ELSEVIER

A Practical Guide to Video and Audio Compression

From Sprockets and Rasters to Macro Blocks

Cliff Wootton

Focal Press

A Practical Guide to Video and Audio Compression

From Sprockets and Rasters to Macroblocks

Cliff Wootton

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
Focal Press is an imprint of Elsevier

Acquisition Editor: Joanne Tracy/Angelina Ward
Project Manager: Brandy Lilly
Editorial Assistant: Becky Golden-Harrell
Marketing Manager: Christine Degon
Cover Design: Eric DeCicco

Focal Press is an imprint of Elsevier
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA
Linacre House, Jordan Hill, Oxford OX2 8DP, UK

Library of Congress Cataloging-in-Publication Data
Application submitted.

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

ISBN-13: 978-0-240-80630-3
ISBN-10: 0-240-80630-1

For information on all Focal Press publications
visit our website at www.books.elsevier.com

06 07 08 09 10 10 9 8 7 6 5 4 3 2
Printed in the United States of America

A Practical Guide to Video and Audio Compression

This book is dedicated to my
friend Bernard Fisk.

Preface

The last few years have been an extraordinary time for the digital video industry. Not long before the turn of the millennium, digital video editing systems were expensive capital items of equipment that only major broadcasters and production companies could afford. To think that now the same capability is available in a laptop that you can buy off the shelf and it comes with the software for something in the region of $1200 is amazing. This is a capability we have dreamed about having on our desktops for 15 years. The price of the hardware and software needed to run an entire TV broadcast service is now within the reach of any organization or individual who cares to get involved.

Recall the boom in publishing that happened when the Apple LaserWriter was launched with Adobe PostScript contained inside and those early page composition programs enhanced what we were able to do with Word version 1 or MacWrite. We are now at that place with digital media and while some people will create an unattractive mess with these powerful tools, they will also enjoy themselves immensely and learn a lot at the same time. Eventually, a few skilled people will emerge from the pack and this is where the next generation of new talent will come from to drive the TV and film industry forward over the next couple of decades.

When Joanne Tracey asked me to prepare a proposal for this book I realized (as had most authors I have spoken to) that I didn't know as much about the topic I was about to write on as I thought I did. So this book has been a journey of exploration and discovery for me, just as I hope it will be for you. And yet, we also don't realize how much we do already know, and I hope you will find yourself nodding and making a mental comment to yourself saying "Yes—I knew that" as you read on.

We excel through the efforts of those around us in our day-to-day interactions with them. I have been particularly lucky to enjoy a few truly inspirational years with a group of like-minded people at the BBC. We all shared the same inquisitive approach into how interactive TV news could work. Now that we have all gone our separate ways I miss those "water cooler moments" when we came up with amazingly ambitious ideas. Some of those ideas live on in the things we engineered and rolled out. Others are yet to develop into a tangible form. But they will, as we adopt and implement the new MPEG-4, 7, and 21 technologies.

We are still at a very exciting time in the digital video industry. The H.264 codec is achieving enormous potential and there is much yet to do in order to make it as successful as it could be. Looking beyond that is the possibility of creating HDTV services and interactive multimedia experiences that we could only dream about until now.

Video compression can be a heavy topic at the best of times and we cover a lot of ground here. I thought the idea of illustrating the concept with a cartoon (see the first illustration in Chapter 1) would be helpful, because this subject can be quite daunting and I have purposely tried not to take it all too seriously. The cartoon is in order to disarm the subject and make it as accessible as possible to readers who haven't had the benefit of much experience with compression.

In some chapters you'll find a gray box with an icon on the left and a briefly encapsulated hot tip. These have been placed so that they are relevant to the topic areas being discussed but also to help you flick through the book and glean some useful knowledge very quickly. They have come out of some of those brainstorming times when I discussed digital video with colleagues in the various places I work and in online discussions. It's a bit of homespun wisdom based on the experiences of many people and intended to lighten the tone of the book a little.

If you are wondering about the face on the cover, it is my daughter Lydia. But if you look more closely at the cover, it tells a story. In fact, it is an attempt to show what the book is all about in one snapshot.

On the left you'll see the sprocket holes from film. Then in the background some faint raster lines should be evident. As you traverse to the right, the detail in the face becomes compressed. This illustrates how an image becomes degraded and finally degenerates into small macroblock particles that waft away in the breeze. Coming up with these illustrative ideas is one of the most enjoyable parts of writing a book.

So there you have it. I've enjoyed working on this project more than any other book that I can recall being involved with. I hope you enjoy the book too and find it helpful, as you become a more experienced compressionist.

In closing I'd like to say that the finer points of this publication are due to the extremely hard work by the team at Focal Press and any shortcomings you find are entirely my fault.

Cliff Wootton
Crowborough, South East England

Acknowledgments

When you write a book a book like this, it is the sum of so many people's efforts and goodwill. I would like to especially thank "J and Lo" (Joanne Tracey and Lothlórien Homet) of Focal Press for guiding me through the process of writing this book. Thanks to Gina Marzilli, who guided us down the right path on the administrative side. The manuscript was skillfully progressed through the production process by Becky Golden-Harrell—thanks, Becky. Let's do it again. Copyediting was ably managed by Cara Salvatore, Sheryl Avruch, and their team of experts. Thanks guys; you really turned this into a silk purse for me.

Of course, without the products in the marketplace, we'd have very little success with our endeavors. I'd like to send warm thanks to the team at Popwire in Sweden. Anders Norström and Kay Johansson have been immensely helpful. Over the last couple of years I've enjoyed getting to know members of the QuickTime team at Apple Computer. Thanks to Dave Singer, Rhondda Stratton, Tim Schaaf, Vince Uttley, and Greg Wallace for their help and inspiration. Guys, you are doing wonderful stuff. Just keep on doing that thing that you do. Also at Apple, I'd like to thank Sal Soghoian for pointing out some really cool stuff that AppleScript does. Thanks go to Envivio for some very thought-provoking and inspiring conversations, especially the time I've spent with Rudi Polednik, Frank Patterson, and Sami Asfour. Greetings also to Diana Johnson, Dave Kizerian, and Matt Cupal of Sorenson and Annie Normandin of Discreet. Thanks for being there when I needed your help. In the latter stages of completeing the book, Janet Swift and Barbara Dehart at Telestream came through with some coolness that enabled me to make Windows Media files effortlessly on a Mac.

To the people who work so hard at the MPEGIF (formerly known as the M4IF), Rob Koenen, Sebastian Möritz, and your team, I thank you for your time and patience explaining things to me. I hope this is a journey we can travel together for many years yet as we see the new MPEG standards being widely adopted.

I have so many friends from my time at the BBC who unselfishly shared their expertise and knowledge. Foremost of these must be Russell Merryman, who produced the elephant cartoon and was also responsible—with Asha Oberoi, Robert Freeman, Saz Vora, and John Nicholas—for the MPEG-4 packaged multimedia concept studies way back in 2002. Thanks also to Julie Lamm, John Angeli, and everyone in the News Interactive department.

Thanks are due also to those individuals, companies, and organizations who graciously permitted me to use their images in this project or spent time talking to me about

their work: Christopher Barnatt from the University of Nottingham; Simon Speight and Mark Sherwood from Gerry Anderson Productions; Guan at Etiumsoft; Jim Cooper at MOTU; David Carew-Jones, Anna Davidson, and Paul Dubery at Tektronix; Diogo Salari at DPI Productions; the folks at M-Audio; the Sales Web team at Apple Computer; Grant Petty and Simon Hollingworth at Black Magic Design; Julie Aguilar of ADC Telecommunications; Victoria Battison of AJA Video Systems; and Amanda Duffield of Pace Micro Technology.

I'd also like to thank Ben Waggoner for his unselfish sharing of many Master Compressionist's secrets at conferences. Ben, I've learned many new things from you whenever I've been at your presentations. Thank you so much for encouraging people the way you do.

Table of Contents

1 Introduction to Video Compression

1.1 Starting Our Journey

We (that is, you and I) are going to explore video compression together. It is a journey of discovery and surprise. Compression might seem daunting at this point, but like the old Chinese proverb says, "Even the longest journey starts with a single step." Let's head into that unknown territory together, taking it carefully, one step at a time until we reach our destination.

1.2 Video Compression Is Like . . .

It really is like trying to get a grand piano through a mailbox slot or an elephant through the eye of a needle. In fact, we thought the elephant was such an appropriate description, my friend Russell Merryman created a cartoon to illustrate the concept:

Video compression is all about trade-offs. Ask yourself what constitutes the best video experience for your customers. That is what determines where you are going to compromise. Which of these are the dominant factors for you?

- Image quality
- Sound quality
- Frame rate
- Saving disk space
- Moving content around our network more quickly
- Saving bandwidth
- Reducing the playback overhead for older processors
- Portability across platforms
- Portability across players
- Open standards
- Licensing costs for the tools
- Licensing costs for use of content

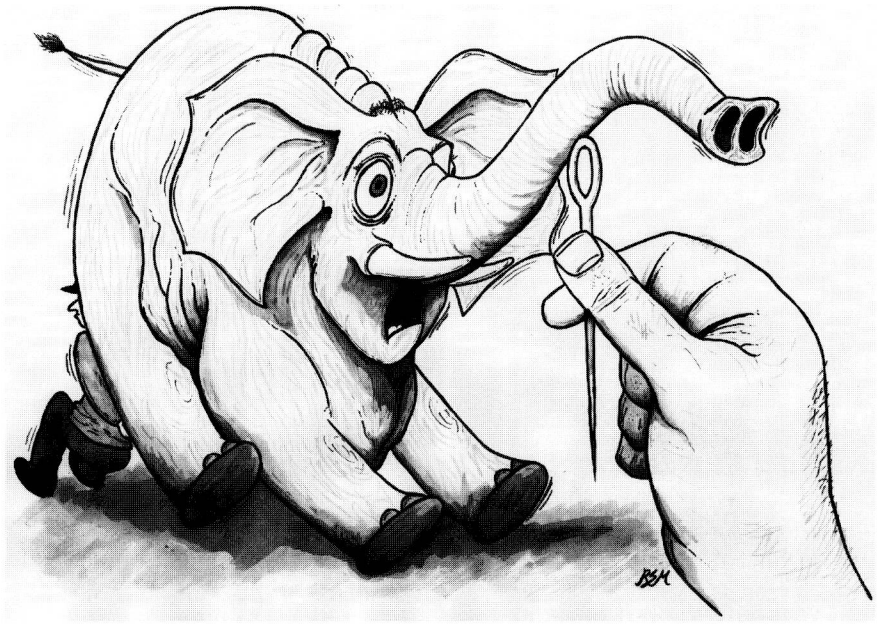

Figure 1-1 How hard can it be?

- Revenue streams from customers to you
- Access control and rights management
- Reduced labor costs in production

You will need to weigh these factors against each other. Some of them are mutually exclusive. You cannot deliver high quality from a cheap system that is fed with low-quality source material that was recorded on a secondhand VHS tape. Software algorithms are getting very sophisticated, but the old adage, "Garbage in, garbage out" was never truer than it is for video compression.

1.3 It's Not Just About Compressing the Video

The practicalities of video compression are not just about how to set the switches in the encoder but also involve consideration of the context—the context in which the video is arriving as well as the context where it is going to be deployed once it has been processed.

Together, we will explore a lot of background and supporting knowledge that you need to have in order to make the best decisions about how to compress the video. The actual compression process itself is almost trivial in comparison to the contextual setting and the preprocessing activity.

1.4 What Is a Video Compressor?

All video compressors share common characteristics. I will outline them here and by the end of the book you should understand what all of these terms mean. In fact, these terms describe the step-by-step process of compressing video:

- Frame difference
- Motion estimation
- Discrete cosine transformation
- Entropy coding

Wow! Right now you may be thinking that this is probably going to be too hard. Refrain from putting the book back on the shelf just yet though. Compression is less complicated than you think. If we take it apart piece by piece and work through it one item at a time, you will see how easy it is. Soon, you will be saying things like, "I am going to entropy code the rest of my day," when what you actually mean is you are going home early because there is nothing to do this afternoon. You can have a secret guffaw at your colleagues' expense because you know all about video compression and they don't.

1.5 The Informed Choice Is Yours

Despite all the arguments about the best technology to use, in the end your decisions may be forced by your marketing department arguing about reaching larger audiences. Those decisions should be backed up by solid research and statistics. On the other hand, they might be based just on hearsay. The consequences of those decisions will restrict your choice of codecs to only those that your selected platform supports. However, you will still have some freedom to innovate in building the production system.

Video compression is only a small part of the end-to-end process. That process starts with deciding what to shoot, continues through the editing and composition of the footage, and usually ends with delivery on some kind of removable media or broadcast system. In a domestic setting, the end-to-end process might be the capture of analogue video directly off the air followed by digitization and efficient storage inside a home video server. This is what a TiVo Personal Video Recorder (PVR) does, and compression is an essential part of how that product works.

There is usually a lot of setting up involved before you ever compress anything. Preparing the content first so the compressor produces the best-quality output is very important. A rule of thumb is that about 90% of the work happens before the compression actually begins. The content of this book reflects that rule of thumb: about 90% of the coverage is about things you need to know in order to utilize that 10% of the time you will actually spend compressing video in the most effective way possible.

1.6 Parlez-Vous Compressionese?

A few readers may be unfamiliar with the jargon we use. Words such as *codec* might not mean a lot to you at this stage. No need to worry—jargon will be explained as we go along. The important buzzwords are described in a glossary at the end of the book. Glossary entries are italicized the first time they are used.

The word codec is derived from coder–decoder and is used to refer to both ends of the process—squeezing video down and expanding it to a viewable format again on playback. Compatible coders and decoders must be used, so they tend to be paired up when they are delivered in a system like QuickTime or Windows Media. Sometimes the coder is provided for no charge and is included with the decoder. Other times you will have to buy the coder separately. By the way, the terms coder and encoder in general refer to the same thing.

1.7 Tied Up With Your Cabling?

Because there are so many different kinds of connectors, where it is helpful, there are diagrams showing how things connect up. In Appendix M, there are pictures of the most common connectors you will encounter and what they are for. Even on a modest, semi-professional system, there could be 10 different kinds of connectors, each requiring a special cable. FireWire and USB each have multiple kinds of connectors depending on the device being used. It is easy to get confused. The whole point of different types of connectors is to ensure that you only plug in compatible types of equipment. Most of the time it is safe to plug things in when the cable in your left hand fits into a socket in the piece of hardware in your right (okay, if you are left-handed it might be the other way around). Knowing whether these connections are "hot pluggable" is helpful, too.

Hot-pluggable connections are those that are safe to connect while your equipment is turned on. This is, in general, true of a signal connection but not a power connection. Some hardware, such as SCSI drives, must never be connected or unconnected while powered on. On the other hand, Firewire interfaces for disk drives are designed to be hot pluggable.

1.8 So You Already Know Some Stuff

Chapters 2 to 7 may be covering territory you already know about. The later chapters discuss the more complex aspects of the encoding process and will assume that you already know what is in the earlier chapters or have read them.

1.9 Video Compression Is Not Exactly New

Video compression has been a specialist topic for many years. Broadband connections to the Internet are becoming commonplace, and consumers are acquiring digital video cameras. Those consumers all have a need for video compression software.

The trick is to get the maximum possible compression with the minimum loss of quality. We will examine compression from a practical point of view, based on where your source material originated. You will need to know how film and TV recreate images and the fundamental differences between the two media. Then you will make optimal choices when you set up a compression job on your system.

You don't have to fully understand the mathematics of the encoding process. This knowledge is only vital if you are building video compression products for sale or if you are studying the theory of compression. Some background knowledge of how an encoder works is helpful though. In a few rare instances, some math formulas will be presented but only when it is unavoidable.

Our main focus will be on the practical aspects of encoding video content. Once you've read this book, you should be able to buy off-the-shelf products and get them working together. However, this book is not a tutorial on how to use any particular product. We discuss compression in a generic way so you can apply the knowledge to whatever tools you like to use.

1.10 This Is Not About Choosing a Particular Platform

We will discuss a variety of codecs and tools, and it is important to get beyond the marketing hyperbole and see these products independently of any personal likes, dislikes, and platform preferences.

My personal preference is for Apple-based technologies because they allow me to concentrate on my work instead of administering the system. I've used a lot of different systems, and something in the design of Apple products maps intuitively to the way I think when I'm doing creative work. You may prefer to work on Windows- or Linux-based systems, each of which may be appropriate for particular tasks. Compression tools are available for all of the popular operating systems.

This book is about the philosophy and process of compression. The platform is irrelevant other than to facilitate your choosing a particular codec or workflow that is not supported elsewhere, although even that problem is becoming obsolete as we move forward with portability tools and wider use of open standards.

Sometimes, lesser-known technology solutions are overlooked by the industry and are worth considering, and I've tried to include examples. But space is limited, so please don't take offense if I have omitted a personal favorite of yours. Do contact us if you find a particularly useful new or existing tool that you think we should include in a later edition.

1.11 Putting the Salesmen in a Corner

You need to be armed with sufficient knowledge to cut through the sales pitch and ask penetrating questions about the products being offered to you. Always check the specifications thoroughly before buying. If you can, check out reference installations and read reviews before committing to a product. If this book helps you do that and saves you from