

现代通信信号处理丛书

通信中的 视频信号处理

|| 郭宝龙 倪伟 闫允一 编著 ||



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>



现代通信信号处理丛书

通信中的视频信号处理

郭宝龙 倪 伟 闫允一 编著

电子工业出版社

Publishing House of Electronics Industry

北京 · BEIJING

内 容 简 介

本书系统地阐述现代通信中视频信号处理的基础理论和关键技术,反映当前视频信号处理技术的最新进展和实用算法。全书介绍数字视频的生成技术和格式类型等基本知识,讨论视频编码的基本原理和主要方法;探讨当前国际上主流的视频编码标准,并对运动估计进行重点分析;论述视频编码的前沿技术和视频通信中的差错控制技术,介绍视频对象分割和视频稳像技术。全书共分为8章,每章前给出中英文摘要,章后附有小结、习题和参考文献。全书系统性强,涉及的技术先进,贴近工程实践。

本书论述层次清晰,结构合理,图文并茂,并配有大量实例算法和源代码。随书光盘内含书中程序源代码和实验图像及视频序列,方便读者学习和使用。本书可作为高等院校相关专业高年级本科生和研究生的教材,也可以供从事视频信号处理、视频通信和图像工程等相关领域的专业人员参考。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有,侵权必究。

图书在版编目(CIP)数据

通信中的视频信号处理 / 郭宝龙等编著. —北京: 电子工业出版社, 2007.5

(现代通信信号处理丛书)

ISBN 978-7-121-03927-0

I. 通… II. 郭… III. 通信系统—视频信号—信号处理 IV. TN919.81

中国版本图书馆CIP数据核字(2007)第026782号

责任编辑: 王春宁

印 刷: 北京市海淀区四季青印刷厂

装 订: 涿州市桃园装订有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路173信箱 邮编 100036

开 本: 787×1092 1/16 印张: 22.5 字数: 595千字

印 次: 2007年5月第1次印刷

印 数: 5000册 定价: 45.00元(含光盘1张)

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010)88254888。

质量投诉请发邮件至 zlt@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线:(010)88258888。

出版说明

随着通信技术的发展,特别是第三代移动通信技术的发展,宽带无线通信网、高速通信网、新一代信息网技术、光通信技术、个人通信技术和智能信息处理技术等已经进入了一个新的高速发展时期,对各种信号处理技术有了更高的要求。为了适应各种现代通信信息传输网络的技术要求,除了继续采用传统的数字信号处理技术外,还应在此基础上提出新的信号处理技术、算法和模型,以满足应用的需求。

随着通信智能化、大数据量、高速实时的多媒体应用需求的不断增多,处理信号的类型已经不仅仅局限于对常规数据的处理,还要处理大量的语音信号和视频信号等。这类信号的特点是数字化、宽频带、大数据量。信号处理技术在通信工程、电子信息工程、电子信息科学与技术、光信息科学与技术、测控技术与仪器、移动通信、无线通信、卫星通信、光通信、网络通信、智能信息系统以及多媒体通信等领域获得广泛的应用,已经成为应用工程的关键技术之一。而目前,国内市场有关通信应用领域信号处理技术系列化的图书种类还非常匮乏,而市场对这方面的需求量又较大。因此,从这个角度出发,我们依托中国通信学会通信理论与信号处理专业委员会,在专业委员会专家、教授的大力支持下,组织出版一套面向 21 世纪的《现代通信信号处理丛书》。

这套丛书从我国现代通信信号处理技术应用现状与发展情况出发,以应用为中心,全面、系统地介绍信号处理领域所涉及的有关关键技术与热点技术,如通信中的智能信号处理、通信中的阵列信号处理、通信中的自适应信号处理、通信中的光信号处理、超宽带技术、通信中的信号检测与估值、量子通信中的量子信号处理、网络信号处理和信号处理算法的实时 DSP 实现等内容。虽然所涉及的领域还不够全面,但我们会一直努力。在确保丛书质量的前提下,不断丰富,不断完善,力求内容的先进性、实用性和系统性,突出理论与应用实践的紧密结合,引导读者将信号处理的原理、技术与应用有机地结合,力争为读者奉献一套可读性与可操作性强的系列丛书。

这套丛书将从 2006 年初陆续推出,主要读者对象是广大从事通信信号处理技术工作的科技研发人员和工程技术人员,也适合高等院校相关学科各专业在校师生及刚刚走上工作岗位的毕业生阅读参考。

在编辑出版这套丛书的过程中,得到了中国通信学会通信理论与信号处理专业委员会的大力支持,参与编著和审定的各位专家都为此付出了大量的心血,可以说,没有他们的支持和帮助,就没有这套丛书的出版,对此,我们表示衷心的感谢。希望广大读者对这套丛书提出宝贵意见和建议,以便今后我们加以改进,为广大读者奉献更多、更好的优秀通信类图书。联系信箱:wchn@phei.com.cn。

电子工业出版社
通信分社

序 言

近 10 年来,随着通信技术、计算机技术和网络技术的不断发展,人们的生活与工作方式随之发生了巨大变化。人们对通信领域的多样化需求越来越多,传统的一维音频信号已经不能满足人们日益增长的需要,人们越来越希望通过网络获得连续的二维图像信号或视频信号。与此同时,目前正蓬勃发展的第三代无线通信系统以提供高速接入和多媒体服务为特征,高速的接入带宽为无线网络多媒体服务提供了物理上的保证,使无线通信与多媒体通信的结合成为可能。未来,互联网与无线通信网的发展将不再是孤立的,通信领域正面临着通信(Telecommunication)、计算(Computing)与互联网(Internet)的结合。可以想象,跨越网络局限,在任何地点、任何时间与任何人自由地传递多媒体信息的梦想即将成为现实。因此,视频信号处理技术将成为影响通信系统性能和通信质量的一项关键技术。

本书系统地讨论了通信中采用的视频信号处理技术和原理,如视频编码技术、运动估计技术、抗误码技术、视频增强技术,并结合具体实例对其中的主流算法和相关标准进行深入分析,内容先进,并贴近工程实践。

本书在编排上注重实用性、简洁性和先进性。在全面介绍各类视频信号处理理论和技术的同时,对视频信号处理中的新技术和新发展给予了充分的关注,书中给出了大量实例分析及源代码,便于读者学习和使用。随书光盘中包含书中所有的源代码、可执行程序以及实验图像和视频序列。结合书中的讲解,读者能够在很快地理解视频信号处理理论和技术的同时,提高自己的编程能力和编程技巧。

全书共分为 8 章。第 1 章给出了视频技术的概述,包括模拟视频和数字视频的产生、转换以及视频格式和颜色空间的相关知识。第 2 章主要针对视频编码中的基本技术,对视频编码的必要性和可行性给出分析,并重点介绍了主要的熵编码方案和变换编码技术。第 3 章针对视频编码中的核心技术——运动估计进行深入探讨,对其各类经典算法和当前发展的新技术进行介绍,这也是第 4 章的基础。第 4 章论述了当前视频编码领域的主流技术,侧重于最新的 H.264/AVC 视频编码标准。第 5 章紧跟目前视频编码的前沿技术,对可伸缩编码、小波视频编码和分布式信源编码进行研究与介绍。第 6 章概述了经典的视频通信抗误码技术,分别是容错编码、差错控制和错误掩盖技术,并对几种视频编码标准中采用的抗误码策略也进行了简单的分析。第 7 章给出了在视频通信中两种实用的信号处理技术,分别是视频对象分割技术和视频稳像技术。第 8 章介绍与分析了视频处理技术在多媒体通信中的具体应用。

本书可以作为高等学校计算机、通信和电子工程等相关专业高年级本科生或研究生课程的教材,也可以作为从事计算机应用、通信系统和视频信号处理等方面工作的科技人员的参考书。本书由郭宝龙统稿,郭宝龙、倪伟撰写了第 1 章~第 5 章,倪伟、闫允一撰写了第 6 章~第 7 章,闫允一撰写了第 8 章。本书在编写过程中参考了大量国内外图像处理和视频编码方面的论文和学术论著,以及西安电子科技大学智能控制与图像工程(ICIE)研究所的论文和学术报告,作者谨向这些文献的作者表示衷心的感谢!本书的出版得到了国家 863 高技术研究发展计划(2006AA01Z127)、国家自然科学基金(60572152)、教育部高等学校博士学科点科研基金(20060701004)的支持,在此表示感谢!最后,感谢电子工业出版社以及本

书责任编辑王春宁老师在本书出版过程中的支持和帮助！

作者衷心希望：本书的出版能够对视频信号处理的发展和多媒体通信技术的应用起到一定的促进作用。由于视频信号处理技术发展迅速，新技术、新算法层出不穷，同时限于作者水平，因此书中难免存在错误和遗漏之处，恳请广大读者批评指正。

编著者

2006年12月于古城西安

目 录

第 1 章 数字视频基础	(1)
1.1 引言	(1)
1.2 视频通信应用	(2)
1.3 数字视频信号	(4)
1.3.1 采样与量化	(4)
1.3.2 颜色空间	(5)
1.3.3 视频格式	(7)
1.3.4 采样格式	(8)
1.4 视频质量评估与方法	(9)
1.4.1 图像质量评价标准	(9)
1.4.2 视频传输质量评估	(11)
小结	(12)
参考文献	(12)
第 2 章 视频编码基本原理与技术	(14)
2.1 视频编码概述	(14)
2.1.1 必要性与可行性	(14)
2.1.1.1 空间冗余	(15)
2.1.1.2 时间冗余	(16)
2.1.1.3 心理视觉冗余	(16)
2.1.1.4 编码冗余	(17)
2.1.2 视频编码的发展与分类	(17)
2.1.3 视频编 / 解码系统结构	(19)
2.2 熵编码	(20)
2.2.1 Huffman 编码	(20)
2.2.1.1 算法简介	(20)
2.2.1.2 编码实例	(22)
2.2.2 游程编码	(26)
2.2.2.1 算法简介	(26)
2.2.2.2 编码实例	(27)
2.2.3 算术编码	(29)
2.2.3.1 算法简介	(29)
2.2.3.2 编码实例	(31)
2.3 变换编码	(37)

2.3.1	DCT 变换	(37)
2.3.1.1	基本原理	(37)
2.3.1.2	快速算法	(38)
2.3.1.3	编码实例	(41)
2.3.2	小波变换	(45)
2.3.2.1	基本原理	(46)
2.3.2.2	小波性能分析	(47)
2.3.2.3	Mallat 算法	(48)
2.3.2.4	二维小波变换原理	(48)
2.3.2.5	小波变换实例	(50)
2.3.3	基于小波变换的图像编码算法	(57)
2.3.3.1	EZW 算法	(58)
2.3.3.2	SPIHT 算法	(60)
2.3.3.3	SPECK 算法	(61)
2.3.3.4	小波图像编码实例	(63)
2.4	量化	(81)
2.4.1	标量量化	(81)
2.4.2	矢量量化	(83)
	小结	(84)
	思考题	(85)
	参考文献	(85)

第 3 章 视频信号运动估计技术 (87)

3.1	运动估计算法简介	(87)
3.1.1	研究现状	(87)
3.1.2	基本原理	(89)
3.1.3	提高搜索效率的主要技术	(89)
3.1.3.1	初始搜索点的选择	(90)
3.1.3.2	块匹配准则	(90)
3.2	经典运动估计算法	(91)
3.2.1	全搜索法	(92)
3.2.2	三步搜索法	(92)
3.2.3	新三步搜索法	(94)
3.2.4	四步搜索法	(95)
3.2.5	基于块的梯度下降搜索法	(96)
3.2.6	菱形搜索法	(97)
3.2.7	六边形搜索法	(99)
3.2.8	运动矢量场自适应搜索算法	(100)
3.2.9	UMHexagonS 算法	(102)
3.2.10	渐进消除算法	(104)

3.3 新型运动估计技术	(105)
3.3.1 亚像素运动估计	(105)
3.3.2 多参考帧运动估计	(116)
3.3.3 可变块运动估计	(117)
3.4 运动估计算法实例	(118)
小结	(129)
思考题	(129)
参考文献	(129)
第4章 视频编码国际标准	(131)
4.1 视频编码标准概述	(131)
4.2 MPEG-2	(133)
4.2.1 档次与等级	(133)
4.2.2 系统层原理	(134)
4.2.3 层次结构	(135)
4.2.4 编 / 解码原理	(138)
4.2.5 关键技术	(139)
4.3 MPEG-4	(141)
4.3.1 档次与等级	(142)
4.3.2 层次结构	(143)
4.3.3 编 / 解码原理	(144)
4.3.3.1 形状信息编码	(145)
4.3.3.2 运动信息编码	(147)
4.3.3.3 纹理信息编码	(148)
4.3.4 编码新技术	(149)
4.3.4.1 视频对象提取技术	(149)
4.3.4.2 可伸缩编码技术	(149)
4.3.4.3 Sprite 对象编码	(150)
4.4 H.263	(151)
4.4.1 编 / 解码原理	(151)
4.4.2 关键技术	(153)
4.4.3 层次结构	(154)
4.4.4 可选模式	(155)
4.4.4.1 H.263	(155)
4.4.4.2 H.263+	(157)
4.4.5 语法元素	(160)
4.5 H.264/AVC	(165)
4.5.1 档次与等级	(165)
4.5.2 编 / 解码原理	(167)
4.5.3 关键技术	(168)

4.5.3.1	整数变换算法	(168)
4.5.3.2	帧内预测技术	(173)
4.5.3.3	高效游程编码	(179)
4.5.3.4	熵编码算法	(179)
4.5.3.5	SP/SI 帧编码	(183)
4.5.3.6	去块滤波器	(187)
4.5.4	语法元素	(188)
4.5.5	H.264 程序实例	(190)
小结	(196)
思考题	(197)
参考文献	(197)
第 5 章	新型视频编码技术	(199)
5.1	可伸缩编码	(199)
5.1.1	基本精细粒度可伸缩视频编码	(201)
5.1.2	渐进精细粒度可伸缩视频编码方案	(204)
5.1.3	运动补偿精细粒度可伸缩视频编码方案	(205)
5.1.4	基于宏块的渐进精细可伸缩性视频编码	(206)
5.1.5	时域可伸缩视频编码方案	(208)
5.2	小波视频编码	(209)
5.2.1	纯三维小波编码方案	(210)
5.2.2	含运动补偿的小波编码方案	(211)
5.2.2.1	基于空域运动补偿的小波视频编码 (MC-DWT)	(211)
5.2.2.2	基于变换域运动补偿的小波视频编码 (DWT-MC)	(212)
5.2.2.3	运动补偿提升视频编码	(213)
5.3	分布式信源编码	(216)
5.3.1	基本原理	(217)
5.3.1.1	分布式无损信源编码	(217)
5.3.1.2	分布式有损信源编码	(218)
5.3.2	分布式视频编码	(219)
5.3.2.1	具有低复杂度编码端的视频编码	(219)
5.3.2.2	分布式信源编码在可伸缩视频编码中的应用	(222)
5.3.2.3	分布式信源编码在视频多描述编码中的应用	(224)
5.3.2.4	分布式信源编码在光场编码中的应用	(225)
小结	(226)
思考题	(227)
参考文献	(227)
第 6 章	视频通信中的抗误码技术	(229)
6.1	概述	(229)

6.2 容错编码	(230)
6.2.1 多描述编码	(230)
6.2.2 信源、信道联合编码	(232)
6.3 传输层差错控制	(234)
6.4 错误掩盖	(235)
6.5 视频编码标准中的抗误码策略	(237)
6.5.1 H.263 的抗误码策略	(237)
6.5.2 MPEG-4 的抗误码策略	(238)
6.5.3 H.264 的抗误码策略	(238)
6.6 错误掩盖实例	(240)
小结	(267)
思考题	(268)
参考文献	(268)
第7章 数字视频专用处理技术	(270)
7.1 视频对象分割技术	(270)
7.1.1 研究背景	(270)
7.1.2 经典视频分割算法	(271)
7.1.2.1 空域分割	(272)
7.1.2.2 时域分割	(276)
7.1.2.3 时-空联合分割	(278)
7.1.3 MPEG-4 中的视频对象分割技术	(279)
7.1.4 视频分割质量评价	(280)
7.2 视频稳像技术	(280)
7.2.1 研究背景	(280)
7.2.2 基本原理	(281)
7.2.3 关键技术	(282)
7.2.3.1 运动估计算法	(282)
7.2.3.2 运动补偿算法	(286)
7.2.3.3 运动参数自适应平滑滤波	(288)
小结	(289)
思考题	(290)
参考文献	(290)
第8章 多媒体通信应用	(292)
8.1 网络通信技术	(292)
8.1.1 通信网络	(292)
8.1.1.1 网络参考模型	(292)
8.1.1.2 网络实例	(296)
8.1.2 网络传输协议	(300)

8.1.2.1	TCP 协议	(300)
8.1.2.2	UDP 协议	(302)
8.1.2.3	RTP 协议	(304)
8.2	多媒体通信终端	(306)
8.2.1	DSP 芯片的特点和结构	(306)
8.2.2	DSP 系统设计的基本过程	(308)
8.2.3	视频编 / 解码的 DSP 实现原理	(309)
8.2.4	H.264 通信终端的 DSP 实现	(311)
8.2.4.1	H.264 编码终端的 DSP 实现——以 TI TMS320DM642 为例	(311)
8.2.4.2	H.264 解码终端的 DSP 实现——以 ADI Blackfin533 为例	(314)
8.2.5	基于专用处理器的图像编 / 解码通信终端	(317)
8.2.5.1	多图像终端的网络互连	(317)
8.2.5.2	视频编 / 解码终端概述	(318)
8.2.5.3	ADV612 基本性能	(320)
8.2.5.4	视频图像编码终端的设计与实现	(322)
8.2.5.5	视频图像解码终端的设计与实现	(324)
8.2.5.6	ADSP2185 同步串口及其属性设置	(325)
8.2.5.7	图像通信终端的数据 I/O 接口设计	(327)
8.3	视频通信系统	(331)
8.3.1	视频会议	(331)
8.3.2	流媒体	(335)
8.3.3	IPTV	(337)
	小结	(340)
	思考题	(340)
	参考文献	(340)
	关键词索引	(342)
	程序索引	(345)

第 1 章 数字视频基础

Chapter I Introduction to Digital Video

本章将重点介绍数字视频的基础知识，为读者奠定后续学习的基础。首先解释数字视频兴起的原因（1.1 节），然后给出数字视频通信的典型应用领域（1.2 节）。随后对图像、视频给出严格的定义，并阐述数字视频的生成方式及其所对应的颜色空间、采样格式等，解释什么是视频信号，视频信号如何获得等一系列问题（1.3 节）。最后在 1.4 节中给出实际应用中
对图像和视频所采用的评估方法，包括主观和客观方法两类。

In this chapter, the basic theory of digital video is emphasized, which establishes the foundation for readers' following study. Firstly, we describe the reason of digital video's emergency (section 1.1) and the typical application domain of digital video communication (section 1.2). In section 1.3, the strict definition of image and video is presented. We also elaborate the digital video's generation and its color model, sampling format, which provide good explanations to what the video signal is and how to get it (section 1.3). Finally, we give the object and subject evaluation methods for image and video used in practical applications.

1.1 引 言

在传统的通信应用中，人们习惯于使用话音和数据传输等单一的媒体业务，如电话、电报、传真等。事实上，视觉信息包含的内容最为丰富，它能够给人们以直观、准确和生动的形象，也是人类获取外界信息最直接和最重要的方式之一。根据统计，人类通过视觉所获得的信息占全部信息的 70% 以上。如何能够实现通信中的视频信息传输一直是人们孜孜追求的目标。随着社会的不断进步，现有的语音和数据通信业务已经不能满足信息时代的通信要求，人们越来越希望无论何时何地都能够方便、快捷、灵活地通过语音、数据、图像等多种方式进行通信。而其中的难点就是如何通过通信网络获得连续的视频信息。近 10 多年来，通信、计算机和互联网技术的不断发展与完善为网络视频服务提供了物理上的保证。流媒体技术的逐渐成熟、IP 网传输带宽的不断提高、ISDN 网络的不断发展、3G 移动通信系统数据通信业务的提供、NGN 以及 4G 网络的研究进展，都预示着视频通信即将迈入一个新的发展时期。未来通信业务发展趋势将是宽带化和多媒体化，从而真正实现人们所期望的“远在天边，近在咫尺”的愿望，而快速高效的视频处理技术正是其中的核心问题。

通信网络为各类信息的通信提供了基本的传输环境，网络带宽、信息交换方式以及高层协议直接决定着传输及服务质量。当前具有代表性的通信网络包括：

- 公众电话交换网 (PSTN)；
- 分组交换远程网 (Packet Switch)；
- 以太网 (Ethernet)；

- 光纤分布式数据接口 (FDDI);
- 综合业务数字网 (ISDN);
- 宽带综合业务数字网 (B-ISDN);
- 异步转移模式 (ATM) 和同步数字序列 (SDH)。

表 1.1 给出了部分网络对视频信息的支持情况。

表 1.1 部分网络对视频的支持情况

通信网络	交换方式	网络特点
电话网	电路	传输速率低, 时延小, 适合语音、低质量图像
LAN	分组	传输速率高, 时延小, 适合数据
DDN	电路	提供 $N \times 64$ kbps, 任意带宽, 半固定连接, 适合于数据、图像
ISDN	电路、分组交换	传输速率高, 时延小, 适合高清晰度图像
B-ISDN	ATM	单向传输, 频带宽, 适合视频传输

1.2 视频通信应用

作为近年来兴起的崭新的应用技术, 视频通信技术大大改变了现代人的生活和工作方式。典型的视频通信的具体应用如下。

可视电话/电视会议: 使用电视或电话在两个或多个地点的用户之间举行会议, 能够实时传送声音和图像信号。除此之外, 用户还可以在终端上获得附加的静止图像、文件、传真等信号。

远程教育: 以计算机和通信系统为基础的教育方式, 教师和学生可以处于不同地点, 使用语言、文字、图像、视频等方式进行学习。

远程医疗: 远距离传送以视频为主的医疗信息, 使专家和医生能远距离对患者进行病情诊断、处理、治疗, 乃至手术和紧急救助。其应用范围可以由国内延伸至国际。

视频点播业务: (Video On Demand, VOD), 根据用户的要求提供交互式视频服务的业务。视频点播具有提供给单个用户对大范围的影片、视频节目、游戏、信息, 以及其他服务进行几乎同时访问的能力。用户和被访问的资料之间高度的交互性使它区别于传统的电视视频节目的接收方式。

视频监控: 在远端实现对现场信息的实时监控, 实现方便、快速、高效的管理。作为安全防范系统的重要组成部分, 视频监控以其直观、准确、及时和信息内容丰富而广泛应用于许多场合, 如教育、交通、金融、政府、商业、工业、农林等。

移动视频业务: 利用现代移动网络, 在个人通信终端上实现如移动接收数字电视节目、移动视频点播、移动可视电话、移动游戏、高速上网的业务。

视频电子邮件: 在 Internet 或移动终端上提供包含高分辨率图像 / 视频的电子邮件。

联合计算机辅助设计: 在公司成员之间利用高速通信网络传输视频、计算数据以及声音信号, 以实现资源共享、联合设计的目标。

数字网络图书馆: 通过按需接入自动文本翻译服务器、文本到语音合成服务器和语音识别服务器, 实时处理数字图书馆服务器捕获的内容。

数字电视: 采用从节目摄制、编辑、制作、发射、传输、接收到节目显示完全数字化的

电视系统，具有清晰度高、音频效果好、抗干扰能力强、占用带宽窄等优点。

图 1.1 和图 1.2 分别给出了远程视频监控和电视会议的网络连接示意图。从图中可以看出，通过终端、通信网络、控制器和服务器的有机结合，能够有效地实现异地视频信息的交互能力。

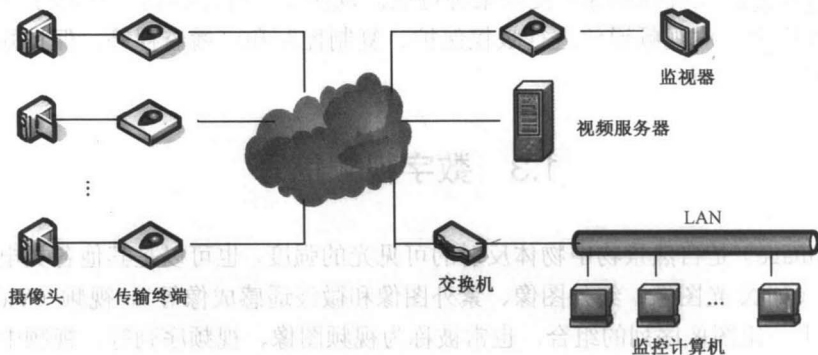


图 1.1 视频监控网络连接示意图

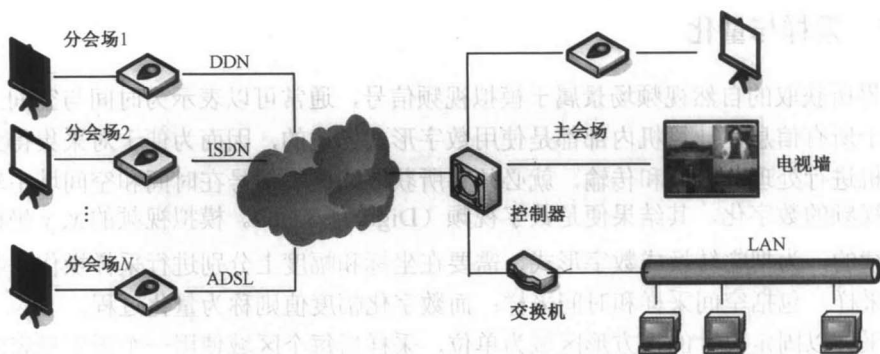


图 1.2 电视会议网络连接示意图

为了在无线移动网络、IP 网络等通信网络上传输视频信号，视频通信业务需要的核心技术是数字视频处理技术。传统的视频处理技术所解决的主要问题是尽可能地提高压缩编码效率，然而视频会议、实时视频广播、远程教育、网络 VOD、实时视频监控、远程医疗以及移动视频通信等多种业务的出现与大规模应用，都对视频编码技术提出了更高、更广的要求，传统的面向存储的视频处理技术已经不能满足实时视频通信的要求。如何实现海量视频数据的实时传输，如何保证视频传输的服务质量，如何面对通信网络上存在的各种丢包、延时、抖动和终端异构性等不利因素，都是当前视频技术中必须解决的问题。在视频通信应用中，视频信号处理的目标从面向存储转到了面向传输，因此面向通信的视频信号处理技术成为视频通信应用中急需解决的关键问题，具有重要的理论与现实意义。

具体来说，通信中的视频信号处理是指对通信系统的具体要求和应用领域对数字视频进行的特定的处理，主要包括如下几个方面。

- **视频编码：**在保证一定图像质量的前提下尽可能地压缩视频图像的数据量。由于视频信号的数据量非常大，因此编码技术也是视频信号处理中最为重要的一环，没有编码技术的突破就不可能有视频通信业务的产生。

- **抗误码**：消除视频信号产生、获取和传输过程中可能带来的失真和干扰，使视频信号尽可能逼真地重构原始景物，如容错编码和错误掩盖。
- **视频增强**：根据主观或客观度量，尽可能地去掉视频图像中的无用信息而突出其主要信息，如视频稳像技术等。
- **视频检索**：从视频图像中提取某种特征，以便对其进行描述、分类与识别。
- **视频安全**：对视频通信进行版权保护、复制控制和广播监视等，保护知识产权不受侵害。

1.3 数字视频信号

图像 (Image) 是自然景物中物体反射的可见光的强度，也可以是其他各类电磁波反射后的强度反映 (如 X 光图像、红外图像、紫外图像和微波遥感成像等)。视频 (Video) 实质上是在时间轴上一组图像序列的组合，也常被称为视频图像、视频序列等。视频中的每幅图像被称为一帧 (Frame)。由于人眼的视觉暂留特性，当以超过每秒 25 帧的速度连续播放静止图像时，在人脑中就形成了连续运动的视频效果。

1.3.1 采样与量化

从外界所获取的自然视频场景属于模拟视频信号，通常可以表示为时间与空间上的连续函数。由于所有信息在计算机内部都是使用数字形式描述的，因而为便于对采集得到的视频使用计算机进行处理、存储和传输，就必须将所获取的模拟信号在时间和空间域中转换为数字量，即视频的数字化，其结果便是数字视频 (Digital Video)。模拟视频的 x, y 坐标及幅度值都是连续的，为把它转换成数字形式，需要在坐标和幅度上分别进行采样操作。数字化坐标值称为采样，包括空间采样和时间采样；而数字化幅度值则称为量化过程。

空间采样以固定尺寸的正方形区域为单位，采样后每个区域使用一个固定量表示。图像质量直接受到采样单元尺寸的影响：采样单元尺寸越小，图像分辨率越高，质量也越好。除了在空间域内进行采样，由于一个自然场景在时间上也是连续的，也需要在时间轴上以固定的间隔对模拟视频信号进行采样，以生成不同的帧 (Frame)。为了保证视频的连续性，一般采样时间需要小于 $1/20$ s。时间域采样频率越高，视频也就越平滑，但也会使得视频数字化后的数据量成倍增加。

在模拟视频信号采样后得到在空间上和时间上离散的视频信号。但是要实现视频的完全数字化，还必须将采集到的视频信号每一帧的幅值都转化为能使用有限位数表示的数值，即量化过程。量化所完成的功能就是按照一定的规则对连续采样值做近似表示，使得量化后输出的幅值为有限个比特。量化输出的整数称为量化级，量化总是将一个范围内的输入值量化为同一个输出级。所以量化必然会造成信息的损失，是一个不可逆的过程。一个常用的时空域采样和均匀量化的例子如图 1.3 所示。图 1.3 (b) 中的一维函数是某图像水平方向扫描得到的连续图像幅度值的曲线，沿线段等间隔采样，采样位置由图 1.3 (b) 中底部垂直线段给出，采样结果如图 1.3 (b) 中圆点所示，若干个圆点的离散位置就给出了采样函数值。图 1.3 (b) 中使用了 8 个离散的量化级，因此所有采样的连续灰度值都被简单量化为 8 个灰度级中的一个。从图像顶部开始逐行执行这一过程便能够产生一幅二维数字图像，对视频中的所有

帧执行便可以最终得到完全数字化的视频序列。

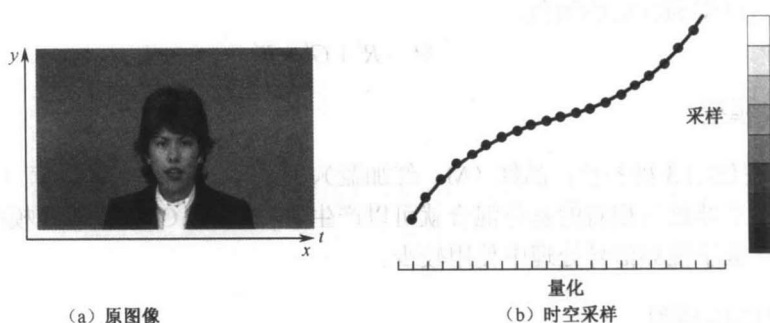


图 1.3 数字视频的时空采样

数字视频一般都表示为一个三维信号 $f(x, y, t)$ 的形式, 给定的 t 定义了某个时刻的一帧图像, 而 x 和 y 表示视频帧中的行和列, 标识图像帧中点的空间位置, 在该处的元素值也就是相应的灰度值, 这个点一般称其为图像元素或像素 (Pixel)。对于彩色图像而言, 要根据需要使用不同的颜色空间加以表示。通常在讨论视频信号处理的时候, 我们更多地是以帧或图像的形式进行, 直接将其表示为如下形式:

$$f(x, y) \quad x=0,1,\dots,M, y=0,1,\dots,N \quad (1-1)$$

这里采样、量化后产生的数字图像为 M 行 N 列, 坐标 (x, y) 变成离散量。对于整幅图像而言, 采样和量化的最终结果可以表示为一个二维矩阵, 则完整的 $M \times N$ 数字图像为:

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & \dots & f(1,N-1) \\ \vdots & \vdots & \ddots & \vdots \\ f(M-1,0) & f(M-1,1) & \dots & f(M-1,N-1) \end{bmatrix} \quad (1-2)$$

数字化过程中, 假设离散灰度级是等间隔的, 并且是区间 $[0, L-1]$ 内的整数, 离散灰度级 L 定义为 $L=2^k$, 存储数字图像所需比特数 $B=M \times N \times k$ 。当一幅图像有 2^k 灰度级时, 通常称该图像是 k bit 图像。例如一幅灰度图像通常使用 256 个灰度级表示, 也称其为 8 bit 图像。

1.3.2 颜色空间

灰度图像仅有空间采样点的亮度信息, 因而只需要采用一个数值即可表示。而对于彩色图像来说, 其表示要更为复杂。由光谱学可知, 白光实质上是不同频率的电磁波混合而成的, 不同频率的电磁波在人眼中呈现出不同的颜色, 混合起来即为白色。如果物体对某些光谱反射较多, 则该物体就呈现出相应的颜色。根据人眼的生理结构, 人们构造了不同的图像颜色模型或颜色空间来表示客观图像的亮度和颜色, 任何一种颜色都可以分解为颜色空间中的一个或多个变量。在视频信号的处理中最为常用的颜色空间是 RGB、YUV 和 $YCbCr$ 。

1. RGB 模型

RGB 空间将所有颜色都表示为 3 个基本颜色 R (红色)、G (绿色)、B (蓝色) 不同强度的组合。在记录及显示彩色图像时, RGB 是最常见的一种方案, 如在彩色阴极射线管显示器 (CRT) 和液晶显示屏 (LCD) 中, 就是根据 RGB 每个分量的强度分别显示像素的红色、