Flicker and unsteadiness compensation for archived film sequences

Guillaume Forbin

Submitted for the Degree of
Doctor of Philosophy
from the
University of Surrey

Centre for Vision, Speech and Signal Processing
Faculty of Engineering and Physical Sciences
University of Surrey
Guildford, Surrey GU2 7XH, U.K.

May 2009

© Guillaume Forbin 2009
Summary

This thesis aims at proposing high-level image processing tools suitable for film restoration. First priority is given to temporal brightness variations (commonly referred to as flicker) and motion unsteadiness, which are the most commonly encountered artefacts in archived film sequences.

A novel approach for the compensation of flicker in archived film sequences is presented. The proposed method is motivated by fundamental principles of photographic image registration and provides a substantial level of adaptation to temporal but also spatial variations of picture brightness. Additionally the proposed scheme provides an efficient mechanism for the adaptive estimation of flicker compensation profile, which makes it suitable for the compensation of long duration film sequences while it addresses problems arising from scene motion and illumination using a novel motion-compensated grey-level tracing approach.

A similar approach is formulated for unsteadiness compensation associated with mechanical tolerances of various components involved in film production, film processing and film-to-video transfer. Sub-pixel accuracy motion unsteadiness between successive frames is estimated using a phase correlation based technique combined with robust parameter estimation. The method designed for the estimation of flicker compensation profile is adapted and a novel algorithm featuring motion vector tracing is formulated, allowing the adaptive estimation and compensation of the two artefacts within a common framework.

We present experimental evidence which suggests that both our methods for flicker and unsteadiness compensation offer high levels of performance and compare favourably with competing state-of-the-art techniques.
**Key words:** Film restoration, Flicker compensation, Flicker non-linear modelling, Unsteadiness compensation, Robust motion estimation, Temporal motion parameters filtering.

**Email:** g.forbin@surrey.ac.uk

**WWW:** http://personal.ee.surrey.ac.uk/Personal/G.Forbin/
Acknowledgements

My first thanks go to my supervisor Dr. Theodore Vlachos for guiding me during the four years of preparation of my PhD thesis. He taught me how to carry out rigorous scientific research. His expertise in the field of film restoration and computer vision contributed greatly to shaping this thesis and to the dissemination of the achievements into the scientific community.

I am also very grateful to Prof. Josef Kittler for giving me the opportunity to do a PhD at the Centre for Vision Speech and Signal Processing (CVSSP) and for supervising me in the last couple of months of my PhD.

I would like to acknowledge all my colleagues at CVSSP, and also all my friends, for their help and for making the PhD an enjoyable experience.

I would like to thank my parents Françoise and André, my sister Aurélie for their constant help, support and encouragement. Finally, my special thanks are for my wife Amanda who constantly supported me throughout the four years of my PhD.

This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) under Research Grant GR/S70098/01.
Contents

1 Introduction ......................................................... 1
   1.1 Context ......................................................... 1
   1.2 Restoration of archived films ................................. 2
   1.3 Objectives ...................................................... 4
   1.4 Contribution .................................................... 6
   1.5 List of Publications ............................................. 8
   1.6 Structure of the thesis .......................................... 9

2 Flicker .............................................................. 13
   2.1 Flicker in archived film sequences ............................ 13
   2.2 A spatially variable artefact .................................. 15

3 Literature review .................................................... 17
   3.1 Global compensation ............................................ 17
   3.2 Spatially-adaptive compensation ............................... 19
   3.3 Compensation for sequences of longer duration .............. 22
   3.4 Commercial solutions for archived film restoration .......... 23

4 Non-linear modelling ................................................ 25
   4.1 Intensity error profile estimation based on the Density versus log-Exposure characteristic ............................... 25
   4.2 Telecine grading ................................................ 30
5 Flicker compensation framework 33
  5.1 Grey-level intensity error reliability weighting 34
  5.2 Motion compensated intensity error profile estimation 37
  5.3 Adaptive estimation of intensity error profile 39
  5.4 Intensity error profile estimation between distant frames using motion-compensated grey-level tracing 41
  5.5 Spatial adaptation 44
     5.5.1 Block-based spatial adaptation 44
     5.5.2 Segmentation-based spatial adaptation 46
  5.6 Algorithms flowchart 50

6 Experimental results 53
  6.1 Test material 53
  6.2 Evaluation protocol 57
  6.3 Results evaluation against competing algorithms 59
  6.4 Impact of the block-based/segmentation-based spatial adaptation 61
  6.5 Complexity of the proposed algorithms 61
  6.6 Conclusions 63

7 Unsteadiness 69

8 Literature Review 73
  8.1 Global motion models 74
  8.2 Global motion estimation 75
     8.2.1 Feature-based motion estimation 76
     8.2.2 Direct motion estimation 77
     8.2.3 Robust motion parameter estimation 80
  8.3 Unsteadiness compensation 81

9 Unsteadiness compensation framework 83
  9.1 Global motion model selection 84
  9.2 Block-based motion estimation using phase correlation 86
     9.2.1 Phase correlation method 86
     9.2.2 Sub-pixel accuracy motion estimation 88
### Contents

9.3 Adaptive estimation of unsteadiness compensation parameters ........................................... 90
9.4 Global motion estimation between distant frames using motion vectors tracing ................................................................. 93
9.5 RANSAC/MSAC for the robust estimation of motion parameters ........................................... 95
  9.5.1 Robust parameter estimators ................................................................................................. 95
  9.5.2 Random Sample Consensus (RANSAC) and M-estimator Sample Consensus (MSAC) ......................................................... 97
  9.5.3 Parametrisation of motion vector fields ............................................................................. 103
9.6 Algorithm flowchart .................................................................................................................. 108

10 Experimental results .................................................................................................................. 109
  10.1 Global motion estimation using quadratic motion model ..................................................... 110
  10.2 Conclusions .......................................................................................................................... 111

11 Conclusion and future research ................................................................................................ 121

A Notations .................................................................................................................................. 127

B Flicker - additional results .......................................................................................................... 131

C Unsteadiness - additional results ................................................................................................ 147

Bibliography .................................................................................................................................. 157
Chapter 1

Introduction

1.1 Context

UK broadcasters, archivists, media librarians and film makers are custodians of huge stocks of moving picture archive material which collectively represent a unique record of the historic, artistic and cultural development of essentially every aspect of the nations life including sport, politics and entertainment. On the other hand, the emergence of new multimedia and broadcasting outlets has the potential of dramatically improving public access to cultural assets of such unique educational and entertainment value.

Against this favourable socio-economic background, the level of public access today is limited by a number of technological factors. Many of the historically significant items are either unavailable in their original format or too fragile to survive any attempt at copy or playback. Very often the only surviving record can be found on a different medium such as videotape on which artefacts associated with multiple generations of copying or film-to-video transfers are embedded in the pictorial information. In addition, old film sequences are likely to have suffered severe degradations as early as the acquisition stage due to inherent technological limitations. Examples are irregular frame rates due to the use of hand-cranked cameras and inconsistent exposure due to imprecise shutter mechanisms. Later in their lifetime, films may suffer further damage due to environmental hazards such as humidity and dust, chemical instabilities,
improper storage and handling practices and even poorly maintained projectors as illustrated in Figure 1.1.

Whatever the case, the complexity and associated cost of manual processes involved in a conventional restoration chain place considerable limitations on processing throughput rendering the restoration of entire collections an unrealistic proposition. Additionally, conventional restoration relies on the use of dedicated equipment such as special copying machines which can only target a limited range of artefacts due to the fact that the unit of manipulation can only be the physical film strip. Ultimately the combined and inevitable effect of the above barriers is to deny the public, indisputably the true owners of these unique cultural assets, all the benefits arising from free access.

1.2 Restoration of archived films

Film restoration, or preservation is defined as the set of operations which can be applied to a film copy to retrieve a version similar to the original. Historical research aimed at estimating how the original scene looked like, film editing, film restoration on the chemical material or by digital process are concerned. Manipulating directly the film strip is a risky operation and must be performed by professionals knowing the constraints associated to this kind of degraded materials. On the other hand, digital image and video processing have recently been applied to film restoration and powerful tools used by professional archivists exist nowadays.
There is growing consensus that automatic restoration is a key enabling technology towards facilitating access to film and television archives for a number of reasons. By improving baseline picture quality and by reducing the perceptual impact of archive-related artefacts restoration can meet viewers aesthetic expectations and enrich the viewing experience. Moreover, the suppression of such artefacts has vital implications on the efficiency of video coding algorithms used in the television and multimedia distribution chains such as MPEG-2 and MPEG-4. Finally, since restoration processes almost invariably result in the enhancement of semantic content, they are also likely to contribute to more efficient management of pictorial databases and archives.

Further evidence of the emerging importance of restoration and its potential benefits can be found in recent collaborative research efforts at a European level such as PRESTOSPACE [Pre09], LIMELIGHT [Lim09], AURORA [Aur09] and BRAVA [Bra09]. For various reasons such projects have mostly focused on the development of systems whose ambitions rarely exceeded the limitations of the technology which was available at the time. In addition, a significant amount of effort was diverted from algorithm development towards the fulfilment of other operational requirements such as real-time processing and improved human-machine interaction.

The added value of supporting a restoration activity at a national level becomes evident by considering the significant number of film archives resident in this country as well as the unprecedented level of talent and expertise in film and television programme-making. These have been identified by the Foresight panel as key elements driving technological development and ultimately having a critical impact on the social and economic life of the nation.

Original recordings are at risk from the irreversible and ever accelerating forces of decay and disintegration. While contemporary film production and storage technology may prolong life expectancy up to 400 years, the type of acetate-base technology often
encountered in archived film can produce acetic acid (vinegar syndrome) after as little as 20 years even in reasonable storage conditions. Only urgent and decisive action can limit the irretrievable damage threatened by this time bomb which is sadly ticking away as these lines are being typed. It is estimated that 90% of silent films and 50% of films shot before 1950 have already disappeared, and that a majority of existing films will vanish in the next couple of centuries [DF97].

The different stages involved during film digital restoration process are illustrated in Figure 1.2. Film is first digitalised at high resolution (4k) to preserve the quality of the original material, using either film scanner or high resolution telecine. Restoration is then performed on digital sequences of images. This stage involves the correction of artefacts such as flicker, unsteadiness, scratches, dirt or dust, but also enhancement regarding the grey-level dynamic or colourimetry of the digital sequences. Finally, the restored sequence is either stored on a digital, video or new film support. Video tapes and digital supports have a low age expectancy, whereas recent film strips are expected to last up to 400 years under optimal conservation and thus are preferred for film preservation.

1.3 Objectives

This thesis aims at proposing high-level image processing algorithms tools suitable for film restoration. First priority is given to flicker and unsteadiness, which are the most commonly encountered artefacts in archived film sequences. These are immediately recognisable by non-expert viewers as signature artefacts of old film sequences. Their perceptual impact can be significant as both interfere substantially with the viewing experience and have the potential of concealing essential detail. These two artefacts have often been categorised as global artefacts in the sense that they usually affect a film frame in its entirety as opposed to so-called local artefacts such as dirt, dust or scratches which manifest themselves only locally on the image plane. In operational restoration systems flicker and unsteadiness are invariably among the top-priority targeted artefacts as their correction offers substantial benefits both with regard to visual
1.3. Objectives

Figure 1.2: Traditional digital film restoration chain ([DF97]).
quality as well as subsequent restoration operations.

Flicker refers to random temporal fluctuations in picture intensity and is a common artefact in archived film sequences. The main contributing cause of film flicker is inconsistent exposure. Other causes may include printing errors in film processing, film ageing, multiple copying, aliasing, mould and dust. Film flicker is a very noticeable impairment especially in cases where film is displayed simultaneously with video or with electronically generated graphics and captions. More importantly it can be quite unsettling to the viewer and may also lead to considerable discomfort and eye fatigue after prolonged viewing. Camera and scene motion can partly mask the effects of flicker and as a consequence the latter is much more noticeable in sequences consisting primarily of still frames or frames of low-motion content.

Unsteadiness (also referred to as instability or shake) has long been identified as one of the most severe artefacts of film scanned for television display. The main contributing causes of film unsteadiness have been associated with mechanical tolerances of various components involved in film production, film processing and film-to-video transfer. Examples are the film sprocket hole positioning accuracy of the camera and telecine mechanisms, the perforation accuracy of positive and negative stock and the relative positioning of positive and negative stock in the printer. Film unsteadiness can be very unsettling to the viewer. As is the case with flicker, unsteadiness is especially noticeable when film is displayed simultaneously with video or electronically generated graphics and captions which is a common occurrence in modern day documentaries.

1.4 Contribution

This thesis presents novel image processing tools for flicker and unsteadiness compensation, competing favourably with methods reported in the literature.
1.4. Contribution

Flicker

Investigation had been carried out at the CVSSP [VSC09] prior to the beginning of this thesis. A non-linear model based on the Density versus log-Exposure characteristic was proposed in [Vla04] to characterise exposure inconsistencies, which are the main source for flicker. In addition, a baseline algorithm was drawn allowing the estimation of an intensity error profile between a reference and a degraded frame. Contributions related to flicker modeling and compensation are presented in Chapter 5 and can be summarised as follows:

- The frequency of occurrence of each grey-level and the reliability of the intensity error estimations are employed as a weight for the estimation of a compensation profile between a reference and a degraded frame. This allows to reduce the influence of poorly represented grey-levels, but also the influence of inaccurate intensity error estimations (Section 5.1). This new feature will be later on referred to as "grey-level intensity error reliability weighting".

- Motion estimation and compensation is employed as a pre-process before intensity error profile estimation between frames. In addition, motion prediction error is used as an input for the profile estimation to lower the influence of poorly compensated pixels (Section 5.2).

- A novel compensation framework is proposed allowing flicker estimation and compensation in long duration sequences presenting scene motion. Temporal filtering is employed for the estimation of an adaptive flicker compensation profile evolving over time (Section 5.3).

- The framework mentioned above requires flicker parameters estimation between distant frames having different content due to large scene motion often present in real film sequences. A novel algorithm featuring motion-compensated grey-level tracing is proposed to estimate flicker parameters by tracing flicker compensation values along trajectories of estimated motion between consecutive frames (Section 5.4).
Chapter 1. Introduction

- Two novel algorithms have been developed to deal with flicker spatial variability. Block-based and segmentation-based approaches are formulated in Section 5.5.

Unsteadiness

Unsteadiness compensation in archived film sequences has been investigated in the second part of the thesis. Achievements for flicker compensation such as the temporal filtering mechanism using grey-level tracing have been successfully adapted to the unsteadiness compensation problem. A new motion vector field temporal filtering algorithm featuring motion vector tracing has been developed and is presented in Chapter 9. Combined with phase correlation based sub-pixel motion estimation and RANSAC/MSAC parametrisation of compensation motion vector fields, the new framework is able to compensate for unsteadiness accurately and competes favourably with state-of-the-art techniques. In addition flicker and unsteadiness can be estimated and compensated within a common framework.

1.5 List of Publications

The results of this research have been reported in a number of publications.

Journals:


Conferences:

1.6 Structure of the thesis

This thesis is split in two parts, namely flicker and unsteadiness compensation which are organised as follows:

Flicker:

- Chapter 2: flicker artefact in archived film sequences is introduced by presenting its origins and characteristics. The non-linearity and spatial variability characteristics of the deterioration are detailed.

- Chapter 3: an overview of global and spatially-adaptive flicker models and compensation methods reported in the literature is given.

- Chapter 4: previous research carried out at the Centre for Vision, Speech and Signal Processing (CVSSP) [VSC09] regarding flicker compensation is reviewed. A non-linear model based on the Density versus log-Exposure characteristic is presented and an algorithm allowing the estimation of flicker parameters between a reference and a degraded frame is formulated [Vla04].

- Chapter 5: author's contributions are formulated as follows:


Chapter 1. Introduction

- Section 5.1: grey-level intensity error reliability weighting allowing either to strengthen or reduce the influence of certain grey-levels during the estimation of intensity error profile between two frames of a sequence is presented. The weighting is based on grey-levels frequency of occurrence, but also on the reliability of intensity error estimations.

- Section 5.2: motion-compensated intensity error profile estimation is formulated, and motion prediction error is employed to reduce the influence of poorly compensated pixels in the intensity error profile estimation.

- Section 5.3: a new framework based of the adaptive estimation of intensity error profile over time is developed allowing the compensation of long duration sequences containing substantial scene motion.

- Section 5.4: motion-compensated grey-level tracing is formulated for the estimation of intensity error profile between distant frames containing inhomogeneous content due to scene motion is real image sequences. Flicker compensation parameters are retrieved by tracing compensation values along trajectories of estimated motion between consecutive frames.

- Section 5.5: two algorithms able to cope with flicker spatial variations are presented. Intensity error profile estimation over a block-partition of the frames or regions of uniform intensity obtained through frame segmentation are investigated.

- Section 5.6: sequential flowcharts of the proposed flicker compensation algorithms are presented.

- Chapter 6: the developed algorithms are evaluated both subjectively and objectively against competing methods reported in the literature. Spatial adaptation algorithms (block/segmentation-based) are compared and finally conclusions related to flicker compensation are drawn.

Unsteadiness:

- Chapter 7: unsteadiness artefact in archived films is introduced. Origins and characteristics are detailed, and several illustrations are provided.
1.6. Structure of the thesis

- Chapter 8: unsteadiness estimation and compensation methods reported in the literature are reviewed. An overview of global motion models and motion estimation algorithms is provided.

- Chapter 9: author's contributions are formulated as follows:
  - Section 9.1: global motion models regarding unsteadiness modelling in archived films are discussed, and two models are found to be realistic and suitable to characterise motion unsteadiness.
  - Section 9.2: block-based motion estimation and sub-pixel phase correlation technique are employed to formulate a sparse motion estimator suitable for unsteadiness modelling in archived films.
  - Section 9.3: a novel unsteadiness compensation framework is elaborated. Compensation parameters are adaptively estimated over time. The method is inspired from the adaptive estimation of flicker parameters formulated in Section 5.3 and is employed to estimate unsteadiness compensation motion vector fields.
  - Section 9.4: motion vector tracing is formulated to estimation displacement vector fields between distant frames presenting different content due to scene motion. Motion vectors between consecutive frames are traced and accumulated over time.
  - Section 9.5: two robust parameter estimators named Random Sample Consensus (RANSAC) and M-estimator sample consensus (MSAC) are reviewed and applied to the parametrisation of unsteadiness compensation motion vector fields.
  - Section 9.6: sequential flowchart of the proposed unsteadiness compensation algorithm is presented.

- Chapter 10: the developed algorithm is evaluated both subjectively and objectively against competing methods reported in the literature and conclusions related to unsteadiness compensation are drawn.

Conclusion:
Chapter 11: conclusion regarding this thesis are drawn.

Appendices:

- Appendix A: mathematical notation used in this thesis.
- Appendix B: experimental results for flicker compensation.
- Appendix C: experimental results for unsteadiness compensation.
Chapter 2

Flicker

Flicker in archived film sequences is presented in this Chapter. An overview of flicker's origins and characteristics is given in Section 2.1. It is then demonstrated in Section 2.2 that flicker is a spatially varying artefact which is a key factor to consider for flicker modelling.

2.1 Flicker in archived film sequences

Flicker refers to random temporal fluctuations in image intensity and is one of the most commonly encountered artefacts in archived film. Inconsistent film exposure at the image acquisition stage is its main contributing cause. Other causes may include printing errors in film processing, film ageing, multiple copying, mould, and dust.

Film flicker is immediately recognisable even by non-expert viewers as a signature artefact of old film sequences. Several frames of three sequences degraded by this artefact are presented in Figure 2.1. Its perceptual impact can be significant as it interferes substantially with the viewing experience and has the potential of concealing essential details. In addition it can be quite unsettling to the viewer, especially in cases where film is displayed simultaneously with video or with electronically generated graphics and captions as is typically the case in modern-day television documentaries. It may
Chapter 2. Flicker

also lead to considerable discomfort and eye fatigue after prolonged viewing. Camera and scene motion can partly mask film flicker and as a consequence, the latter is much more noticeable in sequences consisting primarily of still frames or frames with low-motion content. In addition it must also be pointed out that inconsistent intensity between successive frames reduces motion estimation accuracy and by consequence the efficiency of compression algorithms.

Flicker has often been categorised as a global artefact in the sense that it usually affects all the frames of a sequence in their entirety as opposed to so-called local artefacts such as dirt, dust, or scratches which affect a limited number of frames and are usually localised on the image plane. Nevertheless it is by no means constant within the boundaries of a single frame as explained in the next Section and one of the main aims of this work is to address this issue.

Figure 2.1: Frames 33-36, 13-16 and 31-34 of test sequences Caption, Tunnel and Lumière respectively.
2.2 A spatially variable artefact

Flicker can be spatially variable and can manifest itself in any one of the following ways. Firstly, when flicker affects approximately the same position of all the frames in a sequence. This may occur directly during film shooting if scene lighting is not synchronised with the shutter of the camera. For example, if part of the scene is illuminated with synchronised light while the rest is illuminated with natural light a localised flickering effect may occur. This can also be due to fogging (dark areas in the film strip) which is caused by the accidental exposure of film to incident light, partial immersion or the use of old or spent chemicals on the film strip in the developer bath. In addition there can be other contributing causes, such as drying stains from chemical agents or vignetting (reduction of image brightness in its borders due to imperfection in the camera lens design) although the latter will not be temporally variable.

It is also possible that flicker localisation varies randomly. This is the case when the film strip ages badly and becomes affected by mould, or when it has been charged with static charge generated from mechanical friction. The return to a normal state often produces static marks.

Figure 2.2 shows the first frame of the test sequence Boat. The camera lingers in the same position during the 93 frames of the sequence. There is also some slight unsteadiness. Despite some local scene motion, overall motion content is low. This sequence is chosen to illustrate that the spatial variation of flicker is not perceivable on the top-left part of the shot, while the bottom left part changes from brighter initially to darker later on. On the right hand side of the image, flicker is more noticeable, with faster variations of higher amplitude. This is shown in Figure 2.2, where the median intensities of four manually-selected blocks (16 x 16 pixels) located at different parts of the frame are plotted as a function of frame number.

The selected blocks are motionless, low-textured and have pairwise similar grey-levels

1Our Shrinking World (1946) - Young America Films, Inc. - Sd, B&W. (1946)
Figure 2.2: Left: test sequence *Boat* used to illustrate spatial variability of flicker measured at selected location. Right, evolution of the median intensity of the selected blocks.

(A, B and C, D) at the start of the sequence. As the sequence evolves we can clearly observe that each block of a given pair undergoes a substantially different level of flicker with respect to the other block. This example also illustrates that flicker can affect only a temporal segment of a sequence. Indeed, from the beginning of the shot to frame 40 the evolution of the median intensities for blocks A and B is highly similar, thus degradation is low compared to the segment that follows the first 40 frames.
Chapter 3

Literature review

Flicker compensation techniques broadly fall into two categories. Initial research addressed flicker correction as a global compensation in the sense that an entire frame is corrected in a uniform manner without taking into account the spatial variability issues illustrated previously. Global flicker models are reviewed in Section 3.1. More recent attempts addressing spatial variability are presented in Section 3.2. Finally flicker compensation frameworks available in the literature are reported in Section 3.3.

3.1 Global compensation

Previous research has frequently led to linear models where the corrected frame was obtained by linear transformation of the original pixel values. A global model was formulated which assumed that the entire degraded frame was affected with a constant intensity offset. In [WS95], flicker was modelled as a global intensity shift between a degraded frame and the mean level of the shot to which this frame belongs. In [DF97] flicker was modelled as a multiplicative constant relating the mean level of a degraded frame to a reference frame. Both the additive and multiplicative models mentioned above require the estimation of a single parameter which although straightforward fails to account for spatial variability.
Chapter 3. Literature review

### Global compensation techniques

<table>
<thead>
<tr>
<th></th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y. Wu and D. Suter</td>
<td>Linear compensation - flicker is modelled as a global intensity shift.</td>
</tr>
<tr>
<td>[WS95]</td>
<td></td>
</tr>
<tr>
<td>E. Decencière</td>
<td>Linear compensation - flicker is modelled as a multiplicative constant.</td>
</tr>
<tr>
<td>[DF97]</td>
<td></td>
</tr>
<tr>
<td>P. Richardson and D. Suter</td>
<td>Histogram stretching across the available greyscale.</td>
</tr>
<tr>
<td>[RS95]</td>
<td></td>
</tr>
<tr>
<td>Y. Wu and D. Suter</td>
<td>Histogram-based compensation - histogram stretching across the reference frame greyscale.</td>
</tr>
<tr>
<td>[WS95]</td>
<td></td>
</tr>
<tr>
<td>P. Schallauer, A. Pinz and W Haas</td>
<td>Histogram-based compensation - histogram equalisation with respect to a reference frame.</td>
</tr>
<tr>
<td>[SPI99, NA00]</td>
<td></td>
</tr>
<tr>
<td>T. Vlachos</td>
<td>Non-linear approach : flicker parameters are estimated independently for each grey-level and a compensation profile is obtained.</td>
</tr>
<tr>
<td>[Vla04]</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1: An overview of the global flicker compensation techniques.

In [RS95] it was observed that archive material typically has a limited dynamic range. Histogram stretching was applied to individual frames allowing the available dynamic range to be used in its entirety (typically [0:255] for 8 bits per pixel images). Despite the general improvement in picture quality the authors admitted that this technique was only moderately effective as significant residual intensity variations remained. The concept of histogram manipulation has been further explored in [WS95] where degradation due to flicker was modelled as a linear two-parameters grey-level transformation. The required parameters were estimated under the constraint that the dynamic range of the corresponding non-degraded frames does not change with time.

Work in [SPI99, NA00] presented a non-linear flicker compensation framework featuring histogram equalisation. A degraded frame was first histogram-equalised and then inverse-histogram was performed with respect to a reference frame. Inverse equalisation was carried out in order for the degraded frame to inherit the histogram profile of the reference. Previous work described in [Vla04] used non-linear compensation motivated by principles of photographic image registration. The model and the main features of the algorithm are summarised in Section 4.1. Table 3.1 presents a brief overview of
3.2 Spatially-adaptive compensation

Recent work has considered the incorporation of spatial variability into the previous models. In [VRLB99] a semi-global compensation was performed based on a block-partitioning of the degraded frame. Each block was assumed to have undergone a linear intensity transformation independent of all other blocks. A linear minimum mean-square error (LMMSE) estimator was used to obtain an estimate of the required parameters. A block-based motion detector was also used to prevent blocks containing motion to contribute to the estimation process and thus the missing parameters due to the motion were interpolated using a successive over-relaxation technique. This smooth block-based sparse parameter field was bi-linearly interpolated to yield a dense pixel-accurate correction field.

Research carried out in [OSKS00, KDPD03] has extended the global compensation methods of [WS95, DF97] by replacing the additive and multiplicative constants with two-dimensional second-order polynomials. An illustration of manually induced spatial variability characteristics that can be portrayed by such polynomial modelling is

Figure 3.1: Two examples of flicker model based on additive and multiplicative 2nd order polynomials estimated in [WS95, DF97].
Chapter 3. Literature review

<table>
<thead>
<tr>
<th>Spatially adaptive compensation techniques</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>T. Ohuchi, T. Seto, T. Komatsu and T. Saito [OSKS00]</td>
<td>linear compensation: flicker is modelled as 2-parameter 2nd order polynomials, hierarchical parameters estimation.</td>
</tr>
<tr>
<td>A. C. Kokaram, R. Dahyot, F. Pitie and H. Denman [KDPD03]</td>
<td>linear compensation: flicker is modelled as 2-parameter 2nd order polynomials, parameters estimation based on an unbiased linear regression.</td>
</tr>
<tr>
<td>J. Jung M. Antonini and M. Barlaud [JAB00]</td>
<td>linear compensation: spatio-temporal segmentation isolating the background and the moving objects. Temporal average of the grey levels preserving the edges to reduce the flicker.</td>
</tr>
<tr>
<td>F. Pitie, R. Dahyot, F. Kelly and A. C. Kokaram [PDKK04]</td>
<td>histogram-based compensation: Joint probability density functions (pdfs) estimated locally in several control points. Dense correction function obtained using interpolation splines.</td>
</tr>
<tr>
<td>F. Pitie, B. Kent, B. Collis and A. C. Kokaram [PKCK06]</td>
<td>pixel-based flicker estimation: flicker strength is estimated for each pixel using a ‘mixing model’ of the global illumination.</td>
</tr>
<tr>
<td>P. Bhat, L. Zitnick, M. Cohen and B. Curless [BZCC08]</td>
<td>GradientShop: minimisation of temporal gradient between motion-compensated frames, assuming that the original frames are flicker-free.</td>
</tr>
</tbody>
</table>

Table 3.2: An overview of the spatially adaptive compensation techniques.

provided in Figure 3.1 and matches the visual impression one gets by inspecting actual flicker-impaired material. In [OSKS00] a robust hierarchical framework was proposed to estimate the polynomial functions, ranging from zero-order to second-order polynomials. Parameters were obtained using M-estimators minimising a robust energy criterion while lower-order parameters were used as an initialisation for higher-order ones. Nevertheless, it has to be pointed out that the previous estimators were integrated in a linear regression scheme, which introduces a bias if the frames are not entirely correlated (Regression ‘fallacy’ or regression ‘trap’ [Sti86], demonstrated by Galton [Gal86]). In [KDPD03] an alternative approach to the parameter estimation problem which tried
3.2. Spatially-adaptive compensation

to solve this issue was proposed. A histogram-based method \cite{PDKK04} was formulated later on and joint Probability Density Functions (pdfs) (establishing a correspondence between grey-levels of consecutive frames) were estimated locally in several control points using a maximum-a-posteriori (MAP) technique. Afterwards a dense correction function was obtained using interpolation splines. The same authors proposed recently in \cite{PKCK06} a flicker model able to deal within a common framework with very localized and smooth spatial variations. The flicker model is parametrised with a single parameter per pixel and is able to handle non-linear distortions. A so-called 'mixing model' is estimated reflecting both the global illumination of the scene and the flicker impact.

A method suitable for motionless sequences was described in \cite{JAB00}. It was based on spatio-temporal segmentation, the main idea being the isolation of a common background for the sequence and the moving objects. The background was estimated through a regularised average (preserving the edges) of the sequence frames, while moving objects were motion compensated, averaged and regularised to preserve spatial continuities. Table 3.2 presents a brief overview of the above methods.

Flicker has also been investigated in \cite{BZCC08} where GradientShop is presented. Research carried out on the human visual system demonstrates that the human vision is more sensitive to local pixel gradients than absolute pixel values. Based on this assumption, GradientShop is a perceptually-motivated optimisation-framework, applicable to image but also video processing. The framework is designed to create filters by controlling the gradient of an image. Possible applications are for instance image sharpening, relighting or de-blocking. In addition, the framework is equipped to preserve image pixel values. An optimisation formulation is proposed with a cost-function composed of desired image gradient and data fidelity constraints. The authors noticed that applying such filter on individual frames create a flicker artefact when the resulting sequence is played-back. As a consequence, a third constraint is incorporated in the optimisation problem to preserve the temporal coherence of the sequence, based on temporal gradient between motion-compensated frames. A flicker-free sequence is
obtained if the temporal gradient of the original and filtered frames are equal. This approach is very interesting, but not applicable to film restoration as it relies on the assumption that the original sequence is flicker-free.

### 3.3 Compensation for sequences of longer duration

While the above efforts addressed the fundamental estimation problem with varying degrees of success far fewer attempts were made to formulate a complete and integrated compensation framework suitable for the challenges posed by processing longer sequences. In such sequences the main challenges relate to continuously evolving scene motion and illumination rendering considerably more difficult the appointment of reference frames. In [SPH99] reference frames were appointed and a linear combination of the inverse histogram equalisation functions of the two closest reference frames (forward / backward) was used for the compensation. In [NA00] a target histogram was calculated for histogram-equalisation purposes by averaging neighbouring frames' histograms within a sliding window. This technique was also used in [De106], but there the target histogram was defined as a weighted intermediary between the current frame and its neighbouring histograms, the computation being inspired from Scale-Time Equalisation theory.

In [VRLB99] compensation was performed recursively. Error propagation is likely in this framework as previously generated corrections were used to estimate future flicker parameters. A bias was introduced and the restored frame was a mixture of the actual compensated frame and the original degraded one. In [KDPD03, PKCK06] an approach motivated by video stabilisation described in [DF97] is proposed. Several flicker parameter estimations are computed for a degraded frame within a temporal window and an averaging filter is employed to provide a degree of smoothing of those parameters.

It is worth noting that [Vla04] is a proof-of-concept algorithm that was originally designed to compensate frame-pairs but was never engineered as a complete solution
Figure 3.2: Comparison of mean frame intensity as a function of time between the original, the baseline scheme [Vla04] and the compensation framework proposed in this thesis.

for long-duration sequences containing arbitrary camera and scene motion, intentional scene illumination changes and spatially-varying flicker effects. This is demonstrated in Figure 3.2 where the algorithm in [Vla04] achieves flicker removal by stabilising the global frame intensity over time but only with respect to the first frame of the sequence which is used as a reference. In contrast the proposed algorithm is well-equipped to deal with motion, intentional illumination fluctuations and spatial variations and, together with a shot change detector, it can be used as a complete solution for any sequence irrespective of content and length. This will be demonstrated in this thesis.

### 3.4 Commercial solutions for archived film restoration

Commercial solutions especially engineered for film restoration have been proposed to compensate for flicker, as well as unsteadiness. Diamant [Dia09] and Revival [Rev09], commercialised respectively by Hs-Art [HA09] and Da-Vinci [DV09] are able to compensate for most of the artefacts encountered in archived film materials. This includes dust, dirt, blotches, mould, bacteria, hairs, scratches, unsteadiness, flicker, film grain, noise, bad splices, tears, burned frames, warped images or dead pixels. Correct V7 [v709], proposed by Mti-Film [Fil09] can be shipped with two packages designed for film restoration: "Restoration Artist" and "Restoration Expert". The later one contains
tools for the reduction of flicker and unsteadiness. Digital Vision [Vis09] commercialises
two pieces of software, *DVO Steady* and *DVO flicker* [SF09] designed for these arte-
facts. Snell and Wilcox [SW09] proposes a hardware-based solution called *Archangel
Ph.C* [PH.09] which includes filters for the compensation of flicker and unsteadiness.
Finally, Sarnoff Corporation [Cor09] offers *Video ResolvR* [Res09] which provides tools
for video enhancement and stabilisation.

More generic tools for post-production such as *Combustion* [Com09] from Auto-desk,
*Shake* [Sha09] from Apple or *After Effects* [AE09] from Adobe feature algorithms for
image sequence stabilisation as it is a rather generic problem, as opposed to flicker
which is mainly related to archived film sequences. Finally, there are many plug-ins
(sometime free) for flicker and unsteadiness designed for video editing software such as
Premiere, Final Cut, After Effects or Virtual Dub. This includes *MSU deflicker filter*
[fp09] and *Deshaker* [plu09a] for Virtual Dub or *Film Fix* [plu09b] for After Effects.

It is worth pointing out that the model and the method of estimation of flicker and
unsteadiness employed by most of these solutions is kept hidden, due to the commercial
nature of the products. As a consequence, it is difficult to assess and understand their
output because many parameters are unknown.
Chapter 4

Non-linear modelling

This Chapter summarises the previous work carried out at the CVSSP [IVSC09] reported in [Vla04], which addressed the problem using photographic acquisition principles leading to a non-linear intensity error profile between a reference and degraded frame. The proposed model assumes that flicker is originated from exposure inconsistencies introduced at the acquisition stage or during film-to-video conversion. It is worth mentioning that flicker shares similar properties when arising from these two sources and that it is not possible to differentiate artefacts introduced within a camera or generated by a telecine. Quadratic and cubic models are provided to cope with telecine grading as explained in Section 4.2, which means that the method presented in Chapter 5 is able to compensated for other sources of flicker respecting these constraints.

4.1 Intensity error profile estimation based on the Density versus log-Exposure characteristic

The Density versus log-Exposure characteristic $D(\log E)$ attributed to Hurter and Driffield [Mes54] defines the response of a physical film strip to different amount of exposure and is used in this thesis to characterise exposure inconsistencies and their associated density errors.
Chapter 4. Non-linear modelling

Exposure, expressed in lux-second, refers to number of photon reaching the film. It depends of the time of exposition but also the intensity of the radiation. Film density is a measure of the radiation transmitted through the film. It is a well-documented fact [AH77] that the relationship between density and observed image intensity is logarithmic according to the following relationship:

\[ D = \log \frac{I_R}{I} \]  \hspace{1cm} (4.1)

In this equation \( I_R \) is the intensity of the light incident on the film and \( I \) is the intensity of the light transmitted through the film, or observed intensity. Film density is unitless. The relationship between observed intensity \( I \) and density \( D \) can be expressed as follow:

\[ I \propto \exp(-D) \]  \hspace{1cm} (4.2)

The Density versus log-Exposure characteristic (Figure 4.1) has a sigmoid shape and typically consists of a toe region close to the origin, a linear region for midrange values, and finally a shoulder region. In the toe and shoulder, a large change in exposure is necessary to produce a small change in film density. The slope of the linear region is often referred to as gamma and defines the contrast characteristics of the photosensitive material used for image acquisition. In the region, a small change in exposure results
4.1. **Intensity error profile estimation based on the Density versus log-Exposure characteristic**

in large change in film intensity.

Exposure inconsistencies consist of a translation of this characteristic on the left if the film is under-exposed, or one the right if the film is over-exposed. It is assumed in [Vla04] that the main contributing cause of flicker artefacts is exposure inconsistency at the acquisition stage and deduced that the consequence of an exposure error is a corresponding density error $\Delta D$ as depicted in Figure 4.1.

Using Equation 4.1, this density error will correspond to an observed intensity error $\Delta I$ as follows:

\[
\Delta D = D - D' = \log \frac{I'}{I} = \log \left(1 + \frac{\Delta I}{I}\right) \tag{4.3}
\]

where $D', I', D$ and $I$ are flicker-induced and flicker-free densities and intensities, respectively, and the approximation $\log(1 + x) \approx x$ is assumed to hold true for small values of $x$. The latter is a reasonable approximation. Indeed, even for severely degraded film samples, intensity error rarely exceed 10% of the available greyscale, i.e., 20-25 grey-levels for a sequence quantised at 8 bits per pixel. Combining Equations 4.2 and 4.3, it is demonstrated that intensity error $\Delta I$ can be expressed as:

\[
\Delta I \propto \Delta D \cdot \exp(-D) \tag{4.4}
\]

The mapping $I \rightarrow \Delta I$ relates grey-level $I$ in the reference image and the intensity error $\Delta I$ in the degraded image. In other words this mapping determines the amount of correction $\Delta I$ to be applied to a particular grey-level $I$ in order to undo the flicker error. Combining Equations 4.2 and 4.4, $I \rightarrow \Delta I$ can be expressed as a mapping involving density $D$ and the corresponding density error $\Delta D$ as follows:

\[
I \rightarrow \Delta I : \exp(-D) \rightarrow \Delta D \cdot \exp(-D) \tag{4.5}
\]
As the Hurter-Driffield characteristic is usually film stock dependent and hence unknown, $D$ and $\Delta D$ are difficult to obtain. Because typically this characteristic is determined experimentally in a piecewise fashion, a closed-form solution for the error profile would be difficult to obtain. Nevertheless an intensity error profile $\Delta I$ across the entire greyscale can be estimated numerically, and a normalised version of it is plotted in Figure 4.2. In sharp contrast to most models reported in the literature the profile is highly non-linear, concave, peaking at the midgrey region and decreasing at the extremes of the available scale, as plotted in Figure 4.2. As it does not seem to contain any inflection points, a quadratic polynomial could be chosen to approximate the intensity error profile in a parametrised fashion. Nevertheless, telecine grading (contrast, greyscale linearity and dynamic range adjustments performed during film-to-video transfer) can introduce further non-linearity as discussed in [Vla04] and presented in Appendix 4.2. A cubic polynomial approximation is more appropriate in those cases.

An intensity error profile $\Delta I_{t,\text{ref}}$ is determined between a reference and a degraded frame $F_{\text{ref}}$ and $F_t$ respectively, where $I_{\text{ref}}$ and $I_t = I_{\text{ref}} - \Delta I_{t,\text{ref}}(I_t)$ are grey-levels of co-sited pixels in the reference and degraded frames and $\Delta I_{t,\text{ref}}(I_t)$ is the flicker component for grey-level $I_t$. For monochrome 8-bits-per-pixel images, $I_t, I_{\text{ref}} \in \{0, 1, ..., 255\}$. This compensation profile allows to reduce $F_t$ flicker artefact according to $F_{\text{ref}}$. In this framework, $F_{\text{ref}}$ is chosen arbitrarily, as a non-degraded frame is usually not available. It
4.1. Intensity error profile estimation based on the Density versus log-Exposure characteristic

Figure 4.3: Intensity difference histograms \( H_{t,\text{ref}}(50) \) and \( H_{t,\text{ref}}(60) \) and their maxima for two consecutive frames of test sequence Caption.

is assumed that motion content between those two images is low and does not interfere in the calculations. To estimate \( \Delta I_{t,\text{ref}}(I_t) \), pixel differences between all pixels with intensity \( I_t \) in the degraded frame and their co-sited pixels in position \( \vec{p} = (x, y) \) in the reference frame are computed and a histogram \( H_{t,\text{ref}}(I_t) \) of the error is compiled as follows:

\[
\forall F_t(\vec{p}) = I_t : H_{t,\text{ref}}(I_t) = \text{hist}(F_t(\vec{p}) - F_{\text{ref}}(\vec{p}))
\] (4.6)

An example is shown in Figure 4.3 for the test sequence Caption and two sample grey-levels. The intensity error is given by:

\[
\Delta I_{t,\text{ref}}(I_t) = \arg \max \{H_{t,\text{ref}}(I_t)\}
\] (4.7)

The process is repeated for each intensity level \( I_t \) to compile an intensity error profile for the entire greyscale. As the above computation is obtained from real images, the profile \( \Delta I_{t,\text{ref}} \) is unlikely to be smooth and is likely to contain noisy measurements. Either a quadratic or cubic polynomial least-square fitting can be applied to the compensation profile. Cubic approximation is more complex and more sensitive to noise but is able to cope with non-linearity originated from telecine grading (Appendix 4.2), as discussed in [Vla04]:

\[
\tilde{A} = \arg \min_{\tilde{A}} \sum_{I_t} (P_{t,\text{ref}}(I_t) - \Delta I_{t,\text{ref}}(I_t))^2
\] (4.8)
with $\vec{A} = \{a_0, ..., a_O\}$ and $P_{t,\text{ref}}(I_t) = \sum_{k=0}^{O} a_k \cdot I_t^k$ \hfill (4.9)

$\vec{A}$ and $O$ being respectively the polynomial coefficients and order. An example is shown in Figure 4.4. Finally the correction applied to the pixel at location $\vec{p}$ is:

$$F_t(\vec{p}) = F_t(\vec{p}) + P_{t,\text{ref}}(F_t(\vec{p})) \hfill (4.10)$$

4.2 Telecine grading

This section reports research carried out in [Vla04] regarding the influence of telecine grading on flicker compensation profile estimation. Grading refers to adjustments commonly performed by a telecine operator during film-to-video transfers in order to improve visual quality and partially compensate for the different characteristics of film and video. Typically this is achieved by adjusting contrast characteristics, greyscale linearity, and dynamic range. On the $D\log(E)$ characteristic, such operations usually affect the toe, linear, and shoulder regions and are occasionally referred to as lift, gamma, and gain control respectively. Gain control can be used, for instance, to render whites and highlight tones even lighter (positive gain) or darker (negative gain) or
4.2. Telecine grading

Figure 4.5: Left: Input-output characteristic highlighting the effect of positive gain control. Right: Input-output characteristic highlighting the effect of positive lift control.

Figure 4.6: Left: Normalised intensity error profile due to positive gain control. Right: Normalised intensity error profile due to positive lift control.

change their balance. Lift control has a similar effect on blacks and shadow tones such as crushing thin shadows to black (negative lift) or lightening darker shades to reveal more detail (positive lift). As there are no precise analytical expressions to account for grading operations, their effect can be best explained by means of input-output characteristics. Examples are shown in Figure 4.5.

The important consequence of telecine grading, as far as this study is concerned, is that it further emphasises the nonlinear nature of the greyscale error characteristic due
to flicker. The grading of the input-output characteristics shown in Figure 4.5 changes the shape of the intensity error profiles, as highlighted in Figure 4.6 for positive lift and gain control. It can be seen that, contrary to the profile obtained earlier (Figure 4.2), these error profiles contain one or more inflection points, suggesting that any polynomial approximation would have to be cubic or higher.
Chapter 5

Flicker compensation framework

Author's contributions are formulated in this Chapter as follows:

- Section 5.1: grey-level intensity error reliability weighting allowing either to strengthen or reduce the influence of certain grey-levels during the estimation of intensity error profile between two frames of a sequence is presented. The weighting is based on grey-levels frequency of occurrence, but also on the reliability of intensity error estimations.

- Section 5.2: motion-compensated intensity error profile estimation is formulated, and motion prediction error is employed to reduce the influence of poorly compensated pixels in the intensity error profile estimation.

- Section 5.3: a new framework based of the adaptive estimation of intensity error profile over time is developed, allowing the compensation of long duration sequences containing substantial scene motion.

- Section 5.4: motion-compensated grey-level tracing is formulated for the estimation of intensity error profile between distant frames containing inhomogeneous content due to scene motion is real image sequences. Flicker compensation parameters are retrieved by tracing compensation values along trajectories of estimated motion between consecutive frames.
Section 5.5: two algorithms able to cope with flicker spatial variations are presented. Intensity error profile estimation over a block-partition of the frames or regions of uniform intensity obtained through frame segmentation are investigated.

Section 5.6: sequential flowcharts of the proposed flicker compensation algorithms are presented.

5.1 Grey-level intensity error reliability weighting

The first important improvement to the baseline scheme in [Vla04] is motivated by the observation that taking into account the frequency of occurrence of grey-levels can enhance the reliability of the estimation process. Grey-levels with low pixel representation should be less relied upon and vice versa. In addition, $\Delta I_{t,\text{ref}}$ estimation accuracy can vary for different intensities as illustrated in Figure 4.3. It can be seen for example that $H_{t,\text{ref}}(50)$ is spread around an intensity error of 15 and even if the maximum is reached for 12, many pixels actually voted for a different compensation value. On the other hand the strength of consensus (i.e. height of the maximum) of $H_{t,\text{ref}}(60)$ suggests a more unanimous verdict. Thus the reliability of $\Delta I_{t,\text{ref}}$ depends on the frequency of $I_{\text{ref}}$ but also on $H_{t,\text{ref}}$. A weighted polynomial least square fitting [Hub81] is then used to compute the intensity error profile and the weighting function reflecting grey-level reliability is chosen as:

$$r_{t,\text{ref}}(I_t) = \max\{H_{t,\text{ref}}(I_t)\} \tag{5.1}$$

Indeed, if $I_t$ does not occur very frequently in $F_t$ then $r_{t,\text{ref}}(I_t)$ will be close to 0 and reliability will be influenced accordingly. The polynomial $C_{t,\text{ref}}$ parameters are now obtained as the solution to the following weighted least-squares minimisation problem:

$$\tilde{A}^t = \arg \min_{\tilde{A}^t} \sum_{I_t} r_{t,\text{ref}}(I_t) \cdot (C_{t,\text{ref}}(I_t) - \Delta I_{t,\text{ref}}(I_t))^2 \tag{5.2}$$
5.1. Grey-level intensity error reliability weighting

Figure 5.1: Measured and polynomial approximated (dashed: basic fitting - solid: weighted fitting) intensity error profiles as a function of intensity between the first two frames of test sequence Caption. A quadratic polynomial fitting is used. The histogram below shows the normalised confidence values \( r_{t, ref} \) for each grey-level.

An example of reliability distribution \( r_{t, ref} \) is shown at the bottom of Figure 5.1, and highlights that pixel intensities above 140 are poorly represented. A comparison between the resulting unweighted correction profile \( P_{t, ref} \) (dashed line) and the improved one \( C_{t, ref} \) (solid line) confirms that more densely populated grey-levels have a stronger influence on the fidelity of the fitted profile.

A side benefit of this enhancement is that it allows the proposed scheme to deal with compressed sequences such as MPEG material. The quantisation used in compression may obliterate certain grey-levels. An absent grey-level \( I_t \) implies that \( H_{t, ref}(I_t) = 0 \), thus \( r_{t, ref}(I_t) = 0 \), which means that \( \Delta I_{t, ref}(I_t) \) will not be used at all in the fitting process.

Figure 5.2 illustrates the impact of the intensity error profile reliability weighting and
shows the results of the compensation of the second frame (top-right) of test sequence Caption with (bottom-left) and without (bottom-right) this new feature. The first frame (top-left) of the sequence is considered as a reference. The intensity errors and the estimated polynomial fittings are plotted in Figure 5.1. It can be observed that the reliability weighting provides a satisfactory result, the compensated frame having a grey-level dynamic very close to the original. On the contrary, flicker reduction is incomplete when not using the reliability weighting. For instance, the letters displayed on the frame which are too bright compared to the reference frame. It is worth mentioning that the grey-level of these letters oscillates between 120 and 140 on the reference frame. Figure 5.1 shows that the compensation values for this grey-range are not properly fitted to raw intensity errors estimation and this is reflected by observing
5.2. Motion compensated intensity error profile estimation

The mechanism presented in the previous Section works well if motion variations between a reference and a degraded frame are low. However, motion compensation must be employed to cope with longer duration sequences. This will enable the estimation of a flicker compensation profile between a degraded and a motion-compensated reference frame $F_{t,ref}^c$.

Many motion estimators are based on two assumptions commonly violated in real world image sequences. The data conservation constraint implies that the intensity of a region remains constant over time. This hypothesis is of course violated in the presence of flicker. The spatial coherence constraint relies on the fact that neighbouring pixels in an image are likely to belong to the same object and thus that their displacements are gradual, without any abrupt transition. This constraint is also not respected in real world image sequence.

In our work we use the well-known Black and Anandan dense motion estimator [BA96] which has been designed to address the violation of these two constraints. The algorithm outputs a piecewise-smooth dense motion vector field. Other dense or sparse motion estimators could be used depending on robustness and speed requirements. Robustness is crucial as incorrect motion estimation will fail the flicker compensation. The motion compensation error will provide a key influence towards intensity error profile estimation. Indeed, Equation 4.6 attributes the same importance to each pixel contributing to the histogram. The motion compensation error is employed to decrease the influence of poorly compensated pixels. This is achieved by compiling $H_{t,ref}^c(I_t)$ presented in Equation 4.6 using real-valued (as opposed to unity) increments $e_{t,ref}^c(p)$ for each pixel located at $p$ (i.e. $F_t(p) = I_t$) according to the following relationship:
Figure 5.3: Comparison of mean frame intensity as a function of time for test sequence "Tunnel" and flicker-compensated sequences with and without the motion prediction error weighting feature.

\[ e_{t,ref}^c(\vec{p}) = 1 - \frac{|E_{t,ref}^c(\vec{p})|}{\max\{|E_{t,ref}^c(\vec{p})|\}} \]  

(5.3)

\( E_{t,ref}^c \) being the motion prediction error expressed by intensity differences between frames, i.e. \( E_{t,ref}^c = F_{ref}^c - F_t \). Thus \( e_{t,ref}^c(\vec{p}) \) varies between 0 and 1 and is inversely proportional to \( E_{t,ref}^c(\vec{p}) \), and so high confidence is placed on pixels with a low-motion compensation error and vice versa. In other words, areas where local motion can be reliably predicted (hence yielding low levels of motion compensation error) are allowed to exert high influence on the estimation of flicker parameters. Pixels with poorly estimated motion, on the other hand, are prevented from contributing to the flicker correction process.

Figure 5.3 highlights the improvement brought by the motion compensated intensity error profile estimation by showing the mean frame intensity of the de-flickered frames plotted as a function of time. Test sequence "Tunnel" is uniform in terms of grey-level dynamic and constant mean frame intensity over time would be expected. As a consequence, an efficient de-flickering algorithm should reduce the standard deviation
5.3. Adaptive estimation of intensity error profile

The baseline compensation scheme described in [Vla04] allows the correction of the degraded frame according to a fixed reference frame $F_{\text{ref}}$ (typically the first frame of the shot). This is only useful for the restoration of static or nearly-static sequences as performance deteriorates with progressively longer temporal distances between a compensated frame and the appointed reference especially when considerable levels of camera and scene motion are present. In addition it gives incorrect results if $F_{\text{ref}}$ is degraded by other artefacts (scratches, blotches, special effects like fade-ins or even MPEG compression can damage a reference frame). Restoration of long sequences requires a carefully engineered compensation framework.

Let us denote by $C_{t,R}$ the intensity error profile between frame $F_t$ and flicker-free frame $F_R$. We use an intuitively plausible assumption by considering that the average of intensity errors $C_{t,i}(I_t)$ between frames $F_t$ and $F_i$ within a temporal window centred...
Chapter 5. Flicker compensation framework

I. 40 Chapter 5. Flicker compensation framework

Figure 5.5: Comparison of mean frame intensity as a function of time between the original, the baseline scheme [Vla04] and the compensation framework featuring the adaptive estimation of intensity error profile for test sequence Boat.

at frame $t$ yields an estimate of flicker-free grey-level $I_R$. The intensity error $C_{t,R}(I_t)$ between grey-levels $I_t$ and $I_R$ is estimated using the polynomial approximation $C_{t,i}(I_t)$ which provides a smooth and compact parametrisation of the correction profile (Section 5.1)

$$C_{t,R}(I_t) \approx \frac{1}{N} \sum_{i=t-\frac{N}{2}}^{t+\frac{N}{2}} \Delta I_{t,i}(I_t)$$

(5.4)

In other words a correction value $C_{t,R}(I_t)$ on the profile is obtained by averaging correction values $C_{t,i}(I_t)$ where $i \in [t - N/2; t + N/2]$ i.e. a sliding window of width $N$ centred at the current frame. We incorporate reliability weighting (as obtained from Section 5.1) by taking into account individual reliability contributions for each frame within the sliding window which are normalised for unity:

$$C_{t,R}(I_t) \approx \sum_{i=t-\frac{N}{2}}^{t+\frac{N}{2}} r'_{t,i}(I_t) \cdot C_{t,i}(I_t)$$

(5.5)

with $\sum_{i=t-\frac{N}{2}}^{t+\frac{N}{2}} r'_{t,i}(I_t) = 1$

(5.6)
5.4. Intensity error profile estimation between distant frames using motion-compensated grey-level tracing

The scheme is summarised in the block diagram of Figure 5.4. A reliable correction value $C_{t,t}(I_t)$ will have a proportional contribution to the computation of $C_{t,R}(I_t)$. A reliability measure corresponding to $C_{t,R}(I_t)$ is obtained by summing unnormalised reliabilities $r_{t,t}(I_t)$ of interframe correction values $C_{t,t}(I_t)$ inside the sliding window:

$$r_{t,R}(I_t) = \sum_{i=t-a}^{t+a} r_{t,i}(I_t)$$

(5.7)

The improvement brought by this new feature to the baseline scheme [Vla04] is presented in Figure 5.5 where the evolution of the mean frame intensity is plotted as a function of time between the original, the baseline algorithm and the proposed approach featuring the adaptive estimation of intensity error profile. The baseline scheme operates by estimating flicker between degraded frames and the first frame of the sequences, and as a consequence the brightness of the compensated sequence is always aligned with the reference frame. On the contrary, it can be observed that the adaptive estimation of intensity error profile deals with intentional brightness variations.

5.4 Intensity error profile estimation between distant frames using motion-compensated grey-level tracing

As Frames $F_i$ and $F_t$ can be distant in a film sequence, large motion may interfere and the motion compensation framework presented is Section 5.2 cannot be used directly as it is likely that the two distant frames are entirely different in terms of content. To overcome this we first estimate intensity error profile between motion-compensated consecutive frames. Raw intensity error profiles and associated reliabilities are computed between consecutive frames in both directions yielding values to $\Delta I_{t,t+1}$, $\Delta I_{t+1,t}$ and $r_{t,t+1}, r_{t+1,t}$ for $t = [0; L]$, $L$ being the number of frames of the sequence (flow-chart 5.11, first stage). The mapping functions are then combined as follows:

$$\Delta I_{t,t+2}(I_t) = \Delta I_{t,t+1}(I_t) + \Delta I_{t+1,t+2}(I_t + \Delta I_{t,t+1}(I_t))$$

(5.8)
which can be generalised for $\Delta I_{t,t\pm i}, i > 2$. This amounts to tracing correction values from one frame to the next along trajectories of estimated motion. The associated reliability is computed as follows:

$$r_{t,t+2}(I_t) = \min(r_{t,t+1}(I_t), r_{t+1,t+2}(I_t + \Delta I_{t,t+1}(I_t)))$$

(5.9)

The above generalises for any frame-pair (flow-chart 5.11, second stage). If a specific correction $\Delta I_{t,t\pm i}$ is unreliable then the min operator above ensures that the compound reliability $r_{t,t\pm i}(I_t)$ will also be rendered unreliable.

A numerical example is presented in Figure 5.6 where correction of grey-level 15 between frames $F_t$ and $F_{t+1}$ is estimated as $\Delta I_{t+1,t+2}(15) = -1$. Thus, grey-level 15 is mapped to grey-level 14 in $F_{t+1}$. As $\Delta I_{t+1,t+2}(14) = 2$ we have $\Delta I_{t,t+2}(15) = -1 + 2 = 1$. Nevertheless we know that $r_{t+1,t+2}(14) = 0.1$ which means that $\Delta I_{t+1,t+2}(14)$ is unreliable. As a consequence $\Delta I_{t,t+2}(15) = 1$ is not a trustworthy estimation and its reliability computed as $r_{t,t+2}(15) = \min(r_{t+1,t+1}(15), r_{t+1,t+2}(\Delta I_{t+1,t+2}(15))) = \min(0.9, r_{t+1,t+2}(14)) = \min(0.9, 0.1) = 0.1$ reflects that.

In the same manner we find that $\Delta I_{t,t+2}(20) = \Delta I_{t+1,t+1}(20) + \Delta I_{t+1,t+2}(\Delta I_{t,t+1}(20) + 20) = 5 + \Delta I_{t+1,t+2}(25) = 8$ and $r_{t,t+2}(20) = \min(r_{t+1,t+1}(20), r_{t+1,t+2}(\Delta I_{t,t+1}(20) + 20)) = \min(0.9, 0.1) = 0.1$.
5.4. Intensity error profile estimation between distant frames using motion-compensated grey-level tracing

\[ (I(t), r_{t+1}, r_{t+2}(25)) = \min(0.7, 1) = 0.7 \]

The improvements brought by the motion-compensated grey-level tracing feature are pictured in Figure 5.7. Sequence Greatwall contains a panoramic motion, observable by tracking the location of the Great Wall over the frames presented in the sliding window. The position of the bottom part of the wall is very different between frames \( F_t \) and \( F_{t-7} \). As a consequence, flicker estimation without any motion compensation...
would be inaccurate. In addition, motion estimators reported in the literature are usually able to handle a displacement capped to a defined limit. The displacement between these two frames is large, thus grey-level tracing is necessary. The bottom part of the wall is mainly composed of dark grey-levels and it can be observed that the intensity errors are poorly estimated and compensated for this grey-levels range. On contrary, the new feature is able to cope with large displacement as can been seen on the flicker-compensated frame.

5.5 Spatial adaptation

The compensation scheme presented in Chapter 5 performs well if the degraded sequence is globally affected by flicker artefact. However, as illustrated in Section 2.2 this is not always the case. Two spatially adaptive algorithms are presented in this Section. Block-based partitioning is employed in Section 5.5.1 while spatial adaptation is achieved by taking into account regions of homogeneous intensity in Section 5.5.2.

5.5.1 Block-based spatial adaptation

Spatial adaptation is achieved by means of block-based frame partitioning (similar to [VRLB99]) and the mechanism is illustrated in Figure 5.9. Compensation profiles $C_{t,R,b}$ are computed independently for each block $b$ of frame $F_t$ using the temporal window approach detailed in Section 5.3. The algorithm is summarised in the flow-chart presented in Figure 5.11. As brute force correction of each block would lead to blocking artefacts at block boundaries (Figure 5.8), a weighted bi-linear interpolation is used.

It is assumed initially that flicker is spatially invariant within each block. For each block a correction profile is computed independently between $I_R$ and $I_t$, yielding values for $\Delta I_{t,R,b}, C_{t,R,b}$ and $r_{t,R,b}, b = [1; B], b$ being the block index and $B$ the total number of blocks.
5.5. Spatial adaptation

Figure 5.8: Left: compensation of the frame 20 of the test sequence Boat applied independently on each block of a $3 \times 3$ grid. As expected blocking artefacts are visible. Right: compensation using the spatially adaptive version of the algorithm.

Blocking is avoided by applying bi-linear interpolation of the $B$ available correction values $C_{t,R,b}(F_t(\vec{p}))$ for pixel $\vec{p}$. Interpolation is based on the inverse of the Euclidean distance $c_b(\vec{p}) = \sqrt{(x - x_b)^2 + (y - y_b)^2}$:

$$d_b(\vec{p}) = \frac{1}{c_b(\vec{p}) + 1} \quad (5.10)$$

$d_b(\vec{p})$ being the weight assigned to the compensation value estimated for pixel $\vec{p}$ in block $b$ and $(x_b, y_b)$ being the coordinates of the centre of the block $b$ for which the block-based correction derived earlier is assumed to hold true.

This interpolation smooths the transitions across blocks boundaries. In addition, reliability measurements $r_{t,R,b}$ of $C_{t,R,b}$ detailed in Section 5.3 are also used as a second weight in the bi-linear interpolation. This allows to discard measurements coming from blocks where $F_t(p)$ is poorly represented. Polynomial approximation on blocks with a low grey-level dynamic will only be accurate on a narrow part of the greyscale, but rather unpredictable for absent grey-levels. $r_{t,R,b}$ is employed to lower the influence of such estimation. Intensity error estimation $C_{t,R,b}$ are finally weighted by the product of the two previous terms, giving equal influence to distance and reliability. In general it is possible to apply unequal weighting. If the distance term is favoured, unreliable compensation values will degrade the quality of the restoration. If the influence of the distance term is diminished, blocking artefacts will emerge as shown in Figure 5.8.
Figure 5.9: Block-based partition of the first frame of Boat using a $3 \times 3$ grid. The pixel undergoing compensation and the centre of each block are represented by black and white dots respectively. The black lines represent the Euclidean distances $c_b(p)$. Polynomial correction profiles $C_{t,ref,b}$ and associated reliabilities $r_{t,ref,b}$ are available for each block $b$. Compensation value for pixel $\tilde{p}$ is obtained by a bi-linear interpolation of the block-based compensation values (9 in this example). Bilinear interpolation involves weighting by block-based reliabilities and distances $d_b$.

The optimality of this trade-off has not been investigated, but it has been experimentally observed that equal weights provide a good balance between the two. The final correction value is then given by:

$$F'(\tilde{p}) = F_t(\tilde{p}) - \sum_{b=1}^{B} [d_b(\tilde{p}) \cdot r_{t,R,b}(F_t(\tilde{p}))] \cdot C_{t,R,b}(F_t(\tilde{p}))$$  \hspace{1cm} (5.11)$$

with $$\sum_{b=1}^{B} [d_b(\tilde{p}) \cdot r_{t,R,b}(F_t(\tilde{p}))] = 1$$  \hspace{1cm} (5.12)$$

Figure 5.9 illustrates the bi-linear interpolation scheme. It shows block-partitioning, computed compensation profiles and reliabilities, and distances $d_b$. For pixel $\tilde{p}$ the corresponding compensation value is given by bi-linear interpolation of the block-based compensation values, weighted by their reliabilities and distances $d_b$.

5.5.2 Segmentation-based spatial adaptation

So far entire blocks have been considered for the compensation profile estimation. It was shown that the weighted polynomial fitting and the motion prediction are capable
of dealing with outliers. However, it is also possible to enhance the robustness and the accuracy of the method by performing flicker estimation of regions of homogeneous brightness. The presence of outliers (Figure 4.3) is reduced in the compensation profile estimation and the compensation profile (Figure 5.1) is computed on a narrower grey-level range, improving the polynomial fitting accuracy.

A straightforward approach would be to segment one frame into regions of homogeneous intensity and estimate region-based intensity error profiles. Interpolation of the compensation values would then be performed to compensate each pixel. This method has two drawbacks. First, the interpolation requires the definition of a distance between a pixel and a region. For instance, distance between a pixel and the centre of gravity of a region, or the closest pixel belonging to this region are two possible definitions. It appears unclear to decide the metric to use to perform an interpolation. Second, flicker estimation must be localised to handle spatial variability. Figure 5.10 shows a segmentation map overlaid on top of the 20th frame of test sequence Tunnel and highlights that certain regions are spread over large part of the image.

In our approach we incorporate segmentation information to the block-based spatial
adaptation presented in section 5.5.1. A degraded block is divided into sub-regions of uniform intensity and one compensation profile is estimated per sub-region afterwards. The most reliable sections of the obtained profiles are combined to create a compound compensation profile for a block. The popular unsupervised segmentation algorithm called JSeg [DM01] is used to partition the degraded image $F_t$ into uniform regions (Figure 5.10). The method is fully automatic and operates in two stages. Firstly, grey-level quantisation is performed on a frame based on peer group filtering and vector quantisation. Secondly, spatial segmentation is carried out. A $J$-image where high and low values correspond to possible regions boundaries is created using a pixel-based so-called $J$ measure. Region growing performed within a multi-scale framework allows refining the segmentation map. The choice of segmentation algorithm is not of particular importance and alternative approaches such as Meanshift [CM02] or Statistical region merging [NN04] can also be employed for segmentation with similar results as the ones presented later in this Section.

The segmentation map is then overlaid onto the block grid, generating block-based sub-regions $F_{t,b}^k$, $k$ being the index of the region within the block $b$. Block partitioning allows to deal with flicker spatial variability while grey-level segmentation permits to estimate flicker in uniform regions. Local compensation profiles $C_{t,ref,b}^k$ and associated reliabilities $r_{t,ref,b}^k$ are then computed independently on each sub-region of each block. $k$ compensation values are then available for each grey-level and the aim is to retain the most accurate one. The quality of the region-based estimations is proportional to the frequency of occurrence of grey-levels. Reliability measurement $r_{t,ref,b}^k$ presented in Section 5.1 is employed to reflect the quality of the region-based compensation values estimation. The block-based compensation value associated with grey-level $I_t$ for block $b$ is obtained by maximising the reliability $r_{t,ref,b}^k$ for the $k$ region-based compensation values estimation:

$$
C_{t,ref,b}(I_t) = \max_{r_{t,ref,b}(I_t)} \{C_{t,ref,b}^k(I_t)\} \quad (5.13)
$$
5.5. Spatial adaptation

\[ r_{t,ref,b}(I_t) = \max_k \{ r_{t,ref,b}^k(I_t) \} \]  \hspace{1cm} (5.14)

Finally, \( \max_k \{ r_{t,ref,b}^k(I_t) \} \) is retained as a measure of the block-based compensation value reliability.
5.6 Algorithms flowchart

Figure 5.11: Flow chart of the proposed block-based flicker compensation algorithm
Figure 5.12: Flow chart of the proposed segmentation-based flicker compensation algorithm
Chapter 6

Experimental results

The nonlinear flicker model based on the Density versus Log-exposure characteristic, the adaptive compensation framework using motion-compensated grey-level tracing and the block-based/segmentation-based spatial adaptation methods presented in Chapters 4 and 5 are assessed by comparing our results with competing algorithms on several test sequences.

Test material and competing algorithms retained for evaluation purpose are detailed in Section 6.1 while the protocol and measurements used for comparison are presented in Section 6.2. An objective comparison between the different methods is given in Section 6.3, and the impact of the block-based/segmentation-based spatial adaptation on the results is analysed in Section 6.4. Finally, the complexity of the consider flicker compensation techniques in studied and compared in Section 6.5. Graphical plots employed to assess the efficiency of the proposed method are shown at the end of this Chapter. Additional plots are presented in Appendix B.

6.1 Test material

The proposed two flicker compensation frameworks (using block-based and segmentation-based spatial adaptation) are compared with two spatially-adaptive state-of-the-art
Chapter 6. Experimental results

<table>
<thead>
<tr>
<th></th>
<th>Roosmalen [VRLB99]</th>
<th>Pitié [PDKK04]</th>
</tr>
</thead>
<tbody>
<tr>
<td>flicker model</td>
<td>linear affine model</td>
<td>Probability density functions (pdfs) used as a model</td>
</tr>
<tr>
<td>motion estimation</td>
<td>motion is detected on each blocks. Blocks presenting moving content are discarded</td>
<td>No motion estimation</td>
</tr>
<tr>
<td>spatial adaptation</td>
<td>interpolation of the block-based flicker estimation using a successive over-relaxation technique</td>
<td>pdfs estimated in several control points and interpolated using splines</td>
</tr>
<tr>
<td>compensation framework</td>
<td>Recursive compensation. Resulting frame is a mixture of original and flicker-compensated frame to avoid error accumulation</td>
<td>temporal filtering of flicker parameters using a sliding window concept</td>
</tr>
</tbody>
</table>

Table 6.1: Description of the competing algorithms used for flicker compensation assessment

techniques, detailed respectively in [VRLB99] and [PDKK04]. The competing algorithms are presented in Chapter 3 and their main features are summarised in Table 6.1.

Seven CIF resolution (360 × 288) monochromes test sequences, Caption, Boat, Lumière, Lostworld, Tunnel, Broadway and Greatwall composed of 50, 93, 198, 147, 50, 148 and 141 frames respectively are used for evaluation purposes. A selection of frames of these sequences is presented in Figure 6.1 and the sequences are available for viewing at: http://www.ee.surrey.ac.uk/Personal/G.Forbin/index.html. Each of these sequences represents historical footage and are therefore susceptible to other archive-related artefacts (such as dirt, unsteadiness and scratches) in addition to flicker. The available sequences are not excessively long, but they still provide substantial motion and a reasonable amount of frames to test and highlight the efficiency of the proposed algorithm.

The first three sequences, Caption, Boat, Lumière contain slight unsteadiness but sub-
6.1. Test material

substantial flicker. The impairments are global in the first sequence while they are localised in the last two sequences as explained in Section 2.2. Caption simply shows an unsteady but otherwise static film caption. Sequence Boat was extracted from MPEG-1-coded footage Our Shrinking World (1946) and presents a person repairing a boat. Finally, Lumièrè is also largely static expect for object motion towards the bottom of the image. It depicts people walking in the Trocadero garden during the Universal Exposition of 1900.

Sequences Lostworld, Tunnel, Broadway contain substantial scene motion. Lostworld depicts a woman who remains almost still in the first hundred frames and then leaves the scene. Tunnel shows the conversation between two soldiers, one of them emerging in the scene from a hatch by climbing up a ladder. Finally Broadway depicts three men counting horses in an outdoor ranch. A fade-in opens this sequence, which is a special effect that flicker compensation algorithms should be able to avoid confusion with genuine flicker.

Finally the last sequence Greatwall, extracted from an MPEG-1-coded video contains a panoramic scan of the Chinese Great Wall, thus the entire frame contents undergo global motion.
Chapter 6. Experimental results

Figure 6.1: Frames 33-36, 15-18, 31-34, 133-136, 13-16, 110-113 and 68-71 of test sequences Caption, Boat, Lumière, Lostworld, Tunnel, Broadway and Greatwall.
6.2 Evaluation protocol

Direct measurement using a panel of human observers would be the ideal way to assess the quality of flicker-compensated image sequences. This is a time consuming process which must be performed under controlled conditions. A common alternative is to employ a vision model which provides estimates of the perceived differences between original and compensated frames outputted by flicker compensation algorithms under consideration.

The well-known Sarnoff Just-Noticeable Differences (JND) [Lub97] developed by Sarnoff Corporation [Cor09] has been designed for this purpose. This model is based on known physiological and psychophysical principles of human visual discrimination performance. It generates a JND map on which each point represents an estimate of the magnitude of perceptual differences between two frames. The model is currently integrated in ClearView Shuttle DVI, Shuttle Broadcast and Extreme DVI Solutions commercialised by Video Clarity [Cla09]. Due to fees involved by the commercial nature of these products, alternative solutions designed for flicker compensation assessment have been proposed in the literature.

For each test sequence, a 4 x 4 grid-partitioning (Section 5.5) is employed. In addition, the temporal window length (Section 5.3) is set to 15 frames centred at the current degraded frame. In [VRLB99, OSKS00, KDPD03, Vla04, PDKK04, PKCK06, De10], flicker reduction algorithms are evaluated by examining the variation of the mean frame intensity over time, computed as follows:

\[ M_1(t) = \frac{1}{N} \sum_{\tilde{p}} F_1(\tilde{p}) \]  

Measurements are presented in Figure 6.2, B.1, B.2, B.3 for each of the test sequences. The smoother the curve, the better the compensation is supposed to be.
It is also useful to compare the standard deviation of each frame as a good-quality compensation should not distort the greyscale dynamic range of the original frames. Time-normalised cumulative standard deviation are computed as follows:

\[ M_2(t) = \frac{1}{t} \sum_{k \in [0, t]} \sigma(F_k) \]  

(6.2)

Measurements for the available sequences are presented in Figure 6.3, B.4, B.5, B.6.
A flicker reduction algorithm should preserve the greylevel dynamic of the frames and as a consequence the curves for the original and flicker-compensated sequences should follow the same trend.

These measurements cannot highlight the spatial variation issues discussed earlier in 2.2. Three new visualisation methods are proposed in order to highlight spatial variability. These provide flicker compensation objective measurements for sequences impaired by localised flicker and containing substantial scene motion.

The first representation is similar to the traditional technique described above, the mean frame intensity being presented as a surface plot on a block-by-block basis rather than for the entire frame. Results for test sequence Caption and Tunnel are plotted in Figures B.7, B.8 and B.9, B.10 respectively.

Let us now consider a pair of flicker-compensated frames. In the case of a near-perfect correction, the first frame and the motion-compensated second one should be very similar, the differences being only due to motion estimation inaccuracy. The remaining two of the new visualisation techniques are based on this hypothesis and assess the similarity between those two images as follows:

- The absolute difference between co-sited pixels of the above frames is averaged. In addition this average is weighted for each pixel by considering the motion prediction error introduced in Section 5.2. The better the compensation, the closer to zero this value should be. The measurement is applied to image sequences by
6.3 Results evaluation against competing algorithms

accumulating measurements obtained for pairs of consecutive frames. Normal-
ising the values by the running total number of frames give more clarity to the
plots:

\[
M_3(t) = \frac{1}{t} \sum_{k \in \{0, t-1\}} \left( \frac{1}{N} \sum_{\bar{y}} |F_k(\bar{y}) - F_{k,k+1}(\bar{y})| \cdot e_{k,k+1}(\bar{y}) \right)
\]  

(6.3)

Results are presented in Figures 6.4, B.11, B.12 and B.13. An efficient de-
flickering algorithm should minimise \(M_3(t)\) over time.

- A threshold on the available greyscale (typically between 0 and 255) is applied.

Then the percentage of co-sited pixels having an absolute difference lower than
this threshold is counted. Each pixel's influence is weighted by the motion predic-
tion error. A curve for the entire greyscale is then compiled by suitably moving
the threshold across the scale. This measurement is formulated as follows:

\[
M_4(t) = \frac{1}{t} \sum_{k \in \{0, t-1\}} \rho \left( |F_k(\bar{y}) - F_{k,k+1}(\bar{y})| \right) \cdot e_{k,k+1}(\bar{y})
\]  

(6.4)

with \(\rho(x) = \begin{cases} 
\rho(x) ; x \leq t \\
0 ; x > t
\end{cases} \)

(6.5)

Results are presented in Figure 6.5, B.14, B.15 and B.16 for the seven test se-
quences under consideration. Figure 6.5 shows that, for test sequence Lumière,
89% of co-sited pixels between successive motion compensated frames have an ab-
solute difference lower than 15. Using the proposed approach, this measurement
is increased to 95%, meaning that successive motion compensated frames are on
overall more similar. A high slope characterises an efficient flicker compensation
algorithm.

6.3 Results evaluation against competing algorithms

Overall, our results show that the four competing algorithms perform well both in
terms of measured performances and qualitative assessment. Figure 6.2, B.1, B.2, B.3
demonstrates that a smoothing of frame mean intensity variation is achieved so the
global flicker component is substantially reduced. It must be noticed that plots gen-
erated from Roosmalen method are somehow more noisy than plots generated from
the two others techniques for several test sequences (the compensated frame being a
mixture of the corrected and degraded ones), and this is visually confirmed. This per-
formance difference is significantly more noticeable in Figures 6.3, B.4, B.5, B.6 where
the time-normalised cumulative standard deviation of the frames are plotted. Pitié’s
de-flickering algorithm and the proposed techniques are able to preserve the dynamic
range characteristics of the sequences, whereas a dramatic reduction may be observed
for Roosmalen’s method. Next we assess performance in relation to spatial variability.

Figures B.7, B.8 and B.9, B.10 show the mean block intensity of test sequences Caption
and Tunnel respectively presented as surface plots. The same remarks as before can be
made concerning Roosmalen’s method. In the case of Caption, curves looks smoother
for the proposed technique while for the Tunnel sequence Pitié’s method has an ad-
vantage. Better discrimination can be obtained by examining Figures 6.4, B.11, B.12,
B.13 which shows the average variation between motion-compensated frames. It may
be observed that the proposed technique compares favourably for all test sequences
except Broadway where Roosmalen’s method gives the best results. Nevertheless this
is not visually confirmed and a dynamic range reduction is evident as suggested by our
previous evaluation metric.

Finally the percentage of pixels having a lower absolute difference than a variable
threshold is computed in Figures 6.5, B.14, B.15, B.16. The higher the percentage the
better the performance of the scheme under assessment. Also in this case our method
performs best except for Broadway (for the same reason as above). Test sequences
and results obtained with the different approaches are available at: http://www.ee.
surrey.ac.uk/Personal/G.Forbin/index.html.
6.4 Impact of the block-based/segmentation-based spatial adaptation

For most of the test sequences, Figures 6.4, B.11, B.12, B.13 highlight that both the block-based (Section 5.5.1) and segmentation-based (Section 5.5.2) spatial adaptation algorithms perform well and compete favourably with the two other flicker compensation methods. However, using segmentation information for the estimation of intensity error profile does not meet the expectation as the results are outperformed by the standard block-based spatial adaptation.

Segmentation-based spatial adaptation performs intensity error estimation on homogeneous regions. Only the most representative region for a specific grey-level is considered and thus some relevant pixels can be omitted if they do not belong to this region. On the other hand the block-based spatial adaptation provides a wider pixel support as all the available pixels contribute to the intensity error estimation. The robustness introduced by the histogram voting system exposed in Section 4.1 added to the reliability weighting presented in Section 5.1 filter out properly the invalid estimations. By considering the various measurements presented in 6.2, we can conclude that this filtering method is superior to the pixels rejection mechanism introduced in the segmentation-based spatial adaptation.

In addition, segmentation maps computation between consecutive frames and segmentation-based spatial adaptation have a severe impact on the overall computational load of the flicker compensation algorithm, as demonstrated in Section 6.5. As a consequence the block-based spatial adaptation method must be preferred.

6.5 Complexity of the proposed algorithms

In this Section we discuss computational complexity issues of the proposed algorithms and compare it with competing techniques. Measurements were carried out on a Pentium 4 platform, running at 3GHz and equipped with 512 MB RAM. The algorithm was
Table 6.2: Number of frames processed per second for the different compensation techniques.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>352 x 288 res.</td>
<td>0.62</td>
<td>0.45</td>
<td>0.80</td>
<td>0.55</td>
</tr>
<tr>
<td>720 x 576 res.</td>
<td>0.35</td>
<td>0.26</td>
<td>0.43</td>
<td>0.27</td>
</tr>
</tbody>
</table>

We can see that the proposed algorithm using the block-based spatial adaptation (Section 5.5.1) is located between Pitié's (GPU implementation) and Roosmalen's techniques in terms of computation time. Additional effort required to employ segmentation information is noticeable on this table, the segmentation-based spatial adaptation scheme being slower than the other competing techniques. It has to be pointed out that the figures reported in this table do not take into account the pre-processing stage of computation the segmentation map, so the full algorithm is actually even slower. As mentioned in Section 6.4 segmentation-based spatial adaptation does not improve the quality of flicker compensation so this extra computational cost is not justified.

It has to be pointed out that a GPU implementation is proposed in Pitié's technique [PDKK04]. For clarity in this Chapter we focus only on CPU implementations. An interesting point to highlight is the fact that the relation between frame resolution and
processing time is not linear in our implementation. The weighted bi-linear interpolation used during block-based spatial adaptation (Section 5.5.1, Equation 5.11) is the most time consuming stage of the algorithm and would require a degree of further optimisation (Assembler, GPU programming).

6.6 Conclusions

In the first part of this thesis, a new scheme for flicker compensation was introduced. The approach was based on non-linear modelling introduced in previous work and contains important novel components addressing successfully the challenges posed by the spatial variability of flicker impairments, the unreliability of raw correction parameters, the adaptive estimation of flicker compensation profile for long duration sequences and also scene motion. Our results demonstrate that the algorithm is very effective towards flicker compensation both in qualitative and quantitative terms and compares favourably to state-of-art methods featured in the literature.
Figure 6.2: Comparison of mean frame intensity as a function of time for test sequences \textit{Lumière} and \textit{Lostworld}. 
Figure 6.3: Comparison of time-normalised cumulative standard deviation for test sequences *Lumière* and *Lostworld*.
Figure 6.4: Comparison of time-normalised cumulative average of absolute differences between consecutive motion-compensated frames for test sequences Lumiére and Lost-world.
Figure 6.5: Comparison of percentage of motion-compensated pixels having an absolute difference lower than a variable threshold for test sequences *Lumière* and *Lostworld*. 
Chapter 7

Unsteadiness

Unsteadiness (also referred to as instability or shake) has long been identified as one of the most severe artefacts of film scanned for television display. The main contributing causes of film unsteadiness have been associated with mechanical tolerances of various components involved in film production, film processing and film-to-video transfer. Examples are the film sprocket hole positioning accuracy of the camera and telecine mechanisms, the perforation accuracy of positive and negative stock and the relative positioning of positive and negative stock in the printer. Film unsteadiness can be very unsettling to the viewer. As is the case with flicker, unsteadiness is especially noticeable when film is displayed simultaneously with video or electronically generated graphics and captions which is a common occurrence in modern day documentaries.

Film ageing is a common source of unsteadiness. Over time, storage humidity and tension winding have a severe impact on film strips, and on their perforations as shown of Figure 7.2. Instability is then created during telecine transfer or film scanning as pictured in Figure 7.1, especially if the transport system uses sprockets relying on film perforations. Other deteriorations - shown in Figure 7.3 - such as brittleness, shrinkage, buckle, curl edge weave and spoking [Kod09] are frequent in archived materials as detailed in [EBU01, EBU04a]. Telecines are usually not equipped to deal with this kind of deformations [EBU04b], although preprocessing of the film is achievable using a wet-gate [JKC00] apparatus placed before the telecine gate. This technology is mainly
Figure 7.1: Frames 12, 13 and 14 of test sequences Joan2; 197, 188 and 199 of test sequence Calendar; 7, 8 and 9 of test sequence biscotin.

employed to physically fill-in film scratches with chemical liquid and to remove dust on the film strip. The film is then 'soften' before entering the telecine gate and local deformation are slightly reduced. Nevertheless this technique is unpopular because of the chemical used (perchloroethylene) in the process which tend to leak into the telecine and create ultimately corrosion problems. In addition, perchloroethylene can have a severe impact on the operators health, due to its toxic and carcinogen properties.

Degraded perforations generate bi-dimensional jitter, but can also include small rotation and ratio changes, depending on perforations' alteration. Previous deteriorations introduce local deformations on the video or digital transfer. As a consequence, motion models characterising unsteadiness created by mechanical inaccuracies in film transfer or recording devices and local film deformations are slightly different to the ones usually employed to model instabilities originated by hand held camera motion and thus need
Figure 7.2: Damaged film strips. Left: Local deformations and cracking. Right: damaged perforations.

Figure 7.3: Film strip degraded by local deformations.

to be characterised differently.
Chapter 8

Literature Review

A digital image stabilisation system is usually composed of two units: a motion estimator, which computes the global motion parameters between successive frames of a sequence and a motion compensation unit which generates unsteadiness compensation parameters by low-pass filtering the previous estimations and compensates the frames.

Motion estimation is a widely used technique in computer vision, and is a key element in areas such as video compression, restoration, enhancement, summarisation but also image alignment and registration. Research has been mainly focused on compensation of unsteadiness created by hand-held camera motion during film footage. The acquisition device is located in a three-dimensional space and outputs a two-dimensional perspective projection of the real scene [Cou07]. The z-dimension corresponds to the depth of the scene and is unavailable in projected two-dimensional images. In the special case of archived film materials, the camera is often mounted on a stand and is distant from the scene. Motion is usually small between successive frames. As a consequence, translational displacements (track, boom) are difficult to differentiate from rotational movement (pan, tilt) and approximations must be formulated to estimate global motion parameters between two frames. Possible camera motions are illustrated in Figure 8.2.

Global motion models reported in the literature for camera and film unsteadiness are
Figure 8.1: Basic set of two-dimensional motion transformations [Sze06].

detailed in the Section 8.1 while an overview of the global motion estimation algorithms is given in Section 8.2. Finally, techniques employed to compute unsteadiness compensation parameters from motion estimation between successive frames are reviewed in Section 8.3.

8.1 Global motion models

An exhaustive overview of the global motion models available in the literature is presented in [Sze06]. This research focuses on parametric motion models which establish a mathematical relationship between pixels of a frame and another. This relation can be simple (2D translation for instance), but the degree of complexity can increase easily depending on the nature of the motion between the frames. Two-dimensional motion models are often reported in the literature and applied to film images. They are able to cope with film deformations enumerated above as opposed to three-dimensional motion models which are mainly used for motion generated by a camera located in a three-dimensional space.

Quadratic motion model is employed in [ZM100, GT02, Gu100] to characterise three-dimensional motion of objects projected in a two-dimensional space, assuming that the camera is far from the scene. It is mainly used for applications requiring high accuracy, but it is very sensitive to noise as it requires a precise estimation of the second order polynomial terms.
Very often, this level of accuracy is not required and many applications are based on downgraded version of the quadratic motion model, faster to estimate and less sensitive to motion noise. The affine model is commonly used and provides a good balance between stability and motion complexity. Motion between frames is approximated by considering that it is only composed of a rotation, a zoom and a translation along the x and y axis. Simplified models have been reported in the literature where it is assumed that the rotation and the zoom are equals on both x and y axis (similarity transformation). The Euclidean model characterises only rotation and translation and can be simplified even more with the translational model.

The projective motion model characterises successfully motion created by a camera mounted on a stand. Translations are in-existant and the camera can only rotate along its axis (pan, tilt, zoom). A schematic representation of these transformations is shown in Figure 8.1, while a description of the possible camera motion is presented in Figure 8.2.

8.2 Global motion estimation

Once a suitable global motion model have been selected to characterise motion between two frames, several methods are reported in the literature to estimate its parameters. A global motion estimator aims at retrieving the dominant motion, which is often associ-
ated with camera motion in real image sequences as opposed to local motion describing movement of the characters or objects. It would be straightforward to assume that the dominant motion, or most represented motion between two frames corresponds to the motion of the camera. This hypothesis is often wrong: local motion is usually not negligible in film footage because the camera frequently focuses on characters or specific objects and as a consequence the estimation of camera parameters is far from being trivial. Two main approaches for motion parameters estimation coexist. Direct (commonly referred to as pixel-based) and features-based methods [Sze06] are respectively reviewed in Section 8.2.1 and 8.2.2. Their output is either a dense or sparse motion vector field between two frames which is used as an input to parametric motion model estimation. Parameters estimators applied to sparse motion vector fields reported in the literature are presented in Section 8.2.3.

8.2.1 Feature-based motion estimation

Feature-based methods rely on key points (distinctive features of each image) which are extracted for each frames and mapped between the two frames. Parametric global motion estimation is then performed on feature points' displacement using parameter estimators such as the ones presented in Section 8.2.3. Features-based methods are nowadays very robust as long as feature points are equally distributed spatially over the frames [BSW05]. Robustness of the feature extractor is a key element in this approach and repeatability is crucial as corresponding features must be extracted in the two frames. Features are usually extracted on regions of high cornerness [HS88], but new extractors have been recently proposed able to locate blob-like regions [Low04] as well as uniform areas [TVG04]. In addition, key points located at the edges on the frames are usually preferred as they are likely to be part of the background of the scene commonly associated with camera motion.

Feature-based motion estimation is used in [MC96] for real-time stabilisation (15 frames/seconds) purpose. A small set of points is tracked between two frames using multi-resolution Laplacian pyramid images. In [Leo98] a set of candidates edges are
located within the two images and motion vectors are block-searched in a hierarchical structure. In [XL06], motion estimation is performed on feature-points using circular block matching which is supposed to be invariant to rotation. Stationary wavelet transform is used in [PRB03]. Translational motion is first estimated by projecting vertical and horizontal details of stationary wavelet decomposition in level 2. Translational parameters are used as initial transformation values and a gradient based algorithm relying on feature points is proposed to estimate motion parameters of a similarity transformation. Kalman filtering is employed in [CFR99] to predict feature points' position.

8.2.2 Direct motion estimation

Correlation between frames is estimated on a pixel basis for direct methods and every pixel contributes to the correlation measurement. Methods reported in the literature perform global motion estimation either on entire frames, or using a block-partition of the images (method often referred as block-matching). Pixels or blocks are mapped between two frames based on their similarity, yielding respectively a dense or sparse motion vector field. Global motion parameters are finally estimated with respect to this displacement field as explained in Section 8.2.3.

Global motion estimation on entire frames

Many estimators reported in the literature compute motion parameters using all frames pixels. They are employed for a wide variety of applications, such as real-time calculation of translational motion, or accurate estimation of high-order motion models. In this case the computational speed aspect is neglected.

Integral projection matching is presented in [Rat98, CDK+04]. Frames are projected on coordinate axis where one-dimensional matching is performed to retrieve translational motion parameters. In [ED00], a translational motion model is estimated using phase correlation [KH75] on the entire frames. In [DF97, TCR02], translational motion is estimated by minimising the sum of co-sited pixels differences between between
one frame and a shifted version of the other one. This method is commonly known as
frame matching. The search is performed on a limited space to decrease the overall com-
putation load but is performed in an exhaustive way to avoid falling into local optimum.

A method based on phase correlation [KH75] and log-polar magnitude spectra repre-
sentation ([RC96]) is proposed in [Ert03b]. Phase correlation is employed twice: to
compute rotational and scale parameters on the log-polar representation of these and
to estimate afterwards the translation parameters on the Cartesian representation of
the compensated images.

Optical-flow [LK81, BBPW04] is employed in [CHCC02] and [HAD+94] to respectively
estimate an Euclidean and a similarity transformation. The velocity field is computed
on a pixel basis and least-square regression is applied to the optical-flow field to re-
trieve the motion parameters. A similar approach is proposed in [HAD+94, MOG+06]
where a hierarchical framework based on a pyramids of Laplacian images is developed
to decrease the computational load. In [KDPD03] a combined approach for flicker
and unsteadiness compensation is presented. An affine model coping with similarity
transformations but also stretching and skewing of film materials suffering severe de-
teriorations is estimated. Finally, a robust M-estimator is employed in an iterative
fashion to minimise the residual between a frame and a motion-compensated version
on it.

Global motion estimation on a block-partition of the frames

In this approach, frames are partitioned into blocks and translational motion between
cosited blocks are estimated yielding to a block-based two-dimensional displacement
field between frames. It is assumed that complex motion can be locally simplified to
translations if the size of the block is reasonably small. The motion vector field is
then employed to estimate global motion parameters. Filtered is often applied prior
to the parameters estimation in order to distinguish vector belonging to global and lo-
cal motion, which are treated as outliers and are discarded as explained in Section 8.2.3.
Block-matching consists of maximising the correlation between pixels of a block belonging to a frame and a shifted version of the co-sited block on the other frame. Various implementations are reported in the literature, focusing either on the error metric or the search technique. Full search is guaranteed to provide the best translational vector minimising the error metric but due to calculation and time constraints, some faster search methods are often considered such as three step search [KIII+81], orthogonal search [PHS87] or cross search [Gha90]. Hierarchical coarse-to-fine techniques have also been proposed to speed up the block-based motion estimation. A multi-resolution image pyramid is constructed where search is initially performed at the coarse level. Motion estimation from a coarse level is then used to initialise the search at the next finer level [Ana89, BAHH92]. Sum of Absolute Differences (SAD) or Square Differences (SSD) are often used as an error metric, but some robust functions have also been proposed in [Hub81, BA96, Ste99].

Phase correlation is employed in [KII75, Bro92] to estimate the translational displacements between two blocks. The output of the phase correlation is a set of peaks which represent the translational components between the two sub-images. The larger peak is identified and corresponds to the dominant motion between the two signals. It usually outperforms standard correlation, especially if images are contaminated by noise in a narrow frequency band as explained in [Sze06]. Gradient cross-correlation has recently been introduced in [AV03] and is a promising alternative to phase correlation.

Block-matching and its variants are widely employed in algorithms dedicated to video stabilisation. Full search block-matching is performed on 16 x 16 pixels blocks in [Che00]. MPEG motion vectors are often computed using block-matching algorithms and are employed in [VCMM02] for image sequence stabilisation purpose. Point matching is employed in [HLL05, UMI<sup>+</sup>90]. Implementations aiming at reducing the computational load are also reported in the literature. Template matching is proposed in [OHSK06] where only relevant blocks are selected. Bit-plane and grey-coded bit-plane
matching are presented in [KLL98, SJSHSWES99]. Boolean-based block-matching is performed on four sub-images respectively on bit-plane and grey-coded decomposition of the images.

In [BMT94, Tho87] phase correlation surfaces [KH75) are computed on a block-based partition (64 x 64 pixels) of the frames. Only four sub-images (64 x 64 pixels) are taken into account in [Ert03a] to speed up the motion estimation. The same framework is used in [YE05] but a one-bit-transform of the frames is proposed to decrease even more the computational load.

### 8.2.3 Robust motion parameter estimation

Features-based or block-based motion estimators generate a sparse motion vector field which establish a correspondence between feature points or blocks of two frames. A parametric representation of the motion model requires estimation of motion parameters which fits the best this motion field.

Non-robust estimators have been used for video stabilisation purpose. They are particularly adapted for landscape scene, where local motion is minimum. On the other hand, their output is unreliable if local motion is present. Least Mean Square (LMS) and Weighted Least Mean Square (WLMS) are employed in [HAD+94, MC96, CHCC02] and [MC96] for the respective estimation of a similarity and an Euclidean motion model. Rotational parameter is computed through a least-square estimation on the translational-compensated motion field in [Che00]. A gradient-based method is used in [PRB03] in an Iterative Least Mean Square framework (ILMS) where outliers are rejected at each iteration.

Robust parameter estimator are also widely used for global motion estimation as reported in [Ste99]. For instance median filtering is used in [SJSHSWES99, Che00] to retrieve the translation between successive frames. A robust estimator (soft matching-pixel count (SMPC)) is used instead of the classical Least Mean Square error criterion
8.3 Unsteadiness compensation

in [KS99] to estimation a similarity model. The well-know Least Median of Square (LMedS) [Rou84] estimator is employed in [CFR99, OIIK06] while Least Trimmed Square (TLS) is applied in [XL06] to discard outliers influence in the global motion estimation process. Random Sample Consensus (RANSAC) [Fä81] is employed in [FDW05] for the estimation of an affine or projective motion model using feature points mapping.

M-estimators can also be applied on motion vector fields. Iterative modified residuals estimator is presented in [KDPD03] while maximum-likelihood is employed in [SO00].

8.3 Unsteadiness compensation

Motion estimation between successive frames must be performed to compute unsteadiness compensation transformations. This unit is also supposed to retain smooth intended camera motion and discard unwanted motion. An overview of two unsteadiness compensation schemes is presented in [Ert01], named Motion Vector Integration (MVI) and Frame Position Smoothing (FPS). MVI yields unsteadiness compensation parameters by integrating motion parameters between successive frames of a sequence using a damping coefficient (leaky integrator). It is a first order IIR filter that directly low-pass filter the motion parameters between successive frames to remove interframe shaking. FPS retrieves unsteadiness compensation parameters by accumulating global motion parameters between successive frames. A low-pass filter is then applied of the obtained displacements. MVI can be implemented as a real-time operation, whereas FPS is usually an offline post-processing. It is nevertheless possible to use an IIR filter with a time delay to perform nearly real-time stabilisation.

MVI is employed in [KLL98, SJSHWES99, UMI+90, Rat98, BMT94, SPI199, KS99] while FPS is used in [CDK+04, Leo93, KDPD03]. In [KDPD03] compensation is performed by low-pass filtering global motion parameters computed by accumulating global displacements between consecutive frames. A simple 3 tap median filter is used followed by a 30 tap FIR filter using the Hamming function as the coefficient. The last filter is
nonlinear and is employed because impulsive defects such as large single frame shake are often observed. In [Ert03a, Ert03b, YE05] FPS is employed but absolute frame displacement are Kalman filtered in order to operate on a real-time basis. In [ED00] stabilisation vectors are obtained through DFT filtering of the absolute displacement vectors while in [GE04] the stabilisation scheme is based on an adaptive fuzzy filter.

Nevertheless, it has to be pointed out that these two schemes are based on accumulation of the global motion vectors between successive frames, which is likely to generate accumulation error. FPS and MVI are respectively employed in [Che00] and [VCMM02, HLL05], but the reference frame is dynamically updated to control error propagation. In [Rat98, TCR02, MOG+06] a sliding window is used and stabilisation vectors are obtained by averaging global displacements between the current frames and frames within a limited neighbourhood. This framework is similar to the method proposed in Chapter 5 for flicker compensation, which has been proved to be very efficient for flicker reduction as demonstrated in Chapter 6. It is proposed later in Chapter 9 to employ the same compensation framework for the two artefacts in order to design a combined restoration algorithm.
Chapter 9

Unsteadiness compensation framework

This Chapter presents the author's contributions and is organised as follows:

- Section 9.1: global motion models regarding unsteadiness modelling in archived films are discussed, and two models are found to be suitable for our application.

- Section 9.2: block-based motion estimation and sub-pixel phase correlation techniques are employed to formulate a motion estimator suitable for unsteadiness modelling in archived films.

- Section 9.3: a novel unsteadiness compensation framework is elaborated. Compensation parameters are adaptively estimated over time. The method is inspired from the adaptive estimation of flicker parameters elaborated in Section 5.3 and allows the estimation of unsteadiness compensation motion vector fields in long sequences presentation scene motion.

- Section 9.4: motion vector tracing is formulated to estimate displacement maps between distant frames presenting inhomogeneous content due to scene motion. Motion vector fields between successive frames are traced and accumulated over time.
• Section 9.5: two robust parameter estimators named Random Sample Consensus (RANSAC) and M-estimator sample consensus (MSAC) are reviewed and applied to the parametrisation of unsteadiness compensation motion vector fields.

• Section 9.6: sequential flowchart of the proposed unsteadiness compensation algorithm is presented.

9.1 Global motion model selection

Instabilities in stressed old film are mainly generated during recording, telecine transfer or film scanning and are often caused by damaged perforations and more generally film ageing. In addition, strong deformations of original negative material cause heavy distortions of the representing image sequences.

A key element for unsteadiness compensation is the choice of the global motion model. Our application differs from standard stabilisation systems reported in the literature which are mainly designed for the compensation for instabilities generated by camera motion (hand held device for example). For archived film unsteadiness, it is assumed that camera motion is stable and that unsteadiness is created during recording or telecine transfer. Film instabilities in a telecine transport system based on sprockets will mainly be reflected as two-dimensional translational jitter, which may include small rotation and scale changes depending of the perforation damages. Figure 9.1 shows the transport system of the film scanner 'The Director', developed by LaserGraphics [Las09]. It relies on pin registration transport of the film strip.

State-of-the-Art telecine or film scanners tend not rely anymore on sprockets due to perforations commonly damaged. Capstan-driven transport system controlled by spooling servos are preferred to mechanical pin registration. These solutions are well adapted to archived films, but are usually expensive. In addition, many film materials have been copied over time on different film supports for preservation using sprocket-based devices. Copies are altered by unsteadiness as well, thus film strip motion within the
9.1. Global motion model selection

Figure 9.1: LaserGraphics 'The Director' [Las09]. Motion picture film scanning system. Pin registration-based transport of the film strip.

transport system is a key element to take into account for global motion modeling.

Film stocks tend to lose their flatness over time (curling, dimensional differences between the emulsion layer and the base [Kod09]), and thus the film might not be completely pushed on the transport system, causing variations of the image focus. As a consequence, a scaling factor must be introduced is the global motion model which must cope as well with shrunk materials. Regarding the previous assumptions, the affine motion model is able to describe instabilities arising from film displacement in mechanical devices and shrinkage and is formulated as:

\[
\begin{align*}
x_{t+1} &= a_0 + a_1x_t + a_2y_t \\
y_{t+1} &= b_0 + b_1x_t + b_2y_t
\end{align*}
\]

where a pixel located in coordinates \((x_t, y_t)\) in frame \(t\) is mapped to coordinates \((x_{t+1}, y_{t+1})\) in frame \(t+1\). \((a_0, b_0)\) represent the translational displacement while \((a_1, b_1)\) and \((a_2, b_2)\) model both the rotational and zoom components.
Buckle, curl edge weave and spoking cause local deformations and should be taken into account. Buckle occurs when the edges (along the length) of a film are shorter than the centre, while edge weave or fluting occurs when one or both of the edges (along the length) are longer than the centre. Spoking is caused by loose winding of film that has considerable curl [Kod09]. The quadratic motion model fits all the previously cited requirements. Its linear part can cope with motion of the film in the device transport system, while the high order polynomial terms model the local deformations exposed above. Its complete form is formulated as:

\begin{align*}
    x_{t+1} &= a_0 + a_1 x_t + a_2 y_t + a_3 x_t^2 + a_4 x_t y_t + a_5 y_t^2 \\
    y_{t+1} &= b_0 + b_1 x_t + b_2 y_t + b_3 x_t^2 + b_4 x_t y_t + b_5 y_t^2
\end{align*}

This model is usually employed for applications requiring high accuracy as reported in [ZMI00, GT02, Gul00], but it can become unstable because of the second order polynomial terms which must be well estimated. Quadratic and affine motion model are studied in this thesis, the former offering a high accuracy for local deformations in archived films materials while the later represents a safer option, able to model instabilities arising from film displacement in mechanical devices and shrinkage.

### 9.2 Block-based motion estimation using phase correlation

This Section is organised as following: Section 9.2.1 presents briefly the phase correlation technique for motion estimation between blocks while sub-pixel motion accuracy is detailed in Section 9.2.2.

#### 9.2.1 Phase correlation method

Phase correlation [KII75, Bro92] is well adapted for block-based motion estimation in archived film sequences. It is reasonably fast to compute (real-time implementation
9.2. Block-based motion estimation using phase correlation

Figure 9.2: Phase correlation surface $C$ between two bi-dimensional signals. Coordinates of the main peak give a sub-pixel accuracy estimation of the translational displacement between the two signals.

can be achieved, is robust to local artefacts and is able to deal as well with violation of the brightness consistency over frames and large motion. It is used in a wide range of broadcasting application [Tho87]. It consists of computing the normalised cross power spectrum between the Fourier transform of the two considered blocks. Let us denote $\mathcal{F}\{F_b^t\}$ and $\mathcal{F}\{F_b^{t+1}\}$ the respective Fourier transform of blocks $F_b^t$ and $F_b^{t+1}$ located respectively in frame $t$ and $t + 1$. The normalised cross power spectrum is given by:

$$\mathcal{S}_{t,t+1}^b = \frac{\mathcal{F}\{F_b^t\} \cdot \mathcal{F}^*\{F_b^{t+1}\}}{|\mathcal{F}\{F_b^t\} \cdot \mathcal{F}^*\{F_b^{t+1}\}|}$$

(9.5)

$\mathcal{S}$ contains the phase difference between the two signals. The normalised phase correlation surface is given by calculating the inverse Fourier transform of the cross power spectrum:

$$\mathcal{C}_{t,t+1}^b = \mathcal{F}^{-1}\{\mathcal{S}_{t,t+1}^b\}$$

(9.6)

The normalised phase correlation surface contains several peaks which represent the
translational components between the two blocks. The larger peak refers to the dominant motion, which is commonly associated with global motion. As a consequence, the translational components \([\Delta x^b_{t+1}, \Delta y^b_{t+1}]\) computed for block \(b\) are given by:

\[
[\Delta x^b_{t+1}, \Delta y^b_{t+1}] = \arg \max_{x,y} \{ \varphi^b_{t+1}(x,y) \}
\]

(9.7)

The amplitude of the peak is used as an indicator of the motion estimation reliability:

\[
x^b_{t+1} = \max_{x,y} \{ \varphi^b_{t+1}(x,y) \}
\]

(9.8)

Phase correlation outputs a single high peak for signal containing uniform translation but phase correlation surface can contain several peaks of different amplitude in the presence of multiple translations. Figure 9.2 shows a typical phase correlation surface where the coordinates of the main peak are estimation of the translation between two blocks.

### 9.2.2 Sub-pixel accuracy motion estimation

Spatial sampling operated during acquisition or digitisation of film frames results in diminished accuracy to characterise motion in real image sequences. A key element to consider for motion estimation is sub-pixel accuracy which has a significant impact on motion compensated prediction error. Sub-pixel accuracy motion estimation is nowadays a very popular technique used in the whole computer vision and broadcasting community. Phase correlation technique deliver sub-pixel accuracy at no additional computational cost by fitting an appropriate function around the main peak through interpolation. Coordinate axis are discretised to sub-pixel accuracy and fractional coordinates maximising the interpolated function are retrieved.

Sub-pixel phase correlation is considered in our application. Correlation peak is retrieved for block \(b\), and phase correlation surface is interpolated around its peak. 1/4 pixel accuracy is considered, providing a good trade off between accuracy and computational efficiency. An illustration for uni-dimensional sub-pixel motion estimation
9.2. Block-based motion estimation using phase correlation

Figure 9.3: Sub-pixel motion estimation applied to a uni-dimensional signal. Phase correlation peak is retrieved and curve fitting (red curve) is performed on a limited neighbourhood centred around peak's coordinate $\Delta x^b_{t,t+1}$. Approximations are performed on a fractional basis to retrieve the sub-pixel coordinate $\Delta' x^b_{t,t+1}$ of the function's maximum. $\Delta x^b_{t,t+1}$ is afterwards updated accordingly.

is provided in Figure 9.3, which is extended to bi-dimensional curve fitting for two-dimensional motion vectors.

Reasonable block size must be considered to provide sufficient support to phase correlation surface estimation. It must be pointed out that phase correlation estimation is based on an assumption of continuous edges [KH75] and as a consequence a 'windowing' function (Hamming window for instance) must be applied on the blocks prior to the phase correlation estimation as explained in [OS99]. This allows to reduce boundary discontinuities effects.

Square blocks of $32 \times 32$ pixels are considered. However, overlapping blocks of $96 \times 96$ pixels are employed to provide more support to phase correlation estimation. An overview of the partitioning is provided in Figure 9.4. Sub-pixel accuracy translational motion estimation is performed for all blocks as explained in Section 9.2.2, yielding a motion vector field shown in Figure 9.5. Figure 9.6 offers another representation of the field by considering individually $x$ and $y$ components.
Finally, block-based motion estimation is performed for all frames of the sequence, yielding values to motion vector field $[\Delta x_{t,t+1}, \Delta y_{t,t+1}]$ and vector estimation reliabilities $r_{t,t+1}$, with $t \in [1; L]$, $L$ representing the number of frames composing the sequence. Phase correlation is a symmetric operation, as a consequence we have:

$$[\Delta x^b_{t,t+1}, \Delta y^b_{t,t+1}] = -[\Delta x^b_{t+1,t}, \Delta y^b_{t+1,t}]$$  \hspace{1cm} (9.9)

$$r^b_{t,t+1} = r^b_{t+1,t}$$  \hspace{1cm} (9.10)

And thus uni-directional motion estimation between successive frames provides bi-direction motion vector fields.

### 9.3 Adaptive estimation of unsteadiness compensation parameters

This Section is inspired from Section 5.3 which presents the adaptive estimation of flicker parameters as they evolve over time. In the previous Section, block-based mo-
9.3. Adaptive estimation of unsteadiness compensation parameters

Let us denote $[\Delta x_{t,R}, \Delta y_{t,R}]$ the motion vector field between degraded frame $F_t$ and unsteadiness-free frame $F_R$. It can easily be assumed that the average of translational displacements $[\Delta x_{t,i}^b, \Delta y_{t,i}^b]$ between blocks $F_t^b$ and $F_R^b$ within a temporal window centred at frame $t$ yields an estimation of unsteadiness-free motion vector $[\Delta x_{t,R}^b, \Delta y_{t,R}^b]$. This assumption is credible in the sense that unsteadiness in mainly generated from mechanical impression of periodical nature within film scanner or telecine devices.

As a consequence, a compensation motion vector is estimated by averaging inter-frame block motion vectors $[\Delta x_{t,i}^b, \Delta y_{t,i}^b]$ where $i \in [t - N/2; t + N/2]$, i.e. a sliding window of width $N$ centred at frame $t$. 

Figure 9.5: Estimated motion vector field $[\Delta x_{t,t+1}, \Delta y_{t,t+1}]$ between frames 12 and 13 of test sequence joan2. Estimation have been performed on 16 x 16 pixels block by considering $32 \times 32$ overlapping blocks.
Motion vectors reliability $r_{t,i}^b$ obtained previously are incorporated as well to diminish the influence of poorly estimated block motion vectors, and are normalised for unity. Unsteadiness compensation vector $(\Delta x_{t,R}^b, \Delta y_{t,R}^b)$ for block $b$ of frame $t$ is given by:

$$\begin{bmatrix}
\Delta x_{t,R}^b \\
\Delta y_{t,R}^b
\end{bmatrix} \approx \frac{1}{N} \sum_{i=t-N/2}^{t+N/2} r_{t,i}^b \cdot \begin{bmatrix}
\Delta x_{t,i}^b \\
\Delta y_{t,i}^b
\end{bmatrix}$$  \hspace{1cm} (9.11)

with $\sum_{i=t-N/2}^{t+N/2} r_{t,i}^b = 1$ \hspace{1cm} (9.12)

The process is illustrated in Figure 9.7. More robust methods, such as an histogram voting mechanism of values $\Delta x_{t,i}^b$ and $\Delta y_{t,i}^b$ could also be employed at this stage. However, we observed that averaging inter-frame block motion vector components within a temporal window provides satisfactory results.

Reproducible motion vectors have a proportional influence in the estimation of $(\Delta x_{t,R}^b, \Delta y_{t,R}^b)$. In addition, a reliability measurement $r_{t,R}^b$ related to $(\Delta x_{t,R}^b, \Delta y_{t,R}^b)$ is computed by summing unnormalised reliabilities $r_{t,i}^b$ of interframe block motion vectors within the
Figure 9.8: Vector tracing between motion fields \((\Delta x_{t,t+1}, \Delta y_{t,t+1})\) and \((\Delta x_{t+1,t+2}, \Delta y_{t+1,t+2})\). The closest block centre to the destination of the first motion vector is retrieved and the two motion vectors are added afterwards.

\[
\begin{bmatrix}
\Delta x_{t+1,t+2}^b \\
\Delta y_{t+1,t+2}^b
\end{bmatrix}
= \begin{bmatrix}
\Delta x_{t+1,t+2}^b \\
\Delta y_{t+1,t+2}^b
\end{bmatrix}
+ \begin{bmatrix}
\Delta x_{t,t+1}^b \\
\Delta y_{t,t+1}^b
\end{bmatrix}
\]

(9.14)

with

\[
z = \arg \min_z \{(\Delta x_{t+1,t+1}^b + x_b) - x_z)^2 + ((\Delta y_{t+1,t+1}^b + y_b) - y_z)^2\}
\]

(9.15)

\((x_b, y_b)\) and \((x_z, y_z)\) being respectively centre coordinates of block \(b\) and \(z\). The tracing is illustrated in Figure 9.8 and consists of adding block-based displacements between two frames along trajectory of estimated motion as follows:

- Motion vectors \((\Delta x_{t,t+1}^b, \Delta y_{t,t+1}^b)\) characterises the translational displacement of block \(b\) between frames \(t\) and \(t+1\)
- The position of the centre of block \(b\), \((x_b, y_b)\) in frame \(t\) is motion-compensated using \((\Delta x_{t,t+1}^b, \Delta y_{t,t+1}^b)\) to obtain its position in frame \(t+1\), and more importantly the index of the block \(z\) where it belongs
- Motion vectors \((\Delta x_{t+1,t+2}^z, \Delta y_{t+1,t+2}^z)\) characterises the translational displacement of block \(z\) between frames \(t+1\) and \(t+2\)
- \((\Delta x_{t,t+1}^b, \Delta y_{t,t+1}^b)\) and \((\Delta x_{t+1,t+2}^z, \Delta y_{t+1,t+2}^z)\) are added to obtain the displacement of block \(b\) between frames \(t\) and \(t+2\)

In addition, it is possible to estimate reliabilities \(r_{t,t+2}^b\) of the interframe motion vector \((\Delta x_{t,t+2}^b, \Delta y_{t,t+2}^b)\). Estimation is inspired from the grey-level tracing approach elaborated in Section 5.4 for flicker compensation and is given by:
Figure 9.8: Vector tracing between motion fields \((\Delta x_{t,t+1}, \Delta y_{t,t+1})\) and \((\Delta x_{t+1,t+2}, \Delta y_{t+1,t+2})\). The closest block centre to the destination of the first motion vector is retrieved and the two motion vectors are added afterwards.

\[
\begin{bmatrix}
\Delta x_{t,t+2}^b \\
\Delta y_{t,t+2}^b
\end{bmatrix} =
\begin{bmatrix}
\Delta x_{t,t+1}^b \\
\Delta y_{t,t+1}^b
\end{bmatrix} +
\begin{bmatrix}
\Delta x_{t+1,t+2}^z \\
\Delta y_{t+1,t+2}^z
\end{bmatrix}
\]

(9.14)

with \(z = \arg\min_z \{(\Delta x_{t,t+1}^b + x_b) - x_z)^2 + ((\Delta y_{t,t+1}^b + y_b) - y_z)^2\}\) (9.15)

\((x_b, y_b)\) and \((x_z, y_z)\) being respectively centre coordinates of block \(b\) and \(z\). The tracing is illustrated in Figure 9.8 and consists of adding block-based displacements between two frames along trajectory of estimated motion as follows:

- Motion vectors \((\Delta x_{t,t+1}^b, \Delta y_{t,t+1}^b)\) characterises the translational displacement of block \(b\) between frames \(t\) and \(t+1\)

- The position of the centre of block \(b\), \((x_b, y_b)\) in frame \(t\) is motion-compensated using \((\Delta x_{t,t+1}^b, \Delta y_{t,t+1}^b)\) to obtain its position in frame \(t+1\), and more importantly the index of the block \(z\) where it belongs

- Motion vectors \((\Delta x_{t+1,t+2}^z, \Delta y_{t+1,t+2}^z)\) characterises the translational displacement of block \(z\) between frames \(t+1\) and \(t+2\)

- \((\Delta x_{t+1,t+1}^b, \Delta y_{t+1,t+1}^b)\) and \((\Delta x_{t+1,t+2}^z, \Delta y_{t+1,t+2}^z)\) are added to obtain the displacement of block \(b\) between frames \(t\) and \(t+2\)

In addition, it is possible to estimate reliabilities \(r_{t,t+2}^b\) of the interframe motion vector \((\Delta x_{t,t+2}^b, \Delta y_{t,t+2}^b)\). Estimation is inspired from the grey-level tracing approach elaborated in Section 5.4 for flicker compensation and is given by:
9.5 RANSAC/MSAC for the robust estimation of motion parameters

A parametric representation of the motion model requires estimation of motion parameters which fits the best the motion vector field obtained by phase correlation. However, it is usually contaminated by local motion, and can be altered as well by local artefacts such as scratches and blotches as can be observed in Figures 9.5 and 9.6. Filtering prior to the parameters estimation is necessary in order to retain only vectors belonging to global motion, while discarding outliers and local motion vectors. This Section is organised as follows: Section 9.5.1 presents the different robust parameter estimator commonly used in computer vision while Random Sample Consensus (RANSAC) and M-estimator Sample Consensus (MSAC) estimators are explained in Section 9.5.2. Finally MSAC is applied to parametrise motion vector fields in Section 9.5.3.

9.5.1 Robust parameter estimators

A robust motion estimator is required in to retain only the vectors describing global motion. Robustness of the measurement is usually quantified by considering the breakdown point, i.e. the minimum fraction of outliers that can cause the estimator to diverge from the solution. Least square estimator has a breakdown point equal to 0, as a single outlier can invalidate the estimation and thus least-square is not a robust estimator. An estimator can be qualified as robust when its is not affected by outliers and its robustness can be measured by estimating its breakdown point. The maximum theoretical
breakdown point is 0.5, meaning that if the proportion of outliers represents more than half the data the estimator can lock on outliers [Ste99], as they may optimise the estimator objective function. The objective function, or residual function consists of the distance from the data to the model. In other words, it is the error between raw data and the model estimated by the robust estimator. This function needs to be minimised.

Three major robust estimator are widely employed in computer vision: M-estimators, Least-Median of Square (LMedS) which is a specific type of M-estimator and Random Sample Consensus (RANSAC) [FB81, Ste99, FDW05]. M-estimators are based on iteratively re-weighted least squares, and minimise a robust loss function which is meant to decrease the influence of outliers in the estimation process. The algorithm being iterative, extreme care must be taken for the initialisation in order to avoid local optimums. Least-Square is reported in the literature for the computation of the initial parameters but is commonly replaced by robust estimators such as the one reported afterwards. The choice of the loss function is also very important, and depends of the scale of the residual errors. This information may be known a-priori for certain applications or must be estimated in the opposite case.

Least-Median of Square (LmedS) [Rou84] has a breakdown of 0.5, the highest theoretical possible value. The objective function to minimise is the median of the distances between data and estimated model. As a consequence, the objective function minimum is not affected as long as the outliers proportion is not superior to 0.5. A random sampling technique must be employed as the median function is not differentiable [Rou84, FB81]. A certain number of subsets of the data are extracted and the model is fitted on these subsets. The solution is given by the parameters minimising the objective function. The number of subset to considered is crucial, as it is necessary to have at least one subset containing only inliers. Probabilistic measurements taking into account the a-priori proportion of outliers and the number of data exist to calculate the number of subsets required to ensure that at least one subset contains only inliers.
9.5.2 Random Sample Consensus (RANSAC) and M-estimator Sample Consensus (MSAC)

Random Sample Consensus (RANSAC)

RANSAC [FB81, Ste99, FDW05] has been designed within the computer vision community. As LmedS, a random sampling technique is employed and the number of subset to consider is fundamental. However, it is more flexible as the objective function to minimise is the number of outliers having absolute residuals larger than a certain threshold. In other words, the number of inliers located on a narrow band around the estimated model is maximised. The threshold needs to reflect the scale of the inliers' noise. This feature is interesting for some applications, which requires for example a breakdown point superior to 0.5. RANSAC [FB81] is able to retrieve inliers even if they are in minority in the set of data, as long as the outliers do not follow the same model. This is an interesting feature for global motion estimation. Indeed, it is very common that local motion represents more than half of the motion vectors, if the camera focuses on a character for example. RANSAC is theoretically able to retrieve the motion parameters from vectors belonging to the background, as long as local motion does not follow an affine or quadratic motion model.

RANSAC aims to minimise the cost function:

\[
\begin{align*}
Q_T(\theta) &= \sum_{i=1}^{Q} \rho(d_i(\theta)), \quad \rho(d_i(\theta)) = \begin{cases} 
0 & d_i(\theta) \leq T \\
1 & d_i(\theta) > T 
\end{cases}
\end{align*}
\]

\(f(\theta)\) is minimised by trial-and-error: A minimal subset of data points is selected and model parameters \(\theta\) are estimated, usually through Least-Square regression. Score \(f(\theta)\)
is attributed afterwards. The process is repeated for $\mu$ iterations. Model parameters $\theta$ and inlier set minimising $f(\theta)$ are the results of RANSAC. Figure 9.9 shows an example where RANSAC is employed to estimation a 1-dimensional affine model. The top Figure shows a possible RANSAC-iteration while the bottom one presents the results of the algorithm.

Threshold $T$ and number of iterations $\mu$ are user-dependant. $T$ will be discussed in Section 9.5.3. $\mu$ is closely related to the percentage of outliers $p_0$ expected in the data set, to the size of the data subset $|S|$ (2 points for a line fitting for instance) but also to the probability of failure of RANSAC $P$. RANSAC is a probabilistic technique, and thus is not always successful. In [FDW05], it is demonstrated that the probability $P$ that RANSAC fails is:

$$P(p_0, |S|, \mu) = (1 - (1 - p_0)^{|S|})^\mu$$  \hspace{1cm} (9.18)

Increasing $\mu$ allows to reduce the probability of failure, but comes at extra computational cost. To achieve a maximum error rate of no more that $P$, $\mu$ needs to be greater than:

$$\mu \geq \frac{\log(P)}{\log(1 - (1 - p_0)^{|S|})}$$  \hspace{1cm} (9.19)

Many algorithms based on RANSAC have been proposed aiming at minimising the number of iterations needed for the method to guarantee a given confidence in the optimality of the solution. LO-RANSAC is introduced in [CMO04]. A so-called "LO-step" is performed when a new maximum in the number of inliers in reached. It is observed that the estimated set of inliers can be contaminated, thus it is refined by drawing a constant number of samples from this set and compute their respective scores. The aim is to retrieve an un-contaminated set of inliers.
9.5. **RANSAC/MSAC for the robust estimation of motion parameters**

Randomised RANSAC is presented in [MC04, MC05]. R-RANSAC speeds up the evaluation stage of the model by introducing a two-stage procedure. Statistical tests are performed on a limited set of randomly selected data points. If the test fails, the evaluation of the remaining points is not carried out, which allows to reduce significantly the computational load.

**M-estimator Sample Consensus (MSAC)**

Many research have been carried on RANSAC in order to improve its breakdown point, to estimate automatically the optimum threshold introduced to differentiate inliers and outliers or to select the best data subsets. A cost function where the inlier scores a penalty depending on how well it fits the model has been introduced in MSAC (M-estimator sample consensus) [TZ00] and was found to give better performances than RANSAC. This cost function is adapted in MLESAC (maximum likelihood sampling and consensus) [TZ00] to yield the maximum likelihood estimate. MSAC aims to minimise the cost function:

\[
f(\theta) = \sum_{i=1}^{N} \rho(r_i(\theta)) \\
\rho(d_i(\theta)) = \begin{cases} 
           d_i(\theta) : d_i(\theta) \leq T \\
           T : d_i(\theta) > T 
\end{cases} \quad (9.20)
\]

MSAC is applied to the parametrisation of motion vector fields in Section 9.5.3 and the implemented version is detailed in Algorithm 1.
Figure 9.9: Random Sample Consensus (RANSAC) applied to the estimation of a 1-dimensional affine model. Two points (green points) are randomly selected iteratively and an affine model is computed (blue line). Inliers are defined as having a distance to the model lower than a certain threshold $T$, and must be maximised. In this example, RANSAC will lock on the bottom case.
Algorithm 1 M-estimator sample consensus (MSAC)

Require: data - set of data points
Require: model - model to fit the data points
Require: n - minimum number of points required to fit the model
Require: k - maximum number of iterations allowed
Require: T - Threshold value for determining if a data points fits the model

iteration ← 0
best_score ← +∞
best_model ← ∅
best_inliers_set ← ∅

while iteration < k do
    sample_data ← n randomly selected data points
    model_fitted ← model fitted to sample_data (Least Square)
    model_score ← 0
    model_inliers_set ← ∅
    for all points in data do
        data_distance ← distance(data, model_fitted)
        if data_score ≤ T then
            model_score ← model_score + data_distance
            model_inliers_set ← model_inliers_set ∪ point
        else
            model_score ← model_score + T
        end if
    end for
    if model_score < best_score then
        best_score ← model_score
        best_model ← model_fitted
        best_inliers_set ← model_inliers_set
    end if
    iteration ← iteration + 1
end while
Algorithm 2 Refinement of model parameters obtained with MSAC (algorithm 1)

Require: data - set of data points
Require: model - model to fit the data points
Require: inliers_set - set of inliers obtained with MSAC (algorithm 1)
Require: inliers_set_score - score obtained by inliers_set
Require: inliers_set_model - model estimated for inliers_set
Require: k - maximum number of refinement iterations allowed
Require: T - Threshold value for determining if a data point fits the model

\[
\text{refinement iteration} \leftarrow 0 \\
\text{best_inliers_set} \leftarrow \text{inliers_set} \\
\text{best_score} \leftarrow \text{inliers_set_score} \\
\text{best_model} \leftarrow \text{inliers_set_model} \\
\text{while refinement iteration} < k \text{ do} \\
\quad \text{model_fitted} \leftarrow \text{model fitted to best_inliers_set (Least Square)} \\
\quad \text{model_score} \leftarrow 0 \\
\quad \text{model_inliers_set} \leftarrow \emptyset \\
\quad \text{for all points in data do} \\
\quad \quad \text{data_distance} \leftarrow \text{distance(vector , data)} \\
\quad \quad \text{if data_distance} \leq T \text{ then} \\
\quad \quad \quad \text{model_score} \leftarrow \text{model_score} + \text{data_distance} \\
\quad \quad \quad \text{model_inliers_set} \leftarrow \text{model_inliers_set} \cup \text{point} \\
\quad \quad \text{else} \\
\quad \quad \quad \text{model_score} \leftarrow \text{model_score} + T \\
\quad \quad \text{end if} \\
\quad \text{end for} \\
\quad \text{if model_score} < \text{best_score} \text{ then} \\
\quad \quad \text{best_inliers_set} \leftarrow \text{model_inliers_set} \\
\quad \quad \text{best_score} \leftarrow \text{model_score} \\
\quad \quad \text{best_model} \leftarrow \text{model_fitted} \\
\text{end if} \\
\text{end while} \]
9.5.3 Parametrisation of motion vector fields

MSAC is well adapted to parametrise global motion vector fields. It has been proved to outperform RANSAC, at no extra computational cost [FDW05, TZ00]. In addition, several constraints linked to MSAC can be easily resolved. The main focus being accuracy and not rapidity, the choice of the number of subsets in not particularly important, so the number of MSAC-iterations \( \mu \) (Equation 9.19) can be increased to ensure success of the algorithm.

The number of randomly selected vectors \(|S|\), is respectively set to 3 and 6 for affine and quadratic motion models presented in Section 9.1. This represent the minimum number of vectors necessary to estimate model parameters \( \theta \) through Least-Square regression.

Motion parameters \( \theta \) depends of the motion model selected. For affine motion model, \( \theta \) is defined as:

\[
\theta = \begin{bmatrix} a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \end{bmatrix}
\]  

(9.21)

and is extended for quadratic motion model to:

\[
\theta = \begin{bmatrix} a_0 & a_1 & a_2 & a_3 & a_4 & a_5 \\ b_0 & b_1 & b_2 & b_3 & b_4 & b_5 \end{bmatrix}
\]  

(9.22)

Distance between measured motion vector \([\Delta x_{t,R}^b, \Delta y_{t,R}^b]\) and estimated motion vector \([\theta x_{t}^b, \theta y_{t}^b]\) computed using motion model parameters \( \theta \) is calculated using an Euclidean metric:

\[
d_b(\theta) = (\Delta x_{t,R}^b - \theta x_t^b)^2 + (\Delta y_{t,R}^b - \theta y_t^b)^2
\]  

(9.23)

Threshold \( T \) is employed to separate inliers and outliers and must be proportional to the scale of inliers' noise, and in this case to the scale of the motion estimator noise.
Right selection of $T$ is not critical [FDW05]. If it is set to low valid global motion vectors will be classified as outliers, but it is likely that the set of inliers will still be large enough to give a low score $f(\theta)$ to model $\theta$. In addition, the refinement stage will solve the problem. If it is set too high, then some outliers will be classified as inliers. However, their influence will be negligible as they do not differ much from inliers. $T$ is experimentally set to 1.5 as this value provides relevant discrimination of inliers and outliers in experiments carried out in Chapter 10. An automatic decision rule could be drawn based on the scale of motion vector estimation noise.

Finally, motion vector reliability $r^{b}_{i,R}$ needs to be considered prior to parameter estimation with MSAC. It would be prejudicial to include poorly estimated motion vectors in MSAC data set. This would render the process more complicated by adding extra outliers. A threshold, set experimentally to 0.3 is applied to motion vector reliability $r^{b}_{i,R}$. Irrelevant motion vectors are discarded and are removed from motion vector field processed by MSAC.

Figure 9.10 and 9.10 illustrates the fitting of a quadratic motion model on the x and y components of a motion vector field using MSAC and highlights that outliers do not interfere with motion parameter estimation. Another representation is presented in Figure 9.12 where the raw and fitted motion vector fields are drawn over a frame.
Figure 9.10: First iteration of MSAC for fitting quadratic motion model parameters $\theta$ to the motion vector field. Top: $\Delta x_{t,R}^b$ (blue) and $\delta x_{t,R}^b$ (red) Bottom: $\Delta y_{t,R}^b$ (blue) and $\delta y_{t,R}^b$ (red). The algorithm estimates in an iterative fashion the parameters models $\theta$ fitting raw motion vectors.
Figure 9.11: Final iteration of MSAC for fitting quadratic motion model parameters $\theta$ to the motion vector field. Top: $\Delta x_{t,R}^b$ (blue) and $\delta x_{t,R}^b$ (red) Bottom: $\Delta y_{t,R}^b$ (blue) and $\delta y_{t,R}^b$ (red). The algorithm estimates in an iterative fashion the parameters models $\theta$ fitting raw motion vectors.
Figure 9.12: Top: Raw motion vector field drawn over frame 4 of test sequence \textit{joan2}. Bottom: Fitted motion vector field using a affine model drawn over frame 4 of test sequence \textit{joan2}. 
9.6 Algorithm flowchart

Figure 9.13: Flowchart of the proposed unsteadiness compensation algorithm
Chapter 10

Experimental results

Four test sequences are processed to highlight the effect of the unsteadiness compensation algorithm. Sequences biscotin, calendar, joan2 and presenter are composed respectively of 70, 122, 248 and 281 frames and contain substantial level of unsteadiness. Global motion vectors are computed between successive frames using an affine motion model and accumulated for the original and unsteadiness-compensated sequences. Concerns regarding the quadratic global motion model are formulated in Section 10.1. Compensated sequences are compared with results of [RC96, Ert03b] for objective assessment purpose. In [RC96, Ert03b] phase-correlation-based methods are presented and employed for image sequence stabilisation purposes. As mentioned in Section 8.2.2 phase correlation is robust to local artefacts such as blotches and scratches and is able to deal with temporal brightness fluctuation. It is well adapted to estimate motion in archived film sequences and thus it is reasonable to assess the proposed framework against algorithms drawn in [RC96, Ert03b].

Global motion vector components calculated between successive unsteadiness-compensated frames are accumulated over time and presented in Figures 10.2, 10.3, 10.4 - 10.5, 10.6, 10.7 while additional results are presented in Appendices C.1, C.2, C.3 and C.4, C.5, C.6. Finally numerical values are computed in Table 10.1 by comparing the standard deviation of motion vector components over time for the four test sequences and histogram-based representations are provided in Figures C.7, C.8 and C.9.
Figures enumerated above highlight that unwanted motion is filtered. Indeed, for all the test sequences motion components $a_0, b_0$ (translational motion) and $a_1, a_2, b_1, b_2$ (combination of aspect ratio changes and rotational motion) curves are smoother for the unsteadiness-compensated sequences. The plots illustrate as well that the proposed method achieve better filtering than [RC96, Ert03b]. In these techniques, phase correlation is employed both to estimation global motion parameters. A first stage consists of estimating rotational motion and zoom ratio on a log-polar representation of the images where each axis represent the rotational and zoom ratio parameters. Then phase correlation is employed a second time on a Cartesian representation of the compensated frames to estimate the remaining translational displacement. Experiments have shown that the first stage requires the projected frame dimensions to be extremely large in order to keep a certain degree of accuracy. Rotational and zoom ratio parameters need to be highly accurately estimated, and thus log-polar representation axis grid highly fine. On the other hand, the projected image become blurred when projected onto a large log-polar frame, making the phase correlation estimation less accurate. A correct balance difficult to obtain is required.

The proposed compensation framework is able to discard the unwanted displacements while conserving intentional motion. This is highlighted by the different figures drawn at the end of this Chapter and the additional figures presented in Appendix C and confirmed visually by inspecting the restored sequences. Test sequences and results obtained with the different approaches are available at: http://www.ee.surrey.ac.uk/Personal/G.Forbin/index.html.

10.1 Global motion estimation using quadratic motion model

Graphical plots using a quadratic global motion model are not presented in Appendix C. Experiments have shown that it was difficult using sparse motion vector field to estimate accurately this motion model. Indeed, slight imprecision particularly concerning the
rotation and scale factor estimation are not tolerated and generate important artefact in the unsteadiness-compensated frames as illustrated in Figure 10.1. More research in this area should be carried out in order to model precisely local deformations associated to archived films.

10.2 Conclusions

In the second part of this thesis, an efficient framework for unsteadiness estimation and compensation is proposed. Priority is given to motion unsteadiness in archived film sequences associated with mechanical imprecision of various devices used in film production. Motion estimation is performed on a block-based basis using the well-known phase correlation technique. This method is able to cope with brightness variation and frame deteriorations such as dirt and scratches commonly encountered in archived film sequences. Sparse motion vector fields between consecutive frames are then filtered using a novel motion vector tracing algorithm inspired by the approach developed for flicker parameters estimation. As a consequence, flicker and unsteadiness can be estimated within a common framework. Finally affine or quadratic compensation parameters are obtained through M-estimator Sample Consensus (MSAC) on filtered unsteadiness compensation motion vector field. Our results demonstrate that
the proposed approach performs successfully both in subjective and objective terms and compares favourably to state-of-art methods featured in the literature.
### 10.2. Conclusions

Table 10.1: Standard variation of the affine motion vector components computed over time for the four test sequences.

<table>
<thead>
<tr>
<th></th>
<th>Original</th>
<th>Ertürk [Ert03b]</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a0</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biscotin</td>
<td>1.29</td>
<td>0.42</td>
<td>0.31</td>
</tr>
<tr>
<td>Calendar</td>
<td>1.75</td>
<td>0.57</td>
<td>0.37</td>
</tr>
<tr>
<td>Joan2</td>
<td>0.97</td>
<td>0.24</td>
<td>0.16</td>
</tr>
<tr>
<td>Presenter</td>
<td>25.53</td>
<td>7.03</td>
<td>5.18</td>
</tr>
<tr>
<td><strong>b0</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biscotin</td>
<td>2.17</td>
<td>0.47</td>
<td>0.42</td>
</tr>
<tr>
<td>Calendar</td>
<td>2.25</td>
<td>0.63</td>
<td>0.45</td>
</tr>
<tr>
<td>Joan2</td>
<td>6.78</td>
<td>2.24</td>
<td>2.19</td>
</tr>
<tr>
<td>Presenter</td>
<td>4.09</td>
<td>1.63</td>
<td>1.22</td>
</tr>
<tr>
<td><strong>a1 (\times 10^{-3})</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biscotin</td>
<td>3.31</td>
<td>3.06</td>
<td>2.58</td>
</tr>
<tr>
<td>Calendar</td>
<td>2.74</td>
<td>0.85</td>
<td>0.52</td>
</tr>
<tr>
<td>Joan2</td>
<td>1.53</td>
<td>0.23</td>
<td>0.21</td>
</tr>
<tr>
<td>Presenter</td>
<td>16.98</td>
<td>5.70</td>
<td>3.83</td>
</tr>
<tr>
<td><strong>b1 (\times 10^{-3})</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biscotin</td>
<td>2.14</td>
<td>0.92</td>
<td>0.59</td>
</tr>
<tr>
<td>Calendar</td>
<td>8.20</td>
<td>5.49</td>
<td>4.06</td>
</tr>
<tr>
<td>Joan2</td>
<td>1.51</td>
<td>0.39</td>
<td>0.28</td>
</tr>
<tr>
<td>Presenter</td>
<td>30.11</td>
<td>15.1</td>
<td>10.02</td>
</tr>
<tr>
<td><strong>a2 (\times 10^{-3})</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biscotin</td>
<td>0.44</td>
<td>0.21</td>
<td>0.14</td>
</tr>
<tr>
<td>Calendar</td>
<td>8.19</td>
<td>5.48</td>
<td>4.08</td>
</tr>
<tr>
<td>Joan2</td>
<td>1.53</td>
<td>0.38</td>
<td>0.25</td>
</tr>
<tr>
<td>Presenter</td>
<td>68.52</td>
<td>24.69</td>
<td>19.87</td>
</tr>
<tr>
<td><strong>b2 (\times 10^{-3})</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biscotin</td>
<td>2.11</td>
<td>1.01</td>
<td>0.68</td>
</tr>
<tr>
<td>Calendar</td>
<td>3.25</td>
<td>1.42</td>
<td>1.08</td>
</tr>
<tr>
<td>Joan2</td>
<td>2.05</td>
<td>0.82</td>
<td>0.68</td>
</tr>
<tr>
<td>Presenter</td>
<td>8.56</td>
<td>4.44</td>
<td>2.84</td>
</tr>
</tbody>
</table>
Figure 10.2: Motion vector component $a_0$ and $b_0$ accumulated over frame number for test sequence *Biscotin*. 
Figure 10.3: Motion vector component $a_1$ and $b_1$ accumulated over frame number for test sequence Biscotin.
Figure 10.4: Motion vector component $a_2$ and $b_2$ accumulated over frame number for test sequence Biscotin.
10.2. Conclusions

Figure 10.5: Motion vector component $a_0$ and $b_0$ accumulated over frame number for test sequence *Calendar*.
Figure 10.6: Motion vector component $a_1$ and $b_1$ accumulated over frame number for test sequence *Calendar*. 
Figure 10.7: Motion vector component $a_2$ and $b_2$ accumulated over frame number for test sequence Calendar.
Chapter 11

Conclusion and future research

Digital video technology has developed at exponential speed in the last decade. Scientific progress have been impressive and have allowed public access to a huge range of multimedia content at any time and any place of the world. Web 2.0, mobile communication and digital television are primary examples of technologies which have contributed to the opening and dissemination of broadcaster archives to the audience. TV programs are freely available to download on broadcaster websites and are often repeated on the constantly increasing number of services available on digital television. Television and digital set-up box manufacturers are starting to embed digital recording feature into their products while movies are available to download or stream on commercial website. These multimedia contents can be easily copied into handheld devices which are nowadays able to decode any kind of video compression format and thus can be watch at any location in the world. Public access to this huge source of content has never been so simple.

Nevertheless, multimedia content available on these new platforms is still quite limited as broadcaster and movie providers tend to provide only a selection of their respective copyrighted recent programs and movies. It must be pointed out that these technologies are still in an emerging state and will probably provide a larger selection of multimedia content in the future.
Chapter 11. Conclusion and future research

Archived materials is a perfect candidate of content which would provide multimedia content of quality to these technologies and would please the general public. Archived materials cover an infinite range of multimedia content of historical, cultural and artistic value. Access to the public is very limited nowadays due to several factors related to film processing. Sharing and disseminating digital content is nowadays a trivial task as the recording device, the video processing chain and the receiver are all digital-based. Digital format conversion and digital streaming/broadcasting are the main tools involved and are well-documented and researched scientific topics.

Archived film dissemination is more complex. As mentioned in the introduction of this thesis original copy of the archived have often disappeared or are in such a degraded state that a single playback or copy operation would put at danger the film strip. It is also common that old film have suffered degradation directly at the acquisition stage due to technological limitations. Humidity and dust, chemical instabilities, improper storage and handling practices can damage irreversibly archived materials. Chemical properties of film strip are also very important and can lead to severe degradation even in proper storage conditions. Contemporary storage technology can prolong life expectancy of film materials by 400 years, but old-fashioned technology (acetate-based for instance) have a life expectancy of only 20 years.

Film preservation consists of transferring film materials onto digital support to increase their life expectancy. In the long term, digital film can be re-transferred onto film format for cinema projection. As mentioned above, film may have suffered severe physical damages throughout its lifetime, rendering the digitisation complex without a prior film restoration stage. The cost of manual processes involved in a conventional restoration chain place considerable limitations on processing throughput rendering the restoration of entire collections an unrealistic proposition. Conventional restoration relies on the use of dedicated equipment such as special copying machines which can only target a limited range of artefacts due to the fact that the unit of manipulation can only be the physical film strip. On the other hand, digital image and video processing have recently been applied to film restoration and powerful tools used by professional archivists exist.
nowadays. Risks related to film materials and costs are reduced while restoration can be automated in a faster way.

High-level image processing algorithms tools have been developed in this thesis, targeting brightness variations (flicker) and motion unsteadiness, which are the most commonly encountered artefacts in archived film sequences. Their perceptual impact can be significant as both interfere substantially with the viewing experience and have the potential of concealing essential detail. In operational restoration systems flicker and unsteadiness are invariably among the top-priority targeted artefacts as their correction offers substantial benefits both with regard to visual quality as well as subsequent restoration operations.

In the first part of this thesis, a new flicker compensation algorithm was introduced. The approach was based on previous work carried out at the CVSSP where a non-linear model based on the Density versus log-Exposure characteristic was proposed to describe flicker impairment originated from exposure inconsistencies. The algorithm proposed was able to compensate for flicker between a degraded and a reference frame. Major improvements have been proposed in this thesis to yield a fully-automated flicker compensation algorithm providing satisfactory results and competing favourably against methods reported in the literature. Grey-level intensity error reliability weighting has been introduced to strengthen the influence of certain grey-levels during the flicker compensation profile estimation. Flicker spatial variability has been addressed successfully through local flicker estimation combined with spatial interpolation of flicker compensation values. Local flicker estimation has been perform on a grid partitioning of the frames, but also on regions of homogeneous intensities obtained through image segmentation. Finally, a novel algorithm featuring motion-compensated grey-level tracing has been proposed to estimate flicker parameters between distant frames by tracing flicker compensation values along trajectories of estimated motion between consecutive frames. Combined with temporal filtering of flicker parameters, this method addresses the challenges posed by scene motion in long duration sequences.
The proposed algorithm performs well for real sequences images where flicker is originated from exposure inconsistencies. It can also be pointed out that satisfactory results have been observed when flicker is originated from other sources (fogging, accidental exposure of film to incident light), as long as the artefact respects the proposed model hypothesis. For instance, our model is based on the assumption that flicker can be spatially located, but follows smooth spatial transitions. The algorithm would fail in the presence of abrupt transitions which happen occasionally in archived film sequences. For instance, neon-based lights within the scene could lead to such a situation which would require the estimation of flicker parameters on a extremely fine partition of the frames. Flicker parameters would probably be inaccurate due to the insufficient number of pixels employed.

The proposed algorithm would certainly benefit from a low-level implementation (GPU for instance) to improve computational speed. Indeed, spatial adaptation (block or segmentation based) is achieved through bi-linear implementation of local compensation values where distances from the considered pixel to the centre of each blocks are used as a weights. To speed up the restoration, Euclidian distances have been pre-calculated. Nevertheless, the bi-linear interpolation is still computationally heavy and a low-level implementation would help to achieve a fast implementation. Motion estimation (dense or block-based) could also benefit from such an implementation.

Finally, experiments are currently carried out on colour image sequences. Flicker is estimated on luminosity of the frames using the proposed algorithm. Flicker parameters are then employed to compensate each RGB plane individually. Results are visually satisfactory, however more research would be needed, especially regarding the theory surrounding flicker in colour image sequences. For instance, other colour spaces could be more suitable for flicker parameters estimation and frame luminosity might not be the best plane to estimation flicker parameters. Experimental results obtained so far are very promising and this research path deserves to be explored.
In the second part of this thesis a novel approach for unsteadiness estimation and compensation was proposed. The well-known phase-correlation technique has been employed to estimate local motion between consecutive frames on a block-based basis. This motion estimator is able to cope with brightness variations and frame deteriorations and thus is well-adapted for archived film materials. In addition, it provides sub-pixel accuracy motion estimation at no additional computational costs. Sparse motion estimation between consecutive frames are processed in a novel motion vector fields temporal filtering algorithm featuring motion vector tracing. This method was inspired by the compensation scheme developed for flicker removal. Finally, unsteadiness-compensation motion vector fields are robustly parametrised using a MSAC (M-estimator Sample Consensus) method. Our results demonstrate that the proposed approach for flicker and unsteadiness compensation performs successfully both in subjective and objective terms and compares favourably to state-of-art methods featured in the literature.

The proposed algorithm for unsteadiness estimation compensation is versatile and can be employed with different motion models. Most of the experiments in this thesis have been performed using the affine model, which characterises a wide variety of rigid motion. Results show that the proposed algorithm performs well to estimate and compensate for unsteadiness artefact. Nevertheless, experiments carried out using a quadratic motion model have not been as successful. This model adds more complexity but is also extremely sensitive. It is intended to characterise local film deformations attributed to film ageing and requires the estimation of second-order polynomials. The proposed method is usually able to estimate the model accurately but slight error in the estimation can generate very visible artefact.

Digital restoration tools are nowadays well documented in the literature and a significant amount of efficient algorithms designed for unsteadiness, flicker, scratches or bloches correction have been proposed. Several commercial solutions able to process a wide set of artefacts either in a automatic or semi-automatic fashion are available for film restoration purposes. Digital restoration has become easier and more cost-effective. As stated in the introduction, it is estimated that 90% of silent films and 50% of films
shoted before 1950 have already disapeared, and that a majority of existing films will vanish in the next couple of centuries. The problem must now be tacked urgently at a national or European level to invest in new funds dedicated to film restoration and take decisive measures to prevent archived materials vanishment from the society collective memory.
Appendix A

Notations

The following mathematical notations are used in this thesis:

Flicker part:

- \( F_t \): frame sampled at time \( t \)
- \( F'_t \): flicker compensated frame sampled at time \( t \)
- \( F_{\text{ref}} \): a generic reference frame
- \( L \): total number of frames in a test sequence
- \( \vec{p} = (x, y) \): pixel coordinates
- \( F_t(p) \): grey-level value of frame \( F_t \) at position \( \vec{p} \)
- \( I \): image intensity (grey-level value)
- \( I_t \): intensity \( I \) in frame \( F_t \)
- \( \Delta I_{t,\text{ref}} \): intensity error profile between frames \( F_{\text{ref}} \) and \( F_t \)
- \( \Delta I_{t,\text{ref}}(I_t) \): intensity error for grey-level \( I_t \) between frames \( F_{\text{ref}} \) and \( F_t \)
- \( r_{t,\text{ref}} \): intensity error reliability between frames \( F_{\text{ref}} \) and \( F_t \)
- \( r_{t,\text{ref}}(I_t) \): reliability associated with intensity error \( \Delta I_{t,\text{ref}}(I_t) \)
Appendix A. Notations

- $P_{t,ref}$: polynomial fitted to intensity error profile between frames $F_{ref}$ and $F_t$
- $O$: Polynomial order used for the intensity error profile fitting
- $C_{t,ref}$: weighted polynomial fitted to intensity error profile between frames $F_{ref}$ and $F_t$
- $H_{t,ref}(I_t)$: Histogram of the intensity errors between pixels with intensity $I_t$ in frame $F_{ref}$ and co-sited pixels in frame $F_t$
- $F_{t,ref}^c$: motion-compensated version of $F_{ref}$ relative to $F_t$
- $e_{t,ref}^c$: motion prediction error of $F_{t,ref}^c$
- $E_{t,ref}^c$: error weighting derived from $H_{t,ref}^c(I_t)$
- $H_{t,ref}^c(I_t)$: Histogram of the intensity errors between pixels with intensity $I_t$ in frame $F_t$ and co-sited pixels in motion-compensated frame $F_{t,ref}^c$
- $N$: number of frames in the temporal filtering window
- $F_R$: flicker-free frame $F_t$
- $I_R$: flicker-free intensity $I_t$
- $C_{t,R}$: intensity error profile between flicker-free frames $F_R$ and $F_t$
- $r_{t,R}$: reliability associated with intensity error $C_{t,R}(I_t)$
- $B$: number of blocks considered in the block partitioning scheme
- $C_{t,R,b}$: intensity error profile computed within block $b$ between flicker-free frames $F_R$ and $F_t$
- $r_{t,R,b}$: reliability associated with intensity error $C_{t,R,b}$
- $C_{t,ref,b}^k$: intensity error profile computed within region $k$ between blocks $F_{ref,b}$ and $F_{t,b}$
- $r_{t,ref,b}^k$: reliability associated with intensity error $C_{t,ref,b}^k$
- $d_b(\vec{p})$: inverse of euclidean distance between position $\vec{p}$ and centre of block $b$
Unsteadiness part:

- \( F_t \): frame sampled at time \( t \)
- \( F_R \): unsteadiness-free frame \( F_t \)
- \( F'_t \): unsteadiness compensated frame sampled at time \( t \)
- \( L \): total number of frames in a test sequence
- \((x_t, y_t)\): pixel coordinates in frame \( t \)
- \( B \): number of blocks considered in the block partitioning grid
- \( F^b_t \): block \( b \) of frame \( F_t \)
- \( \mathcal{F} \): Discrete Fourier transform
- \( \mathcal{F}^* \): complex conjugate of \( \mathcal{F} \)
- \( \mathcal{P} \): normalised cross power spectrum
- \( \mathcal{C} \): cross power spectrum
- \( \Delta x^b_{t,t+1} \): translational displacement on the \( x \) axis of block \( b \) between frames \( F_t \) and \( F_{t+1} \), measured by phase correlation method
- \( \Delta y^b_{t,t+1} \): translational displacement on the \( y \) axis of block \( b \) between frames \( F_t \) and \( F_{t+1} \), measured by phase correlation method
- \([\Delta x^b_{t,t+1}, \Delta y^b_{t,t+1}]\): motion vector computed for block \( b \) between frames \( F_t \) and \( F_{t+1} \)
- \( r^b_{t,t+1} \): confidence attributed on motion vector \([\Delta x^b_{t,t+1}, \Delta y^b_{t,t+1}]\)
- \([\Delta x_{t,t+1}, \Delta y_{t,t+1}]\): motion vector field computed between frames \( F_t \) and \( F_{t+1} \)
- \( r_{t,t+1} \): confidence of each vector belong to motion vector field \([\Delta x_{t,t+1}, \Delta y_{t,t+1}]\)
- \([\Delta x_{t,R}, \Delta y_{t,R}]\): motion vector field computed between frames \( F_t \) and unsteadiness-free frame \( F_R \)
130 Appendix A. Notations

- $r_{t,R}$ : confidence of each vector belong to motion vector field $[\Delta x_{t,R}, \Delta y_{t,R}]$
- $Q$ : number of vectors in a motion vector field
- $N$ : number of frames in the temporal filtering window
- $\theta$ : motion model parameters
- $f(\theta)$ : score attributed to motion model parameters $\theta$
- $\rho$ : cost function
- $d_i(\theta)$ : distance from vector $i$ to motion model characterised by $\theta$
- $T$ : threshold used by RANSAC to discriminate inliers and outliers
- $\mu$ : number of iterations performed within RANSAC
- $|S|$ : minimal number of motion vectors necessary to estimate motion model parameters
- $[\theta x_t, \theta y_t]$ : motion vector field computed using motion model parameters $\theta$
Appendix B

Flicker - additional results

Figure B.1: Comparison of mean frame intensity as a function of time for test sequences

Caption.
Figure B.2: Comparison of mean frame intensity as a function of time for test sequences *Boat* and *Greatwall.*
Figure B.3: Comparison of mean frame intensity as a function of time for test sequence 
*Tunnel* and *Broadway.*
Figure B.4: Comparison of time-normalised cumulative standard deviation for test sequences Caption and Boat.
Figure B.5: Comparison of time-normalised cumulative standard deviation for test sequences Tunnel and Broadway.
Figure B.6: Comparison of time-normalised cumulative standard deviation for test sequence *Greatwall*. 
Figure B.7: Mean block intensity for frames 0-49 of test sequenceCaption for original and Roosmalen flicker compensation methods.
Figure B.8: Mean block intensity for frames 0-49 of test sequence. Caption for Pitié and the proposed flicker compensation methods.
Figure B.9: Mean block intensity for frames 0-49 of test sequence Tunnel for original and Roosmalen flicker compensation methods.
Figure B.10: Mean block intensity for frames 0-49 of test sequence *Tunnel* for Pitié and the proposed flicker compensation methods.
Figure B.11: Comparison of time-normalised cumulative average of absolute differences between consecutive motion-compensated frames for test sequences Caption and Boat.
Figure B.12: Comparison of time-normalised cumulative average of absolute differences between consecutive motion-compensated frames for test sequences *Tunnel* and *Broadway.*
Figure B.13: Comparison of time-normalised cumulative average of absolute differences between consecutive motion-compensated frames for test sequence *Greatwall.*
Figure B.14: Comparison of percentage of motion-compensated pixels having an absolute difference lower than a variable threshold for test sequences Caption and Boat.
Figure B.15: Comparison of percentage of motion-compensated pixels having an absolute difference lower than a variable threshold for test sequences *Tunnel* and *Broadway.*
Figure B.16: Comparison of percentage of motion-compensated pixels having an absolute difference lower than a variable threshold for test sequence Greatwall.
Appendix C

Unsteadiness - additional results
Figure C.1: Motion vector component $a_0$ and $b_0$ accumulated over frame number for test sequence *joan2.*
Figure C.2: Motion vector component $a_1$ and $b_1$ accumulated over frame number for test sequence $j Ocean$. 
Figure C.3: Motion vector component $a_2$ and $b_2$ accumulated over frame number for test sequence *joan2*. 
Figure C.4: Motion vector component $a_0$ and $b_0$ accumulated over frame number for test sequence \textit{presenter}. 
Figure C.5: Motion vector component $a_1$ and $b_1$ accumulated over frame number for test sequence presenter.
Figure C.6: Motion vector component $a_2$ and $b_2$ accumulated over frame number for test sequence presenter.
Figure C.7: Standard deviation of motion components $a_0$ and $b_0$ calculated over time for the four test sequences.
Figure C.8: Standard deviation of motion components $a_1$ and $b_1$ calculated over time for the four test sequences expressed in a $10^{-3}$ basis.
Figure C.9: Standard deviation of motion components $a_2$ and $b_2$ calculated over time for the four test sequences expressed in a $10^{-3}$ basis.
Bibliography


