TURING 图灵原版计算机科学系列

ELSEVIER 爱思唯尔

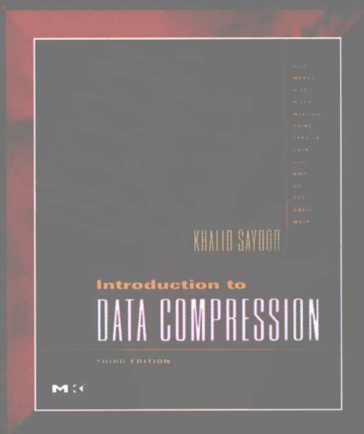
Introduction to Data Compression

Third Edition

数据压缩导论

（英文版·第3版）

[美] Khalid Sayood 著



人民邮电出版社
POSTS & TELECOM PRESS

TURING 图灵原版计算机科学系列

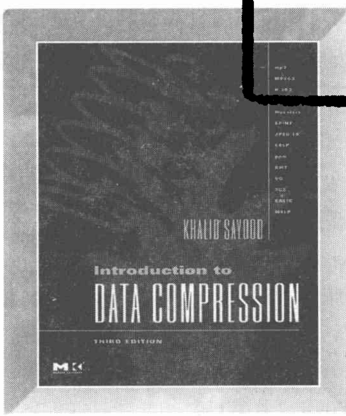
Introduction to Data Compression

Third Edition

数据压缩导论

（英文版·第3版）

[美] Khalid Sayood 著



人民邮电出版社
北京

图书在版编目（CIP）数据

数据压缩导论：第3版：英文/（美）萨尤得（Sayood, K.）著. —北京：人民邮电出版社，2009.2
（图灵原版计算机科学系列）
书名原文：Introduction to Data Compression, Third Edition
ISBN 978-7-115-19520-3

I. 数… II. 萨… III. 数据压缩－英文 IV. TP274

中国版本图书馆CIP数据核字（2008）第201713号

内容提要

本书是数据压缩方面的经典著作，介绍了各种类型的压缩模式。书中首先介绍了基本压缩方法（包括无损压缩和有损压缩）中涉及的数学知识，为常见的压缩形式打牢了信息论基础，然后从无损压缩体制开始，依次讲述了霍夫曼编码、算术编码以及字典编码技术等，对于有损压缩，还讨论了使用量化的模式，描述了标量、矢量以及微分编码和分形压缩技术，最后重点介绍了视频加密。本书不但分析了各种压缩模式及其优缺点，而且还说明了它们最适合处理哪种内容。

本书非常适合从事数据压缩相关工作的专业技术人员、软硬件工程师、学生等阅读，数字图书馆、多媒体等领域的技术人员也可参考。

图灵原版计算机科学系列

数据压缩导论（英文版·第3版）

◆ 著　[美] Khalid Sayood
责任编辑　杨海玲
◆ 人民邮电出版社出版发行　北京市崇文区夕照寺街14号
邮编　100061　电子函件　315@ptpress.com.cn
网址　http://www.ptpress.com.cn
北京顺义振华印刷厂印刷
◆ 开本：800×1000　1/16
印张：43.75
字数：840千字　2009年2月第1版
印数：1-2 000册　2009年2月北京第1次印刷
著作权合同登记号　图字：01-2008-5829号
ISBN 978-7-115-19520-3/TP

定价：99.00元

读者服务热线：（010）88593802　印装质量热线：（010）67129223
反盗版热线：（010）67171154

Preface

Within the last decade the use of data compression has become ubiquitous. From *mp3* players whose headphones seem to adorn the ears of most young (and some not so young) people, to cell phones, to DVDs, to digital television, data compression is an integral part of almost all information technology. This incorporation of compression into more and more of our lives also points to a certain degree of maturation of the technology. This maturity is reflected in the fact that there are fewer differences between this and the previous edition of this book than there were between the second and first editions. In the second edition we had added new techniques that had been developed since the first edition of this book came out. In this edition our purpose is more to include some important topics, such as audio compression, that had not been adequately covered in the second edition. During this time the field has not entirely stood still and we have tried to include information about new developments. We have added a new chapter on audio compression (including a description of the *mp3* algorithm). We have added information on new standards such as the new video coding standard and the new facsimile standard. We have reorganized some of the material in the book, collecting together various lossless image compression techniques and standards into a single chapter, and we have updated a number of chapters, adding information that perhaps should have been there from the beginning.

All this has yet again enlarged the book. However, the intent remains the same: to provide an introduction to the art or science of data compression. There is a tutorial description of most of the popular compression techniques followed by a description of how these techniques are used for image, speech, text, audio, and video compression.

Given the pace of developments in this area, there are bound to be new ones that are not reflected in this book. In order to keep you informed of these developments, we will periodically provide updates at *http://www.mkp.com*.

Audience

If you are designing hardware or software implementations of compression algorithms, or need to interact with individuals engaged in such design, or are involved in development of multimedia applications and have some background in either electrical or computer engineering, or computer science, this book should be useful to you. We have included a large number of examples to aid in self-study. We have also included discussion of various multimedia standards. The intent here is not to provide all the details that may be required to implement a standard but to provide information that will help you follow and understand the standards documents.

Course Use

The impetus for writing this book came from the need for a self-contained book that could be used at the senior/graduate level for a course in data compression in either electrical engineering, computer engineering, or computer science departments. There are problems and project ideas after most of the chapters. A solutions manual is available from the publisher. Also at *http://sensin.unl.edu/idc/index.html* we provide links to various course homepages, which can be a valuable source of project ideas and support material.

The material in this book is too much for a one semester course. However, with judicious use of the starred sections, this book can be tailored to fit a number of compression courses that emphasize various aspects of compression. If the course emphasis is on lossless compression, the instructor could cover most of the sections in the first seven chapters. Then, to give a taste of lossy compression, the instructor could cover Sections 1–5 of Chapter 9, followed by Chapter 13 and its description of JPEG, and Chapter 18, which describes video compression approaches used in multimedia communications. If the class interest is more attuned to audio compression, then instead of Chapters 13 and 18, the instructor could cover Chapters 14 and 16. If the latter option is taken, depending on the background of the students in the class, Chapter 12 may be assigned as background reading. If the emphasis is to be on lossy compression, the instructor could cover Chapter 2, the first two sections of Chapter 3, Sections 4 and 6 of Chapter 4 (with a cursory overview of Sections 2 and 3), Chapter 8, selected parts of Chapter 9, and Chapter 10 through 15. At this point depending on the time available and the interests of the instructor and the students portions of the remaining three chapters can be covered. I have always found it useful to assign a term project in which the students can follow their own interests as a means of covering material that is not covered in class but is of interest to the student.

Approach

In this book, we cover both lossless and lossy compression techniques with applications to image, speech, text, audio, and video compression. The various lossless and lossy coding techniques are introduced with just enough theory to tie things together. The necessary theory is introduced just before we need it. Therefore, there are three *mathematical preliminaries* chapters. In each of these chapters, we present the mathematical material needed to understand and appreciate the techniques that follow.

Although this book is an introductory text, the word *introduction* may have a different meaning for different audiences. We have tried to accommodate the needs of different audiences by taking a dual-track approach. Wherever we felt there was material that could enhance the understanding of the subject being discussed but could still be skipped without seriously hindering your understanding of the technique, we marked those sections with a star ($\star$). If you are primarily interested in understanding how the various techniques function, especially if you are using this book for self-study, we recommend you skip the starred sections, at least in a first reading. Readers who require a slightly more theoretical approach should use the starred sections. Except for the starred sections, we have tried to keep the mathematics to a minimum.

Learning from This Book

I have found that it is easier for me to understand things if I can see examples. Therefore, I have relied heavily on examples to explain concepts. You may find it useful to spend more time with the examples if you have difficulty with some of the concepts.

Compression is still largely an art and to gain proficiency in an art we need to get a "feel" for the process. We have included software implementations for most of the techniques discussed in this book, along with a large number of data sets. The software and data sets can be obtained from *ftp://ftp.mkp.com/pub/Sayood/*. The programs are written in C and have been tested on a number of platforms. The programs should run under most flavors of UNIX machines and, with some slight modifications, under other operating systems as well. More detailed information is contained in the README file in the *pub/Sayood* directory.

You are strongly encouraged to use and modify these programs to work with your favorite data in order to understand some of the issues involved in compression. A useful and achievable goal should be the development of your own compression package by the time you have worked through this book. This would also be a good way to learn the trade-offs involved in different approaches. We have tried to give comparisons of techniques wherever possible; however, different types of data have their own idiosyncrasies. The best way to know which scheme to use in any given situation is to try them.

Content and Organization

The organization of the chapters is as follows: We introduce the mathematical preliminaries necessary for understanding lossless compression in Chapter 2; Chapters 3 and 4 are devoted to coding algorithms, including Huffman coding, arithmetic coding, Golomb-Rice codes, and Tunstall codes. Chapters 5 and 6 describe many of the popular lossless compression schemes along with their applications. The schemes include LZW, *ppm*, BWT, and DMC, among others. In Chapter 7 we describe a number of lossless image compression algorithms and their applications in a number of international standards. The standards include the JBIG standards and various facsimile standards.

Chapter 8 is devoted to providing the mathematical preliminaries for lossy compression. Quantization is at the heart of most lossy compression schemes. Chapters 9 and 10 are devoted to the study of quantization. Chapter 9 deals with scalar quantization, and Chapter 10 deals with vector quantization. Chapter 11 deals with differential encoding techniques, in particular differential pulse code modulation (DPCM) and delta modulation. Included in this chapter is a discussion of the CCITT G.726 standard.

Chapter 12 is our third mathematical preliminaries chapter. The goal of this chapter is to provide the mathematical foundation necessary to understand some aspects of the transform, subband, and wavelet-based techniques that are described in the next three chapters. As in the case of the previous mathematical preliminaries chapters, not all material covered is necessary for everyone. We describe the JPEG standard in Chapter 13, the CCITT G.722 international standard in Chapter 14, and EZW, SPIHT, and JPEG 2000 in Chapter 15.

Chapter 16 is devoted to audio compression. We describe the various MPEG audio compression schemes in this chapter including the scheme popularly known as *mp3*.

Chapter 17 covers techniques in which the data to be compressed are analyzed, and a model for the generation of the data is transmitted to the receiver. The receiver uses this model to synthesize the data. These analysis/synthesis and analysis by synthesis schemes include linear predictive schemes used for low-rate speech coding and the fractal compression technique. We describe the federal government LPC-10 standard. Code-excited linear prediction (CELP) is a popular example of an analysis by synthesis scheme. We also discuss three CELP-based standards, the federal standard 1016, the CCITT G.728 international standard, and the relatively new wideband speech compression standard G.722.2. We have also included a discussion of the mixed excitation linear prediction (MELP) technique, which is the new federal standard for speech coding at 2.4 kbps.

Chapter 18 deals with video coding. We describe popular video coding techniques via description of various international standards, including H.261, H.264, and the various MPEG standards.

A Personal View

For me, data compression is more than a manipulation of numbers; it is the process of discovering structures that exist in the data. In the 9th century, the poet Omar Khayyam wrote

> The moving finger writes, and having writ,
> moves on; not all thy piety nor wit,
> shall lure it back to cancel half a line,
> nor all thy tears wash out a word of it.
>
> *(The Rubaiyat of Omar Khayyam)*

To explain these few lines would take volumes. They tap into a common human experience so that in our mind's eye, we can reconstruct what the poet was trying to convey centuries ago. To understand the words we not only need to know the language, we also need to have a model of reality that is close to that of the poet. The genius of the poet lies in identifying a model of reality that is so much a part of our humanity that centuries later and in widely diverse cultures, these few words can evoke volumes.

Data compression is much more limited in its aspirations, and it may be presumptuous to mention it in the same breath as poetry. But there is much that is similar to both endeavors. Data compression involves identifying models for the many different types of structures that exist in different types of data and then using these models, perhaps along with the perceptual framework in which these data will be used, to obtain a compact representation of the data. These structures can be in the form of patterns that we can recognize simply by plotting the data, or they might be statistical structures that require a more mathematical approach to comprehend.

In *The Long Dark Teatime of the Soul* by Douglas Adams, the protagonist finds that he can enter Valhalla (a rather shoddy one) if he tilts his head in a certain way. Appreciating the structures that exist in data sometimes require us to tilt our heads in a certain way. There are an infinite number of ways we can tilt our head and, in order not to get a pain in the neck (carrying our analogy to absurd limits), it would be nice to know some of the ways that will generally lead to a profitable result. One of the objectives of this book is to provide you with a frame of reference that can be used for further exploration. I hope this exploration will provide as much enjoyment for you as it has given to me.

Acknowledgments

It has been a lot of fun writing this book. My task has been made considerably easier and the end product considerably better because of the help I have received. Acknowledging that help is itself a pleasure.

The first edition benefitted from the careful and detailed criticism of Roy Hoffman from IBM, Glen Langdon from the University of California at Santa Cruz, Debra Lelewer from California Polytechnic State University, Eve Riskin from the University of Washington, Ibrahim Sezan from Kodak, and Peter Swaszek from the University of Rhode Island. They provided detailed comments on all or most of the first edition. Nasir Memon from Polytechnic University, Victor Ramamoorthy then at S3, Grant Davidson at Dolby Corporation, Hakan Caglar, who was then at TÜBITAK in Istanbul, and Allen Gersho from the University of California at Santa Barbara reviewed parts of the manuscript.

For the second edition Steve Tate at the University of North Texas, Sheila Horan at New Mexico State University, Edouard Lamboray at Oerlikon Contraves Group, Steven Pigeon at the University of Montreal, and Jesse Olvera at Raytheon Systems reviewed the entire manuscript. Emin Anarım of Boğaziçi University and Hakan Çağlar helped me with the development of the chapter on wavelets. Mark Fowler provided extensive comments on Chapters 12–15, correcting mistakes of both commission and omission. Tim James, Devajani Khataniar, and Lance Pérez also read and critiqued parts of the new material in the second edition. Chloeann Nelson, along with trying to stop me from splitting infinitives, also tried to make the first two editions of the book more user-friendly.

Since the appearance of the first edition, various readers have sent me their comments and critiques. I am grateful to all who sent me comments and suggestions. I am especially grateful to Roberto Lopez-Hernandez, Dirk vom Stein, Christopher A. Larrieu, Ren Yih Wu, Humberto D'Ochoa, Roderick Mills, Mark Elston, and Jeerasuda Keesorth for pointing out errors and suggesting improvements to the book. I am also grateful to the various instructors who have sent me their critiques. In particular I would like to thank Bruce Bomar from the University of Tennessee, Mark Fowler from SUNY Binghamton, Paul Amer from the University of Delaware, K.R. Rao from the University of Texas at Arlington, Ralph Wilkerson from the University of Missouri–Rolla, Adam Drozdek from Duquesne University, Ed Hong and Richard Ladner from the University of Washington, Lars Nyland from the Colorado School of Mines, Mario Kovac from the University of Zagreb, and Pierre Jouvelet from the Ecole Superieure des Mines de Paris.

Frazer Williams and Mike Hoffman, from my department at the University of Nebraska, provided reviews for the first edition of the book. Mike read the new chapters in the second and third edition in their raw form and provided me with critiques that led to major rewrites. His insights were always helpful and the book carries more of his imprint than he is perhaps aware of. It is nice to have friends of his intellectual caliber and generosity. Rob Maher at Montana State University provided me with an extensive critique of the new chapter on

audio compression pointing out errors in my thinking and gently suggesting corrections. I thank him for his expertise, his time, and his courtesy.

Rick Adams, Rachel Roumeliotis, and Simon Crump at Morgan Kaufmann had the task of actually getting the book out. This included the unenviable task of getting me to meet deadlines. Vytas Statulevicius helped me with LaTex problems that were driving me up the wall.

Most of the examples in this book were generated in a lab set up by Andy Hadenfeldt. James Nau helped me extricate myself out of numerous software puddles giving freely of his time. In my times of panic, he was always just an email or voice mail away.

I would like to thank the various "models" for the data sets that accompany this book and were used as examples. The individuals in the images are Sinan Sayood, Sena Sayood, and Elif Sevuktekin. The female voice belongs to Pat Masek.

This book reflects what I have learned over the years. I have been very fortunate in the teachers I have had. David Farden, now at North Dakota State University, introduced me to the area of digital communication. Norm Griswold at Texas A&M University introduced me to the area of data compression. Jerry Gibson, now at University of California at Santa Barbara was my Ph.D. advisor and helped me get started on my professional career. The world may not thank him for that, but I certainly do.

I have also learned a lot from my students at the University of Nebraska and Boğaziçi University. Their interest and curiosity forced me to learn and kept me in touch with the broad field that is data compression today. I learned at least as much from them as they learned from me.

Much of this learning would not have been possible but for the support I received from NASA. The late Warner Miller and Pen-Shu Yeh at the Goddard Space Flight Center and Wayne Whyte at the Lewis Research Center were a source of support and ideas. I am truly grateful for their helpful guidance, trust, and friendship.

Our two boys, Sena and Sinan, graciously forgave my evenings and weekends at work. They were tiny (witness the images) when I first started writing this book. Soon I will have to look up when talking to them. "The book" has been their (sometimes unwanted) companion through all these years. For their graciousness and for always being such perfect joys, I thank them.

Above all the person most responsible for the existence of this book is my partner and closest friend Füsun. Her support and her friendship gives me the freedom to do things I would not otherwise even consider. She centers my universe and, as with every significant endeavor that I have undertaken since I met her, this book is at least as much hers as it is mine.

Contents