Advances in Image Communication

Volume 9

E.B. Bellers G. de Haan

De-interlacing: A Key Technology for Scan Rate Conversion

ELSEVIER

Series Editor: J. Biemond, Delft University of Technology, The Netherlands

---

Volume 1	Three-Dimensional Object Recognition Systems (edited by A. K. Jain and P. J. Flynn)
Volume 2	VLSI Implementations for Image Communications (edited by P. Pirsch)
Volume 3	Digital Moving Pictures – Coding and Transmission on ATM Networks (J.-P. Leduc)
Volume 4	Motion Analysis for Image Sequence Coding (G. Tziritas and C. Labit)
Volume 5	Wavelets in Image Communication (edited by M. Barlaud)
Volume 6	Subband Compression of Images: Principles and Examples (T.A. Ramstad, S.O. Aase and J.H. Husøy)
Volume 7	Advanced Video Coding: Principles and Techniques (K.N. Ngan, T. Meier and D. Chai)
Volume 8	Image and Video Databases: Restoration, Watermarking and Retrieval (A. Hanjalic, G.C. Langelaar, P.M.B. van Roosmalen, J. Biemond and R.L. Lagendijk)
Volume 9	De-interlacing: A Key Technology for Scan Rate Conversion (E.B. Bellers and G. de Haan)

ADVANCES IN IMAGE COMMUNICATION 9

De-interlacing

A Key Technology for Scan Rate Conversion

E.B. Bellers and G. de Haan

Philips Research
Prof. Holstlaan 4, 5656 AA Eindhoven, The Netherlands

The work described in this book has been carried out at Philips Research Laboratories Eindhoven, The Netherlands, as part of the Philips Research Programme.

2000

ELSEVIER
Amsterdam – Lausanne – New York – Oxford – Shannon – Singapore – Tokyo

ELSEVIER SCIENCE B.V.
Sara Burgerhartstraat 25
P.O. Box 211, 1000 AE Amsterdam, The Netherlands

First edition 2000

Library of Congress Cataloging in Publication Data
A catalog record from the Library of Congress has been applied for.

ISBN: 0-444-50594-6
ISSN: 0928 1479

∞ The paper used in this publication meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper).

Printed in The Netherlands.

INTRODUCTION TO THE SERIES
"Advances in Image Communication"

Dear Colleague,

Image Communication is a rapidly evolving multidisciplinary field on the development and evaluation of efficient means for acquisition, storage, transmission, representation, manipulation and understanding of visual information. Until a few years ago, digital image communication research was still confined to universities and research laboratories of telecommunication or broadcasting companies. Nowadays, however, this field is witnessing the strong interest of a large number of industrial companies due to the advent of narrow band and broadband ISDN, GSM, the Internet, digital satellite channels, digital over-the-air transmission and digital storage media. Moreover, personal computers and workstations have become important platforms for multimedia interactive applications that advantageously use a close integration of digital compression techniques (JPEG, MPEG), Very Large Scale Integration (VI-SI) technology, highly sophisticated network facilities and digital storage media.

At the same time, the scope of research of the academic environment on Image Communication has further increased to include model- and knowledge-based techniques, artificial intelligence, motion analysis, and advanced image and video processing techniques. The variety of topics on Image Communication is so large that no one can be a specialist in all the topics, and the whole area is beyond the scope of a single volume, while the requirement of up-to-date information is ever increasing.

This was the rationale for Elsevier Science Publishers to approach me to edit a book series on 'Advances in Image Communication', next to the already existing and highly successful Journal: "Signal Processing: Image Communication". The book series was to serve as a comprehensive reference work for those already active in the area of Image Communication. Each author or editor was asked to write or compile a state-of-the-art book in his/her area of expertise, including information until now scattered in many journals and proceedings. The book series therefore would help Image Communication specialists to gain a better understanding of the important issues in neighbouring areas by reading particular volumes. It would also give newcomers to the field a foothold for doing research in the Image Communication area.

in order to produce a quality book series, it was necessary to ask authorities well known in their respective fields to serve as volume editors, who would in turn attract outstanding contributors. It was a great pleasure to me that ultimately we were able to attract such an excellent team of editors and authors.

Elsevier Science and I, as Editor of the series, are delighted that this book series has already received such a positive response from the image communication community. We hope that the series will continue to be of great use to the many specialists working in this field.

Jan Biemond
Series Editor

Preface

THE human visual system is less sensitive to flickering details than to large–area flicker. Television displays apply interlacing to profit from this fact, while broadcast formats were originally defined to match the display scanning format. As a consequence, interlace is found throughout the video chain. If we describe interlacing as a form of spatio-temporal sub–sampling, then *de–interlacing*, the topic of our book, is the reverse operation aiming at the removal of the sub–sampling artefacts.

The major flaw of interlace is that it complicates many image processing tasks. Particularly, it complicates scanning–format conversions. These were necessary in the past mainly for international programme exchange, but with the advent of high-definition television, videophone, Internet, and video on PCs, many scanning formats have been added to the broadcast formats, and the need for conversion between formats is increasing.

This increasing need, not only in professional but also in consumer equipment, has restarted the discussion *'to interlace or not to interlace'*. Particularly, this issue divides the TV and the PC communities. The latter seems biased towards the opinion that present–day technologies are powerful enough to produce progressively scanned video at high rate and do not need to trade–off vertical against time resolution through interlacing. On the other hand, the TV world seems more conservative, and biased towards the opinion that present–day technologies are powerful enough to adequately de–interlace video material, which reduces, or even eliminates, the need to introduce incompatable standards and sacrifice the investments of so many consumers.

It appears that the two camps have had disjunct expertises for a long time. In a world where the two fields are expected by many to be converging, it becomes inevitable to appreciate and understand each other's techniques to some extent. Currently, the knowledge in the PC community on scan rate conversion in general, and on de–interlacing in particular, seems to be lagging behind on the expertise available in the TV world. Given the availablility of advanced motion–compensated scan rate conversion techniques in consumer TV–sets since some years, it is remarkable that the PC community still relies on techniques developed for use in the television chain in the seventies.

The question, 'to interlace or not to interlace', touches various issues. Whether present–day technologies are powerful enough to produce progressively scanned video at a high rate and a good signal to noise ratio is not evident. Moreover, a visual–communication system also involves display and transmission of video signals. The issue translates for the transmission channel into the question: 'Is interlacing and de-interlacing still the optimal algorithm for reducing the signal bandwidth with a factor of two?' Before answering this question, it is necessary to know what can be achieved with de–interlacing techniques nowadays. Although the literature provides evidence that an all–progressive chain gives at least as good an image quality as an all–interlaced chain with the same channel bandwidth, recent research suggests that modern motion–compensated de–interlacing techniques, used in todays consumer electronics products can improve the efficiency of even highly efficient compression techniques. It seems appropriate, therefore, to evaluate the available options in de–interlacing, before jumping to conclusions.

As a consequence of the many related issues, the scope of our book is relatively broad.

Chapter 1 reviews the historical background of interlace, the meaning and significance of the reversed process called de–interlacing, and the motivation for the research that formed the basis of this book.

Chapter 2 presents an overview of de–interlacing techniques. Over the last two decades, many de–interlacing algorithms have been proposed. They range from simple spatial interpolation, via directional dependent filtering, up to advanced motion–compensated interpolation. Some methods are already available in products, while the more recent ones will appear in products when technology economically justifies their complexity. Chapter 2 outlines the most relevant algorithms, available either in TV and PC products, or in recent literature, and compares their performance. This comparison provides figures of merit, but also screen photographs are included showing the typical artifacts of the various de–interlacing methods. Although the evaluation shows good results with motion–compensated de–interlacers, it also reveals that there is room for improvement, that can result from modifications in the de–interlacing algorithm, or from improved motion estimator accuracy.

Chapter 3, therefore, introduces motion estimation techniques developed during roughly the last thirty years for different applications, such as motion compensated (MC) filtering for noise reduction, MC prediction for coding and MC interpolation for video format conversion. MC de–interlacing is probably the most demanding application of motion estimation, as it requires estimation of the true motion with a sub–pixel accuracy. This chapter is focussed on motion estimation algorithms that enabled the breakthroughs required for consumer priced MC de–interlacing. A relative comparison of the performance of the most relevant ME algorithm is part of this chapter.

In Chapter 4, we present the research aiming at further improvement of accuracy of the best motion estimation algorithm found in Chapter 3. Particularly, we aimed at eliminating the preferences for particular fractional values of the motion vector, resulting from the use of simple sub-pixel interpolation filters.

In Chapter 5, we present the research aiming at further improvement of the best de–interlacing algorithm found in Chapter 2. In the evaluation section of this chapter we conclude that the resulting algorithm, the majority selection de–interlacer, indeed gives the best overall de–interlacing quality.

The combination of the best de–interlacer, obtained in Chapter 5, with the best motion–estimator as proposed in Chapter 4, offers a solid basis for investigating, in Chapter 6, the MPEG–2 coding efficiency of interlaced and progressive video. In contrast to research published earlier, we include a *subjective assessment* for the relevant bit rates. We also present a comparison in terms of the Block Impairment Metric which is more relevant than the commonly used peak–signal–to–noise–ratio. Finally, we use a more balanced test set than found in earlier publications. Our improved evaluation of interlaced and progressive coding in various scenarios, enables a better judgement of the current value of interlace in video standards, and shows that still many modern video chains profit from this old technique.

In Chapter 7 we further explored the comparison of interlaced versus progressive video with focus on the *display format*. This comparison is of particular interest for the display of highly detailed pictures as text, Internet pages, and for resizing of pictures. (Picture resizing is for example required for the so–called 'dual–window' television, and for the so–called picture–in–picture feature). It was demonstrated that the interlaced format yields subjectively an improved vertical resolution, unless line flickering becomes predominant.

In Chapter 8, we draw our final conclusion that interlace is not a relic in the digital age, but is still a relevant ingredient of modern video formats. Therefore, de–interlacing remains a key technology for future image quality improvements.

We cannot hope that this book shall silence the discussions on interlace. We do hope, however, that it serves to provide a common knowledge basis for the divided camps. It can be a starting point for further experiments that will contribute to the final *technical* answer. The debate is unlikely to end even there, as introducing incompatible new TV standards in the past proved difficult, and balancing technical and non–technical issues may prove to be difficult.

Erwin B. Bellers
Gerard de Haan
July 2000

Acknowledgments

We would like to thank the management of Philips Research for providing the opportunity to perform the interesting research which forms the basis of this book, and giving the permission to publish this unique material.

We would also like to thank our colleagues for their help with the research that forms the basis of this book. In particular, we are indebted to Anthony Ojo, Robert-Jan Schutten, Frits de Bruijn, Mihaela van der Schaar-Mitrea, Bram Riemens, Rimmert Wittebrood, Christian Hentschel and Ton Kalker for their support for some parts of this book.

Last, but not least, we would like to express our gratitude for the critical review of a major part of this book by Jan Biemond of the Delft University of Technology.

Contents

1 Introduction **1**
- 1.1 Historical background of interlace . . . 3
- 1.2 De–interlacing . . . 8
- 1.3 Relation with superresolution . . . 10
- 1.4 Relation with MPEG–2 coding . . . 12
- 1.5 Motivation and scope of this book . . . 13

I Basic technology **17**

2 Overview of de–interlacing algorithms **19**
- 2.1 The de–interlacing problem . . . 21
  - 2.1.1 Spatio–temporal sampling . . . 21
  - 2.1.2 Motion and its spatio–temporal representation . . . 23
  - 2.1.3 Progressive scanning and reconstruction . . . 24
  - 2.1.4 Interlaced scanning and reconstruction . . . 28
  - 2.1.5 Psycho–visual effects . . . 32
  - 2.1.6 Problem statement . . . 35
- 2.2 Non–motion–compensated de–interlacing . . . 36
  - 2.2.1 Linear techniques . . . 36
  - 2.2.2 Non–linear techniques . . . 41
- 2.3 Motion–compensated de–interlacing . . . 53
  - 2.3.1 Direct methods . . . 54
  - 2.3.2 Hybrids . . . 57
  - 2.3.3 Temporal Backward Projection . . . 58
  - 2.3.4 Time–Recursive de–interlacing . . . 59
  - 2.3.5 Adaptive–Recursive de–interlacing . . . 60
  - 2.3.6 'Transversal' Generalized sampling . . . 62

2.3.7 'Recursive' generalized sampling 68
2.4 Evaluation 71
2.4.1 Objective performance measurement 72
2.4.2 Complexity 73
2.4.3 Test set 75
2.4.4 Results 76
2.5 Conclusions 87

3 Overview on motion estimation techniques 91
3.1 Historical developments in motion estimation 92
3.2 Pel–recursive estimators 94
3.3 Block–matching algorithms 97
3.3.1 The match criterion 98
3.3.2 Efficient search strategies 100
3.4 True–motion estimation 105
3.4.1 Hierarchical motion estimation 108
3.4.2 Phase plane correlation 109
3.4.3 Recursive search block–matching 112
3.5 Global motion models 115
3.5.1 Upgrading an efficient block matcher with a global motion model 116
3.6 Object based motion estimation 118
3.6.1 Brief overview of methods 118
3.6.2 An example object based estimator 120
3.7 Evaluation of motion estimation methods 128
3.7.1 Estimator performance testing 129
3.7.2 Evaluation results 137
3.7.3 Subjective evaluation of vector fields 142
3.8 Conclusion 149

II System optimization 151

4 Accurate motion estimates from interlaced video 153
4.1 Accuracy of the motion vectors 154
4.1.1 Improving the vertical resolution 154
4.1.2 Theoretical and practical accuracy limits 156
4.2 Improving block–based motion estimation 159
4.2.1 Cost function 160
4.2.2 Symmetrical versus asymmetrical motion estimation . 166
4.3 Interpolation to improve the motion vector accuracy 166

4.3.1 Linear interpolators 169
4.3.2 Non–linear interpolators 172
4.3.3 Interpolation and generalized sampling 173
4.4 Evaluation 174
4.4.1 Test set 175
4.4.2 Objective performance measures 177
4.4.3 Results of asymmetrical motion estimation 179
4.4.4 Results of symmetrical motion estimation 187
4.5 Conclusions 190

5 On the optimization of de–interlacing 195
5.1 Evaluation of the performance on detailed images 196
5.1.1 Experimental setup 197
5.1.2 Experimental results and evaluation 199
5.2 Evaluation of the performance on edges 201
5.2.1 Experimental setup 201
5.2.2 Experimental results and evaluation 201
5.3 Evaluation of the robustness 203
5.3.1 Experimental setup 205
5.3.2 Results and evaluation 206
5.4 The Majority–Selection de-interlacer 208
5.4.1 Combining de–interlacing strengths 209
5.4.2 MS–hypothesis validation 210
5.4.3 'Optimized' MS de–interlacer 212
5.5 Evaluation 215
5.5.1 Quality criteria 215
5.5.2 Results 216
5.6 Conclusions 219

III The future of interlace 221

6 The efficiency of interlaced versus progressive video on an MPEG–2 digital channel 223
6.1 Introduction 224
6.2 Summary of the MPEG–2 video–coding standard 225
6.2.1 Group Of Pictures 227
6.2.2 Intra/inter–frame/field coding 227
6.2.3 Field and frame prediction 228
6.2.4 Macroblock 229
6.2.5 Motion vectors 229

6.2.6 Discrete Cosine Transform Coding 229
6.2.7 Profiles and levels 230
6.3 The experiments 230
6.3.1 Test sequences 231
6.3.2 Subjective assessment 235
6.3.3 Objective quality criteria 235
6.3.4 Algorithms 237
6.4 Results and evaluation 242
6.4.1 All progressive–coding chain versus interlaced–coding chain 242
6.4.2 All interlaced–coding chain versus progressive–coding chain 248
6.4.3 Receiver–side de–interlacing versus transmitter–side de–interlacing 251
6.4.4 Receiver–side interlacing versus transmitter–side interlacing 253
6.4.5 All progressive–coding chain versus progressive–coding chain 254
6.5 Discussion 256

7 Towards an optimal display format 259
7.1 Display format options 260
7.2 Exploiting the source picture resolution 263
7.2.1 High resolution pictures 263
7.2.2 Standard definition video 264
7.3 Evaluation 267
7.3.1 First subjective assessment 268
7.3.2 Second subjective assessment 274
7.4 Conclusions 278

8 Conclusions 281

A Cycles per degree and cycles per picture width 287

B Motion and alias 289

C Interpolation for sub–pixel de–interlacing 291

D Example: derivation of a 4 taps TGST filter 295

E Robustness of the TGST de–interlacer 297

F MS de–interlacer optimization for a programmable architecture 301

F.1 TriMedia architecture 301

F.2 Majority–Selection de–interlacer on the TriMedia 303

F.3 Computational requirements 305

References 309

Glossary of special definitions, symbols and notations 323

List of Tables 327

List of Figures 329

Biography 345

Index 347

This Page Intentionally Left Blank

CHAPTER 1

Introduction

FOR centuries, mankind has been creating paintings to portray real or imagined scenes. The oldest paintings in the world, found in a cave in the Ardeche Valley of France, are estimated to go back to about 30,000 years ago. Paintings rather than text written in characters were the first means of communication, and as an old Chinese proverb 'a picture is worth ten thousand words' indicates, an efficient one. It is, therefore, not surprising that man is highly interested in looking at pictures. It took centuries for the next step to be taken; motion pictures. The first movie pictures shown to the public, by the Lumiere brothers, date from 1895 in the Grand Cafe in Paris (France). The early years of the film industry was a time of exploration. Of course no preconceived idea about how to make films yet existed, so filmmakers had to learn by trial and error.

The idea of *'vision at a distance'*, i.e. scenes reproduced far from their origin, can be traced back to the 19th century, and it is not unlikely it originated from an even earlier time. However, it took till the late thirties before *television* (TV), as a first realization of this concept was introduced to the public, which took place at the World's Fair in 1939. However, from that time onwards, the television industry did not take long to grow to a multi–billion dollars industry.

The penetration of the TVs in U.S. households was about 9% in 1950 [1]. Within five years the percentage went up to 64.5%. The 1999 penetration is at a level of 98.2%. The U.S. television households with two or more sets accounted for about 1% in 1950, and grew to 74.3% in 1999 [1]. So we may conclude that television has become a major product for entertainment, communication and information.

Webster's dictionary defines television as

> *'an electronic system for transmitting images of fixed or moving objects together with sound over a wire or through space by an apparatus that converts light and sound into electrical waves and reconverts them into visible light rays and audible sound'.*

The process of converting light into electrical signals was enabled by the discovery of the photoelectric effect from *selenium bars* in 1873. Exposed to light, these bars show a variation in resistance. As such, variation in light can be transformed into a variation of an electrical signal, and therefore, be transmitted.

One of the earliest methods of *scanning* a picture to generate a corresponding electrical signal is described in a patent granted to the German Paul Nipkow. He invented an electromechanical scanning technique based on a rotating disk with series of holes arranged in a spiral. The light sensitive selenium bars behind this perforated disk captured the picture. This disk became known as the *Nipkow disk.* However, Nipkow could not put his idea into practice with the materials and knowledge available at that time.

Another scientific development in the end of the 19th century offered an alternative; the usage of the *electron.* A tiny particle of negative charge with almost negligible inertia became a main focus of research. Karl Ferdinand Braun of the University of Strasbourg had, in 1897, the idea of using two electromagnets to make the electron beam move in the horizontal and vertical direction. To demonstrate his idea, he built the oscilloscope. The cathode rays of electrons were illuminated by fluorescent materials at the end of the tube. This system became known as the *Cathode Ray Tube* (*CRT*). His idea still forms the basis of the scanning system in most of the television sets of today.

With the introduction of television in the 1930s, standardization was required, i.e. rules or constraints for transmitting and receiving pictorial information, similar to e.g. the rules how to read a paper; in many countries the commonly accepted rules are: read from the top to the bottom of a paper, and from left to right. Common TV display use the same scanning direction.

Next to economical constraints, technical and psycho–visual criteria mainly formed the core of the standardization for television signals. Although many television standards evolved over time (as e.g. PAL, NTSC, SECAM), some elementary characteristics remained common in several stan-