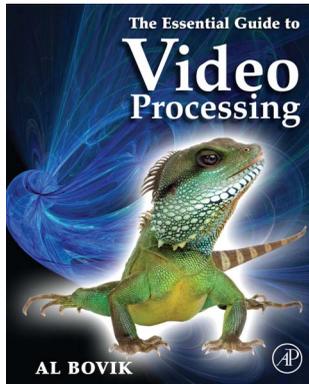


The Essential Guide to Video Processing

Al Bovik, 778 pp., ISBN 13: 978-0-12-374456-2, Academic Press, (2009), \$89.95 paperback.

Reviewed by Amit Singhal, Kodak Research Labs, Eastman Kodak Company, Rochester, New York.



Wikipedia describes the field of video processing as a particular case of signal processing, where the input and output signals are video files or video streams. This is, perhaps, an overly simplified statement—while technically correct, it does not do justice to one of the most interesting and timely topics in electrical engineering and computer science today. Digital video is fast becoming, if not already,

ubiquitous in our life today, from television broadcast to the Internet, and in consumer electronics such as cameras. It is fast becoming the cornerstone of various applied fields such as surveillance and security, medicine, and entertainment, among others. *The Essential Guide to Video Processing*, edited by Al Bovik, serves as a comprehensive resource for learning about digital video processing, whether you're a novice or an expert in the field. The book is a compendium of 21 chapters, authored by leading experts in each area of the field, and covers topics ranging from the basics of video processing to advanced topics such as the H.264 standard and applications of digital imaging to surveillance and mobile imaging.

The book is roughly organized into four sections: fundamentals of video processing (Chapters 1–7), video coding and standards (Chapters 8–13), video communications (Chapters 14–18), and applications (Chapters 19–22). In organizing the book in this manner, Bovik has produced a complete reference to the vast field of digital video imaging. Each of the four sections is a complete and independent study of that subtopical area, and even the chapters within each section can be read independently of each other. Given that the book is over 750 pages long, this is a very good thing. If your interest is in the field of video encoding, you can skip the beginning chapters and start with Chapter 8, “Basic Transform Video Coding.” If you are an expert in that field and are interested in learning more about embedded video coding, you can skip straight to Chapter 13. Each of the Chapters in the book is written by leaders in that subtopical area, drawing upon the expertise of almost 50 individuals and their associated research groups, laboratories, departments, universities, and companies.

The first section of the book is a comprehensive overview of digital video-processing techniques, beginning with an introduction to the field in Chapter 1, and including topics such as video sampling, motion estimation and tracking, video enhancement, stabilization, and segmentation. Following the introductory chapter, Chapter 2 covers the basic ideas of sampling and interpolation of time-varying imagery, such as video signals. Chapter 3 provides a discussion of motion from a perspective of

video processing and compression, focusing on methods for motion detection and motion estimation (five algorithms are described in the text, based on various models, estimation criteria, and search strategies). Chapter 4 describes methods for enhancing and restoring corrupted video and image sequences, including detection and removal of impairments such as noise, coding artifacts, blotches, scratches, vinegar syndrome, flicker, and moiré. The techniques and tools described in Chapter 4, while focused on the detection and removal of the aforementioned impairments, are general in nature and can be used to develop enhancement and restoration methods for other types of degradation.

Chapter 5 covers the topics of video stabilization, registration, and video mosaicing. Video stabilization refers to the compensation of motion of pixels in a recorded sequence of images captured from a moving camera. Mosaicing refers to the construction of high-resolution images from video sequences based on registration of image frames in the video sequence. Chapter 6 describes various techniques for segmentation, focusing primarily on extracting objects present in a video sequence, or separating the regions of a video sequence into objects versus background regions. Chapter 7 builds upon the fundamentals of motion estimation in Chapter 3 and extends them into 2-D and 3-D object-tracking algorithms. Object tracking concerns itself with deriving the trajectory over time of moving objects in a video sequence.

Digital video inherently requires handling of high volumes of data. This makes work in the area of video compression extremely important to a video-processing engineer. Chapters 8 through 13 cover various aspects of video compression, including coding schemes and the various standards for video coding in the industry. Chapter 8 is an introduction to video compression and covers the fundamental techniques and methods used in every standard video-compression algorithm. Chapter 9 covers two of the early video-coding standards developed by the Moving Picture Experts Group, MPEG-1 and MPEG-2. MPEG-1 and MPEG-2 standards are primarily used for coding of videos transmitted over the Internet and for DVD releases, respectively. Chapter 10 continues the discussion from Chapter 9, describing the MPEG-4 suite of video-coding standards, including H.264. H.264 is considered to be one of the greatest advances in video compression and error resilience and is fast becoming the coding standard of choice for all digital video applications. Chapter 11 goes beyond the MPEG video-coding standards and describes a family of motion-compensated subband/wavelet coders that exploit temporal correlation and are highly scalable in bit rate, resolution, and frame rate.

As evidenced by the discussions on video coding techniques and standards in Chapters 9 through 11, there are a large number of format options available to a video-processing engineer. This has resulted in significant diversity of content and associated coding schemes. However, most receiving devices only support a subset of these coding schemes, requiring the use of video-transcoding algorithms that convert unsupported video formats to one supported by the receiver. Chapter 12 provides an overview of some of the techniques used to transcode videos from one format to another, primarily using bit-rate control algorithms. This is a highly active area of research, with many open

questions, many of which are posed to the reader, along with information on resources available to fully understand this fast-changing field. Finally, Chapter 13 describes embedded video-coding schemes that describe how to implement the techniques and algorithms covered in the previous chapters on special-purpose processors or digital signal processing (DSP) chips.

The third section loosely covers the area of video communications, including topics in privacy and security, wireless transmission, archival and retrieval, etc. Chapter 14 describes various methods used to measure the quality of videos based on some quantitative metric. The chapter covers the most common of these methods, those based on the human visual system, but also includes extensive discussion of newer feature-based and motion-based visual-quality-assessment algorithms. Chapter 15 addresses the problem of archival, search, and retrieval in large databases of video content including multimodal analysis, representation, summarization, indexing, and browsing of video collections. Most of the techniques described focus on visual-based methods for organizing and working with videos. Chapter 16 focuses on efficient transmission and routing of video streams over a variety of communication networks based on different protocols. Lastly, Chapter 17 covers the very important topic of security and privacy, addressing the concern that multimedia information be accessed only by authorized users and for authorized purposes. Topics include encryption for confidentiality and access control, authentication for content integrity, and watermarking for tracing misuse and illicit distribution of content.

The final chapters in the book describe a diverse set of video-application areas. These include video surveillance (Chapter 19), face tracking and recognition (Chapter 20), medicine (Chapter 21), and the use of video for speech processing (Chapter 22). These chapters provide a detailed overview of

each of these application areas, focusing not just on video imaging, but also on other topics that are essential to the pursuit of these technologies.

According to Bovik, the objectives of this guide are, primarily, to serve as a complete and comprehensive resource in video processing for a diverse group of users, from novices to experts, from various technical backgrounds and disciplines. In that sense, the guide format adopted for this book is ideally suited to the independent reference manual this book represents. The reader does not need to read the book cover to cover, or even in chapter order, and can instead access the specific material he or she is interested in. A secondary objective of this book is to serve as a textbook for upper-level undergraduate and graduate classes. The topical organization of this book (video-processing techniques, video-coding schemes and standards, and video communications) certainly lends itself to coursework focused in those three areas. A course instructor would have to do some work to create appropriate exercises and supplementary material to aid the student's understanding of the topics. However, the extensive list of references provided at the end of each chapter makes this book ideally suited for graduate reading seminars and similar discussion formats.

Amit Singhal received his MS and PhD in computer science from the University of Rochester, Rochester, New York. He is a principal research scientist with the Kodak Research Labs at Eastman Kodak Company. He is also an adjunct faculty member of computer science at the University of Rochester. His current research interests include 3-D and stereo imaging, image understanding, and knowledge engineering. He has authored over 40 journal and conference papers and holds 5 patents.