieved with the densest grid tested ($n = 50$), this suggests that in some images, where the use of a regular grid over the entire image can not achieve optimal results, the binding region needs to be treated separately, either in an initial page flattening stage, or with added constraints in the local warp refinement.

Figures 6.9 and 6.10 show results for an example image pair where the proposed local refinement stage does not perform well. In this example, there is a large region of background only in the bottom half of both images, and faint bleed-through in the reference image. The proposed crop and outline warp both perform reasonably well, as can be seen in the improvement in results from Method 2 to Methods 3 and 4, though with some larger errors close to the binding. The poor performance of the proposed method is likely due to the large background region - in this area the content preserving triangle weights are relatively low due to the low intensity variance, and as a relatively high weighting is given to the data term in the selected warp parameters, this region has no constraint to prevent large warp displacements, which then affect the connected text regions which have higher content preserving weight due to the higher variance. Increasing the number of grid divisions, and decreasing the gradient difference weight improves the result on the image, and as this effectively increases the influence of the content preserving energy over the whole image, it suggests that the content preserving weight may need to be modified in such cases with large regions of background.

Figures 6.11 and 6.12 show results for an example image pair where the visible width of both sides of the page is different. In this example the global Method 1 performs well - the global affine model is sufficient for the required warp. Method 3 performs worse than Method 2, due to the fact that in this case the crop stage actually reduces the alignment between the two sides compared to the original uncropped images. The initial affine warp stage of Method 3 then results in undesirable text stretching and large misalignment and the local refinement stage is unable to recover the original text proportions. This highlights the fact that local warp refinement methods rely on good initial global warps. In most cases the crop stage actually results in a rough alignment between the two pages, but where this is not the case, as discussed in Section 6.1, an initial global warp constrained to uniform scaling will produce better results than a full affine warp.

This dependence of local refinement methods on good initial warps is further emphasised in Figures 6.13 and 6.14, which show results for the example image pair without interleaves (number 31), and which contain large amounts of binding information, non-uniform page surroundings and some non-textual interference on the pages themselves. The outline warp still performs well in this case as the surrounding region is mostly darker than the page background, with brighter regions small enough to be removed by the morphological opening and closing. This is evident in the improvement in results between Method 3, with no outline warp, and Method 4 after

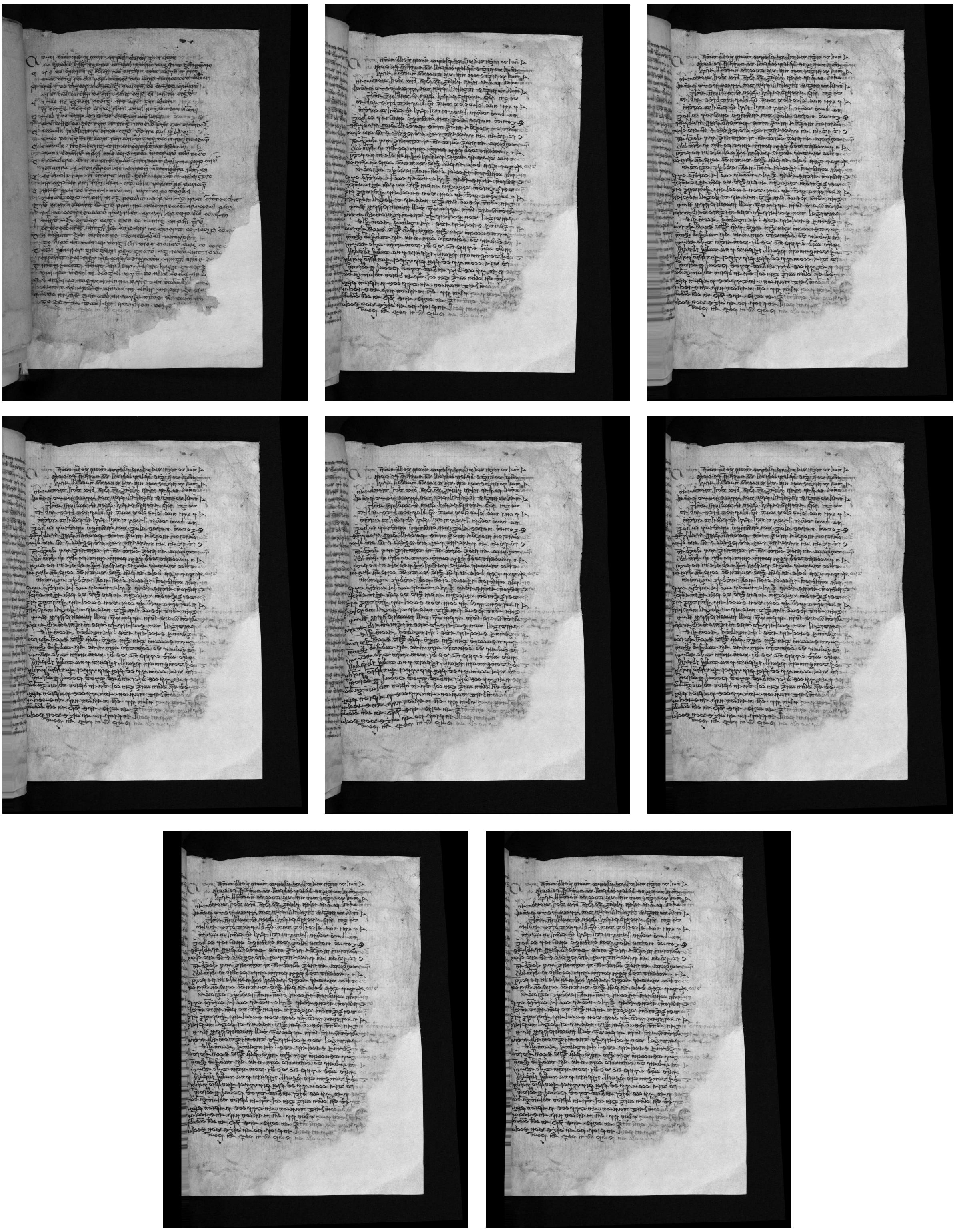


Figure 6.7: Comparison of registered results for image 1, **M39**, where the curvature at the binding leads to registration errors in each method. Top row left to right: Original reference image J, original warped image I, manually registered image. Middle row left to right: results for Methods 1, 2, and 3. Bottom row: Results for Methods 4 and 5.

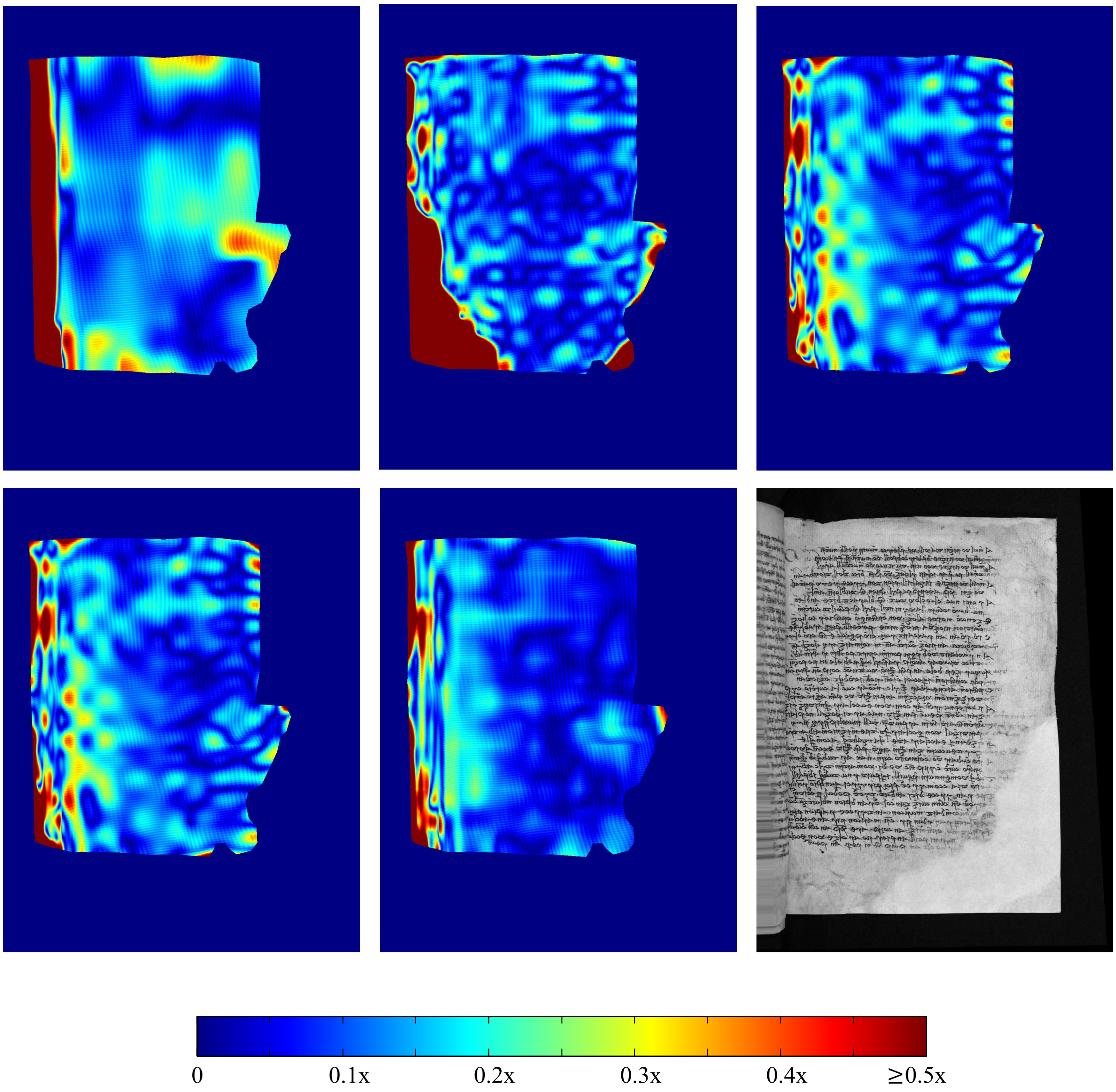


Figure 6.8: Comparison of displacement errors corresponding to the results in Figure 6.7. Top row left to right: displacement errors for Methods 1, 2, and 3 Bottom row: Displacement errors for Methods 4 and 5, and manually registered warped image for reference. **M39**.

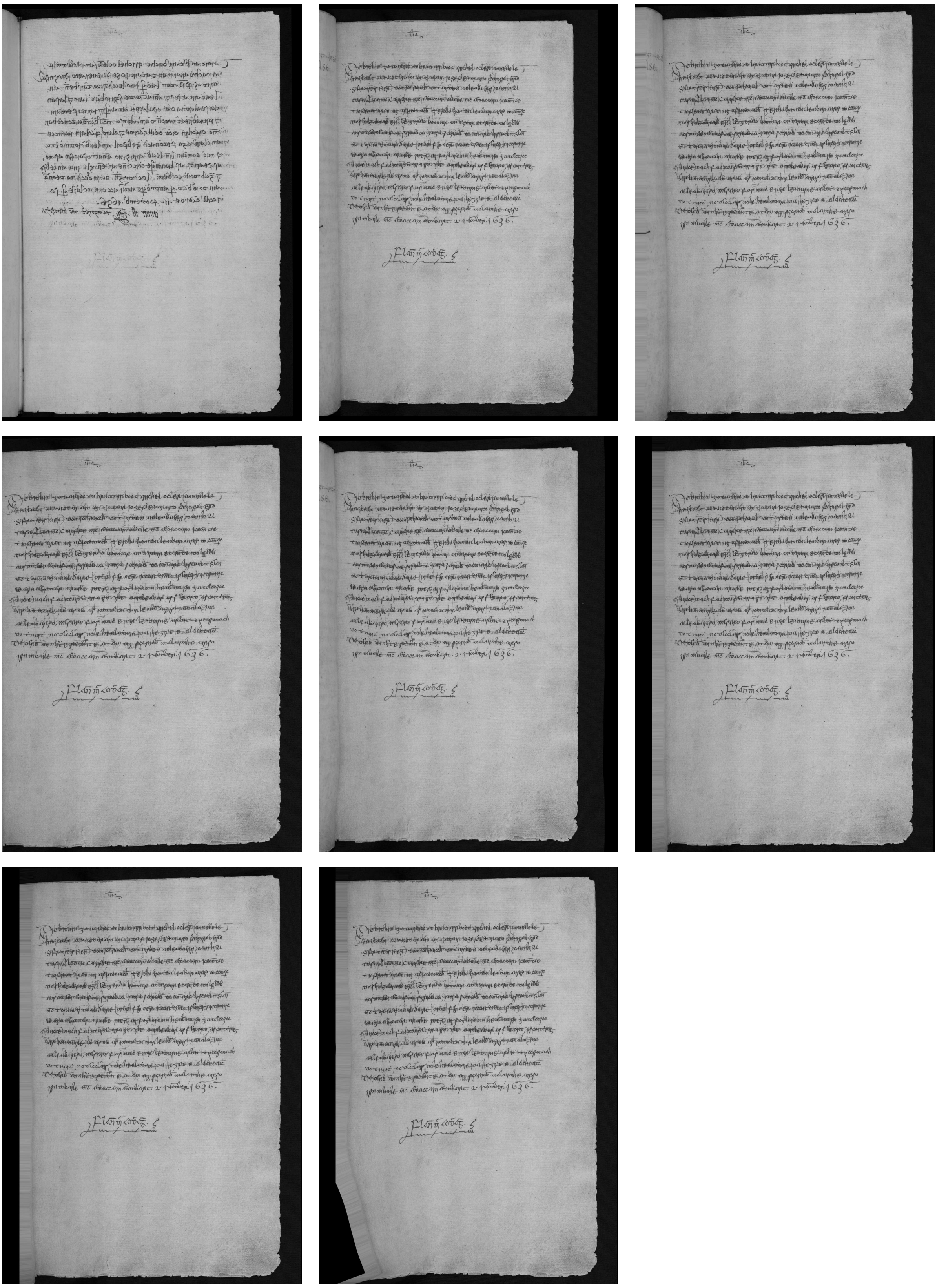


Figure 6.9: Comparison of registered results for image 8, **M32**, where the proposed Method 5 performs poorly. Top row left to right: Static image J, warped image I, manually registered image. Middle row left to right: Results for Methods 1, 2, and 3. Bottom row: Results for Methods 4 and 5.

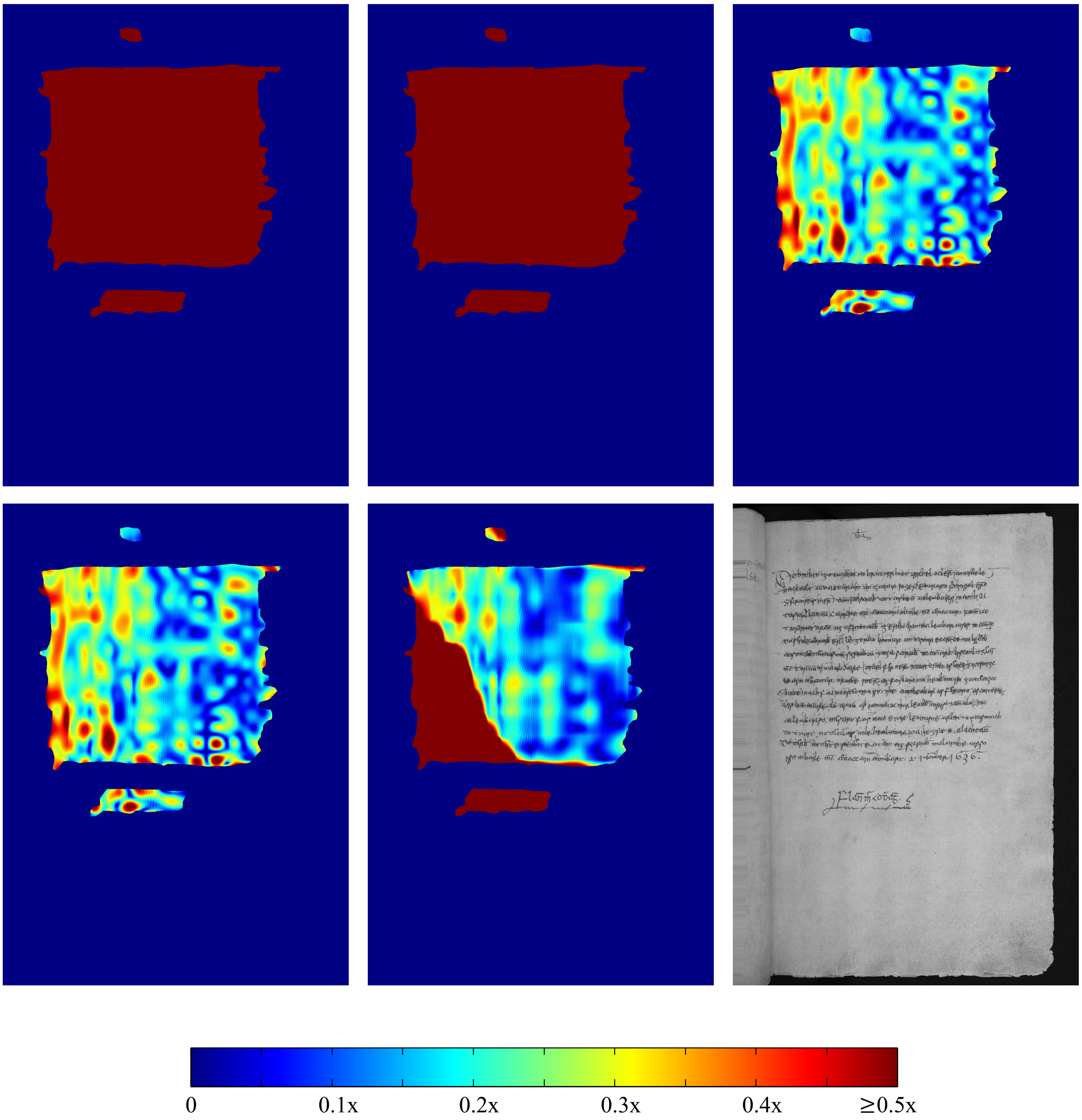


Figure 6.10: Comparison of displacement errors corresponding to the results in Figure 6.9. Top row left to right: Displacement errors for Methods 1, 2, and 3 Bottom row: Displacement errors for Methods 4 and 5, and manually registered warped image for reference. **M32**.

Figure 6.11: Comparison of registered results for image 21, **M02**, where the visible width of each side of the page is different. Top row: reference image J (left), warped image I. 2nd row: manually registered image(left), result for Method 1. 3rd row: Results for Methods 2 (left) and 3. Bottom row: Results for Methods 4 and 5.

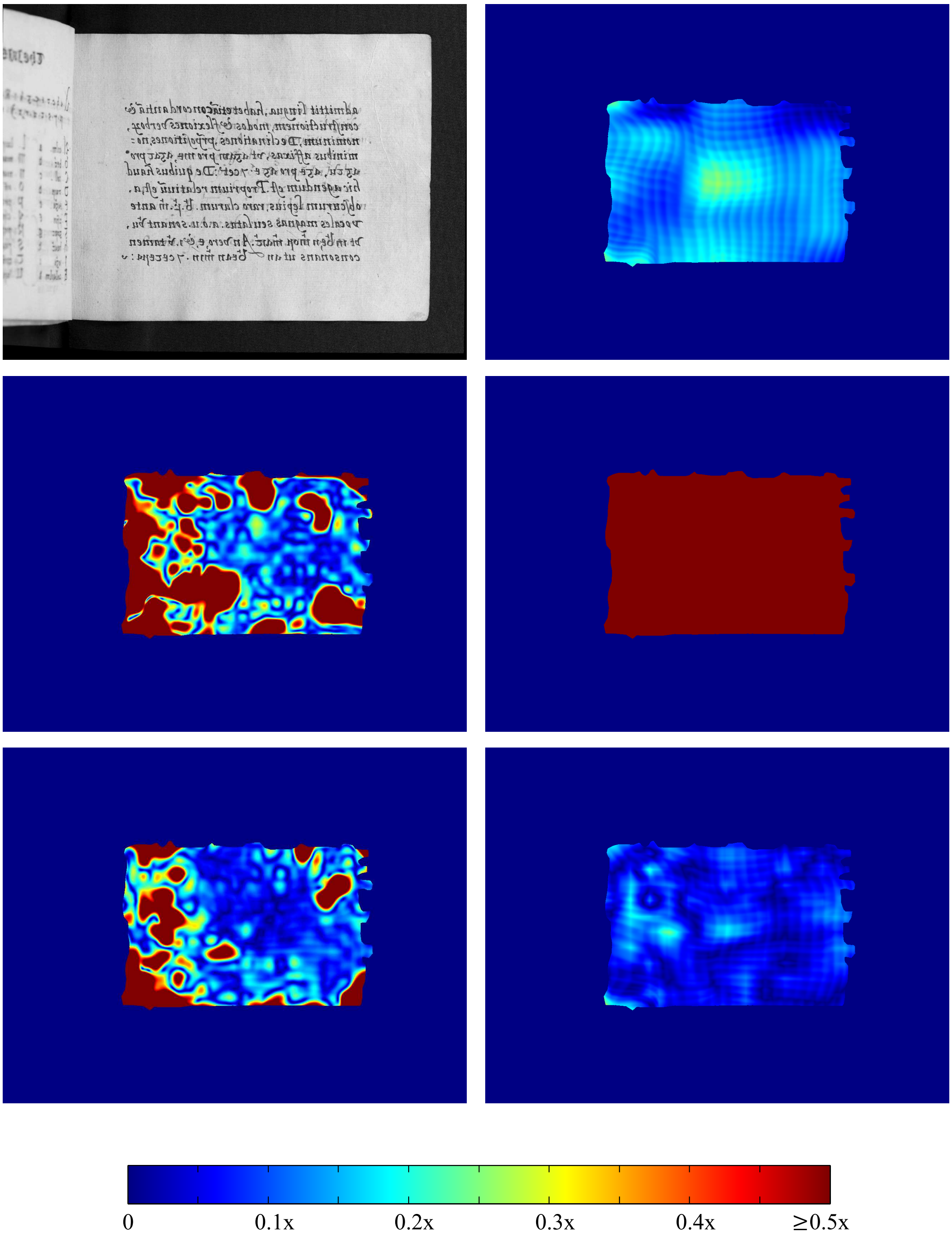


Figure 6.12: Comparison of displacement errors corresponding to the results in Figure 6.11. Top row: manually registered warped image(left), displacement error for Method 1. Middle row: Displacement errors for Methods 2 (left) and 3. Bottom row: Displacement errors for Methods 4 and 5. **M02**.

the outline warp, and also the results of the proposed method. This result is promising as it suggests that the proposed outline warp method is robust to a larger number of document types than only those imaged with interleaves.

Though the implementation of each approach was not optimised for speed, it was noted that the optimisation of Methods 2, 3, and 4 converge extremely slowly due to the evaluation of the DCT of the difference in intensities between reference and warped image at each iteration in order to calculate the residual complexity. In contrast, the proposed approach converged significantly faster.

6.5 Summary

This chapter has presented an automatic two stage registration method for recto-verso images of documents suffering from bleed-through degradation. In the initial stage binarised images of two sides are automatically cropped to remove any binding or opposite page information, and then the outlines are registered using a global similarity transformation. In the second stage a local grid-based warp is applied, using both the sum of the squared difference in image intensities and also gradient fields as a warp metric, with a content preserving smoothness penalty. A numerical evaluation approach for bleed-through registration results has also been presented, using manually registered manuscript images for ground truth. Test results show that the proposed initial page outline registration can improve the robustness and performance of previous approaches, and that the proposed local refinement stage is significantly faster and achieves better registration accuracy in most cases.

In order to further test the robustness of the proposed approach, as for bleed-through removal, the ground truth test set would need to be increased to include a greater variety of different document types based on size, writing media, and image capture, focussing especially on documents without interleaves.

A further application of the initial outline crop and warp stage, unrelated to bleed-through restoration, has been found in automatically creating browsable ‘flipbook’ display copies of imaged documents. The binding crop stage allows adjacent pages to be joined correctly at the binding, and the outline warp can be used firstly to align all corresponding recto and verso images, and then to resize all of the pages to one template page size rather than cropping all of the images and risk losing some page information.

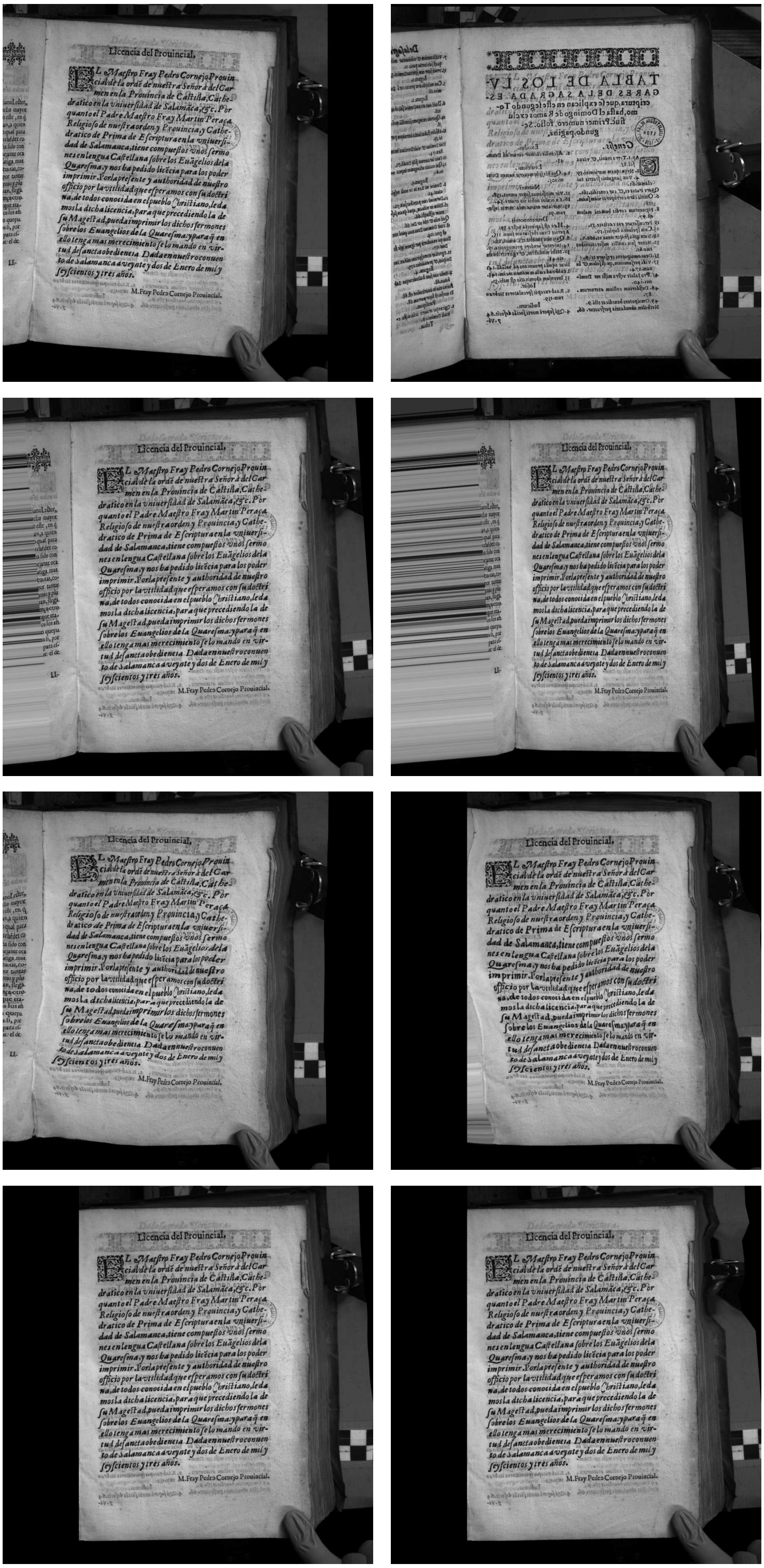


Figure 6.13: Comparison of registered results for image 24, **M08**, where there is no interleaf and outside information causes registration errors.

Top row: reference image J (left), warped image I.

2nd row: manually registered image(left), result for Method 1.

3rd row: Results for Methods 2 (left) and 3.

Bottom row: Results for Methods 4 and 5.

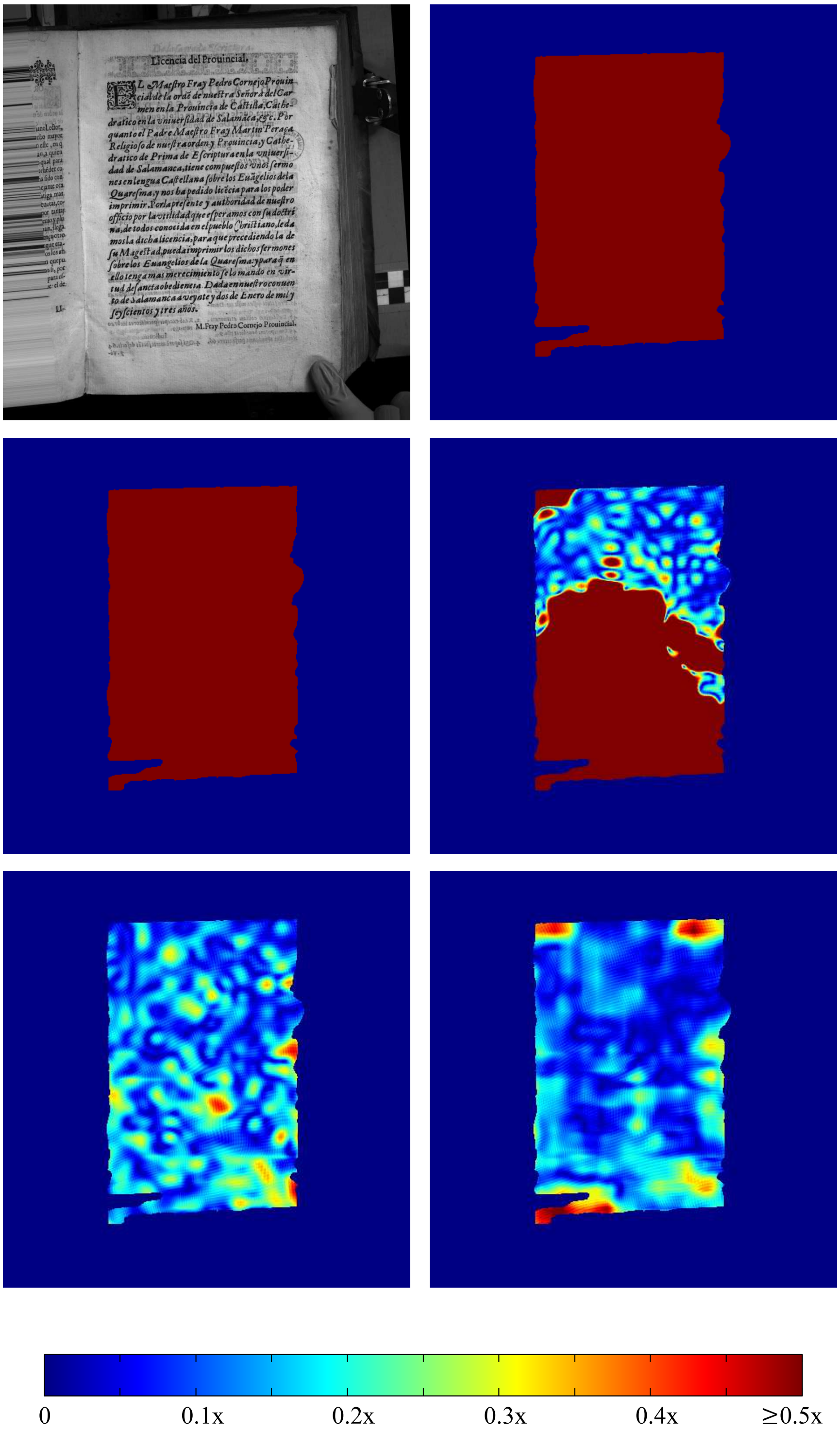


Figure 6.14: Comparison of displacement errors corresponding to the results in Figure 6.13. Top row: Manually registered warped image (left), displacement error for Method 1. Middle row: Displacement errors for Methods 2 (left) and 3. Bottom row: Displacement errors for Methods 4 and 5. **M08**.

7

Conclusion

This thesis has focused on automatic degraded document bleed-through restoration. In what follows, a brief review of the work presented in this thesis is provided, and some possible directions for future work are discussed.

7.1 Restoration

Chapter 2 presented a review of previous work towards bleed-through restoration. The different approaches were classified based on the intensity and spatial models used, whether they operate on one or both sides of the pages, and the type of output that they generate. The most challenging aspects of bleed-through restoration were shown to be the overlapping bleed-through and foreground intensity ranges, and overlapping regions of foreground on both sides of the page. It was further noted that very little emphasis has been placed on preservation of the document appearance.

Chapter 3 introduced a linear model based approach to bleed-through restoration that aims to improve legibility whilst maintaining the document appearance. The observed degraded images are modelled as a per pixel linear combination of the original clean pages, plus some proportion of the corresponding reverse side, and bleed-through explicitly constrained to non-foreground regions only. Visual results showed that this approach performs well on images with relatively light bleed-through degradation, preserving the character of the documents whilst improving legibility. However, in more severe cases, with large regions of dark bleed-through, the approach can not work as well, though legibility is not affected detrimentally - the foreground text is preserved. This highlights the dependence on good initialisations, but also the limitations of per pixel based spatial models in removing large regions of severe bleed-through.

For these reasons, a second bleed-through restoration approach was presented in Chapter 4, in which the classification and restoration stages are decoupled to focus explicitly on bleed-through identification. In this approach classification is performed jointly on both sides of the page by performing joint intensity histogram segmentation, followed by connected component based refinement on the corresponding image labelling. Restoration then follows, using exemplar based inpainting in order to preserve the document appearance. This approach is more powerful at removing severe bleed-through than the previous work mainly due to the connected component approach to label field refinement, which avoids the limitations of pixel based spatial information by extending the neighbourhood over a larger region. The use of exemplar-based inpainting for the restoration has the advantage over the estimation based approach firstly as no assumption needs to be made about the distribution of the background medium, and secondly, as the inpainted regions are samples of background texture this adds a further level of document appearance preservation.

A database of 25 registered recto-verso pairs with ground truth binary foreground text masks was presented in Chapter 5, which, at the time of writing, is the only such database freely available in the public domain. A numerical evaluation approach was proposed based on the database to provide a fair comparison between different output types from different method, where all output result images were binarised and subsequently compared to the ground truth. The two proposed approaches from Chapters 3 and 4, were evaluated visually and numerically in comparison with three previous non-blind methods [29,35,44], using the database and evaluation methodology. The results showed that overall the approach from Chapter 4 performs best in terms of obtaining a balance between bleed-through removal and foreground text conservation. It was further noted that the classification based approach has the advantage over model based in that it can be modified to produce binary, pseudo-binary, and textured outputs, using the generated label fields, depending on the required output.

User Assisted vs Automatic

In Chapter 5, it was noted that a significant improvement in the results for some image pair examples could be achieved with manually tuned parameters. The user assisted paradigm is a common theme in many document analysis and restoration areas, as an interactive approach to restoration can potentially guarantee optimal results within the bounds of what is feasible with the technology available. However, as previously discussed, two distinct end user groups can be identified for use of document restoration techniques: those pursuing individual study or working on relatively small imaging projects, and large scale digitisation projects. It is clear that large scale projects, such as Google Books [5], necessarily require efficient and fully automated restoration procedures, as per-image tuning is simply infeasible due to the quantity of documents that are processed. In smaller scale projects, for example digitisation of a single collection for a new scholarly edition, or simply for people studying individual documents, full automation is not an absolute requirement, and a better approach would be a semi automated system, where the user is free to select certain parameters and inputs, and to interact with the output in order to achieve the optimal results for their own purposes.

7.2 Registration

Chapter 6 presented a framework for full page recto-verso registration including initial coarse global alignment using the page outline, followed by local grid point based refinement using the sum of squared differences (SSD) between image intensities and gradients as a similarity measure. The most significant contribution of this approach is the use of the page outline and uniform scaling for the initial global alignment stage, as this is more robust to examples with little or no bleed-through degradation than previous methods. A numerical evaluation approach specifically for bleed-through registration was also presented, where the estimated displacement fields are compared against manually generated ground truth displacement fields. The proposed error measure was calculated in terms of the smallest bleed-through feature width in order to be able to compare performance over a range of different image resolutions and document types. Comparisons against previous methods showed that the proposed approach is more robust to a variety of different document types, and that the initial binding crop and outline global registration can improve the performance of previous approaches by providing a good initial alignment.

7.3 Future work

Database and Evaluation

The database described in Chapter 5 was designed to overcome the obstacle of obtaining access to good quality bleed-through example images, and to be of use in the initial testing and development of bleed-through removal approaches. However, though the database does contain a variety of degradation strengths, and script styles/sizes, all the images are of the same resolution, and manuscript only, on paper or vellum. As bleed-through degradation is a common problem in a wide range of documents, and as new approaches to restoration are continually being developed, the best way forward for objective evaluation would be to expand the current database into a larger benchmarking dataset that is representative of the full range of different document types that suffer from bleed-through. The benefit of this would be to test the robustness of proposed approaches to a larger range of documents. Extension to full colour images would also be beneficial. Further evaluation methodologies could then be developed to take into account the different motivations and goals of restoration, in order to create a standard similar to the DIBCO series for binarisation evaluation. There would need to be collaboration between different research groups in order to make this feasible, as this would provide a better perspective on the different motivations for performing bleed-through restoration, and ensure that the database and evaluation would be relevant to each of these.

Restoration

Following the previous discussion, to proceed in the direction of full automation, it is necessary for the proposed approaches to be tested on a wider range of images in order to develop a more robust content based parameter selection. As discussed in Chapter 5, gradient information could be incorporated to the framework in Chapter 4 to improve the connected component refinement stage, enabling the splitting components at obvious character edges in one or both images, where they may contain both true foreground and dark bleed-through. The gradient information could further be incorporated to counter the 'halo' effect in textured restored results where bleed-through has spread horizontally; examining the gradient of restored images at the edge of bleed-through regions could give an indication of this effect. A further direction would be the use of full colour information in the classification stage - many manuscripts use colouration to emphasise certain words or phrases, and this information is often lost in the conversion to grayscale.

An interesting application of the framework presented in Chapter 4 that could be explored, is extracting high quality images of watermarks as they are found in old paper. Collections of watermarks are very important for studying the history of paper manufacture, and for dating old documents. A common digitisation approach to highlight watermarks is to image the page both normally and with back-lighting. The two images are then registered and foreground information removed to reveal the watermark [12, 32]. There are obvious parallels with bleed-through restoration, and as verso information is often made visible on the recto side as a result of back-lighting, the framework presented in Chapter 4 could be extended to incorporate two further back-lit images for watermark extraction. The registration approach proposed would also be useful in this instance for registering all four images based on the paper outline.

Registration

Similarly to the case of bleed-through registration, approaches need to be robust to a wide variety of different document types. The main limitation of the approach proposed in Chapter 6 is the assumption that the document pages have been imaged with interleaves. This precludes the necessity of detecting the page outline as the contrast with the interleaf makes it distinct. In the absence of interleaves, there are potential sources of interference that could result in the binarised image and page outlines not coinciding, for example the edges of other pages underneath. Though the test images did contain an example pair without interleaves on which the proposed approach performed well, further testing would be required to determine the robustness to the method to such cases, and whether a page outline detection step would be necessary. It may be the case that the top and bottom edges only are required for the global registration due to the enforcement of uniform scaling. The dataset used for testing therefore needs to be expanded to include a wider variety of document image types. A further direction identified for improvement of the local refinement stage would be to consider separate treatment for regions close to the binding, especially where there is a high degree of curvature. These regions could be automatically detected based on the curvature of the page outline.

As mentioned in Chapter 6, an application of the outline warp approach that could be explored is in improving the digital 'flipbook' representations of documents by warping all of the pages to a set template. A further application in a more closely related area is that of multispectral document image registration. In order to perform analysis of images taken at different wavelengths, they must first be registered. This is, to a certain extent, simpler than bleed-through registration as the images to be registered are of the same side of the page, however different information is often visible under different wavelengths, especially in palimpsest images, so the proposed outline warp stage and local refinement could be more robust than a feature based approach in such cases.

7.4 Final Remark

Image based document enhancement and restoration is a thriving area of research. As the prevalence of digitisation of archive and library materials increases, so too does the demand for image processing based tools with which to analyse the resulting digital resources. However, the number of researchers studying original sources who could benefit from such techniques far outweighs the number who know that they could exist. Though this situation is improving [27], it is important to raise the profile of the potential of image processing based document enhancement and restoration techniques amongst the 'non-technical' research community, encouraging collaboration so that subsequent developments can be more end user driven.

A

Full Results for Chapter 5

Results for the comparative analysis of different bleed-through methods were presented in Section 5.3. Comparison of *BgError* and *FgError* results over all images in the database were shown in Figure 5.8. Due to the variation in results per image, the remaining metrics were compared via ranking performance in Figures 5.9, and 5.10. Full results of the *TotError* and DIBCO metrics are included here for completeness. Figure A.1 shows the *TotError* results, Figure A.2 shows the F_1-score and F_{1ps}-score results, and Figure A.3 shows the $PSNR$ and DRD results.

To summarise the comparative results between methods, the overall performance of each method was ranked for each metric using Ranked Pairs Voting (RP) shown in Table 5.1. The pairwise ranking results used to obtain the RP ranks for each metric are shown in Tables A.1 and A.2, where each entry represents the percentage of images that the method listed vertically was ranked higher than the method listed horizontally.

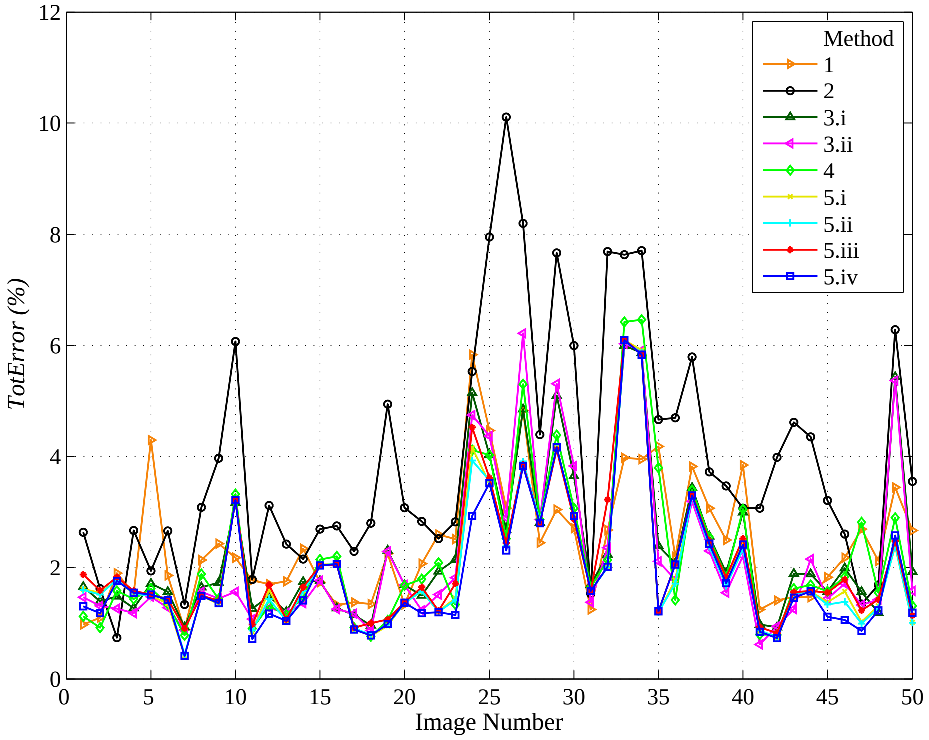


Figure A.1: *TotError* for all 50 images restored with each method.

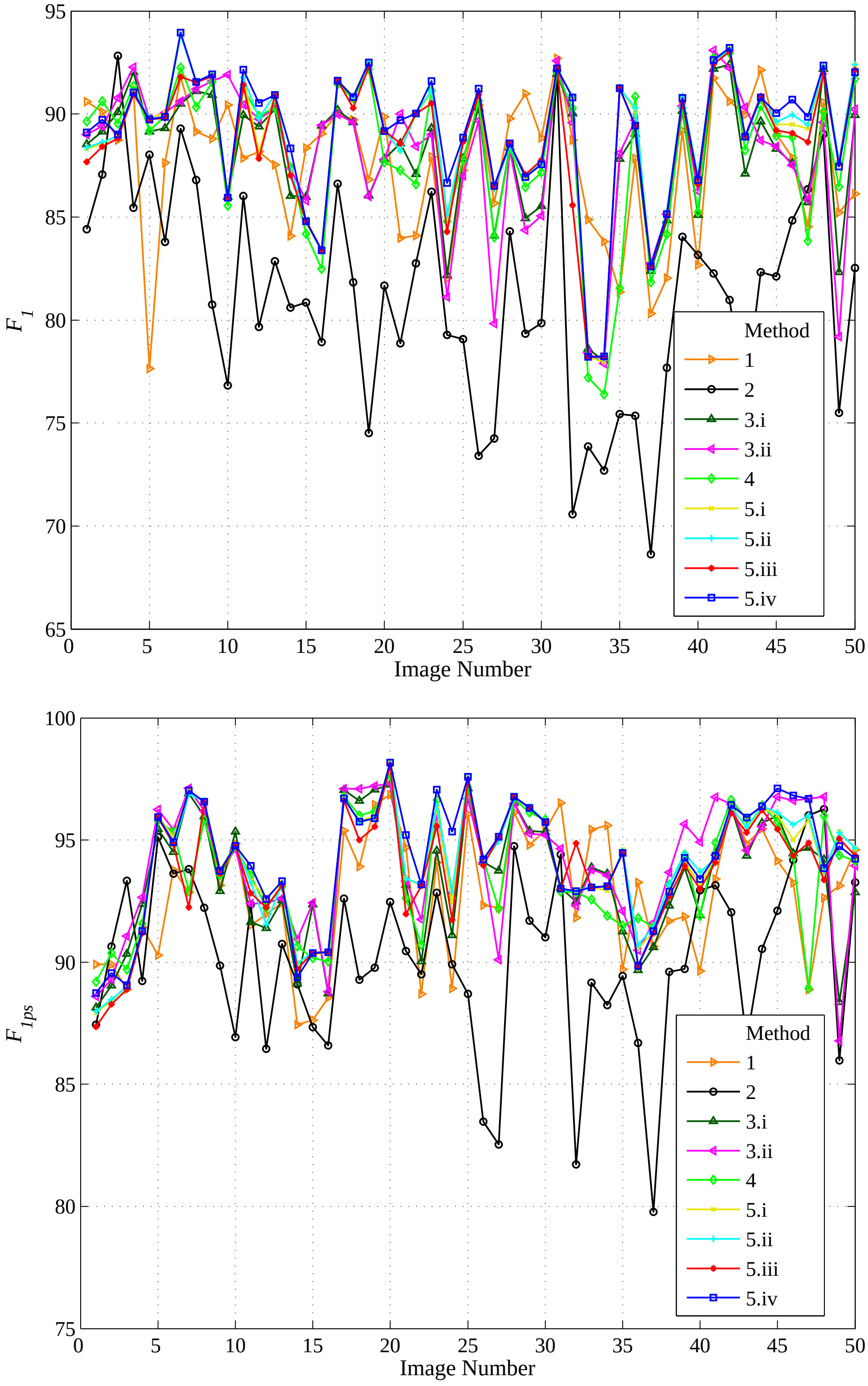
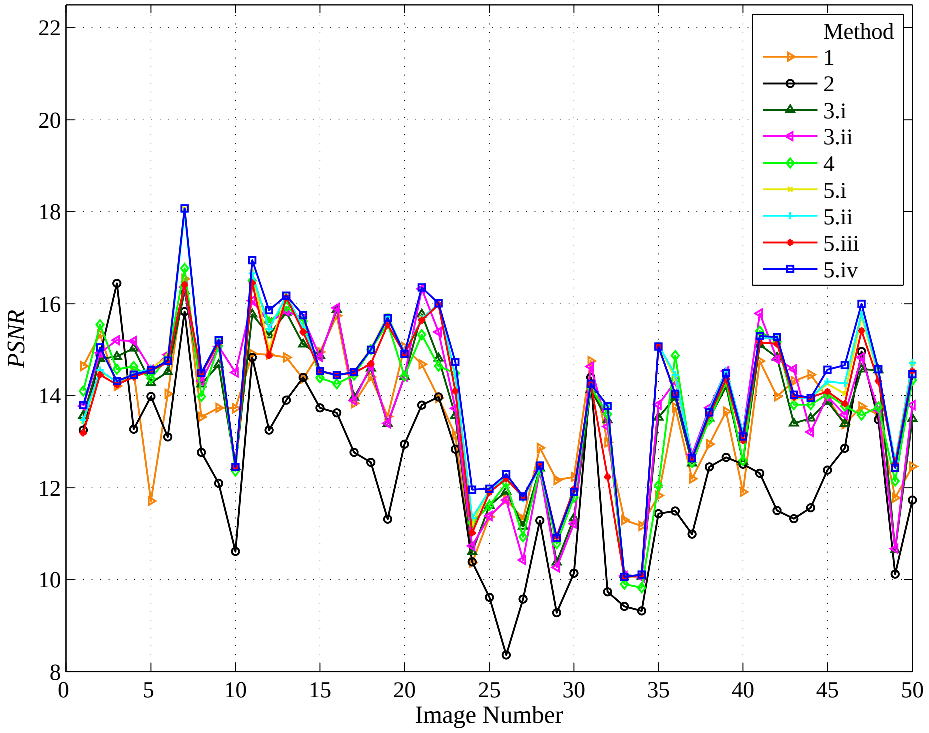


Figure A.2: F_1-Score (top) and F_{1ps}-score (bottom) for all 50 images in the database restored with each method.



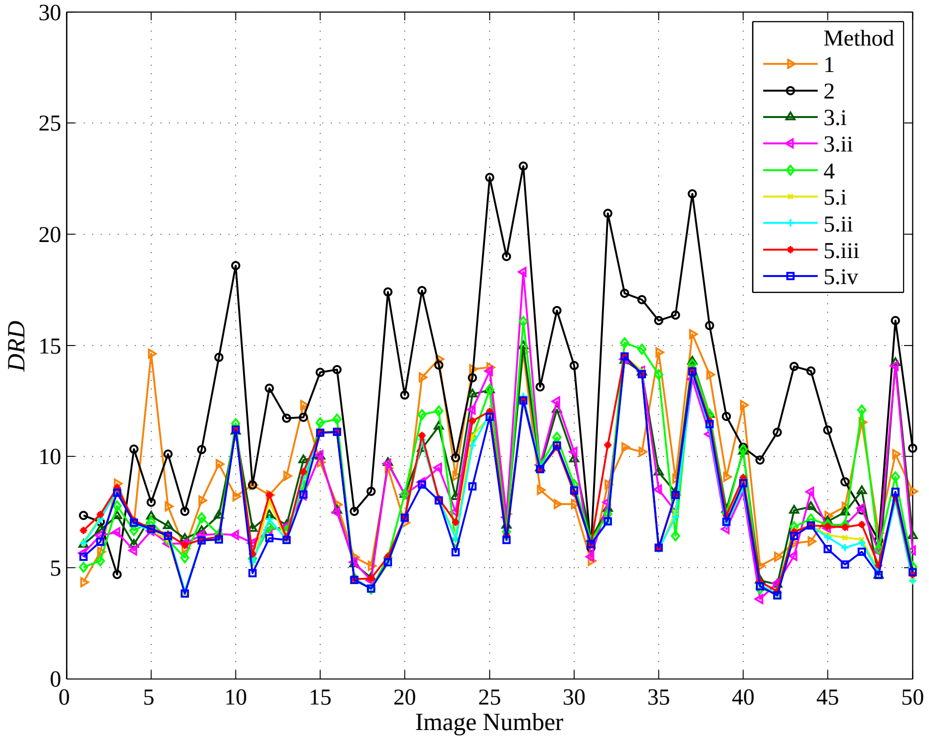


Figure A.3: $PSNR$ (top) and DRD (bottom) for all 50 images in the database restored with each method.

Table A.1: Pairwise method rank comparison (%) for the bleed-through error metrics.

	FgError								
	1	2	3.i	3.ii	4	5.i	5.ii	5.iii	5.iv
1	0	98	54	68	52	48	46	48	46
2	2	0	0	0	0	0	0	0	0
3.i	46	100	0	94	58	30	24	24	18
3.ii	32	100	6	0	40	16	14	14	12
4	48	100	42	60	0	8	6	8	12
5.i	52	100	70	84	92	0	38	12	48
5.ii	54	100	76	86	92	62	0	18	48
5.iii	52	100	76	86	92	86	82	0	76
5.iv	54	100	82	88	88	52	52	24	0

	BgError								
	1	2	3.i	3.ii	4	5.i	5.ii	5.iii	5.iv
1	0	4	28	12	14	22	18	24	20
2	96	0	100	94	96	96	96	96	96
3.i	72	0	0	2	28	26	28	48	22
3.ii	88	6	98	0	68	74	76	82	70
4	86	4	72	32	0	74	70	82	66
5.i	78	4	74	26	26	0	58	94	42
5.ii	82	4	72	24	30	40	0	88	38
5.iii	76	4	52	18	18	6	12	0	14
5.iv	80	4	78	30	34	58	62	84	0

	TotError								
	1	2	3.i	3.ii	4	5.i	5.ii	5.iii	5.iv
1	0	84	34	28	30	28	28	34	28
2	16	0	6	2	6	4	4	4	4
3.i	66	94	0	30	36	24	22	26	14
3.ii	72	98	70	0	54	42	40	54	34
4	70	94	64	46	0	24	22	36	16
5.i	72	96	76	58	76	0	40	84	38
5.ii	72	96	78	60	78	60	0	82	40
5.iii	66	96	74	46	64	16	18	0	18
5.iv	72	96	86	66	84	62	60	82	0

Table A.2: Pairwise method rank comparison (%) for the DIBCO error metrics.

	$\mathbf{F_1}$								
	1	2	3.i	3.ii	4	5.i	5.ii	5.iii	5.iv
1	0	92	42	46	34	32	30	34	30
2	8	0	6	4	6	2	2	2	2
3.i	58	94	0	42	38	24	18	22	12
3.ii	54	96	58	0	46	30	30	40	28
4	66	94	62	54	0	20	20	28	14
5.i	68	98	76	70	80	0	42	72	40
5.ii	70	98	82	70	80	58	0	74	40
5.iii	66	98	78	60	72	28	26	0	24
5.iv	70	98	88	72	86	60	60	76	0

	$\mathbf{F_{1ps}}$								
	1	2	3.i	3.ii	4	5.i	5.ii	5.iii	5.iv
1	0	76	30	26	22	22	24	28	22
2	24	0	16	4	14	10	8	14	8
3.i	70	84	0	24	36	30	32	40	24
3.ii	74	96	76	0	58	54	54	64	42
4	78	86	64	42	0	46	42	52	36
5.i	78	90	70	46	54	0	46	86	42
5.ii	76	92	68	46	58	54	0	86	44
5.iii	72	86	60	36	48	14	14	0	16
5.iv	78	92	76	58	64	58	56	84	0

	PSNR								
	1	2	3.i	3.ii	4	5.i	5.ii	5.iii	5.iv
1	0	86	36	32	30	30	30	36	30
2	14	0	8	4	6	6	6	8	4
3.i	64	92	0	30	34	22	20	22	10
3.ii	68	96	70	0	44	38	34	44	32
4	70	94	66	56	0	20	22	34	14
5.i	70	94	78	62	80	0	44	82	38
5.ii	70	94	80	66	78	56	0	78	40
5.iii	64	92	78	56	66	18	22	0	18
5.iv	70	96	90	68	86	62	60	82	0

	DRD								
	1	2	3.i	3.ii	4	5.i	5.ii	5.iii	5.iv
1	0	84	34	30	30	30	30	34	30
2	16	0	6	4	6	6	6	6	4
3.i	66	94	0	30	38	26	22	24	14
3.ii	70	96	70	0	54	42	40	48	34
4	70	94	62	46	0	22	22	34	16
5.i	70	94	74	58	78	0	42	84	40
5.ii	70	94	78	60	78	58	0	82	40
5.iii	66	94	76	52	66	16	18	0	18
5.iv	70	96	86	66	84	60	60	82	0

B

Full Results for Chapter 6

Results for the comparative analysis of different warp parameter variations were presented in Section 6.4. Comparison of results averaged over all images tested were shown in Figures 6.5 and 6.4. Full results of all parameter variations on all images are included here for completeness. Figures B.1 and B.2 show the percentage displacement error $< 0.5x$, with each plot displaying results for a fixed gradient error weight λ, and fixed content preserving triangle weight variation w_L. Similarly, Figures B.3 and B.4 show the percentage displacement error $< 0.25x$, and Figures B.5 and B.6 show the mean pixel displacement error.

Results for the comparative analysis of different registration methods were also presented in Section 6.4. To summarise these comparative results, the performance of each method was ranked for each error metric examined using Ranked Pairs Voting (RP) shown in Table 6.1. The pairwise ranking results used to obtain the RP ranks for each metric are shown in Table B.1, where each entry represents the percentage of images that the method listed vertically was ranked higher than the method listed horizontally.

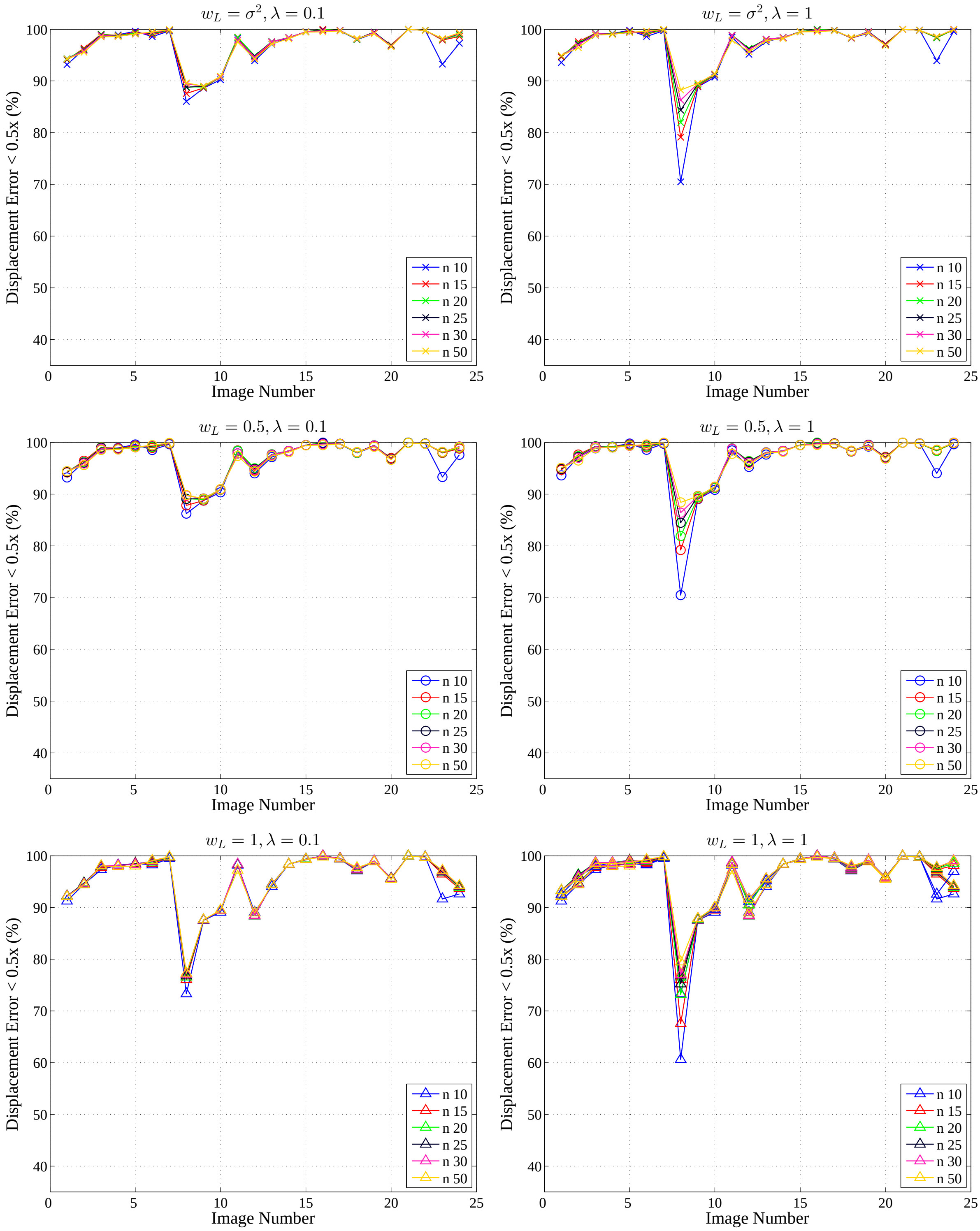


Figure B.1: Percentage displacement error $< 0.5x$ for all 24 image registration results using the method proposed in Chapter 6, with each plot showing different gradient error weight, λ, and content preserving triangle weight, w_L, combinations. Left column: $\lambda = 0.1$. Right column: $\lambda = 1$. Top row: $w_L = \sigma^2$. Middle row: $w_L = 0.5$. Bottom row: $w_L = 1$.

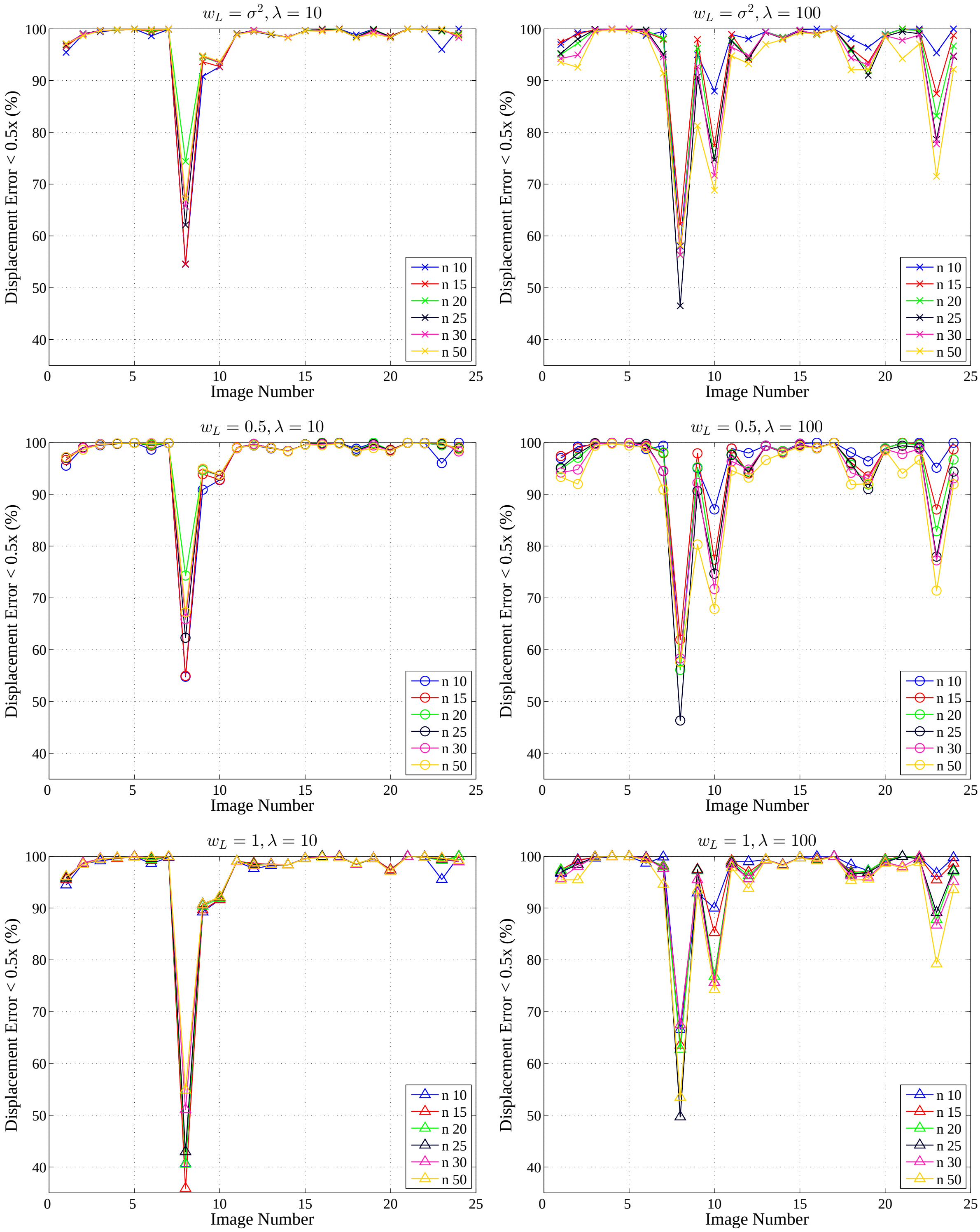


Figure B.2: Percentage displacement error $< 0.5x$ for all 24 image registration results using the method proposed in Chapter 6, with each plot showing different gradient error weight, λ, and content preserving triangle weight, w_L, combinations. Left column: $\lambda = 10$. Right column: $\lambda = 100$. Top row: $w_L = \sigma^2$. Middle row: $w_L = 0.5$. Bottom row: $w_L = 1$.

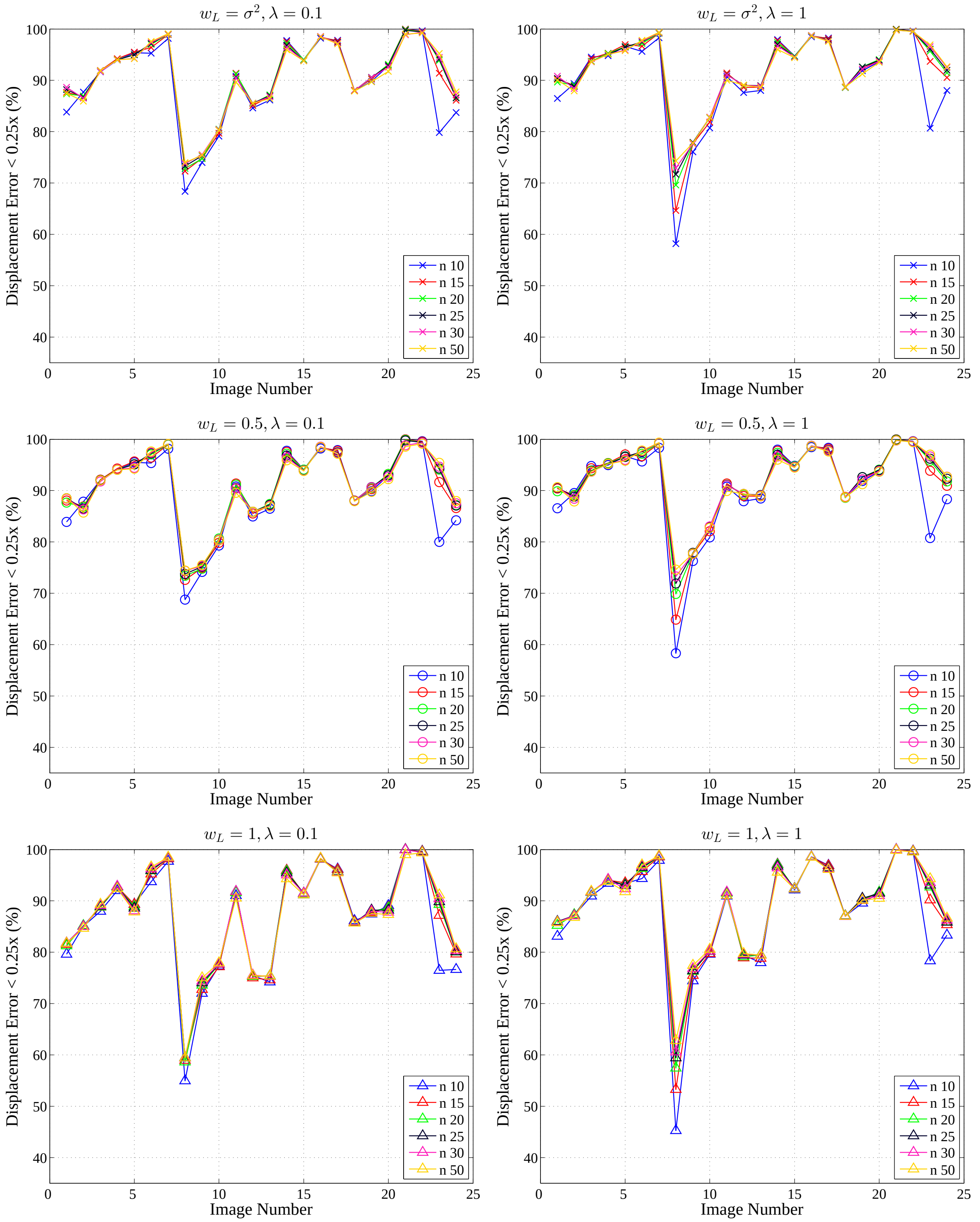


Figure B.3: Percentage displacement error $< 0.25x$ for all 24 image registration results using the method proposed in Chapter 6, with each plot showing different gradient error weight, λ, and content preserving triangle weight, w_L, combinations. Left column: $\lambda = 0.1$. Right column: $\lambda = 1$. Top row: $w_L = \sigma^2$. Middle row: $w_L = 0.5$. Bottom row: $w_L = 1$.

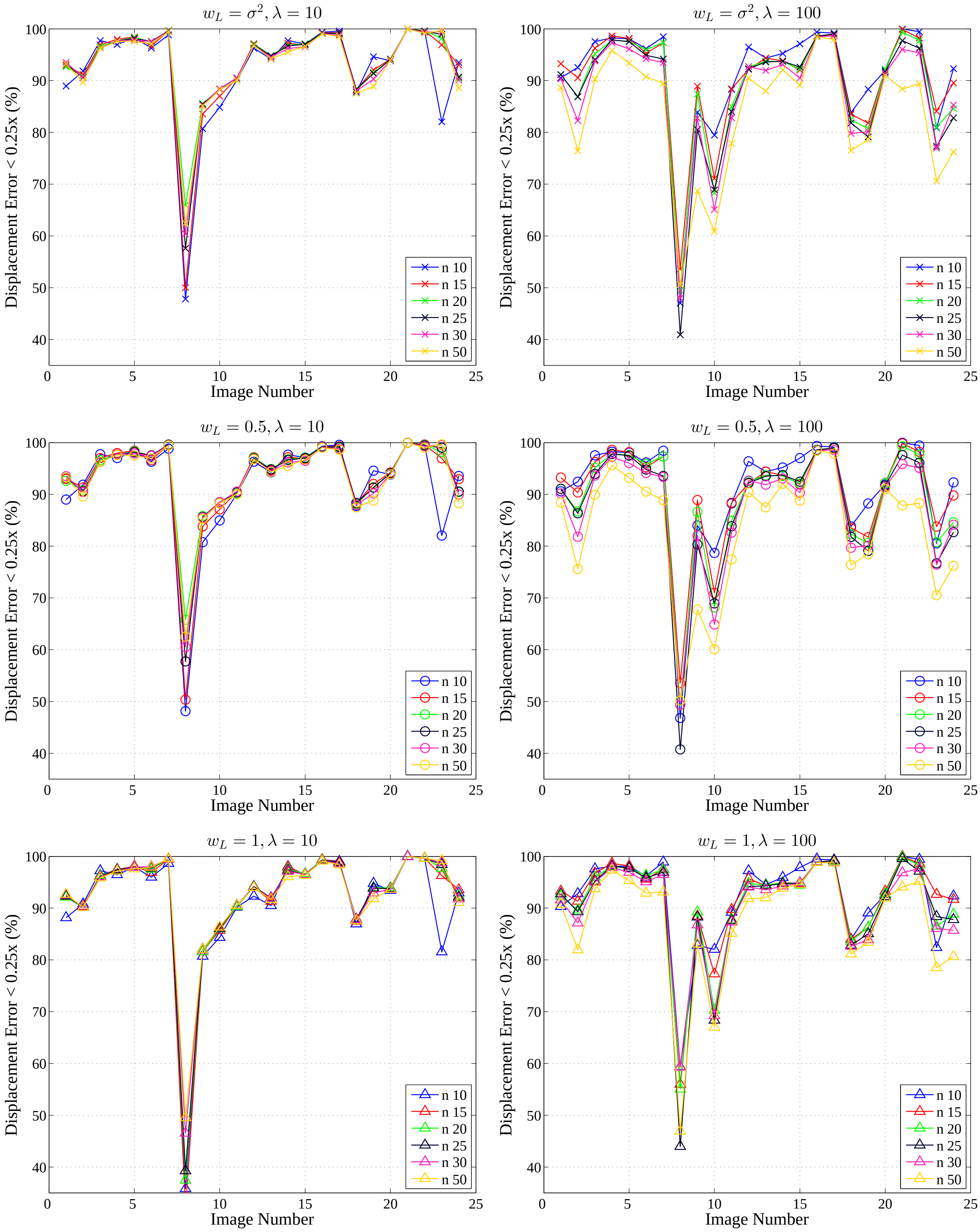


Figure B.4: Percentage displacement error $< 0.25x$ for all 24 image registration results using the method proposed in Chapter 6, with each plot showing different gradient error weight, λ, and content preserving triangle weight, w_L, combinations. Left column: $\lambda = 10$. Right column: $\lambda = 100$. Top row: $w_L = \sigma^2$. Middle row: $w_L = 0.5$. Bottom row: $w_L = 1$.

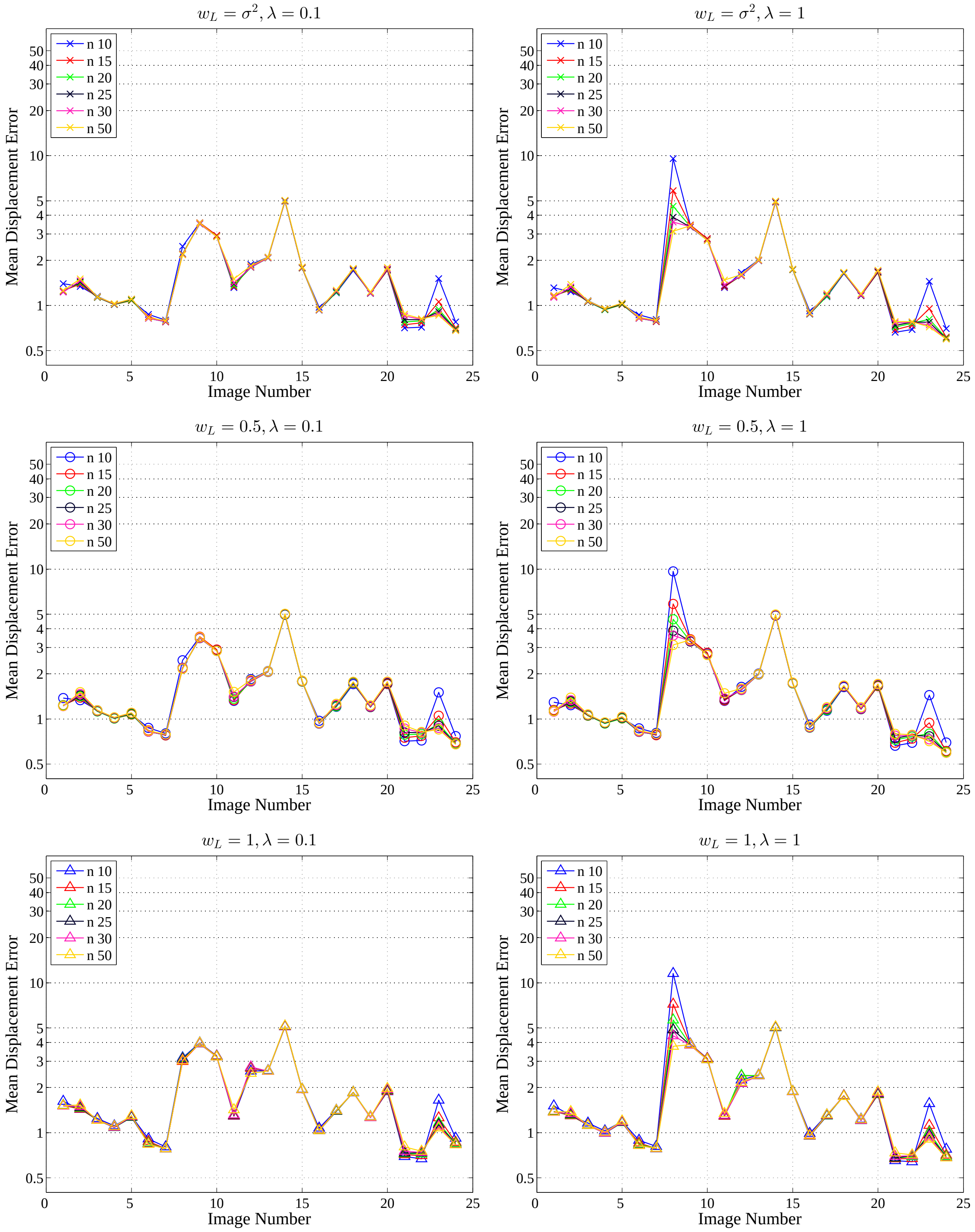


Figure B.5: Mean pixel displacement error for all 24 image registration results using the method proposed in Chapter 6, with each plot showing different gradient error weight, λ, and content preserving triangle weight, w_L, combinations. Left column: $\lambda = 0.1$. Right column: $\lambda = 1$. Top row: $w_L = \sigma^2$. Middle row: $w_L = 0.5$. Bottom row: $w_L = 1$.

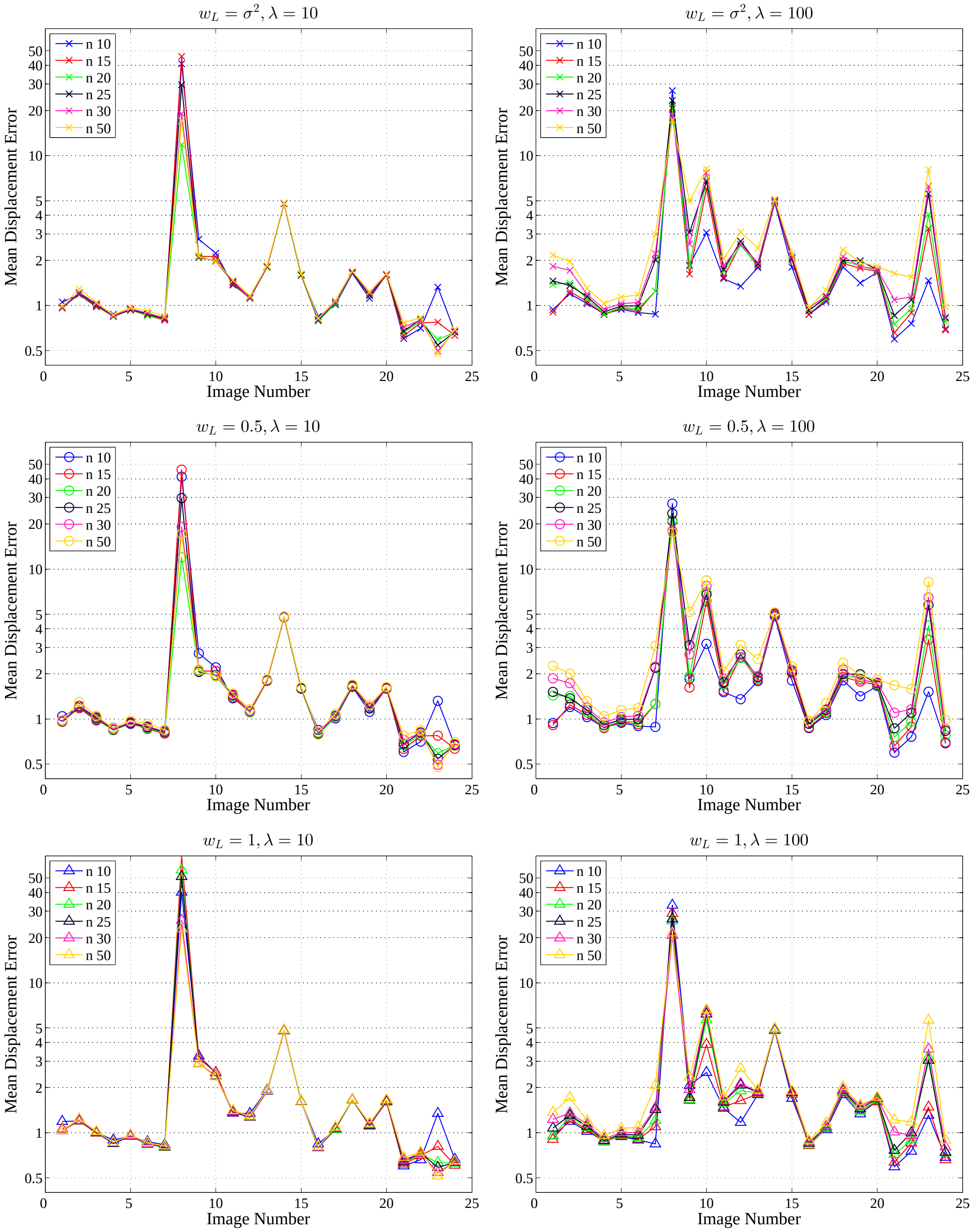


Figure B.6: Mean pixel displacement error for all 24 image registration results using the method proposed in Chapter 6, with each plot showing different gradient error weight, λ, and content preserving triangle weight, w_L, combinations. Left column: $\lambda = 10$. Right column: $\lambda = 100$. Top row: $w_L = \sigma^2$. Middle row: $w_L = 0.5$. Bottom row: $w_L = 1$.

Table B.1: Pairwise method rank comparison (%) for the registration error metrics.

$< 0.5x$	1	2	3	4	5
1	0	50	12.5	12.5	4.17
2	33.33	0	12.5	8.33	0
3	87.5	87.5	0	16.67	20.83
4	87.5	91.67	66.67	0	25
5	91.67	100	79.17	75	0

$< 0.25x$	1	2	3	4	5
1	0	45.83	12.5	16.67	0
2	37.5	0	8.33	4.17	0
3	87.5	91.67	0	37.5	12.5
4	83.33	95.83	58.33	0	20.83
5	95.83	100	87.5	79.17	0

Mean	1	2	3	4	5
1	0	66.67	16.67	16.67	4.17
2	33.33	0	16.67	12.5	0
3	83.33	83.33	0	37.5	8.33
4	83.33	87.5	62.5	0	8.33
5	95.83	100	91.67	91.67	40

Bibliography

[1] The Archimedes Palimpsest Project. `http://www.archimedespalimpsest.org`.

[2] Digital Archive of Medieval Music - digital restoration. `http://www.diamm.ac.uk/about/technial-overview/digital-restoration`.

[3] The Dublin Institue for Advanced Studies. `http://www.dias.ie`.

[4] Early Manuscripts at Oxford University. `http://image.ox.ac.uk`.

[5] Google Books Scanning Project. `http://www.google.com/googlebooks/about`.

[6] The Irish Script On Screen Project. `http://www.isos.dias.ie`.

[7] S. Baker and I. Matthews. Lucas-Kanade 20 years on: A unifying framework. *International Journal of Computer Vision*, 56(3):221–255, 2004.

[8] A. L. Bertozzi, S. Esedoglu, and A. Gillette. Inpainting of binary images using the Cahn-Hilliard equation. *IEEE Transactions on Image Processing*, 16(1):285–291, 2007.

[9] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society. Series B (Methodological)*, 48(3):259–302, 1986.

[10] J. C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms.* Kluwer Academic Publishers, 1981.

[11] C. Bingyu and C. Youbin. Reduction of bleed-through effect in images of Chinese bank items. In *Proceedings of the International Conference on Frontiers in Handwriting Recognition (ICFHR'12)*, pages 174–178, 2012.

[12] R. D. Boyle and H. Hiary. Watermark location via back-lighting and recto removal. *International Journal of Document Analysis and Recognition (IJDAR)*, 12(1):33–46, 2009.

[13] M. P. Brown. *Understanding Illuminated Manuscripts: A Guide to Technical Terms.* J. Paul Getty Museum, 1994.

[14] J. A. Burgoyne, J. Devaney, L. Pugin, and I. Fujinaga. Enhanced bleedthrough correction for early music documents with recto-verso registration. In *Proceedings of the International Conference on Music Information Retrieval (ICMIR'08)*, pages 407–412, Philadelphia, PA, 2008.

[15] P. Castro, R. J. Almeida, and J. R. C. Pinto. Restoration of double-sided ancient music documents with bleed-through. In L. Rueda, D. Mery, and J. Kittler, editors, *Proceedings of the 12th Iberoamerican Congress on Pattern Recongition (CIARP'07)*, volume 4756 of *LNCS*, pages 940–949. Springer Berlin/Heidelberg, 2007.

[16] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.

[17] A. Criminisi, P. Perez, and K. Toyama. Region filling and object removal by exemplar-based image inpainting. *IEEE Transactions on Image Processing*, 13(9):1200–1212, 2004.

[18] F. Deng, Z. Wu, Z. Lu, and M. S. Brown. BinarizationShop: A user-assisted software suite for converting old documents to black-and-white. In J. Hunter, C. Lagoze, C. L. Giles, and Y.-F. Li, editors, *Proceedings of the 10th ACM/IEEE Joint Conference on Digital Libraries (JCDL'10)*, pages 255–258. ACM, 2010.

[19] F. Drira, F. Le Bourgeois, and H. Emptoz. Restoring ink bleed-through degraded document images using a recursive unsupervised classification technique. In H. Bunke and A. Spitz, editors, *Proceedings of the 7th IAPR International Workshop on Document Analysis Systems (DAS'06)*, volume 3872 of *LNCS*, pages 38–49. Springer Berlin / Heidelberg, 2006.

[20] E. Dubois and P. Dano. Joint compression and restoration of documents with bleed-through. In *Proceedings of the 2nd IS&T Archiving Conference*, volume 2, pages 170–174, Washington, DC, USA, 2005.

[21] E. Dubois and A. Pathak. Reduction of bleed-through in scanned manuscript documents. In *Proceedings of the IS&T Image Processing, Image Quality, Image Capture Systems Conference (PICS'01)*, volume 4, pages 177–180, Montreal, Canada, 2001.

[22] R. Estrada and C. Tomasi. Manuscript bleed-through removal via hysteresis thresholding. In *Proceedings of the 10th International Conference on Document Analysis and Recognition (ICDAR'09)*, pages 753–757, Barcelona, Spain, 2009.

[23] B. Gatos, K. Ntirogiannis, and I. Pratikakis. ICDAR 2009 Document image binarization contest (DIBCO 2009). In *Proceedings of the 10th International Conference on Document Analysis and Recognition (ICDAR'09)*, pages 1375–1382, 2009.

[24] B. Gatos, I. Pratikakis, and S. J. Perantonis. Adaptive degraded document image binarization. *Journal of Pattern Recognition*, 39(3):317–327, 2006.

[25] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6(6):721–741, 1984.

[26] A. Gnanadesikan. *The Writing Revolution: Cuneiform to the Internet.* Wiley, 2011.

[27] M. Greengrass and L. Hughes. *The Virtual Representation of the Past.* Digital Research in the Arts and Humanities. Ashgate, 2008.

[28] N. A. Gumerov and R. Duraiswami. Fast radial basis function interpolation via preconditioned Krylov iteration. *SIAM Journal on Scientific Computing (SISC)*, 29(5):1876–1899, 2007.

[29] G. A. Hanasusanto, W. Zheng, and M. S. Brown. Ink-bleed reduction using functional minimization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'10)*, pages 825–832, 2010.

[30] R. Haselden. *Scientific Aids for the Study of Manuscripts.* Oxford University Press, 1935.

[31] J. He, Q. D. M. Do, A. C. Downton, and J. H. Kim. A comparison of binarization methods for historical archive documents. In *Proceedings of the 8th International Conference on Document Analysis and Recognition (ICDAR'05)*, pages 538–542 Vol. 1, 2005.

[32] H. Hiary and K. Ng. A system for segmenting and extracting paper-based watermark designs. *International Journal on Digital Libraries*, 6(4):351–361, 2007.

[33] Y. Huang and M. S. Brown. User-assisted ink-bleed correction for handwritten documents. In *Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL'08)*, pages 263–271, Pittsburgh PA, PA, USA, 2008. ACM.

[34] Y. Huang, M. S. Brown, and D. Xu. A framework for reducing ink-bleed in old documents. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'08)*, pages 1–7, Anchorage, AK, 2008.

[35] Y. Huang, M. S. Brown, and D. Xu. User-assisted ink-bleed reduction. *IEEE Transactions on Image Processing*, 19(10):2646–2658, 2010.

[36] M. Kite and R. Thomson. *Conservation of Leather and Related Materials.* Taylor & Francis, 2012.

[37] V. Kolmogorov and R. Zabin. What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2):147–159, 2004.

[38] F. Liu, M. Gleicher, H. Jin, and A. Agarwala. Content-preserving warps for 3D video stabilization. *ACM Transactions on Graphics*, 28(3):1–9, 2009.

[39] H. Lu, A. C. Kot, and Y. Q. Shi. Distance-reciprocal distortion measure for binary document images. *IEEE Signal Processing Letters*, 11(2):228–231, 2004.

[40] M. Lyons. *Books: A Living History.* Thames & Hudson, Limited, 2011.

[41] C. D. Manning, P. Raghavan, and H. Schütze. *Introduction to Information Retrieval.* Cambridge University Press, 2008.

[42] B. Meehan. *The Book of Kells.* Thames & Hudson, Limited, 2012.

[43] R. F. Moghaddam and M. Cheriet. Low quality document image modeling and enhancement. *International Journal on Document Analysis and Recognition (IJDAR)*, 11(4):183–201, 2009.

[44] R. F. Moghaddam and M. Cheriet. A variational approach to degraded document enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(8):1347–1361, 2010.

[45] R. F. Moghaddam, D. Rivest-Henault, I. Bar-Yosef, and M. Cheriet. A unified framework based on the level set approach for segmentation of unconstrained double-sided document images suffering from bleed-through. In *Proceedings of the 10th International Conference on Document Analysis and Recognition (ICDAR'09)*, pages 441–445, 2009.

[46] A. Myronenko and S. Xubo. Intensity-based image registration by minimizing residual complexity. *IEEE Transactions on Medical Imaging*, 29(11):1882–1891, 2010.

[47] W. Niblack. *An introduction to digital image processing.* Strandberg Publishing Company, 1985.

[48] K. Ntirogiannis, B. Gatos, and I. Pratikakis. An objective evaluation methodology for document image binarization techniques. In *Proceedings of the 8th IAPR International Workshop on Document Analysis Systems (DAS'08)*, pages 217–224, 2008.

[49] J. K. O Ruanaidh and W. J. Fitzgerald. *Numerical Bayesian Methods Applied to Signal Processing.* Statistics and Computing. Springer, 1996.

[50] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.

[51] J. Pinto, P. Vieira, and J. Sousa. A new graph-like classification method applied to ancient handwritten musical symbols. *International Journal on Document Analysis and Recognition (IJDAR)*, 6(1):10–22, 2003.

[52] I. Pratikakis, B. Gatos, and K. Ntirogiannis. H-DIBCO 2010 - Handwritten document image binarization competition. In *Proceedings of the International Conference on Frontiers in Handwriting Recognition (ICFHR'10)*, pages 727–732, 2010.

[53] I. Pratikakis, B. Gatos, and K. Ntirogiannis. ICDAR 2011 Document image binarization contest (DIBCO 2011). In *Proceedings of the 11th International Conference on Document Analysis and Recognition (ICDAR'11)*, pages 1506–1510, 2011.

[54] I. Pratikakis, B. Gatos, and K. Ntirogiannis. ICFHR 2012 Competition on handwritten document image binarization (H-DIBCO 2012). In *Proceedings of the International Conference on Frontiers in Handwriting Recognition (ICFHR'12)*, pages 817–822, 2012.

[55] V. Rabeux, N. Journet, and J. P. Domenger. Document recto-verso registration using a dynamic time warping algorithm. In *Proceedings of the 11th International Conference on Document Analysis and Recognition (ICDAR'11)*, pages 1230–1234, 2011.

[56] C. Rother, V. Kolmogorov, V. Lempitsky, and M. Szummer. Optimizing binary MRFs via extended roof duality. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07)*, pages 1–8, 2007.

[57] R. Rowley-Brooke and A. Kokaram. Bleed-through removal in degraded manuscripts. In *Proceedings of the 22nd IET Irish Signals and Systems Conference (ISSC'11)*, Trinity College, Dublin, Ireland, 2011.

[58] R. Rowley-Brooke and A. Kokaram. Bleed-through removal in degraded documents. In C. Viard-Gaudin and R. Zanibbi, editors, *Proceedings of the IS&T/SPIE Electronic Imaging Symposium Document Recogntion and Retrieval XIX (DRR'12)*, volume 8297, San Francisco, CA, 2012.

[59] R. Rowley-Brooke, F. Pitié, and A. Kokaram. A ground truth bleed-through document image database. In P. Zaphiris, G. Buchanan, E. Rasmussen, and F. Loizides, editors, *Proceedings of the 2nd International Conference on the Theory and Practice of Digital Libraries (TPDL'12)*, volume 7489 of *LNCS*, pages 185–196, Paphos, Cyprus, 2012. Springer.

[60] R. Rowley-Brooke, F. Pitié, and A. Kokaram. A non-parametric framework for document bleed-through removal. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'13)*, Portland, OR, 2013.

[61] R. Rowley-Brooke, F. Pitié, and A. Kokaram. Nonrigid recto-verso registration using page outline structure and content preserving warps. In *Proceedings of the 2nd International Workshop on Historical Document Imaging and Processing (HIP'13)*, Washington DC, USA, 2013.

[62] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes. Nonrigid registration using free-form deformations: application to breast mr images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999.

[63] J. Sauvola and M. Pietikäinen. Adaptive document image binarization. *Pattern Recognition*, 33(2):225–236, 2000.

[64] M. Sezgin and B. Sankur. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1):146–168, 2004.

[65] M. Sokolova and G. Lapalme. A systematic analysis of performance measures for classification tasks. *Information Processing and Management*, 45(4):427–437, 2009.

[66] P. Stathis, E. Kavallieratou, and N. Papamarkos. An evaluation survey of binarization algorithms on historical documents. In *Proceedings of the 19th International Conference on Pattern Recognition (ICPR'08)*, pages 1–4, 2008.

[67] C. L. Tan, R. Cao, and P. Shen. Restoration of archival documents using a wavelet technique. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(10):1399–1404, 2002.

[68] T. N. Tideman. Independence of clones as a criterion for voting rules. *Social Choice and Welfare*, 4(3):185–206, 1987.

[69] A. Tonazzini, L. Bedini, and E. Salerno. Independent component analysis for document restoration. *International Jounal on Document Analysis and Recognition (IJDAR)*, 7(1):17–27, 2004.

[70] A. Tonazzini, L. Bedini, and E. Salerno. A Markov model for blind image separation by a mean-field EM algorithm. *IEEE Transactions on Image Processing*, 15(2):473–482, 2006.

[71] A. Tonazzini, G. Bianco, and E. Salerno. Registration and enhancement of double-sided degraded manuscripts acquired in multispectral modality. In *Proceedings of the 10th International Conference on Document Analysis and Recognition (ICDAR'09)*, pages 546–550, 2009.

[72] A. Tonazzini, I. Gerace, and F. Martinelli. Multichannel blind separation and deconvolution of images for document analysis. *IEEE Transactions on Image Processing*, 19(4):912–925, 2010.

[73] A. Tonazzini, E. Salerno, and L. Bedini. Fast correction of bleed-through distortion in grayscale documents by a blind source separation technique. *International Jounal on Document Analysis and Recognition (IJDAR)*, 10(1):17–25, 2007.

[74] A. Tonazzini, E. Salerno, M. Mochi, and L. Bedini. Bleed-through removal from degraded documents using a color decorrelation method. In S. Marinai and A. Dengel, editors, *Proceedings of the 6th IAPR Internatinoal Workshop on Document Analysis Systems (DAS'04)*, volume 3163 of *Lecture Notes in Computer Science*, pages 250–261. Springer Berlin / Heidelberg, 2004.

[75] J. Wang, M. S. Brown, and C. L. Tan. Accurate alignment of double-sided manuscripts for bleed-through removal. In *Proceedings of the 8th IAPR International Workshop on Document Analysis Systems (DAS'08)*, pages 69–75, 2008.

[76] J. Wang and C. L. Tan. Non-rigid registration and restoration of double-sided historical manuscripts. In *Proceedings of the 11th International Conference on Document Analysis and Recognition (ICDAR'11)*, pages 1374–1378, 2011.

[77] Q. Wang, T. Xia, L. Li, and C. L. Tan. Document image enhancement using directional wavelet. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'03)*, volume 2, pages II–534–II–539 vol.2, 2003.

[78] C. Wolf. Document ink bleed-through removal with two hidden Markov random fields and a single observation field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3):431–447, 2010.

[79] L. Zheng, W. Zheng, and M. S. Brown. Directed assistance for ink-bleed reduction in old documents. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'09)*, pages 88–95, 2009.

[80] L. Zheng, W. Zheng, and M. S. Brown. Interactive degraded document binarization: An example (and case) for interactive computer vision. In *Proceedings of the IEEE Workshop on Applications of Computer Vision (WACV'09)*, pages 1–8, 2009.