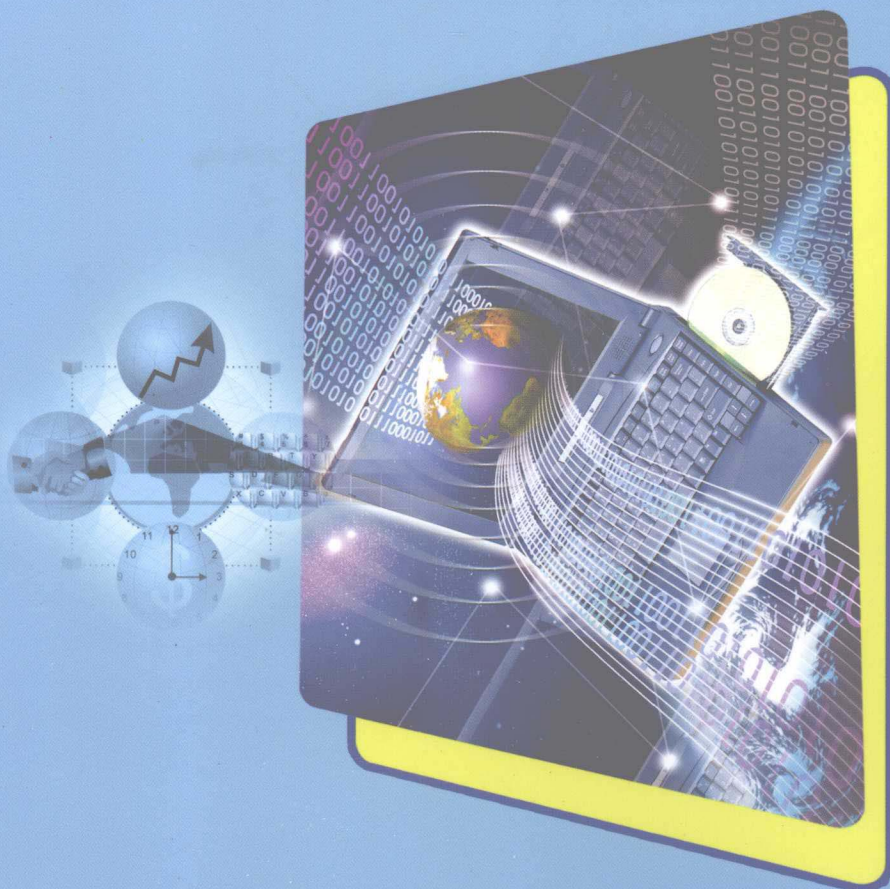




21世纪普通高等院校计算机专业规划教材

数字视频技术

戴辉 卢益民主编



北京邮电大学出版社
www.buptpress.com

21 世纪普通高等院校计算机专业规划教材

数字视频技术

戴 辉 卢益民 主 编



北京邮电大学出版社
[www. buptpress. com](http://www.buptpress.com)

内 容 简 介

本书深入浅出地介绍了数字视频技术设计与开发的基础知识,包括4篇,共15章,内容涉及音频、图形与图像、视频与编码标准、JPEG编码、MPEG编码、H.264-AVC编码、AVS视频编码以及数字视频应用等最新技术。全书理论与实践相结合,包含大量应用实例,强调实际操作技能的培训。为适合教学需要,各章末尾均配有习题,并配有电子课件。

本书面向数字视频技术的中级用户,可以使读者较为全面地了解数字视频技术的基本原理和相关应用开发,为将来更深入地学习数字视频技术奠定基础。

本书适合作为高等院校计算机、电子与信息工程、通信工程、数字媒体等相关专业高年级本科生和研究生的“数字视频技术”课程教材或教学参考书,也可作为工程技术人员的参考资料或培训教材。

图书在版编目(CIP)数据

数字视频技术/戴辉,卢益民主编.--北京:北京邮电大学出版社,2012.12

ISBN 978-7-5635-3125-7

I. ①数… II. ①戴…②卢… III. ①数字视频系统 IV. ①TN941.3

中国版本图书馆CIP数据核字(2012)第141231号

书 名:数字视频技术

主 编:戴 辉 卢益民

责任编辑:张珊珊

出版发行:北京邮电大学出版社

社 址:北京市海淀区西土城路10号(邮编:100876)

发 行 部:电话:010-62282185 传真:010-62283578

E-mail:publish@bupt.edu.cn

经 销:各地新华书店

印 刷:北京联兴华印刷厂

开 本:787 mm×1 092 mm 1/16

印 张:17.5

字 数:434千字

印 数:1—3 000册

版 次:2012年12月第1版 2012年12月第1次印刷

ISBN 978-7-5635-3125-7

定 价:36.00元

· 如有印装质量问题,请与北京邮电大学出版社发行部联系 ·

前 言

人类获取信息的 75%来自于人的视觉,在计算机网络上传输的视频信息只能是数字化的视频。随着科学技术的进步和人机界面技术的引入,计算机变得越来越友好和人性化。视听娱乐的普及、万维网的兴盛和计算机游戏的火爆,大大促进了数字视频技术的应用和发展。

本书介绍了数字视频基础、数字视频编码、数字视频信息处理、数字视频信息传输等内容,较为全面地介绍了数字视频技术及应用。

数字视频技术的内容十分广泛,涉及多种学科和领域。数字视频技术的课程和教材的内容选取和编排也各不相同,主要分为两类:

- 应用型——简单介绍媒体的基本属性,重点讲解各种用户级的数字视频素材和著名工具软件的使用,如 Cool Edit、Cakewalk、Corel Draw、Photoshop、Flash、3DS 等。
- 原理型——也介绍各类媒体的基本属性和数字视频的主要应用,但重点放在压缩算法、编码标准、编程开发和系统应用上,不介绍应用软件的具体使用方法。

本课程属于后者——原理型。强调基本概念的了解、具体方法的掌握和实际动手能力的培养。

采用数字视频教学手段,通过课堂讲解、平时练习和大作业,使学生掌握数字视频技术的基本内容和主要方法。注重数字视频技术的最新发展及与实际应用的紧密结合。

特别强调学生自己动手上机实习。通过 Windows 的 MFC 编程,实现主要的压缩算法和常见图像、音频、视频文件的读写、显示或播放,加深对各种媒体的特性、数据压缩、编码标准及文件格式等内容的理解。

本书第 1 章由卢益民编写,第 2~4 章由魏应彬编写,第 5~6 章由聂梦遥编写,第 7~15 章由戴辉编写。陈晓薇博士、袁晓辉博士、李毓蕙博士、高燕博士等也参与了本书的辅助性工作。全书由戴辉统稿,卢益民教授主审。

本书编写过程中得到了许多人的帮助和支持。本课题受国家自然科学基金(60874116)、海南省自然科学基金(610227)和华中科技大学“新世纪教学改革工程”教材建设基金资助。感谢国家自然科学基金委员会信息科学部、华中科技大学武汉光电国家实验室(筹)、华中科技大学电子与信息工程系以及北京邮电大学出版社领导和老师的大力支持;感谢骆清铭教授、朱光喜教授、喻莉教授的指导和帮助;感谢我的同事、我的学生对本书的建议。

感谢作者家人的大力支持和理解。

由于数字视频技术知识繁杂,作者水平有限,编写时间仓促,本书中错误或不妥之处难免,敬请读者批评指正。

作 者

于华中科技大学

目 录

第一篇 绪 论

第 1 章 数字视频技术概论	3
1.1 视频技术概论	3
1.1.1 视频的概念	3
1.1.2 视频技术的应用	12
1.2 视频相关技术	13
1.2.1 视频信号的获取与显示	13
1.2.2 视频信号的数字化	15
1.2.3 视频信号的处理	20
1.2.4 视频信号的压缩	20
1.2.5 视频信号的传输	20
1.2.6 视频信号的存储	21
1.3 电视基本原理	21
1.3.1 黑白电视原理	21
1.3.2 彩色电视原理	25
1.4 视频信号的数字化	30
1.4.1 图像信号的数字化	30
1.4.2 视频数字化标准建议	35
1.5 图像信号的统计特性	36
1.5.1 空间域上的统计特性	37
1.5.2 频率域上的统计特性	38
1.5.3 差值信号的统计特性	38
1.6 图像质量的评价	39
1.6.1 主观评价	39
1.6.2 客观评价	40
复习思考题	41



第二篇 压缩与编码

第 2 章 压缩与编码概论	45
2.1 压缩编码的必要性	45
2.1.1 压缩基本术语	45
2.1.2 压缩的需要	45
2.2 压缩编码的可行性	46
2.2.1 图像的统计特性	46
2.2.2 图像的视觉特性	46
2.3 压缩算法的特点与分类	48
2.3.1 压缩算法的特点	48
2.3.2 压缩算法的分类	49
2.3.3 压缩比的定义	50
2.3.4 压缩算法优劣的重要指标	50
2.3.5 视频压缩技术的新进展	50
2.4 压缩编码基本模型	51
2.5 压缩编码过程	51
第 3 章 熵编码	53
3.1 熵的定义	53
3.2 Shannon-Fano 编码	53
3.2.1 Shannon-Fano 算法描述	54
3.2.2 Shannon-Fano 算法举例	54
3.3 Huffman 编码	55
3.3.1 Huffman 算法描述	55
3.3.2 Huffman 算法举例	55
3.3.3 Huffman 解码	57
3.3.4 误码扩散及解决办法	59
3.3.5 香农-范诺编码与哈夫曼编码	59
3.3.6 H 表与自适应哈夫曼编码	61
3.4 算术编码	61
3.4.1 编码算法	61
3.4.2 解码算法	63
3.4.3 若干问题	64
3.5 RLE 编码	64
3.5.1 RLE 算法描述	64



3.5.2	BMP 中的 RLE 算法	65
3.6	LZW 编码	66
3.6.1	LZW 算法描述	66
3.6.2	LZW 编码算法	67
3.6.3	GIF 文件格式	70
	复习思考题	74
	作业	77
第 4 章	变换编码	79
4.1	变换编码的概念	79
4.1.1	变换编码的基本概念	79
4.1.2	变换编码的基本组成	79
4.2	正交变换的基本图像	81
4.2.1	变换域	81
4.2.2	基本图像	82
4.3	正交变换的物理意义	83
4.3.1	能量保持定理	83
4.3.2	能量重新分配	84
4.4	变换域的压缩编码	85
4.4.1	8×8 的子图像变换	85
4.4.2	变换域的压缩编码	85
4.5	编解码过程	87
4.5.1	编码准备	87
4.5.2	编解码过程	87
4.6	离散余弦变换	88
4.6.1	余弦变换	88
4.6.2	一维离散余弦变换	88
4.6.3	二维离散余弦变换	89
第 5 章	预测编码	90
5.1	预测编码的基本原理	90
5.1.1	预测编码的基本概念及其分类	90
5.1.2	图像中抽样点的位置	90
5.1.3	预测值、预测系数、预测误差	90
5.1.4	数据压缩基本原理	91
5.1.5	压缩原理方框图	91
5.1.6	帧内预测	91

5.1.7	帧间预测	92
5.1.8	帧内二维预测编码器例子	92
5.2	预测器设计	92
5.2.1	最佳线性预测器	92
5.2.2	自适应预测	94
5.3	量化器设计	95
5.3.1	均匀量化	95
5.3.2	最佳量化	95
5.3.3	最佳矢量量化器	97
5.3.4	自适应量化器	97
第6章 运动补偿与预测编码		99
6.1	帧间预测编码	99
6.1.1	预测编码的基本概念	99
6.1.2	预测编码的主要方法	99
6.2	运动补偿编码	99
6.2.1	运动补偿编码的概念	99
6.2.2	运动补偿编码方框图	100
6.2.3	块匹配方法	100
6.2.4	运动补偿预测编码的几种形式	103
复习思考题		103
作业		103

第三篇 数字视频国际标准

第7章 H.261 标准		107
7.1	H.261 标准概述	107
7.2	图像格式	109
7.2.1	CIF 和 QCIF 格式	109
7.2.2	主要技术参数	109
7.2.3	Y、C 采样点的位置	109
7.2.4	帧结构	109
7.3	数据结构	110
7.4	源编码器	111
7.4.1	源编码器方框图	111
7.4.2	帧内、帧间编码判断方法	112
7.4.3	量化	113



7.4.4 熵编码	113
7.4.5 附加信息的复用规定	113
7.4.6 BCH 编码	115
复习思考题	116
第 8 章 JPEG 标准	117
8.1 JPEG 系列标准概述	117
8.2 JPEG 标准基本系统概述	120
8.2.1 编、解码方框图	120
8.2.2 DCT 系数量化	121
8.2.3 熵编码	121
8.3 JPEG 编码	122
8.3.1 压缩算法	122
8.3.2 编码步骤	123
8.3.3 算法举例	129
8.4 JPEG 文件格式	133
8.4.1 图像准备	133
8.4.2 文件格式框架	134
8.4.3 文件格式内容	134
8.4.4 文件的一般顺序	136
复习思考题	136
作业	139
第 9 章 MPEG 标准	141
9.1 MPEG-1/2 的视频压缩算法	141
9.1.1 MPEG-1/2 的视频压缩算法简介	141
9.1.2 I 帧压缩编码算法	142
9.1.3 P 帧压缩编码算法	142
9.1.4 B 帧压缩编码算法	146
9.1.5 视频帧结构	146
9.1.6 MPEG-1	147
9.1.7 MPEG-2	152
9.2 MPEG-4 视频编码算法	155
9.2.1 MPEG-4 编码算法简介	155
9.2.2 MPEG-4 视频对象平面的概念	156
9.2.3 MPEG-4 视频编码方案	157
9.2.4 视频分辨率可变编码	158



9.3	MPEG 音频编码算法	159
9.3.1	MPEG-1Audio	159
9.3.2	MPEG-2Audio	168
9.3.3	MPEG-2ACC	170
9.3.4	MPEG-4Audio	172
	复习思考题.....	175
	作业.....	179
第 10 章	H. 264/AVC 编码.....	180
10.1	H. 264/AVC 的特点与结构	180
10.1.1	H. 264/AVC 技术特点	180
10.1.2	H. 264/AVC 编码结构与格式	182
10.2	H. 264/AVC 的预测编码	186
10.2.1	H. 264/AVC 帧内预测	186
10.2.2	H. 264/AVC 帧间预测与运动补偿	188
10.3	H. 264/AVC 的块编码	192
10.3.1	4×4 整数 DCT	192
10.3.2	量化.....	195
10.3.3	CAVLC	198
10.3.4	CABAC	203
	复习思考题.....	206
	作业.....	207
第 11 章	AVS 标准	208
11.1	AVS 标准简介	208
11.1.1	AVS 标准原则与特点	208
11.1.2	AVS 标准系列	209
11.1.3	AVS 产业	210
11.2	AVS 视频编码技术	211
11.2.1	AVS 编码框架	211
11.2.2	AVS 关键技术	212
11.3	AVS 视频编码与 MPEG-2 及 H. 264/AVC 的比较	215
11.3.1	编码技术的比较.....	215
11.3.2	计算复杂性的对比.....	218
11.3.3	编码效率对比.....	219
11.4	AVS-P7 概述	220
11.4.1	AVS-P7 系统结构	220



11.4.2 AVS-P7 主要技术	220
复习思考题	222
作业	222

第四篇 数字视频技术应用

第 12 章 数字视频会议系统	225
12.1 视频会议系统概述	225
12.1.1 视频会议系统的定义	225
12.1.2 视频会议系统的优点	225
12.1.3 视频会议系统的应用	225
12.2 视频会议系统的基本组成	225
12.2.1 视频会议系统组成	225
12.2.2 视频会议系统终端设备	228
12.2.3 多点控制设备	228
12.2.4 传输网络	228
12.3 视频会议系统的分类	229
12.4 视频会议主要国际标准	230
12.4.1 H.320 系列标准	231
12.4.2 H.324 系列标准	235
12.5 多点控制单元 MCU	236
12.5.1 MCU 的作用	236
12.5.2 MCU 的基本组成	237
12.5.3 二级形式电视会议网络	237
12.5.4 MCU 的三种控制方式	238
12.5.5 图像显示方式	239
第 13 章 H.324 终端及其实现	241
13.1 H.324 终端的构成	241
13.2 H.324 终端的实现	242
第 14 章 基于 H.323 协议的 IP 可视电话终端	248
14.1 H.323 终端模型	248
14.2 H.323 终端的硬件结构	250
14.2.1 概述	250
14.2.2 视频输入芯片 SAA7111 简介	251
14.2.3 以太网控制芯片 S8900A 简介	253

14.3 H. 323 终端软件	256
第 15 章 DVB 技术简介	259
15.1 DVB 概述	259
15.1.1 DVB 核心系统	259
15.1.2 DVB 分类	259
15.2 H. 324 数字机顶盒	261
15.2.1 概述	261
15.2.2 DVB 数字机顶盒的设计与实现	262
参考文献	267

第一篇

绪 论

视觉是人类最重要的感觉,也是人类获取信息的主要来源。据统计,人类从外界获取的信息中,75%来自视觉。视频信息同其他的信息形式相比,具有直观、具体、生动等诸多显著优点,并且视频所包含的信息量很大。

视频是一组在时间轴上有序排列的图像,是二维图像在一维时间轴上构成的图像序列,又称为动态图像、活动图像或者运动图像。它不仅包含静止图像的内容,还包含场景中目标运动的信息和客观世界随时间变化的信息。电影、电视等都属于视频的范畴。早期的视频主要指模拟的视频信号,随着全球数字化进程的不断推进,视频的采集设备和采集方式都有了很大的进展。而且,压缩算法、多媒体通信协议的不断发展都为数字视频技术的应用奠定了良好的基础,相信在不久的将来,数字视频应用将渗透到我们工作、生活的方方面面。

本篇介绍数字视频的基本概念、视频相关技术、电视基本原理、数字化与编码以及图像信号的统计特性与评价等。

第 1 章 数字视频技术概论

人们获取的信息 75% 来自视觉系统,也就是图像和视频信息。在静止图像的基础上,考虑时间因素,就形成视频,因此,视频也称为时基媒体。从早期的模拟视频发展到如今的数字视频,无疑是一次质的飞跃。数字视频具有易存储、易编辑等特性,正在获得越来越广泛的应用。

本章首先引入有关数字视频技术的基本概念,然后介绍数字视频相关技术,最后讨论图像信号的统计特性及其质量评价标准等。

1.1 视频技术概论

1.1.1 视频的概念

1. 视频

视频是人眼视觉器官所感知(重现)自然景物(物体)的信息。常以图像形式来表示。视频一词译自英文 Video,我们看到的电影和电视都属于视频范畴。

视频为活动图像,又称序列图像,由一幅幅静止图像组成,每幅图像称为一帧。帧是构成视频信息的最小和最基本单元。根据视觉惰性,每秒钟 24 帧的连续静止图像就能形成活动视频的感觉。

2. 彩色电视

(1) 广播频段的划分

广播频段的划分如表 1-1 所示。

表 1-1 广播频段的划分

种类	频段	频率范围(Hz)	波段频道	带宽/Hz	特点
调幅广播	MF	526.5~1 606.5k	中波	10k	国内广播,地/天波可传百/千余公里
	TF	2 300~5 060k	中短波		热带地区的国内广播
	HF	3 900~26 100k	短波		国际广播,电离层反射可传数千公里
电视广播	I(VHF)	48.5~92k	1~5	8M	直线传播,电视广播
调频广播	II(VHF)	87~180M	超短波	200k	直线传播,调频广播
电视广播	III(VHF)	167~223M	6~12	8M	直线传播,电视广播
	IV(UHF)	470~566	13~24		
	V(UHF)	606~988	25~68	8.68M	

其中:MF=Medium Frequency 中频;TF=Tropic Frequency 热带频;HF=High Frequency 高频;VHF=Very High Frequency 甚高频;UHF=UltraHigh Frequency 超高频。

(2) 彩色电视制式

目前世界上现行的模拟彩色电视制式有三种:NTSC制、PAL制和SECAM制。如表1-2所示。这里不包括模拟的高清晰度彩色电视。

- 国家电视系统委员会(National Television Systems Committee, NTSC)彩色电视制是1952年美国国家电视标准委员会定义的彩色电视广播标准,称为正交平衡调幅制,1954年开始广播。美国、加拿大等大部分西半球国家,以及日本、韩国、菲律宾等国和中国台湾采用这种制式。
- 由于NTSC制存在相位敏感造成彩色失真的缺点,因此德国(当时的西德)于1962年制定了相位逐行交变(Phase-Alternative Line, PAL)制彩色电视广播标准,称为逐行倒相正交平衡调幅制,1967年开始广播。德国、英国等一些西欧国家,以及中国、朝鲜等国家采用这种制式。
- 法国1957年起制定了顺序颜色传送与存储(法文:Sequential Couleur Avec Memoire, SECAM)彩色电视广播标准,称为顺序传送彩色与存储制,1967年开始广播。法国、苏联及东欧国家采用这种制式。世界上约有65个地区和国家使用这种制式。

NTSC制、PAL制和SECAM制都是与黑白电视兼容制制式,即黑白电视机能接收彩色电视广播,显示的是黑白图像;而彩色电视机也能接收黑白电视广播,显示的也是黑白图像。为了既能实现兼容性而又要有彩色特性,彩色电视系统应满足下列两方面的要求:

- 必须采用与黑白电视相同的一些基本参数,如扫描方式、扫描行频、场频、帧频、同步信号、图像载频、伴音载频等。
- 需要将摄像机输出的三基色信号转换成一个亮度信号,以及代表色度的两个色差信号,并将它们组合成一个彩色全电视信号进行传送。在接收端,彩色电视机将彩色全电视信号重新转换成三个基色信号,在显像管上重现发送端的彩色图像。

表 1-2 彩色电视制式(宽:高=4:3 隔行扫描)

制式	制定国家	制定/广播时间(年)	(有效)扫描线数/帧数(场频)	使用范围
NTSC	美国	1952/1954	525(480)/30(60)	美国、日本、加拿大、韩国、中国台湾
PAL	西德	1962/1967	625(575)/25(50)	西欧(法国除外)、中国内地、中国香港、朝鲜
SECAM	法国	1957/1967		法国、俄国、东欧、中东

(3) 电视扫描

扫描有隔行扫描(interlaced scanning)和逐行扫描(non-interlaced scanning/progressive scanning)之分。图1-1表示了这两种扫描方式的差别。电视发展的初期,由于技术水平不高,数据传输率受到限制。在低数据传输率下,为了防止低扫描频率的画面所产生的闪烁感,黑白电视和彩色电视都采用了隔行扫描方式,通过牺牲扫描密度来换取扫描频率。而现在已经没有了这些限制,所以计算机的CRT显示器一般都采用非隔行扫描。

- 在非隔行扫描中,电子束从显示屏的左上角一行接一行地扫到右下角,在显示屏上扫一遍就显示一幅完整的图像,如图1-1(a)所示。
- 在隔行扫描中,电子束扫完第1行后回到第3行开始的位置接着扫,如图1-1(b)所

示,然后在第 5、7……行上扫,直到最后一行。奇数行扫完后接着扫偶数行,这样就完成了一帧(frame)的扫描。由此可以看到,隔行扫描的一帧图像由两部分组成:一部分是由奇数行组成,称奇数场,另一部分是由偶数行组成,称为偶数场,两场合起来组成一帧。因此在隔行扫描中,无论是摄像机还是显示器,获取或显示一幅图像都要扫描两遍才能得到一幅完整的图像。

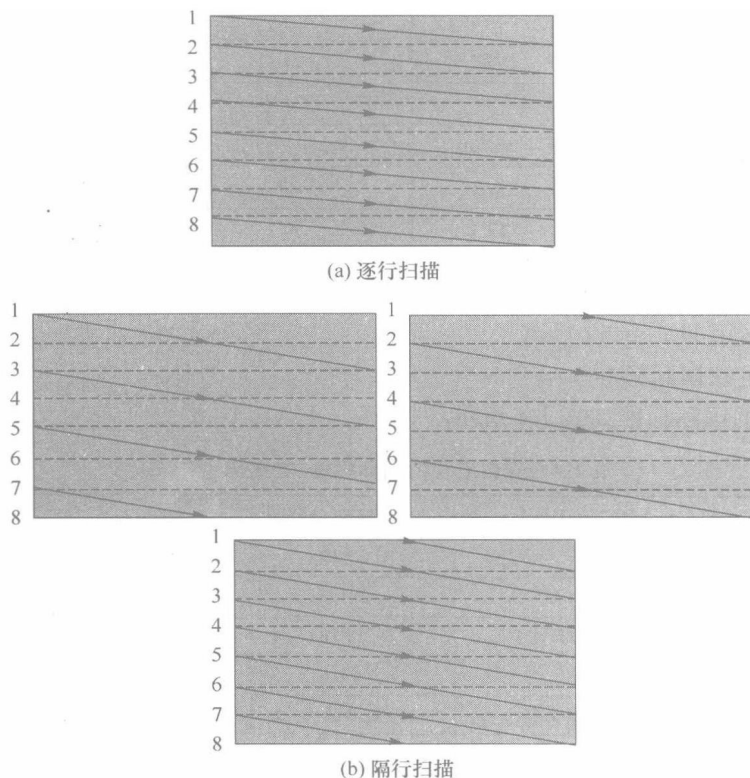


图 1-1 图像的光栅扫描

在隔行扫描中,扫描的行数必须是奇数。如前所述,一帧画面分两场,第一场扫描总行数的一半,第二场扫描总行数的另一半。隔行扫描要求第一场结束于最后一行的一半,不管电子束如何折回,它必须回到显示屏顶部的中央,这样就可以保证相邻的第二场扫描恰好嵌在第一场各扫描线的中间。正是这个原因,才要求总的行数必须是奇数。

每秒钟扫描多少行称为行频 f_H ;每秒钟扫描多少场称为场频 f_f ;每秒扫描多少帧称帧频 f_F 。 f_f 和 f_F 是两个不同的概念。

电视的扫描频率之所以取为 50 场/秒(25 帧/秒)或 60 场/秒(30 帧/秒),一个重要的原因是,受当时技术的限制,电视信号还不能完全避免交流电的干扰,因此才将电视的扫描场频与电源的交变频率取成一致。例如,美日交流电的频率是 60 Hz,所以它们的电视场频也取为 60 Hz(30 帧/秒);而中国和欧洲的交流电频率是 50 Hz,所以我们的电视场频就取为 50 Hz(25 帧/秒)。虽然现在的技术已经有了很大发展,交流电的干扰问题早就获得了解决,但是为了与传统的电视信号兼容,同时也可以避免技术上的复杂性,所以即使是最新的