Video Technology

视频技术

By
Weiping Li
Shipeng Li
Chun Wang

李卫平 李世鹏 王 纯 著

University of Science and
Technology of China Press
中国科学技术大学出版社

内容简介

本书阐述视频技术中的基本概念和关键技术。除了回顾视频技术的背景和历史外，重点介绍视频编码和通信的基本原理和方法，并对相关的标准、应用系统、及未来的研究方向予以讨论。全书共分八部分，依次为引言、早期的视频技术、模拟视频信号、数字视频信号、视频编码、视频编码标准、视频应用、以及对未来的展望。本书可作为信息科学技术领域高年级本科生和研究生针对视频领域的入门教材，也可以供从事该领域科研和技术开发的人员参考。

Video Technology
Weiping Li, Shipeng Li, Chun Wang

Published by University of Science and Technology of China Press
96 Jinzhai Road, Hefei 230026, P. R. China

图书在版编目(CIP)数据

视频技术＝Video Technology：英文/李卫平，李世鹏，王纯著. —合肥：中国科学技术大学出版社，2009. 5
（当代科学技术基础理论与前沿问题研究丛书：中国科学技术大学校友文库）
"十一五"国家重点图书
ISBN 978－7－312－02221－0

Ⅰ. 视…　Ⅱ. ①李…　②李…　③王…　Ⅲ. 视频系统—技术—英文　Ⅳ. TN94

中国版本图书馆 CIP 数据核字(2009)第 052228 号

出版发行　中国科学技术大学出版社
地址　安徽省合肥市金寨路 96 号，230026
网址　http://press.ustc.edu.cn
印　　刷　合肥晓星印刷有限责任公司
经　　销　全国新华书店
开　　本　710mm×1000mm　1/16
印　　张　10.5
字　　数　150 千
版　　次　2009 年 6 月第 1 版
印　　次　2009 年 6 月第 1 次印刷
印　　数　1—1500 册
定　　价　38.00 元

编 委 会

总　　序

侯建国
（中国科学技术大学校长、中国科学院院士、第三世界科学院院士）

大学最重要的功能是向社会输送人才。大学对于一个国家、民族乃至世界的重要性和贡献度，很大程度上是通过毕业生在社会各领域所取得的成就来体现的。

中国科学技术大学建校只有短短的50年，之所以迅速成为享有较高国际声誉的著名大学之一，主要就是因为她培养出了一大批德才兼备的优秀毕业生。他们志向高远、基础扎实、综合素质高、创新能力强，在国内外科技、经济、教育等领域做出了杰出的贡献，为中国科大赢得了“科技英才的摇篮”的美誉。

2008年9月，胡锦涛总书记为中国科大建校五十周年发来贺信，信中称赞说：半个世纪以来，中国科学技术大学依托中国科学院，按照全院办校、所系结合的方针，弘扬红专并进、理实交融的校风，努力推进教学和科研工作的改革创新，为党和国家培养了一大批科技人才，取得了一系列具有世界先进水平的原创性科技成果，为推动我国科教事业发展和社会主义现代化建设做出了重要贡献。

据统计，中国科大迄今已毕业的5万人中，已有42人当选中国科学院和中国工程院院士，是同期（自1963年以来）毕业生中当选院士数最多的高校之一。其中，本科毕业生中平均每1,000人就产生1名院士和700多名硕士、博士，比例位居全国高校之首。还有众多的中青年才俊成为我国科技、企业、教育等领域的领军人物和骨干。在历年评选的“中国青年五四奖章”获得者中，作为科技界、科技创新型企业界青年才俊代表，科大毕业生已连续多年榜上有名，获奖总人数位居全国高校前列。鲜为人知的是，有数千名优秀毕业生踏上国防战线，为科技强军做出了重要贡献，涌现出20多名科技将军和一大批国防科技中坚。

为反映中国科大五十年来人才培养成果，展示毕业生在科学研究中的最新进展，学校决定在建校五十周年之际，编辑出版《中国科学技术大学校友文

库》，于2008年9月起陆续出书，校庆年内集中出版50种。该《文库》选题经过多轮严格的评审和论证，入选书稿学术水平高，已列为国家"十一五"重点图书出版规划。

入选作者中，有北京初创时期的毕业生，也有意气风发的少年班毕业生；有"两院"院士，也有IEEE Fellow；有海内外科研院所、大专院校的教授，也有金融、IT行业的英才；有默默奉献、矢志报国的科技将军，也有在国际前沿奋力拼搏的科研将才；有"文革"后留美学者中第一位担任美国大学系主任的青年教授，也有首批获得新中国博士学位的中年学者；……在母校五十周年华诞之际，他们通过著书立说的独特方式，向母校献礼，其深情厚意，令人感佩！

近年来，学校组织了一系列关于中国科大办学成就、经验、理念和优良传统的总结与讨论。通过总结与讨论，使我们更清醒地认识到，中国科大这所新中国亲手创办的新型理工科大学所肩负的历史使命和责任。我想，中国科大的创办与发展，首要的目标就是围绕国家战略需求，培养造就世界一流科学家和科技领军人才。五十年来，我们一直遵循这一目标定位，有效地探索了科教紧密结合、培养创新人才的成功之路，取得了令人瞩目的成就，也受到社会各界的广泛赞誉。

成绩属于过去，辉煌须待开创。在未来的发展中，我们依然要牢牢把握"育人是大学第一要务"的宗旨，在坚守优良传统的基础上，不断改革创新，提高教育教学质量，早日实现胡锦涛总书记对中国科大的期待：瞄准世界科技前沿，服务国家发展战略，创造性地做好教学和科研工作，努力办成世界一流的研究型大学，培养造就更多更好的创新人才，为夺取全面建设小康社会新胜利、开创中国特色社会主义事业新局面贡献更大力量。

是为序。

2008年9月

Preface

Video technology has been a very active research area for a long time ever since the first television was invented. It is also a technology that has had a great impact to the society. In the recent years, digital video technology has enabled a variety of applications that used to be only possible in our imagination and in science fiction movies. This book is to introduce video technology to a wide range of readers who may or may not be actively conducting research in the field of video technology.

There are many distinguished researchers and industrial practitioners in this area who are USTC alumni. Three of us happen to be all working in this exciting field. Weiping Li (class of 776) worked in academia for 10 years on video technology research and now is working in industry on development of world-first large-scale distributed multimedia communication system. He is an IEEE Fellow for contributions to image and video coding algorithms, standards, and implementations and has served as Editor-in-Chief of IEEE Transactions on Circuits and Systems for Video Technology from 1999 to 2002. He has made significant contributions to international video coding standards and continued to make great impact in video communication systems. Shipeng Li (class of 836) has worked in industry research lab on this subject for over 13 years. He is a world renowned expert on video coding, processing, content analysis, streaming, and wireless and networking. He is a Senior Member of IEEE and an Associate Editor of IEEE Transactions on Circuits and Systems for Video Technology. He is a Principal Researcher of Microsoft Research Asia - the hottest computer science lab on the planet featured in MIT Technical Review. He has contributed significantly to image and video coding standards, algorithms and systems. Chun Wang (class of

846) has over 10 years of experience in the VLSI signal processing industry and is a veteran in designing communication systems. She is a Senior Member of IEEE and is currently the Director of Algorithms and Communication Systems at Amlogic, Inc, an IC company specializing in video technology. Their combined expertise brings a unique yet comprehensive coverage of video technology in this book with emphasis on video coding and communication.

This book on video technology is one in a series of books in electrical/electronic engineering and computer science to promote USTC internationally, report the state-of-the-art research results from alumni of USTC, and bring USTC alumni together to commemorate USTC's 50th anniversary.

A brief introduction of each chapter in this book is given as follows. Chapter 1 kicks off the discussion by introducing the background of video technology. Chapter 2 reviews the history of video technology from the first television to the latest HDTV. Chapter 3 discusses the analog video and the difference between video and film. It also explains what is interlaced video and its impact to the TV industry. Chapter 4 introduces some basic concepts in digital video technology, such as video sampling, pixel quantization and color sub-sampling. Chapter 5 then focuses on video coding that has been and will continue to be a fascinating area of video technology for research and development. The principles of video coding and three-stage video coding model are introduced. Topics such as signal processing for compression, quantization, entropy coding, predictive coding with quantization, symbol formation, bit allocation and rate control, preprocessing and post-processing are discussed. Advanced topics of video coding such as fractal compression, scalable video coding, transcoding, object-based video coding, model-based video coding, hybrid natural and synthetic coding and error-resilient video coding are also addressed. As part of communication technology, video coding standards are important and this book also introduces some major video coding standards including H.261, H.263, MPEG-1, MPEG-2/H.262, MPEG-4 part 2, H.264/MPEG-4 part 10/AVC, AVS, and H.264/MPEG-4 AVC scalable extension. Practical implementation of video coding and processing algorithms are also discussed. Last but not the least, some applications of video technology are included in Chapter 7 such as analog TV, video tape, video discs, digital television, video over IP, etc. In Chapter 8 some future directions of video technology are conjectured. Topics include new types of video contents, such as multi-view video and high dynamic range video; new video coding trends such as knowledge-based video coding and distributed video coding; video delivery such as joint source and channel coding, multi-source and multi-path video delivery, internet-infrastructure assisted video delivery; mobile video such as content creation and management, content processing and adaptation, and content analysis and repurposing; video retrieval and search, and finally new trends of internet video - the new media 2.0 concept.

In summary, we hope this book will be helpful to senior students and young researchers who want to get into the video technology field and prepare them to move into more advanced topics in this area. We also hope this book will be useful for readers who do not work in video technology, but would like to gain some knowledge about it.

The authors would like to acknowledge their Alma Mater USTC for the support, coaching, training, and care they received while studying at USTC. They would like to acknowledge their teachers for unselfishly sharing knowledge, expertise, methodology, philosophy, and experience with them. They would like to thank their families for continuous support during the writing of this book. Finally they would like to acknowledge various supports from the USTC colleagues especially USTC Press.

Weiping Li　Shipeng Li　Chun Wang

Contents

1 Introduction

Video technology has been a very active research area for a long time ever since the first television was invented. It is also a technology that has had a great impact to the society. In the recent years, digital video technology has enabled a variety of applications that used to be only possible in our imagination and in science fiction movies. This book is to introduce video technology to a wide range of readers who may or may not be actively conducting research in the field of video technology. It starts with some basic concepts in video technology, such as the difference between video and film. It continues with video coding that has been and will continue to be a fascinating area of video technology for research and development. As a part of communication technology, video coding standards are important and this book also introduces some major video coding standards. Practical implementation of video coding and processing algorithms is also discussed. Last but not the least, some applications of video technology are included and some future directions of video technology are conjectured.

Video technology may be considered from several perspectives. It may be considered as a part of communication with visual information. It may also be considered as a part of signal processing to extract certain information from visual data. Yet, it may be considered as a part of computer vision for understanding visual information. To some extent, it may also be considered as a part of computer graphics for creating visual animation. In many applications, video technology is applied across multiple areas. For example, in a video surveillance application, video data captured from a video camera may be processed first to detect whether there are object motions or not. If object motion is detected, it may trigger communication of the captured video from the camera location to a video monitoring location. In addition to human monitors, the received video data may be analyzed by computer vision software to understand what objects are contained in the video data. We do not intend to cover all aspects of video technology, but focus on video technology as a part of communication and, to a lesser extent, as a part of signal processing.

Communication is a basic human need. Modern telecommunication started with telephone that enabled real-time voice communication over a long distance. Visual communication is different from voice communication in several aspects. First of all, voice, as the source of voice communication, is human generated while the source of visual communication is not. Human generates voice signal using vocal cord and receives voice signal using ears. By studying the working mechanism of these two human organs, we can gain a good understanding on how voice is generated and received, thus model the voice signal more precisely. On the other hand, human does not generate visual signal, but only receives it with eyes. In addition, voice signal is one-dimensional while visual signal is four dimensional. Visual signal contains much more information than voice signal for a given time period. Therefore, it is much harder to model visual signal than voice signal. Due to the lack of a good mathematical model for visual signal, it is difficult to perform theoretic research in video technology. For example, video coding is a type of source coding but research approach of video coding is very different from that of source coding in the information theory. Therefore, research in video technology is a combination of science and art.

2 Early Days of Video Technology

Development of video technology has been driven by human's desire to extend its visual capability over distance and time. With the eyes, human can see the surrounding scenes and remember them in a short period of time. Visual communication is to transmit visual information over a long distance and/or to store visual information over a long time.

More than 100 years ago, people started inventing something called "television" with a hope to send visual information over a long distance. The first television used a simple mechanical device that turned a disk with a set of holes arranged into a spiral to scan a picture into a series of pixels and transmitted electrically. Figure 1 illustrates how a mechanical television works.

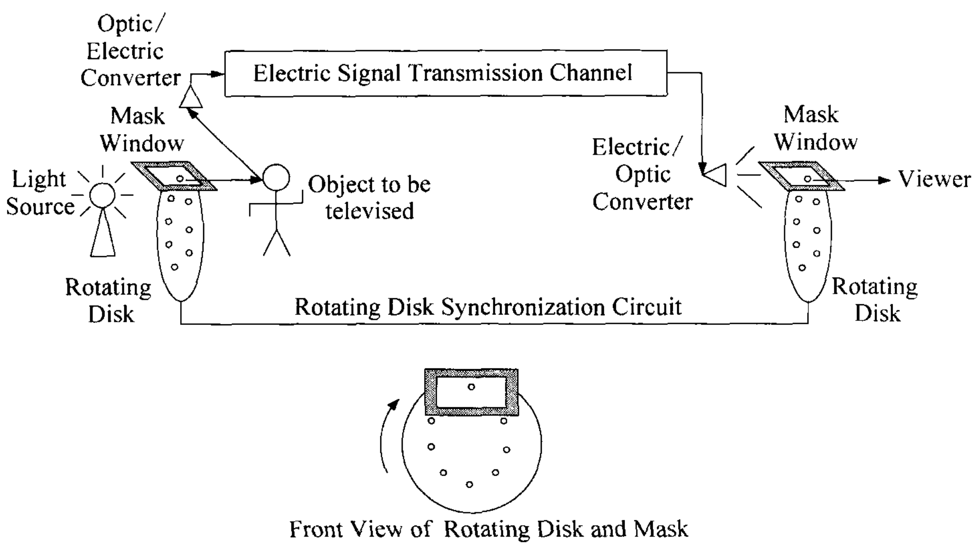

Figure 1 Illustration of a mechanical scanning TV system

The left side of the figure is the television transmission side and the right side is the television receiving side. A light bulb on the left provides the light source. A rotating disk with holes in a spiral allows the light through the holes. A mask with a rectangular window is in between the rotating disk and the object to be "televised". The size of the mask window and the distance between the holes on the disk are designed in such a way that only one hole is in the window at any given moment. Therefore, the light goes through only one hole at any given moment and this lights up a "pixel" at any given moment. Depending on the reflective brightness of the object being scanned, the light intensity of the "pixel" varies. Using an optic-electric converter, the light intensity of the pixel is converted into electric voltage modulated onto a radio signal that is transmitted over a long distance to the receiving side. Upon receiving the radio signal and demodulating the electric voltage, the receiving side converts the electric voltage to light intensity using an electric-optic converter. A similar mask and a disk with holes in a spiral are used to let the light through a single hole at a given moment for the viewer to see. When the disk on the transmission side is turned, the holes in the spiral scan an object from left to right and from top to bottom. The disk on the receiving side has to turn in sync with the disk on the transmission side so that the viewer sees the object at the correct location within the rectangular window. Strictly speaking, the "horizontal" scanning line is actually an arc due to the circular rotation of the disk.

Electronic scanning systems were developed later. An electronic video camera uses an array of sensors, such as charge coupled device (CCD), to convert optic intensity into electric signal that is to be transmitted or stored. On the receiving side, a device such as the cathode-ray tube (CRT) is used to convert the electric signal into an electronic beam that hits the phosphor on the glass screen to emit light. This became the modern television set. The first public television broadcast in the US was at the New York World's Fair in 1939, presented by RCA, the leading US television pioneer.

The basic television technology worked so well that it is still the basis of television today. One major advance in television was from black and white to color in 1954. Now, we are in the middle of another major advance in television — from analog to digital and from the standard definition (that has been with us for so long) to high definition television (HDTV). The readers who would like to learn more about the history of television may want to check out the following references [1][2][3][4][5][6].

Recording video electronically is another aspect of early video technology. The first electronic video recorder was made in 1956 for professional broadcast television industry. The motivation was to have the flexibility of editing or re-shooting. Without electronic recording, the same program could not be broadcast again unless a film crew shot the program simultaneously on 35 mm film, which was way too costly for the television industry. At that time, audio recording was already possible using magnetic tape. A new type of magnetic tape was developed for recording video.

In 1970's, Sony introduced the U-Matic system that dominated the professional video field for a long time. In 1975, Sony again introduced the Betamax system for consumer video recording. A year later, JVC introduced a new video recording format called VHS (Vertical Helical Scan, later changed to Video Home System) and started a video format war between Beta and VHS. Eventually, VHS won the market share for various reasons although Beta video quality was arguably better than VHS. To learn the impact of VCR, the readers may want to check out [7].

3 Analog Video

To start understanding video technology, we have to look at analog video first. There are many aspects of analog video. In this section, we mainly focus on the scanning aspect of analog video. First we discuss the difference between video and film as an introduction on why scanning is used for video. Then we discuss interlaced scanning as an engineering tradeoff between bandwidth, vertical resolution, and refresh rate.

3.1 Difference between Video and Film

It is important to understand the difference between video and film in order to appreciate some of the issues about video. Let us exam how film works first. The capture process of a film frame is to expose the chemicals on the film frame to the light source of the scene to be captured by briefly opening up the shutter of the film camera. Because this is an optical process, all pixels of the film frame capture the scene of the same moment in parallel. The display process of a film frame is to project the film frame to a screen with a light source behind the film frame. This is an optical process too and all pixels are projected to the screen at the same moment in parallel too.

If we were to achieve the same effect as the film using an electronic means, we would have to connect all electronic sensors of an electronic camera with parallel wires as illustrated in Figure 2.

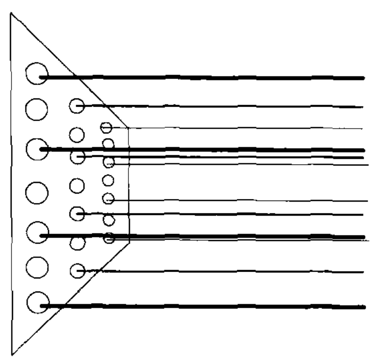

Figure 2 Electronic means for parallel capture of all pixels at the same moment

Such a massively parallel electronic circuit is almost impossible nowadays, not to mention in the early days of video being invented. On the display side using a phosphor screen, we would have to use parallel electron beams to